Masaryk University

Faculty of Science

HABILITATION THESIS

Nucleic acid interactions and genome stability

Ctirad Hofr

Brno 2014

Acknowledgements

On the first place, I thank to my supervisors and mentors who contributed to my scientific education. In particular, my thanks go to Prof. Jiří Fajkus, who supported me all the time and helped me with all most important decisions, with his very friendly family-like approach and great sense of humor. I appreciate the contribution of my PhD supervisor Prof. Viktor Brabec, and Assoc. Prof. Olga Nováková to my understanding of what is the most important part of scientific work. I am indebted to hosts of my long-term research stays Prof. Kenneth J. Breslauer, who showed me the significance of positive attitude in research and Prof. Thomas R. Cech, who fully fills the quotation by Mark Twain: "*Really great people make you feel that you, too, can become great.*" I thank all my colleagues and students for their inspiring thoughts, great effort during our common work and for creative and friendly environment that they created.

Table of Contents

Preface 5

Introduction 6

1. Effects of heavy metal drugs on DNA thermodynamics and structure 7

1.1. Structure of platinum-based anticancer drugs 7

1.2. Mechanism of action of platinum drugs 8

1.3. Formation of cisplatin adducts on DNA 9

1.4. Structural changes of DNA induced by cisplatin 10

1.5. Effect of cisplatin adducts DNA thermodynamic stability 11

1.6. Recognition of platinum-DNA adducts by proteins 14

1.7. Strategies of new platinum drugs development 15

1.8. DNA interactions with new ruthenium compounds 18

1.9. Future strategies of anticancer treatment using metal-based drugs 20

1.9.1. Novel platinum drugs overcoming resistance to cisplatin 21

1.9.2. Improved delivery of platinum drugs to tumors 22

1.9.3. Platinum resistance modulation 23

1.9.4. Combination of platinum drugs with new therapies targeting resistance mechanism 23

1.10. Significance and impact I 24

1.11. Publications related to above topics I 26

2. Genome safe guards: Telomeric proteins recognizing and protecting chromosome ends 27

2.1. Composition of human shelterin 27

2.2. Functions of shelterin in mammals 28

2.3. The mechanism of human Rap1 contribution to selective recognition of telomeric DNA by shelterin 30

2.4. Model organisms in telomere maintenance research 31

2.5. SMH proteins recognize telomeric DNA in plants 32

2.6. Yeast telomeric proteins: is it better to have two DNA binding domains? 34

2.7. Significance and impact II 36

2.8. Publications related to above topics II 37

3. Mechanism of RNA recognition in DNA transcription: An effective combination of quantitative interaction analysis with structural data 38

3.1. Structure of proteins recognizing RNA 39
3.2. Mechanism of RNA recognizing protein ADAR via RNA-binding domains: Double-stranded RNA recognition based on shape or sequence? 41
3.3. Significance and impact III 43
3.4. Publications related to above topics III 43
4. Future prospects 44
5. References 46
Summary 54
Abstrakt 55
List of publications included in this habilitation work 56

Preface

This habilitation work is a compilation of scientific publications to which I have contributed as the first author, corresponding author, and co-author. The articles were published between 2000 and 2014. All these publications have a common theme related to quantitative description of nucleic acid interactions in relation to the stability of the genome. The habilitation thesis text is divided into three main parts:

1. The first part is dedicated to the description of changes in structure and thermodynamics of DNA induced by the binding of new and clinically used anticancer drugs derived from metal compounds.

2. The second part describes interaction of telomeric proteins that selectively bind DNA on chromosomes ends. This part is focused on understanding the mechanism of selective recognition of telomeric DNA by proteins. The characterization of binding of telomeric proteins from different species (human, yeast, plant) is presented.

3. In the final part, it is demonstrated how functional binding studies and molecular structure data are combined to obtain a more comprehensive view of mechanism of double-stranded RNA recognition.

A comprehensive information on the studied nucleic acid interactions can be found in the enclosed original publications.

The accompanying text highlights the author's contribution to the field of quantitative description of nucleic acids and protein interactions and also contains a brief introduction to the studied topic. At the end of each part, publications, where I am either the first, corresponding or collaborative author, relevant to the research area are presented together with basic principles of the methodical approaches that have been used for investigation of the presented scientific objectives. To make the text more enjoyable for reader, the text is equipped with external links for publications and interactive materials that could help to make a well-rounded view on studied research areas.

Introduction

Nucleic acids have been in the center of scientific interest since their discoveries. The interest has been intensified after the disclosure of DNA structure and genetic code conservation. The studies of DNA were motivated by a possibility to cure diseases by modification of DNA structure and sequence. Gradually, proteins that modulate genetic information transfer were disclosed. The proteins that interact with nucleic acids directly have been proven to be essential for regulation of nucleic acid functions. The desire to understand biological mechanisms of how DNA transcription and translation are regulated led to experimental effort that brought discovery of new protein groups that have sometimes surprising relations among DNA binding sites and their functional impacts and phenotypes. The biological studies that described basic rules governing the transfer of genetic information were accompanied by studies focused on actual molecular mechanisms that take part in genome conservation and DNA replication. The detailed studies of interaction network among biomolecules were carried out using biophysical and biochemical approaches. The advantage of approaches based on determination of physical properties is that we obtain "a number". Then the obtained numerical values could be used easily when comparing different molecular partners and when identifying the possible order of interaction cascades. Thus, quantitative biophysical methods enable a detailed description of biomolecular processes that take place during formation of highly ordered nucleoprotein complexes. Additionally, quantitative approaches shed more light into protein-mediated regulation mechanisms of essential functions of nucleic acids.

1. Effects of heavy metal drugs on DNA thermodynamics and structure

1.1. Structure of platinum-based anticancer drugs

Cisplatin is a platinum-based drug which is widely used in cancer chemotherapy. Overall cure rate of cisplatin exceeds 90% for testicular cancer (Bosi 2011). Cisplatin and carboplatin (Figure 1) have been in widespread use for many years to treat several other forms of cancer, including ovarian, cervical, head and neck, and non-

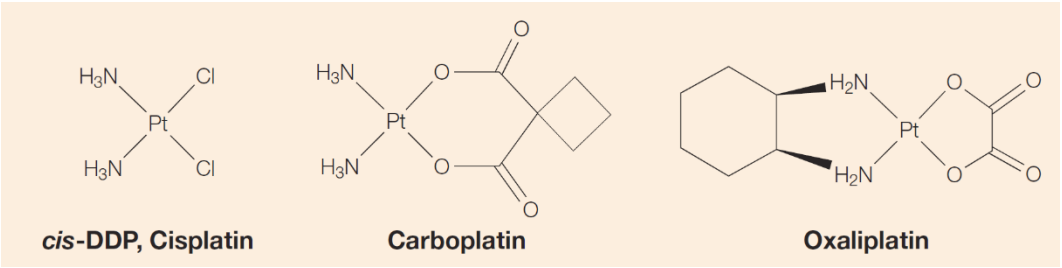


Figure 1. Platinum-based anticancer drugs. Cisplatin, *cis*-diamminedichloroplatinum(II); carboplatin, *cis*-diammine(cyclobutane-1,1-dicarboxylato)- platinum(II); oxaliplatin, 1,2-diamino-cyclohexaneoxalato platinum (II) (taken from (Wang and Lippard 2005)).

small-cell lung cancer (Twelves et al. 2006). The basic chemical structure of all clinically used platinum drugs is the same; amine carrier ligands with chlorido, dicarboxylato or oxalate leaving groups.

The treatment by cisplatin and from cisplatin derived drugs is limited, by side effects including nephrotoxicity, emetogenesis and neurotoxicity. Nephrotoxicity and emetogenesis can be substantially managed by hydration combined with the use of diuretics (2011). A significant dose-limiting toxic effect of cisplatin is still neurotoxicity, which can result in high-frequency hearing loss. In addition, both inherent and acquired resistance to the drugs limit their applications. Thousands of platinum analogues have been synthesized and screened for anticancer activity in an attempt to overcome these limitations and to broaden the range of treatable tumors. A very good overview of platinum anticancer drugs in the clinic and clinical trials has been published recently by Wheate et al. (Wheate et al. 2010). In order to develop new therapeutic strategies and the rational design of more efficient platinum-based drugs the understanding of the cellular processing of cisplatin was urgently needed.

1.2. Mechanism of action of platinum drugs

Globally viewed, cisplatin activates cellular pathways involved in regulating drug uptake, the signaling of DNA damage, cell-cycle checkpoints and arrest, DNA repair and cell death (Wang and Lippard 2005). However, if we look at cisplatin action from molecular point of view, we have to take into consideration local environment. In the blood stream where the chloride concentration is relatively high (100 mM) the chloride ligands stay attached to the drug although binding to serum proteins, such as human serum albumin, does occur (Fig. 1) (Kelland 2007). When it reaches the tumor, cisplatin is thought to be taken up into the cells by three possible

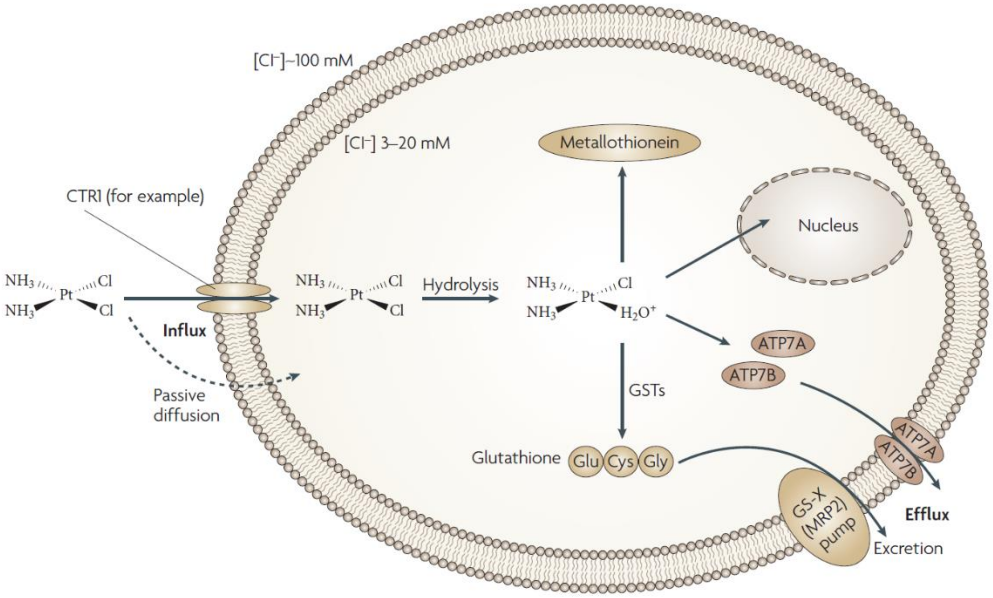


Figure 2. Mechanisms of cisplatin uptake and efflux. Platinum enters cells using either a copper transporter (mainly CTR1) or by passive diffusion. Once inside cells, cisplatin is activated by the addition of water molecules to form a chemically reactive aqua species. This is facilitated by the relatively low chloride concentrations that are found within cells. In the cytoplasm, the activated aqua species preferentially reacts with species containing high sulphur levels by virtue of their containing many cysteine or methionine amino acids. These species include the tripeptide glutathione or metallothioneins. In some platinum-resistant cancer cells, glutathione and metallothionein levels are relatively high, so activated platinum is effectively 'mopped up' in the cytoplasm before DNA binding can occur, thereby causing resistance. Finally, active export of platinum from the cells through the copper exporters ATP7A and ATP7B as well as through the glutathione S-conjugate export GS-X pump (also known as MRP2 or ABCC2) can contribute to platinum drug resistance. GSTs, glutathione S-transferases (Kelland 2007).

mechanisms: passive diffusion, copper transporter proteins (e.g. CTR1) and organic cation transporters. Once inside the cell, the lower chloride concentration (~15 mM) results in drug aquation with the loss of one or both of the chloride ligands. When aquated, cisplatin can go on to bind to its target, DNA.

1.3. Formation of cisplatin adducts on DNA

Cisplatin molecules bind DNA at the N7 position of guanine, and to a lesser extent adenine, through the formation of a covalent coordinate bond with the lone pair of the nitrogen atom. Cisplatin forms either monofunctional (via one covalent bond) or bifunctional adducts (via two covalent bonds) with DNA (Fichtinger-Schepman et al. 1985). There is continued debate as to which of the various platinum–DNA adducts might be more biologically significant. Most adducts occur on the same DNA strand and involve bases adjacent to one another, and are therefore known as intrastrand crosslinks, namely 1,2-d(GpG) intrastrand (60–65% of all adducts) and 1,2-d(ApG) intrastrand (20–25%). Other less frequently produced same-strand adducts are the 1,3-d(GpXpG) intrastrand crosslink (where there is another base in between the two platinated guanines; approximately 2%) and monofunctional adducts on guanines (approximately 2%). In addition, around 8 % of adducts link guanines on opposite DNA strands, so-called G–G interstrand crosslinks. In all cases, the two ammine groups remain bound to platinum.

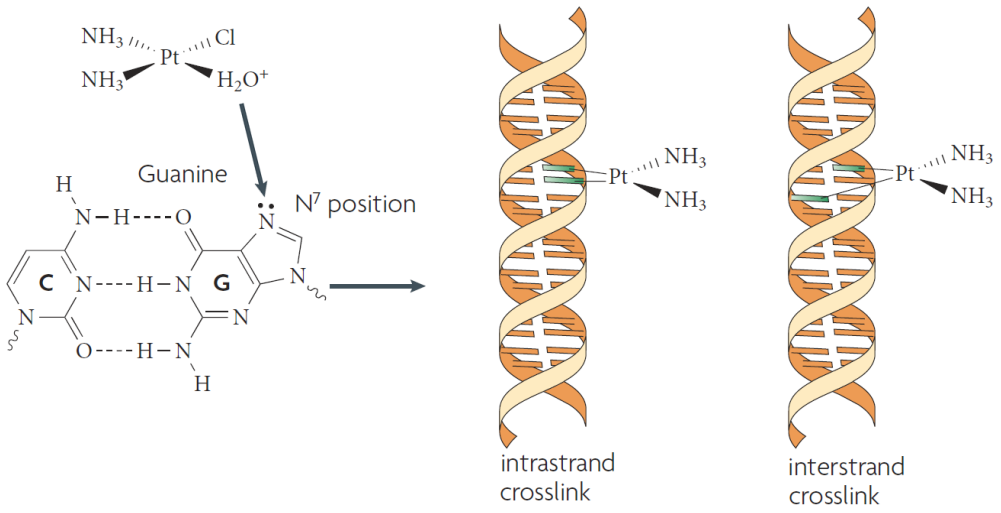


Figure 3. Formation of cisplatin adducts on DNA. Once the activated aqua platinum species (see Fig. 1 and note that this is the same for cisplatin, carboplatin and oxalipatin) has entered the nucleus, preferential covalent binding to the nitrogen on position 7 of guanine occurs. The major covalent bis-adduct that is formed involves adjacent guanines on the same strand of DNA (the intrastrand crosslink); a minor adduct involves binding to guanines on opposite DNA strands (the interstrand crosslink); figure was taken and modified from (Kelland 2007).

1.4. Structural changes of DNA induced by cisplatin

Cisplatin and related drugs bound to DNA have been thoroughly studied by both X-ray crystallography and NMR spectroscopy, yielding plentiful information about the structural effect of DNA modification platinum drugs. Cisplatin adducts cause distortions in DNA, including unwinding and bending. The major cisplatin adduct 1,2-d(GpG) intrastrand crosslink bends DNA helix significantly towards the major groove, exposing a wide, shallow minor groove surface to which several classes of proteins bind. The unwinding angle observed for the 1,2-d(GpG) intrastrand crosslink was 20 degree (Takahara et al. 1995; Wu et al. 2007). The bend angle of DNA after showed by NMR studies and X-ray structure differs, however it falls to interval 20-40 degree (Takahara et al. 1995; Wu et al. 2007). Examination of NMR structures of duplex containing 1,2 GpG-intrastrand crosslink revealed significant distortion of the base pair step to the 5´side of the adduct (Figure 1, A).

The effects of 1,2-d(GpG) intrastrand crosslink of oxaliplatin on DNA structure (Figure

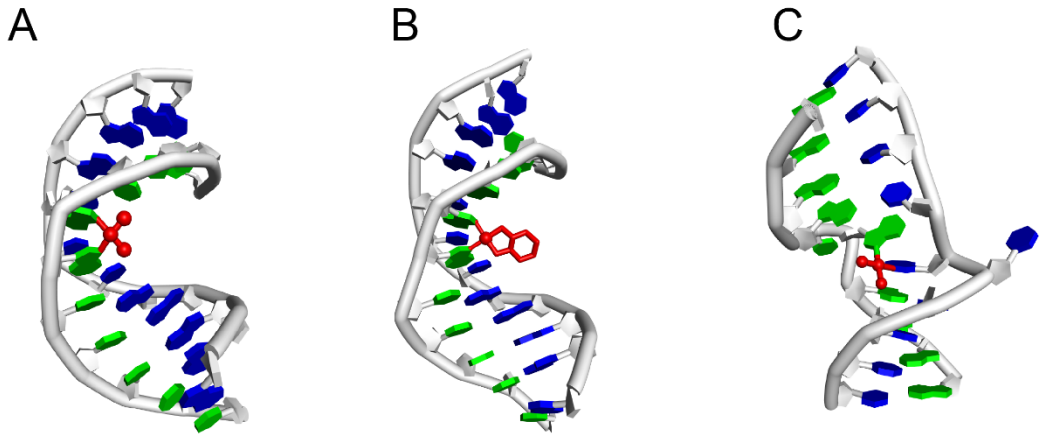


Figure 4. Structures of double stranded DNA containing adducts of cisplatin and oxaliplatin.(A) Cisplatin 1,2 –d(GpG) intrastrand cross-link (1AIO), (B) Cisplatin interstrand cross-link (1A2E), (C) Oxaliplatin 1,2 -d(GpG) intrastrand cross-link (1PG9). PDB accession codes are given in parentheses.

1, B) differs only subtle from the effect of the same adduct of cisplatin, however the bulkier platinum adduct could be important for greater pronouncement of the structural irregularities of the DNA in cells (Wu et al. 2007).

The interstrand crosslink of cisplatin (Figure 1, C) causes DNA to bend toward the minor groove by 20 degree. The structural analysis of the interastrand crosslink of

cisplatin revealed that the cross-linked guanosine residues are not paired with hydrogen bonds to the complementary deoxyribocytidines, which are located outside the duplex and not stacked with other aromatic rings (Huang et al. 1995). On the contrary, the interstrand crosslink of therapeutically ineffective transplatin, formed between complementary guanine and cytosine residues, does not significantly alter the DNA structure (Paquet et al. 1999).

The effect of platinum adducts on DNA structure are well reviewed by Todd, Lippard and Jung (Jung and Lippard 2007; Todd 2010). It has been assumed that the structural changes induced by platinum–DNA adducts can cause instability of double-helix structure that mediate the cytotoxicity of the platinum drugs.

1.5. Effect of cisplatin adducts DNA thermodynamic stability

It has been postulated that the antitumor properties of cisplatin are mediated by sensing the irregularity of DNA helix by proteins involved in DNA damage response (Ohndorf et al. 1999). It has been also shown that the protein recognition of adducts formed on DNA by platinum complexes is dependent on the extent of thermal or thermodynamic destabilization imposed on the duplex by the adduct (Pilch et al. 2000). In addition, the thermal stability of DNA modified by various platinum compounds, which differ in their antitumor effects, has been also studied (Cohen et al. 1980; Ishida et al. 2002; Holzer et al. 2004). These studies have revealed that the important factors influencing the thermal stability of platinated DNA are also interstrand cross-links, which contribute to the global stabilization of DNA.

Consistently with the suggested effect on cisplatin adducts on DNA, initial biophysical experiments showed that cisplatin binding results in a loss of helix stability, as demonstrated by initial calorimetric studies (Maeda et al. 1990). Further calorimetric experiments with site specific cisplatin DNA adducts revealed a duplex destabilization of 6.3 kcal/mol associated with 1,2-d(GpG) adduct formation (Poklar et al. 1996). The thermodynamic destabilization is mainly enthalpic in origin. The enthalpic nature of the destabilization is directly connected with disturbation of DNA base pairing and

stacking interaction induced by formation of cisplatin adduct. Later, it has been shown that the destabilization extent depends on the sequence surrounding cisplatin adduct (Pilch et al. 2000). The presence of flanking A.T base-pairs, relative to T.A or C.G base-pairs, enhances the extent of cross-link-induced alteration to an A-like conformation and dampens the extent of cross-link-induced duplex destabilization.

After thermodynamic characterization of 1,2-d(GpG) intrastrand adduct of cisplatin the research focus was directed to other ciplatin adducts. The most attention was attracted by interstrand adducts of cisplatin, which links two guanine residues in complementary DNA strands. For a long time, Laboratory of Molecular Biophysics and Pharmacology at the Institute of Biophysics has been focused on biophysical subsequences of incorporating of cisplatin adducts, mainly interstrand crosslinks, to DNA. The frequent interstrand crosslink of cisplatin is formed between two guanines in complementary dimeric sequence 5' GC 3'-GC (Figure 4C).

Interestingly, clinically ineffective *trans* isomer of cisplatin, called transplatin, forms DNA interstrand crosslinks on DNA as well (Malinge and Leng 2006). Interstrand crosslinks of transplatin (12%) occur more frequently than the same crosslinks of cisplatin (8%) (Brabec and Leng 1993). As has been shown earlier, transplatin forms intrastrand cross-links between complementary guanine and cytosine nucleotides preferentially (Eastman et al. 1988).

In order to describe the difference in thermodynamic effects of interstrand crosslink of cisplatin and transplatin, the adducts were formed on short DNA duplex. The energetics and thermal stability effect of the single, site-specific interstrand crosslink formed by cisplatin or transplatin were investigated by differential scanning calorimetry, temperature-dependent UV, and CD spectroscopy (Hofr and Brabec 2001). The cross-link of both platinum isomers increased the thermal stability of the modified duplexes by changing the molecularity of denaturation. The covalent platinum link between complementary DNA strands decreased number of hybridizing strands from two to one from entropic point of view. Thus, the structural perturbation resulting from the interstrand cross-link of cisplatin increases entropy of the duplex and in this way entropically stabilizes the duplex. This entropic cross-link-induced stabilization of the duplex is partially but not completely compensated by

the enthalpic destabilization of the duplex, which was caused by disruption of base pairing and stacking interactions around cisplatin adduct (Huang et al. 1995). The structural analysis of 1,2-d(GpG) interastrand crosslink of cisplatin revealed that the cross-linked deoxyriboguanosine residues are not paired with hydrogen bonds to the complementary deoxyribocytidines, which are located outside the duplex and not stacked with other aromatic rings (Huang et al. 1995). All other base residues are paired, but distortion extends over at least four base pairs at the site of the cross-link (Coste et al. 1999). The disruption of interactions that mediate complementary hybridization of DNA strands is the main source of entropic destabilization. The net result of these enthalpic and entropic effects is that the structural perturbation resulting from the formation of the interstrand cross-link by cisplatin induces a decrease in duplex thermodynamic stability, with this destabilization being enthalpic in origin.

Due to the *trans* orientation of reacting bonds of transplatin the interstrand cross-link of transplatin induces only subtle structural changes (Paquet et al. 1999). Hence, the effect of transplatin linking covalently complementary DNA strands on enthalpic stability of the double-helix should be delicate. Indeed, when we performed calorimetric measurements of 15-mer DNA duplex site specifically interstrand crosslinked by transplatin we observe that the effect of transplatin adduct on DNA stability was enthalpicaly neutral; with no significant marks of energetic destabilization (Hofr and Brabec 2005).

The differences in thermodynamic changes of DNA induced by cisplatin and transplatin are consistent with distinct conformational distortions induced by the interstrand cross-links of the two isomers. Importantly, for the duplex cross-linked by cisplatin relative to that cross-linked by transplatin, the compensating enthalpic and entropic effects almost completely offset the difference in cross-link-induced energetic destabilization. It has been proposed that the results of the present work further support the view that the impact of the interstrand cross-links of cisplatin and transplatin on DNA is different for each adduct and might also be associated with the distinctly different antitumor effects of these platinum compounds. In accordance

with the original hypothesis, the thermodynamic destabilization of DNA by cisplatin adduct might regulate recognition and specific binding of proteins to platinated DNA.

1.6. Recognition of platinum-DNA adducts by proteins

Several proteins have been identified to bind platinum-DNA adducts with selectivity over unmodified DNA. Proteins recognizing platinum-DNA adducts include High Mobility Group proteins, proteins associated with DNA repair, and transcription factors (Zhang et al. 2004). Mainly, we are going to focus on proteins that participate in eukaryotic transcription, as the anticancer effect of platinum drugs is closely connected with transcription inhibition and later genome destabilization.

Transcription factors that bind platinum-DNA adducts include human upstream

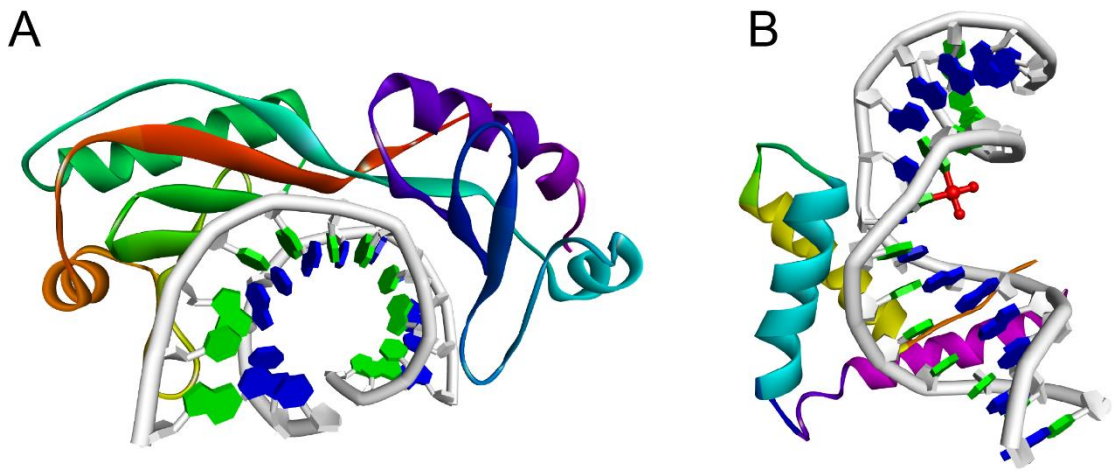


Figure 5. Protein recognition and binding to platinum-DNA adducts. (A) X-ray crystal structure of TBP-bound DNA (1TGH). Note that the DNA is shaped and bent similarly to the DNA containing 1,2-d(GpG) intrastrand crosslink of cisplatin. (B) X-ray crystal structure of HMGB1 domain A bound to a cisplatin 1,2-d(GpG) intrastrand crosslink. PDB accession codes are given in parentheses.

binding factor (hUBF), TATA-binding protein (TBP) and Y-box binding protein (YB-1). High Mobility Group box protein 1 (HMGB1) is an abundant non-histone chromosomal protein that binds cisplatin-DNA adducts tightly with significant selectivity. HMGB1 plays a role in cisplatin mechanism of action in a variety of ways. HMGB1 has been implicated to prevent recognition of cisplatin adducts by sensitive reparation mechanisms (Jung and Lippard 2007). The observed synergetic function of HMGB1 in cisplatin cytotoxicity has focused research on mechanism of binding and recognition of platinated DNA by HMGB1 protein. HMGB1 contains two tandem domains, A and B, with different binding affinity to DNA. The full length protein binds

cisplatin 1,2-d(GpG) intrastrand crosslink primarily through A domain with a dissociation constant of 120 nM (Ohga et al. 1996). HMGB1 also recognizes the interstrand crosslink of cisplatin with 5-fold lower affinity (Kasparkova et al. 2003). The major cisplatin adduct 1,2-d(GpG) intrastrand crosslink bends DNA helix significantly towards the major groove, exposing a wide, shallow minor groove surface to which several classes of proteins bind. Structural analysis of HMGB1 bound to the intrastrand crosslink of cisplatin showed that HMGB1 binds to widened minor groove on 3' side of platinated strand (Figure 5B). A phenylalanine residue intercalates into a hydrophobic notch created by cisplatin crosslink and stabilizes protein-DNA complex (Ohndorf et al. 1999).

Another HMG-domain containing protein is Structure specific recognition protein 1 (SSRP1), which is a subunit of FACT (facilitates chromatin transcription) protein complex. FACT is essential chromatin remodeling factor that is involved in transcription of nucleosomal DNA. PARP-1 is a multi-functional protein with many roles in cells. PARP-1 controls transcriptional levels (Kraus and Lis 2003), binds selectively to cisplatin-DNA crosslinks and is activated when cells are treated by platinum drugs (Aguilar-Quesada et al. 2007).

If we look at the affinity of proteins recognizing platinum adducts from quantitative point of view, we find that the highest binding affinity to platinum adduct so far has been detected for human ribosomal RNA transcription factor hUBF that binds 1,2 GpG intrastrand crosslink of cisplatin with Kd 60 pM. TATA-binding protein binds the same cisplatin adduct with Kd being around 1 nM. The observed high affinities suggest that the platinum adducts are recognized by DNA binding proteins early after their formations, and could initiate DNA reparation cascades that contribute to the cell resistance towards cisplatin treatment.

1.7. Strategies of new platinum drugs development

After the initial promising clinical trial data with cisplatin, attention was aimed at determining how tumor resistance was acquired during platinum drug therapy, and why some tumors were intrinsically resistant. Conversely, other studies investigated causes of the hypersensitivity of testicular cancer to cisplatin (Kelland et al. 1992).

The experimental observations made in cell lines demonstrate that resistance might be mediated through two main mechanisms: first, a failure of a sufficient amount of platinum to reach the target DNA; and, second, a failure to achieve cell death after platinum–DNA adduct formation. We now have answer why testicular tumors are extremely sensitive for cisplatin. The hypersensitivity relates directly to a reduced DNA-repair capacity in response to platinum–DNA adducts. The studies reveal that testicular cancer cells often possess a low DNA-repair capability, and so, upon exposure to cisplatin, will undergo relatively more apoptosis than cells of other cancers (Koberle et al. 1999).

In order to increase sensitivity of cancer cells to drug treatment, there has been a tremendous effort spent on developing novel platinum drugs to overcome drug resistance or increase the range of drug activity by substitution of the ammine ligands or chloride leaving groups of cisplatin.
Another strategy was to change the central atom of Platinum for another heavy metal. This way the complexes of ruthenium and other metals have been synthetized. The complexes have been thoroughly studied and compared with cisplatin binding modes to DNA and cisplatin spectrum of adducts (Chapter 1.8). We have joined the effort for finding new more effective and specific candidates of anticancer drugs.

In collaboration with Chemical Department of University in Bari (Italy), we analyzed thermal and thermodynamic properties of newly synthetized platinum drugs derived from cisplatin. Analogs of antitumor cisplatin containing enantiomeric amine ligands, such as cis-[PtCl(2)(RR-DAB)] and cis-[PtCl(2)(SS-DAB)] (DAB = 2,3-diaminobutane), were studied by various method of molecular biophysics (Malina et al. 2000). The major differences resulting from the modification of DNA by the two enantiomers was the thermodynamic destabilization and conformational distortions induced in DNA by the 1,2-d(GpG) intrastrand cross-link (Malina et al. 2000). We suggested that these differences are associated with a different biological activity of the two enantiomers observed previously (Fanizzi et al. 1987). In addition, the results of the present work are also consistent with the view that formation of hydrogen bonds between the carbonyl oxygen of the guanine residues and the "quasi equatorial"

hydrogen of the *cis* amine in the 1,2-d(GpG) intrastrand cross-link plays an important role in determining the character of the distortion induced in DNA by this lesion.

In a search for new platinum anticancer drugs there were several approaches applied. One of them was driven by the hypothesis that platinum compounds which bind to DNA in a fundamentally different manner to that of cisplatin will have altered biological properties, including the spectrum and intensity of antitumor activity. In agreement with this argument, a new class of anticancer platinum compounds has been designed based on polynuclear platinum complexes by collaborating laboratory of Prof. Farrell from Department of Chemistry, Virginia Commonwealth University, USA (Farrell 2000).

A typical example of one of these compounds is the bifunctional dinuclear agent [{*trans*-$PtCl(NH_3)_2$}$_2$$H_2N(CH_2)_4NH_2$]$Cl_2$ (1,1/t,t). It forms various types of DNA crosslinks capable of terminating DNA replication (Zaludova et al. 1997). Interestingly, it also forms minor 1,2-d(GpG) intrastrand adducts (Kašpárková et al. 1996) which are major DNA lesions of cisplatin (Fichtinger-Schepman et al. 1985).

In our study, we employed differential scanning calorimetry and spectroscopic techniques to characterize the influence of this crosslink on the thermal stability and energetics of 20 bp DNA duplexes site-specifically modified by 1,1/t,t (Hofr et al. 2001). Thermal denaturation data revealed that the crosslink of 1,1/t,t reduced thermal and thermodynamic stability of the duplex noticeably more than that of 'classical' cisplatin. Our thermodynamic data revealed that a 1,2-d(GpG) intrastrand crosslink of 1,1/t,t enthalpically destabilizes the host DNA duplex three-fold more than the same adduct of cisplatin (Hofr et al. 2001). The different effect of 1,1-t,t and cisplatin crosslinks on DNA melting enthalpy could be explained by diverse influence of the adducts on conformation and arrangement of nucleotide residues and the flexibility of DNA in the binding area of platinum compound. We demonstrated that the 1,2-d(GpG) intrastrand crosslink of 1,1/t,t distorts DNA in a different way than 'classical' cisplatin (Kašpárková et al. 1996). The major structural difference of the studied adducts has been found in different DNA bending. The 1,2-d(GpG) intrastrand adduct of 1,1/t,t results in a flexible non-directional bend, which is not recognized by HMG domain proteins. This was in a fundamental contrast to the same type of

adduct formed by cisplatin, which produces a rigid, directed bend into the major groove of DNA (Takahara et al. 1995; Gelasco and Lippard 1998) The rigid bend is than recognized by HMG domain proteins (Ohndorf et al. 1999). It has been suggested that the results of the present work are consistent with different DNA binding modes of cisplatin and polynuclear bifunctional DNA-binding drugs, which might be relevant to their distinct biological effectiveness (Hofr et al. 2001).

When analyzing the effect of platinum compounds on DNA stability, melting temperature obtained from the temperature dependence of absorbance in UV spectrum range was often used in initial studies. Such obtained extent of thermal destabilization was assessed for correlations with biological consequences of platinum adduct formation on DNA. In our subsequent studies we assessed the effects of monofunctional adducts of $[PtCl(NH_3)_3]Cl$ or chlorodiethylenetri-amineplatinum(II) complexes on thermodynamic stability and energetics of DNA duplexes (Bursova et al. 2005). Using differential scanning calorimetry we analyzed melting temperatures and thermodynamic profiles of the studied site-directed platinum adduct formed on short DNA duplexes. Thus, the results of this work represent an experimental support for the view that thermodynamic rather than thermal characterization of DNA modified by monofunctional platinum(II) compounds is a property involved in the modulation of downstream effects such as protein recognition and repair. Importantly, our studies shed more light into the properties of adducts that determine the platinum adduct recognition on DNA and possibly the way how it is going to be processed once recognized by DNA reparation pathways in cells (Bursova et al. 2005).

1.8. DNA interactions with new ruthenium compounds

Another approach how to change reactive properties and thus biological effects of metal based compounds is to exchange the central atom. The research interest has been directed towards compounds containing metal atoms from the platinum group. Very promising were initial findings that suggested that several ruthenium compounds had been found to inhibit DNA replication, exhibit mutagenic activity, induce the SOS repair mechanism, bind nuclear DNA, and reduce RNA synthesis, which is consistent with DNA binding of these compounds *in vivo* (Corda et al. 1993).

It has been shown that in cell-free media ethylelenediamine Ru(II) arene compounds, in which arene =biphenyl, dihydroanthracene, tetrahydroanthracene,p-cymene, or benzene, bind preferentially to guanine residues in natural double-helical DNA. In addition, DNA binding of the complexes containing biphenyl, dihydroanthracene, or tetrahydroanthracene ligands can involve combined coordination to guanine residue in N7 position and noncovalent, hydrophobic interactions between the arene ligand and DNA, which may include arene intercalation and minor groove binding (Chen et al. 2002; Novakova et al. 2003).

We analyzed DNA duplexes modified at central quinine residues by monofunctional Ru(II) arene complexes [(h6-arene)Ru(II)(en)(Cl)]+ (arene = tetrahydroanthracene or p-cymene, Ru-THA or Ru-CYM, respectively) (Novakova et al. 2005). These two complexes were chosen as representatives of two different classes of Ru(II) arene compounds for which initial studies revealed different binding modes: one that may involve DNA intercalation (tricyclic-ring in Ru-THA) and the other (mono-ring Ru-CYM) that may not. Ru-THA is twenty-fold more toxic to cancer cells than Ru-CYM. The adducts of Ru-THA and Ru-CYM have contrasting effects on the conformation, thermodynamic stability, and polymerization of DNA *in vitro*. We used isothermal titration calorimetry (ITC) to determine the thermodynamic parameters of the duplex formation from its two complementary single strands over a range of temperatures (Novakova et al. 2005). The determined thermodynamic parameters revealed that the exothermic formation of the single monofunctional adduct in the duplex TGT(15) by Ru-THA or Ru-CYM resulted in a large decrease of the change in the enthalpy of duplex formation by 4.4 and 7.4 kcal/mol, respectively. In other words, the monofunctional adduct of Ru-THA or Ru-CYM enthalpically destabilizes the duplex relative to its nonmodified counterpart. The adduct of Ru-CYM destabilized DNA significantly more than the adduct of Ru-THA, whose DNA binding mode, additionally, involves noncovalent, hydrophobic interactions between the arene ligand and DNA, such as arene intercalation. Various intercalators thermodynamically stabilize DNA since they lengthen and unwind DNA, increasing the phosphate spacing along the helix axis (Maeda et al. 1990; Bjorndal and Fygenson 2002). Thus, our results suggest that the higher thermodynamic stability of DNA containing Ru-THA adducts is

associated with their intercalations. We observed the low efficiency of the mammalian nucleotide excision repair (NER) systems employed in the present work to excise Ru(II) arene adducts and especially that of Ru-THA. From additional comparisons we found that, Ru-CYM adducts are removed from DNA more efficiently than those of Ru-THA. Moreover, the mammalian nucleotide excision repair system has low efficiency for excision of ruthenium adducts compared to cisplatin intrastrand crosslinks. Thus, the character and extent of DNA distortion induced in DNA by the adducts of Ru(II) arene complexes and resulting thermodynamic destabilization of DNA control the biological effects of this class of ruthenium complexes. The results of the work afford further details, which allow for improving the structure- pharmacological relationship of Ru(II) arene compounds, and provided a more rational basis for the design of new antitumor ruthenium drugs and chemotherapeutic strategies.

1.9. Future strategies of anticancer treatment using metal-based drugs

New approaches need to be developed based on the knowledge of cisplatin anticancer mechanisms. However, equipped with the new knowledge concerning

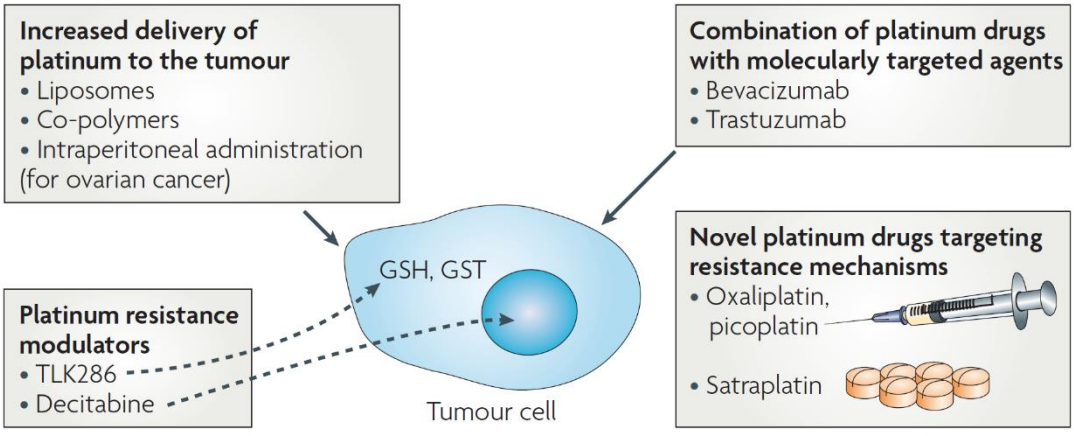


Figure 6. Major ongoing strategies to circumvent cisplatin resistance. Resistance can be tackled by: increasing the levels of platinum reaching tumours (for example, liposomal platinum products) thereby resulting in greater killing; combining existing platinum drugs with molecularly targeted drugs (for example, bevacizumab); using novel platinum drugs such as oxaliplatin that are capable of circumventing cisplatin-mediated resistance mechanisms; and using other drugs either alone (for example, TLK286) or in combination (for example, decitabine), which exploit particular cisplatin-mediated resistance mechanisms. GSH, reduced glutathione; GST,glutathione S-transferase (Kelland 2007) .

combination of platinum drugs with new therapies targeting resistance mechanisms of metal based drugs action and tumor resistance, four major strategies are being proposed to circumvent platinum-drug resistance in cancer patients:

1. Novel platinum drugs targeting resistance mechanisms
2. Improved delivery of platinum drugs to tumors
3. Platinum resistance modulation
4. Combining platinum drugs with new therapies targeting resistance mechanism

Each of the proposed strategies will be commented in subsequent chapters.

1.9.1. Novel platinum drugs overcoming resistance to cisplatin

Many tens of platinum analogues have entered the clinic in the past more than three decades. This strategy seems to finally be proving successful. Clinical trials with new drugs oxaliplatin, satraplatin and picoplatin have shown that the dose-limiting toxicities of these new platinum drugs are generally different from those observed with cisplatin.

Oxaliplatin (1R,2R-diaminocyclohexane oxalatoplatinum (II)) is based on the 1,2-diaminocyclohexane (DACH) carrier ligand and is a more water-soluble derivative of cisplatin. Independent large phase III trials demonstrated that oxaliplatin is active against metastatic colon cancer when used in combination with 5-fluorouracil and leucovorin (Levi et al. 1992). Before oxaliplatin was introduced to clinical use, colon cancer had been widely acknowledged in the medical community as being insensitive to platinum drugs.

A big step towards improvement of drug application comfort for patients was achieved by development of drugs that could be administered orally. Initial clinical trials with satraplatin demonstrated the feasibility of administering a platinum drug by the oral route, especially when dosed on a daily basis for 5 consecutive days (McKeage et al. 1997). Satraplatin showed promising clinical activity in prostate cancer (McKeage 2007).

Picoplatin (cis-amminedichloro, 2-methylpyridine, platinum (II); JM473) was rationally designed to provide steric bulk around the platinum centre. The steric

hindrance of methylpyridine reduced the inactivation of the drug activity by thiol-containing species in comparison to cisplatin. Picoplatin retains activity against a wide range of cisplatin-resistant (Holford et al. 1998) and oxaliplatin-resistant (Sharp et al. 2002) cells *in vitro*, which was independent of whether resistance was due to reduced transport, increased cytoplasmic detoxification or increased DNA repair. It also possesses antitumor activity *in vivo* by both the intravenous and oral routes (Raynaud et al. 1997).

Another interesting platinum drug that entered phase I clinical trials was the cationic trinuclear agent BBR3464 (Brabec et al. 1999). This drug forms interstrand crosslinks over a much longer range than cisplatin or carboplatin — that is, across up to six DNA base pairs — and encouraged the concept of designing platinum drugs with different modes of DNA binding. However, the clinical development of BBR3464 was stopped after initial phase II trials showed a relatively poor response rate and high cytotoxicity.

In summary, the strategy leading to rational design of platinum drugs with modes of binding different form classical cisplatin has proven to be successful and has brought three novel drugs clinically effective against cancer that are resistant to cisplatin.

1.9.2. Improved delivery of platinum drugs to tumors

There are three main possibilities how to improve delivery into tumor tissues.
The first possibility is to include drug molecules into liposomal vesicles that would improve the transmembrane transport and decreases side toxicity of the platinum drugs. The disadvantage of the liposome drug delivery is modest antitumor activity, which might reflect the delivery of drug to tumor in a relatively inactive form. A second approach, is to link a platinum-based drug to a water-soluble, biocompatible co-polymer, such as hydroxypropylmethacrylamide, in order to exploit the enhanced permeability and retention effect of macromolecules in tumors. Finally, in case of patients with ovarian cancer, localized platinum-drug administration through intraperitoneal injection might be applied.

1.9.3. Platinum resistance modulation

The glutathione-mediated detoxification pathway is an important determinant of platinum-drug sensitivity and resistance. The synergy effect of inhibitors of glutathione synthesis were tested in phase I clinical trials (Morgan et al. 1998). Another approach exploits deactivation of the DNA mismatch repair through hypermethylation of the MutL homologue- 1 (MLH1) gene, which has been shown to lead to resistance to cisplatin and carboplatin, and predicts for poor survival of patients with ovarian cancer (Gifford et al. 2004). This has led to the idea of using a DNA demethylating agents such as 2'-deoxy-5-azacytidine in combination with cisplatin or carboplatin to overcome this resistance mechanism (Plumb et al. 2000). The approaches connected with modulation of resistance to platinum drugs have been studied intensively.

1.9.4. Combination of platinum drugs with new therapies targeting resistance mechanism

Most present cancer drug development involves the targeting of specific molecular abnormalities that are characteristic of cancer, described in terms of various phenotypic markers (Hanahan and Weinberg 2000). However, in some cases, these agents might possess no spectacular activity as monotherapy, but are used optimally in combination with existing cytotostatics. The first example was bevacizumab, a humanized monoclonal antibody that targets vascular endothelial growth factor, did not possess marked single-agent antitumor efficacy, but it has subsequently been shown to significantly improve responses and survival of patients with non-small-cell lung cancer when added to carboplatin–paclitaxel combination chemotherapy (Sandler et al. 2006).

Usually, the clinically used platinum drugs have no significant effect on advanced breast cancer. However, under some circumstances, platinum drugs might be useful. Promisingly, clinical data showed that a humanized antibody against ERBB2 (also known as HER2) synergized with cisplatin through a mechanism involving inhibition of the repair of platinum induced DNA damage (Pietras et al. 1994).

Additionally, it has been shown that, platinum drugs might be particularly useful in breast cancers harboring BRCA1 or BRCA2 mutations (approximately 5–10% of cases) as mouse-derived Brca1- negative cell lines have been shown to be fivefold hypersensitive to cisplatin, compared with wild-type cells (Bhattacharyya et al. 2000). This seems to be related to a lower DNA-repair capacity, and clinical trials are ongoing to test this hypothesis (Turner et al. 2005).

My recent PhD. graduate Michal Zimmermann has found connection between BRCA 1 deficiency and DNA repair pathway selection through factor Rif1 (Zimmermann et al. 2013). Based on Michal's results, it has been suggested that Rif1 gene inactivation contributes to the more efficient DNA repair and suppresses toxic nonhomologous end joining of double-strand breaks induced by PARP inhibition in Brca1-deficient cells. The suppression of erroneous DNA breaks means a significantly lower sensitivity of Rif1 deficient tumors to therapeutic agents that induce DNA breaks. Thus, Rif1 gene activity is an important marker that might help to choose the most appropriate therapeutic strategy in breast cancer treatment. The DNA repair pathways regulation is closely connected with telomeric proteins that are object of our recent research and are discussed in the next chapters.

1.10. Significance and impact I

Results of our studies published in the above mentioned and below listed papers revealed that the thermodynamic profile of drug-DNA complex formation for the effective anticancer drug, cisplatin, differs completely from its ineffective isomer. Additionally different effects of newly discovered classes of anticancer compound on DNA and DNA-binding proteins was described and contributed to explanation of different processing of the metal-DNA adducts by reparation pathways. These findings have been used in the design of more effective anticancer drugs that have structural and thermodynamic properties different from clinically used platinum drugs. Based on the published studies, I was invited for a postdoc stay at the Department Chemistry and Chemical Biology at Rutgers, The State University of New Jersey, USA (2002-2003).

1.11. Publications related to above topics I

1. **Hofr C.** and V. Brabec. Thermal and thermodynamic properties of duplex DNA containing site-specific interstrand cross-link of antitumor cisplatin or its clinically ineffective *trans* isomer. *J.Biol. Chem.* 2001. 276: 9655-9661 (IF =7.7)
2. **Hofr C**, Brabec V. Thermal stability, energetics and melting cooperativity of DNA interstrand cross-linked by trans-diamminedichloroplatinum(II). *Biopolymers.* 2005 77: 222-9 (IF 2.7)
3. Novakova O, Kasparkova J, Bursova V, **Hofr C**, Vojtiskova M, Chen H, Sadler PJ, Brabec V. Conformation of DNA modified by monofunctional Ru(II) arene complexes: recognition by DNA binding proteins and repair. Relationship to cytotoxicity. *Chem. Biol.* 2005. 12:121-9. (IF 5.9)
4. Bursova V,Kasparkova J, **Hofr C**, Brabec V,. Effects of Monofunctional Adducts of Platinum(II) Complexes on Thermodynamic Stability and Energetics of DNA Duplexes. *Biophysical Journal.* 2005. 88:1207-14 (IF 4.5)
5. **Hofr C**, Farrell N, Brabec V. Thermodynamic properties of duplex DNA containing site-specific d(GpG) intrastrand cross-link of antitumor dinuclear platinum komplex *Nucleic Acids Research*. 2001. 29:2034-40. (IF = 5.7)
6. Nováková O, **Hofr C**, Brabec V. Modification of natural, double-helical DNA by antitumor cis- and trans-[Cl2(Me2SO)4Ru] in cell-free media. *Biochem. Pharmacol.* 2000. 60:1761- 1771 (IF = 2.8)
7. Malina J, **Hofr C**, Maresca L, Natile G and Brabec V. DNA interactions of antitumor cisplatin analogs containing enantiomeric amine ligands. *Biophysical Journal.* 2000. 78, 2008-2021(IF = 4.6)
8. Buchlovič M, Kříž Z, **Hofr C**, Potáček M. New PAH derivatives functionalized by cyclic nitrone framework: synthetic design, anti-proliferative activity and interaction with DNA. *Bioorg Med Chem.* 2013. 21:1078-81. (IF 3.2)
9. **Hofr C**, Mikrokalorimetrie biologicky významných molekul. Československý časopis pro fyziku, 2006. 56: 288-292.

2. Genome safe guards: Telomeric proteins recognizing and protecting chromosome ends

The maintenance of linear chromosomes requires resolution of two biological complications: the chromosome end-protection problem and the chromosome end-replication problem. Firstly, the ends of linear chromosomes, such as those in humans and other eukaryotes, must be distinguished from broken DNA ends that require repair. If there were no mechanism for the chromosomal end distinction, linear chromosomes would have been disposed to unwanted DNA end-joining that result in crucial consequences such as end-to-end chromosomal fusions. It means that there must be a mechanism that solves the chromosome end-protection problem (Palm and de Lange 2008).

Secondly, the chromosome end-replication problem is defined by the gradual loss of DNA sequence and thus possibly also genetic information at the extreme end of chromosomes because of incomplete replications by DNA polymerases. As the polymerases synthesize DNA in the 5‘ to 3‘ direction, they are unable to fill in the gap left behind by the 5‘ most RNA primer (Levy et al. 1992).

The solution of both chromosome end-related problems is facilitated by telomerase. Telomerase is a two-subunit enzyme that contains telomerase reverse transcriptase (TERT) and a template-containing RNA component (TR). Telomerase synthesizes multiple tandem repeats of telomeric DNA encoded by its RNA template. Thus, telomerase compensates for the erosion of DNA ends during replication and provides the docking sites for telomeric proteins that bind specifically to the ends of chromosomes to distinguish them from broken DNA ends. The access of telomerase and other DNA binding enzymes and modulators is regulated by proteins specifically associated with telomeres.

2.1. Composition of human shelterin

The mammalian chromosome end is capped by a protein complex called shelterin. The name shelterin was introduced by Titia de Lange (de Lange 2005) who identified the first shelterin component in 1992 (Zhong et al. 1992). The first identified mammalian telomeric protein was **T**elomeric **R**epeat-binding **F**actor 1, referred to as

TRF1 ; also known as TERF1, was isolated based on its in vitro specificity for double-stranded TTAGGG repetitive sequence, which is typical of vertebrate telomeres (Zhong et al. 1992; Chong et al. 1995). The other five components of shelterin protein complex were discovered subsequently (de Lange 2005). The other subunits of shelterin are: telomeric repeat-binding factor 2 TRF2 (also known as TERF2), repressor and activator protein 1 (RAP1; also known as TERF2IP), TRF1-interacting nuclear protein 2 (TIN2; also known as TINF2), protection of telomeres 1 (POT1) and TPP1 (also known as ACD). The mammalian shelterin complex brings together telomeric ssDNA- and dsDNA-binding proteins by linking the protein partners to protect the natural chromosome ends from being recognized as sites of DNA damage (Figure 7A). TIN2 is a crucial bridging component that not only links TRF1 and TRF2 but also connects TRF1 and TRF2 to POT1– TPP1 (Kim et al. 1999; Houghtaling et al. 2004; Ye et al. 2004a). TIN2 uses an N-terminal region to bind TRF2, and a FxLxP amino acid motif at the C-terminus of TIN2 facilitates TRF1 binding (Ye et al. 2004a). Bridging of TRF1 and TRF2 by TIN2 is crucial for the assembly of a complete shelterin complex because TRF1 does not heterodimerize or otherwise interact with TRF2 (Fairall et al. 2001). TIN2 uses a region in its N-terminal domain to bind to the C-terminus of TPP1 to recruit it to telomeres (Ye et al. 2004b). As the TRF2- and TPP1-binding sites reside in the N-terminal domain of TIN2, TIN2 binding to TRF2 and TPP1 could potentially occur in a mutually exclusive manner, but data showing stabilization of TRF1–TIN2–TRF2 in the presence of TPP1 support a model where the four shelterin components coexist in a single complex (O'Connor et al. 2006).

2.2. Functions of shelterin in mammals

Shelterin enables cells to distinguish their natural chromosome ends from DNA breaks, represses DNA repair reactions, and regulates telomerase-based telomere maintenance (Palm and de Lange 2008). The components of shelterin are functionally defined by following properties. Shelterin proteins specifically localize to telomeres and three shelterin subunits, TRF1, TRF2 and POT1, recognize TTAGGG; shelterin components are abundant at telomeres throughout the cell cycle; and they do not have function elsewhere in the nucleus (Palm and de Lange 2008). Shelterin determines the structure of the telomere terminus, it is implicated in the generation

of telomeric loops (t-loops), where the 3' overhang is strand-invaded into the adjacent duplex telomeric repeat array, forming a D-loop. Thus shelterin controls the access of telomerase to very ends of telomeric DNA (Figure 7B).

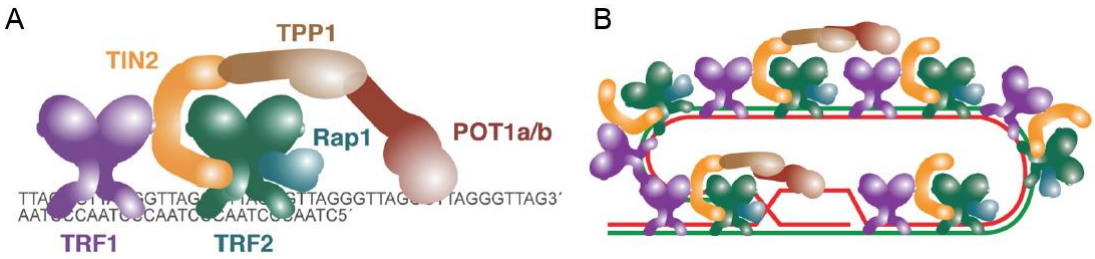


Figure 7. Shelterin protein complex protects telomeric DNA. (A) Components of human shelterin and their arrangements on the end of chromosomal DNA. (B) Relative abundance of the shelterin components and how they might be bound to telomeres in the t-loop configuration (deLange 2010).

The importance of shelterin in protecting chromosome is demonstrated by the following results of experiments performed in live cells. Knockdown of individual components TRF2 and POT1 of shelterin complex leads to deprotection of chromosome ends and the activation of DNA damage response pathways at telomeres (Denchi and de Lange 2007). A recent study using knockout mice in which the shelterin complex was completely removed from telomeres revealed that the shelterin complex prevents six independent DNA damage response pathways from their occurring at chromosome ends (Sfeir and de Lange 2012).

More than two decades of genetic research have identified and assigned main biological functions of shelterin proteins that safeguard telomeres. However, a molecular mechanism of how each protein subunit contributes to the protecting function of the whole shelterin complex remains elusive. The following section shows the strategy that was employed recently to shed more light into the molecular mechanism of basic shelterin functions. The results of experiments discussed below contributed to understanding how shelterin complexes recognize and bind telomeric DNA and thus represses DNA repair signaling pathways including nonhomologous end joining and homology-directed repair.

2.3. The mechanism of human Rap1 contribution to selective recognition of telomeric DNA by shelterin

Rap1-TRF2 complex is a critical part of shelterin as it suppresses homology-directed repair (Kabir et al. 2010). To understand how Rap1 affects key functions of TRF2, full-length Rap1 binding to TRF2 and Rap1-TRF2 complex interactions with double-stranded DNA were investigated by quantitative biochemical approaches, such as fluorescence anisotropy, isothermal titration calorimetry, gel retardation analysis and surface plasmon resonance. We quantified how human Rap1 contributes to DNA binding affinity and selectivity of TRF2 (Janoušková et al. 2014). Our data showed that Rap1 reduces the overall DNA binding affinity of TRF2 but increases the selectivity of TRF2 to telomeric DNA. Additionally, we observed that Rap1 induces a partial release of TRF2 from a preformed TRF2-DNA complex. The improved selectivity of TRF2 to telomeric DNA is caused by less pronounced electrostatic attractions between TRF2 and DNA in Rap1 presence. The neutralization effect of Rap1 was revealed by set of DNA affinity measurement of TRF2 in different salt conditions (Janoušková et al. 2014).

For analysis of the electrostatic component of TRF2 interactions with telomeric DNA in presence or absence of protein Rap1, two independent sets of binding affinity measurements were performed. Dissociation constants of TRF2 binding to DNA were determined using buffers containing different salt concentrations. We concluded that Rap1 serves as a selectivity enhancer of TRF2 with newly found ability to partially remove bound TRF2 from telomeric DNA. The observed Rap1 release activity suggests that Rap1 plays an important role in tuning the interactions of TRF2 with DNA. Our data showing the release of TRF2 from telomeric DNA in presence of Rap1 suggest that protein Rap1 might prompt the relocation of TRF2 to the preferred single-strand/double-strand region junction of telomeric DNA. For the first time here, we used combination of quantitative biophysical approaches to describe and explain molecular origins of Rap1 contribution to selective TRF2 recognition of telomeric DNA.

Based on our quantitative analyses, the known domain arrangement and the available structural data (Hanaoka et al. 2001; Court et al. 2005; Hanaoka et al. 2005), we speculate that Rap1, after binding to TRF2, might shield the positively charged basic domain on the N-terminus of TRF2 from mainly non-specific electrostatic interaction with DNA (Figure 8). Additional TRF2 binding to DNA via basic N-terminal domain may occur and is sterically allowed when Rap1 is absent. Thus, Rap1 prompts more accurate and selective recognition of telomeric DNA by TRF2. The results of our quantitative functional studies contribute to the understanding of the selective recognition of telomeric DNA by the whole shelterin.

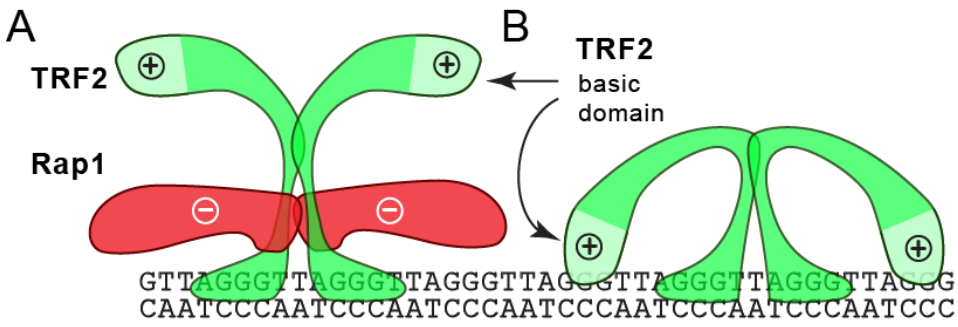


Figure 8. A possible mechanism how Rap1 increases specificity of TRF2 binding of the telomeric DNA. (A) Negatively charged Rap1 induces the neutralization of positively charged TRF2 and prevents nonspecific interactions of TRF2 basic domain with DNA. (B) Additional nonspecific TRF2 binding to DNA via basic N-terminal basic domain may occur and is sterically allowed when Rap1 is absent (Janoušková et al. 2014).

Future studies focused on shelterin protein dynamics using single molecule approaches and *in vivo* interaction analysis of Rap1 and TRF2 will be of particular interest as they can reveal relocation dynamics of TRF2 induced by Rap1 and the mechanism of regulation of shelterin and distribution on telomeres.

2.4. Model organisms in telomere maintenance research

Although a tremendous effort has been developed in the field of telomeric research, there are certain limitations that make the research challenging. The gene for telomerase is silenced in normal human cells in order to minimize the risk of accumulation of mutations leading to a cancer development (Wright et al. 1996). This makes human telomerase research focused on cancer cells or artificially immortalized cells, but these cells could be heavily altered by neoplastic transformations or genetic modifications in comparison with normal cells. On the other hand, telomerase activity in plant cells is reversibly regulated in response to proliferation activity. Moreover, plant meristem cells show similar differential

plasticity as human embryonic stem cells. These advantageous properties make model organisms to be indispensable for telomere research.

Our laboratory identified and characterized telomere repeat binding proteins from the single-myb-histone domain (SMH) protein family in the genome of *Arabidopsis thaliana* (Kuchar and Fajkus 2004; Mozgova et al. 2008a). SMH proteins have double-stranded DNA binding affinity, similarly as human TRF1 and TRF2. In order to reveal interconnections and to describe interactions of proteins participating in the process of telomere synthesis, quantitatively oriented studies have been successfully performed in our laboratory (Mozgova et al. 2008a; Hofr et al. 2009). These studies have clearly demonstrated benefits of quantitative approaches in the research of binding mechanism of telomeric proteins.

2.5. SMH proteins recognize telomeric DNA in plants

The human shelterin proteins TRF1, TRF2 and their homologues in other organisms possess a well conserved DNA-binding structural motif similar to the c-Myb-family of transcriptional activators (Bilaud et al. 1996). The Myb domain of human TRFs is C-terminally positioned and consists of three helices connected in a helix–turn–helix manner. The third helix contains a conserved amino acid sequence called a 'telobox', which has been shown to be important for recognition of telomeric double-stranded DNA (Bilaud et al. 1996). Numerous TRF-like proteins have been identified in plants (Schrumpfova et al. 2005; Kuchar 2006; Peska et al. 2011), and, in a few cases, the influence of these proteins on telomere length homoeostasis has been demonstrated (Hwang and Cho 2007). Interestingly, besides the TRF-like proteins, a plant specific family of other telobox proteins has been described. The novel group of proteins, termed SMH (single-Myb-histone) family, is characterized by a triple-domain structure consisting of an N-terminal Myb domain, central globular histone H1/5 domain, and a C-terminal coiled-coil domain. In *Arabidopsis thaliana* (At), five SMH proteins were identified (AtTRB1–AtTRB5, where TRB is telomere-repeat-binding factor) (Marian et al. 2003), and three of them have been characterized (Schrumpfova et al. 2004; Mozgova et al. 2008b). These proteins show not only specific interactions with telomeric DNA, but also a number of protein–protein interactions functionally related to telomeres. In addition to their ability to form homodimers (similarly to human TRFs), they can also form heterodimers and both homo- and hetero-multimers (Schrumpfova et al. 2004; Mozgova et al. 2008b;

Schrumpfova et al. 2008) via their H1/5 histone domain. It has been revealed that they also interact (using the same H1/5 domain) with one of the POT1 proteins in *A. thaliana*, AtPOT1b (Mozgova et al. 2008b; Schrumpfova et al. 2008). It has been shown that AtTRB (Arabidopsis thaliana telomere repeat-binding factor) proteins from the SMH (single-Myb-histone) family selectively bind double-stranded telomeric DNA and interact with the telomeric protein AtPOT1b (Arabidopsis thaliana protection of telomeres 1b). Protein AtPOT1b is one of the Arabidopsis homologues of mammalian POT1 protein which is involved in telomere capping (Shakirov et al. 2005).

We performed the first quantitative DNA-binding study of this plant-specific family of SMH proteins. Interactions of full-length AtTRB1 and AtTRB3 with telomeric DNA were analyzed by electrophoretic mobility-shift assay, fluorescence anisotropy and surface plasmon resonance to reveal their binding stoichiometry and affinity. We found that AtTRB1 and AtTRB3 bind telomeric DNA with high affinity and specificity. Comparison of dissociation constants thus demonstrated more than ten-fold higher affinity and binding specificity of AtTRB1 to DNA bearing telomeric sequences. AtTRB3 showed seven-fold higher affinity to telomeric DNA duplex than to non-telomeric DNA. Moreover, binding affinity assays were used to determine that the minimal length of DNA for an appropriate binding of SMH proteins. We found that minimal DNA fragment bound by SMH proteins comprises at least two telomeric repeats. Furthermore, the protein-DNA binding ratio one protein monomer per one DNA telomeric repeat was revealed by a native gel retardation assay.

Binding affinity analyses in different salt conditions enabled us to estimate the electrostatic component of binding and explain the different affinities of AtTRB1 and AtTRB3 to telomeric DNA. We designed the following putative model explaining the binding stoichiometry and the protein arrangement on telomeric DNA (Figure 9).

On the basis of results of the present study and the data available, we suggest that interactions of the two AtTRBs with telomeric DNA occur simultaneously with two binding sites. Therefore the minimal length of duplex DNA required for the proper binding of full-length AtTRB1 and AtTRB3 should harbor at least two putative binding sites that are bound by two dimers of AtTRBs. Consequently, our data suggest that SMH proteins are able to distinguish between short (<10 bp) telomere-like sequences that are dispersed throughout the genome, e.g. in promoter regions (Regad et al. 1994), and longer tracts of telomere repeats occurring in telomeres.

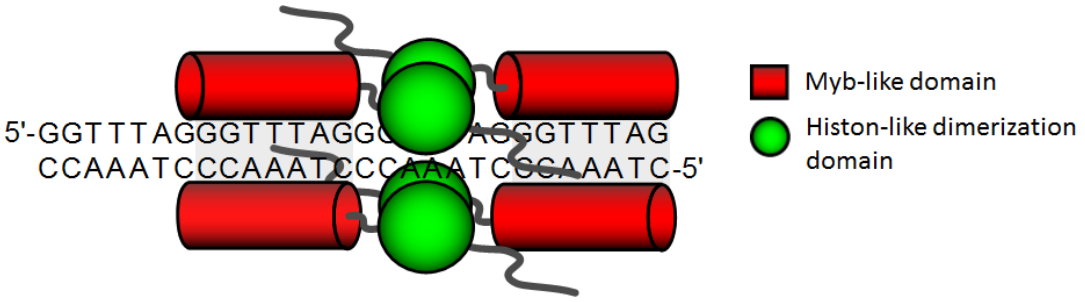


Figure 9. A speculative model of SMH interaction with telomeric DNA. SMH proteins bind telomeric DNA in dimeric form. Shaded DNA regions denominate putative spatially allowed binding sites.

2.6. Yeast telomeric proteins: is it better to have two DNA binding domains?

Recently, a novel type of telomere-binding protein YlTay1p from the yeast *Yarrowia lipolytica* was identified by laboratory of Prof. Tomaska from Departments of Genetics and Biochemistry at Comenius University. YlTay1p contains two Myb domains (Myb1, Myb2) very similar to the Myb domain of mammalian TRF1 and TRF2 (Kramara et al. 2010). Moreover it has been shown that YlTay1p exhibits a preference for telomeric DNA in vitro and binds to the substrate DNA as a dimer (Kramara et al. 2010).

We examined the contribution of each Myb domain of YlTay1p to overall binding affinity and selectivity to telomeric DNA (Visacka et al. 2012). Truncated versions of YlTay1p lacking Myb1, Myb2, or both Myb domains were prepared. The quantification of binding properties of YlTay1p carrying either Myb domain

preferential affinity to both *Yarrowia lipolytica* (GGGTTAGTCA)n and human (TTAGGG)n telomeric sequences. Quantitative measurements of the protein binding to telomeric DNA revealed that the presence of both Myb domains is required for a high affinity of YlTay1p to either telomeric repeat. Additionally, we performed detailed thermodynamic analysis of the YlTay1p interaction with its cognate telomeric DNA, which is, to our knowledge, the first energetic description of a full-length telomeric-protein binding to DNA. Measurements of thermodynamic parameters of YlTay1p binding to DNA by isothermal titration calorimetry revealed its high affinity binding to telomeric repeats in dimeric form. Interestingly, when compared with human TRF1 and TRF2 proteins, YlTay1p exhibited higher affinity not only for *Y. lipolytica* telomeres but also for human telomeric sequences.

The duplication of the Myb domain region in YlTay1p thus produces a synergistic effect on its affinity toward the cognate telomeric sequence, enlightening the need for homodimerization observed in TRF-like proteins possessing a single Myb domain.

The observations that YlTay1p exhibits higher affinity for human telomeres than to *Y. lipolytica* telomeres and that it shows lower dissociation constants for binding to human telomeres than human telomeric proteins TRF1/TRF2 underline an important general evolutionary principle: natural selection does not necessarily lead to maximization of the affinity of a particular DNA-binding protein to its cognate DNA substrate. Perhaps the binding properties of a DNA-binding protein are tuned to inferior values, thus, enabling its dynamic association with the target DNA loci.

Consequently, the question whether it is better for telomeric proteins to have two DNA binding domains can be answered in two ways. It is definitely better in respect that two Myb-like domains prompt a higher DNA binding affinity. More complicated is the answer in respect of the DNA binding selectivity. It has been shown that in case of Tay1 proteins the higher number of DNA binding domains diminish the selective recognition of yeast telomeric DNA compared to human telomeric DNA. If we try to generalize the findings, from the evolutional point of view a selective binding is more important than a higher affinity. This is supported by the fact that human and mammalian telomeric proteins contain only one Myb domain and bind to telomeric DNA as dimers. Additionally, the dimeric arrangement of single-Myb-domain protein

forms a basis for higher levels of regulation that may be important for a proper protections function of shelterin and other essential DNA processing complexes in mammalian cells.

2.7. Significance and impact II

The above described studies were the first fully quantitative studies focused on the binding selectivity of full-length human, plant, and yeast telomeric proteins. The studies demonstrated that the detailed quantification of protein–DNA interactions provide new insights into the mechanism of protein containing single or multiple Myb-like domains to the specific recognition of telomeric DNA. The studies demonstrated that the detailed quantification of protein–DNA interactions may provide new insights into the structural dynamics of telomeres.

On the basis of the published studies, I was invited to Warwick University (UK) to give a seminar talk entitled *Protein binding to telomeric DNA: quantitative approaches in modern experimental biology* (2009). The data were also used to design suitable experiments *for* following *in vivo* binding studies performed in our laboratory. In these studies, the dynamics of AtTRB1 interaction with chromatin was characterized in living cells using fluorescence microscopy (Dvorackova et al. 2010; Schrumpfova et al. 2014). On the basis of the most recently published studies, I was invited to Colorado State University (USA) to give a seminar talk entitled *Telomeric proteins across species: Quantitative approaches in protein function analysis (2013).*

2.8. Publications related to above topics II

1. Janoušková E, Nečasová I, Pavloušková J, Zimmermann M, Hluchý M, Marini V, Nováková M, and **Hofr C**. Human Rap1 modulates TRF2 attraction to telomeric DNA. Under review in *Nucleic Acids Research*.
2. Visacka K, **Hofr C**, Willcox S, Necasova I, Pavlouskova J, Sepsiova R, Wimmerova M, Simonicova S, Nosek J, Fajkus J, Griffith JD, and Tomaska L. Synergism of the two Myb domains of Tay1 protein results in high-affinity binding to telomeres. *J Biol Chem.* 2012; 287:32206-15. (IF 5.4)
3. **Hofr C**, Sultesová P, Zimmermann M, Mozgová I, Procházková Schrumpfová P, Wimmerová M, Fajkus J. Single-Myb-histone proteins from Arabidopsis thaliana: a quantitative study of telomere-binding specificity and kinetics. *Biochem J.* 2009; 419:221-8. (IF 5.2)
4. Mozgová I, Schrumpfová PP, **Hofr C**, Fajkus J. Functional characterization of domains in AtTRB1, a putative telomere-binding protein in Arabidopsis thaliana. *Phytochemistry*. 2008;69:1814-9. (IF 3.1)
5. Zimmermann M, **Hofr C**, Šultesova P, Mozgová I, Procházková Schrumpfová P, Fajkus J. Mechanismus vazby SMH proteinů na telomerickou DNA. *Chemické listy*. 2009; 5:455.

3. Mechanism of RNA recognition in DNA transcription: An effective combination of quantitative interaction analysis with structural data

Methods of structural analysis, such as NMR and X-ray crystallography, can provide us with detailed information about spatial distribution of atoms and thus mutual positions and orientations of whole biomolecules. Moreover, the structural methods identify residues that might be essential for the binding interaction of proteins and nucleic acids. However, from these two methods, only NMR can be used to describe interaction dynamics and binding affinity, but in a very limited extent. On the other hand, if we combine structural data with studies using quantitative biochemical and biophysical approaches, we can reveal the mechanisms that drives the observed biomolecular interactions. The binding interactions of biomolecules give rise to the biological functions that are then expressed and observed as unique phenotypes. In summary, the precise description of relationships between structure and function of molecules is necessary for the understanding of the origins of function and mechanisms of the processes that are essential for role of nucleoprotein complexes in the context of whole organisms. The crucial processes are mediated by large ribonucleoprotein complexes. The ribonucleoprotein complexes are parts of molecular machineries that take place during protein synthesis involved in RNA transcription and subsequent processes, for example splicing machinery. The processes of the assembly and disassembly of RNA-binding proteins and their substrates are determined by the thermodynamics and the kinetics of their interactions. The majority of RNA-binding proteins of large ribonucleoprotein complexes bind RNA tightly. In contrast, many multidomain RNA-binding proteins with regulatory functions bind RNA substrates during cotranscriptional and posttranscriptional processing less tightly. Usually, the binding occurs with sequence flexibility and in cooperative modes (Mackereth and Sattler 2012).

We contributed to the intensive effort of research community to reveal the molecular mechanism by which RNA-binding proteins recognize and bind RNA. We analyzed the RNA recognition mechanism of the Adenosine deaminase ADAR2, recognition of transcription termination signal by the nuclear polyadenylated RNA-

binding (Nab)3 protein (Hobor et al. 2011) and studied how RNA Polymerase II controls binding of an RNA processing factor.

3.1. Structure of proteins recognizing RNA

The association of RNA-binding proteins (RBPs) with RNA transcripts initiates during transcription. Some of the initially-binding RBPs remain bound to the RNA until it is degraded, whereas others recognize and transiently bind to RNA at later stages for specific processes such as splicing, processing, transport and localization (Dreyfuss et al. 2002).

Surprisingly, if we consider the vast variety of roles that RNA-binding proteins play, there are only a few RNA-binding domains. Here, we will focus on current structural and functional understanding of protein–RNA recognition mediated by the two most abundant RNA-binding domains: the RNA-recognition motif (RRM) and the double-stranded RNA-binding motif (dsRBM (Stefl et al. 2005). It has been suggested that dsRBM domains can recognize RNA in two distinct ways. Firstly, domains bind single-stranded RNA by direct readout of the primary sequence. Secondly, the dsRBM domains might recognize the shape of the RNA. There was also question whether the RNA is recognized based on shape and sequence simultaneously. RNA-binding domains, such as the K-homology (KH) domain or the oligonucleotide/oligosaccharide-binding (OB) fold and the most abundant nucleic-acid-binding motif, the CCHH-type zinc-finger domain are not discussed here but they are comprehensively reviewed in (Messias and Sattler 2004; Stefl et al. 2005; Yin et al. 2013).

The RNA-binding domains are often tandemly arranged and tethered by flexible or structured linkers. The advantage of such a modular architecture arises from the resulting versatility. The multidomain RNA-binding proteins can bind RNA with higher specificity and affinity (similarly to DNA binding proteins, Chapter 2.6) when compared to individual RNA-binding domains. On the other hand, one-domain proteins usually bind only short RNA stretches with relatively low affinities. The concept of modularity also enables that RNA-binding proteins can bind RNA with

poorly conserved sequence features, which are often observed, for example, in 3'-end processing and splicing sites (Mackereth et al. 2011).

In the case of proteins with enzymatic activity, the RNA-binding domains define the specificity of the enzymes by recruiting them to their substrates. Examples of such enzymes are Adenosine deaminases that act on RNA 2 (ADAR2). These enzymes recode genomic information by the site-selective deamination of adenosine. In mammals, RNA editing by site-selective adenosine deamination regulates key functional properties of eurotransmitter receptors in the central nervous system. Glutamate receptor subunit B is nearly 100% edited at one position (the QR-site), which is essential for a normal receptor function. Its significance is apparent from mouse models in which a slightly reduced rate of QR-site editing is associated with early onset epilepsy and premature death (Maas et al. 2001). ADARs from all characterized species have a modular domain organization consisting of one-to-three dsRBMs followed by a conserved C-terminal catalytic adenosine deaminase domain. ADARs can edit RNA substrates either specifically or nonspecifically depending upon the structures of the RNA substrates (Bass, 2002). Despite the known biological function ADAR the mechanism of recognition of RNA remained unknown. The independent structures of the two dsRBMs (free in solution) and of the isolated catalytic domain of ADAR2 have been determined previously (Stefl et al. 2006).

To better understand RNA substrate recognition by ADAR2, the structure of the RNA helix surrounding the editing site and the two dsRBMs of ADAR2 bound to the GluR-2 R/G site were determined. Complementarily, we used combination of structural and biophysical approaches to address the mechanisms of double-stranded RNA recognition by dsRBMs of ADAR2 enzymes.

3.2. Mechanism of RNA recognizing protein ADAR via RNA-binding domains: Double-stranded RNA recognition based on shape or sequence?

Sequence-specific recognition of dsRNA by proteins was elusive, despite its importance in RNA maturation pathways. Adenosine deaminases that act on RNA (ADARs) recode genomic information by the site-selective deamination of adenosine. ADAR2 contains two dsRBMs and a catalytic deaminase domain (Bass, 2002; Macbeth et al., 2005; Stefl et al., 2010). We participated in this study reporting the structure and molecular mechanism of function of the ADAR2 double-stranded RNA-binding motifs (dsRBMs) bound to a stem-loop pre-mRNA encoding the R/G editing site of GluR-2 (Stefl et al. 2010). In the determined structure, the two dsRBMs bind one face of the RNA covering approximately 120 degrees of the space around the RNA helix (Figure 10). This suggests that the binding of ADAR2 molecules as dimers would be sterically possible. The possible dimeric binding is consistent with studies indicating that ADAR2 dimerization is necessary for RNA editing (Cho et al. 2003; Valente and Nishikura 2007).

Our functional studies together with structure information provided a basis for molecular mechanism how dsRBMs recognize the shape and, surprisingly, also the sequence of the dsRNA. The unique direct recognition of the RNA primary sequence by dsRBMs is achieved via the minor groove of the dsRNA. More generally, our findings suggest a solution to the sequence-specific paradox faced by many dsRBM-containing proteins that are involved in post-transcriptional regulation of gene expression. We performed a quantitative characterization of dsRBMs binding to RNA substrates with different primary sequences. We designed two sets of mutations, one set was designed to change the recognition sequence on RNA. The second set was designed to maintain the recognition sequence, but the RNA shape was changed via single-site mismatches of RNA bases to measure their effect on overall binding affinity. In other words, the sequence of RNA substrates was modified compared to the canonical sequence in order to assess the influence of alterations to the sequence or spatial arrangement on dsRBMs binding affinity.

We performed fluorescence anisotropy (FA) experiments by titrating dsRBM1 and

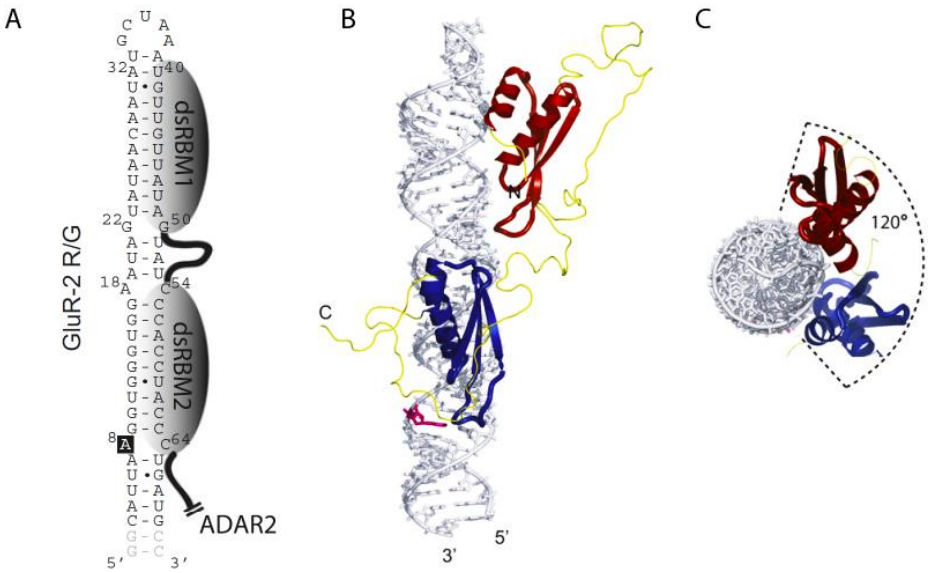


Figure 10. Scheme and structure of ADAR2 dsRBM12 bound to the RNA substrate GluR-2 R/G (A) Secondary structure of GluR-2 R/G RNA with the indicated binding regions for the dsRBMs. (B) The structure of the ADAR2 dsRBM12 bound to GluR-2 R/G. The RNA is represented as a stick model (in gray; the edited adenosine is highlighted in pink) and the protein is shown as a ribbon model (dsRBM1 in red; dsRBM2 in blue; linker in yellow). (C) Top view of the protein-RNA complex (Stefl et al. 2010).

dsRBM2 against labeled RNA substrates. Unlabeled wild-type and mutant RNAs were used for competition experiments. The equilibrium dissociation constants were calculated from the displacement of the binding curves. If the bases of the recognition site for dsRBM1 and dsRBM2 were mutated the apparent affinity was reduced significantly compared to the wild-type. However, when we assessed different single point mismatches in the same position of the RNA sequence, the affinity was almost identical to wild-type RNA, confirming that dsRBM1 recognizes the sequence rather than the shape of the RNA helix.

Altogether, the FA data strongly support the idea that the sequence-specific interactions observed in the structures of ADAR2 dsRBMs-dsRNA are important for the affinity of both dsRBMs and that they finely tune the preferential binding to these recognition motifs. The results of quantitative binding assays with different RNA substrates confirmed a sequence-specific RNA recognition of dsRBMs.

3.3. Significance and impact III

The study above was the first study that quantitatively demonstrated the sequence specific recognition of RNA binding proteins containing double-stranded RNA-binding motifs (dsRBMs). The published findings are of great significance for all proteins that contain dsRBMs and take part in RNA editing. Controlled modification of RNA is eminently important for the posttranscriptional regulation of gene expression. The importance of this study was confirmed by its acceptance for publication in one of the most prestigious scientific journals *Cell*. The publication has been cited more than fifty times (May 2014).

3.4. Publications related to above topics III

1. Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, **Hofr C**, Emeson RB, Allain FH. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell.* 2010;143:225-37. (IF 31.2)
2. Hobor F, Pergoli R, Kubicek K, Hrossova D, Bacikova V, Zimmermann M, Pasulka, J, **Hofr C**, Vanacova S, Stefl R. Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (Nab)3 protein. *J Biol Chem.* 2010; 286: 3645-3657. (IF 5.4)
3. Kubicek K, Cerna H, Holub P, Pasulka J, Hrossova D, Loehr F, **Hofr C**, Vanacova S, and Stefl R. A Phosphoserine-Proline Switch in RNAP II CTD Controls Binding of an RNA Processing Factor, *Genes and Development*. 2012; 26:1891-6. (IF 13.9)

4. Future prospects

The present state of knowledge of interaction landscape of nucleic acids and fast development of new research approaches promise exciting future discoveries and their effective applications in gene therapy and molecular therapeutic approaches.

Maybe the most exciting areas of research are connected with detailed understanding the cancer development and cell aging at molecular level. One of the most challenging tasks in cancer genomics is uncovering genes with an active role in tumorigenesis from a potentially large pool of mutated genes across patient samples (Ghersi and Singh 2014). The recent effort focused on the identification of mutations that are crucial for cancer development have shown that the residues involved in these interactions are more frequently affected by mutations observed in large-scale cancer genomic data than are other residues (Ghersi and Singh 2014). By focusing on and characterizing functionally relevant domains of proteins—specifically those that are known to be involved in molecular interactions, we will be able to address mechanisms that lead to cancer development in earlier stages than it has been known now. Additionally, recent studies show a great potential of sequencing methods that promise a great expansion of our functional understanding of the genomic landscape of cancer.

Similarly, whole-genome analyses showed that combined epigenetic changes may contribute to decreased differentiation and increased stem cell self-renewal and this way contribute to slower cell aging (Sun et al. 2014). Newly emerging views based on results of human stem cell research suggests that the aging clock can be experimentally manipulated. The possibility of setting back the molecular clocks in mammalian cell is also supported with results of studies performed on mice. It has been shown that that constitutive expression of enzymatic subunit of telomerase provides antiaging activity in the context of a mammalian organism (Tomas-Loba et al. 2008). The latest studies show that the treatment inducing telomerase activity in mammalian have potential to extend life span and improve regeneration by tens of percents.

Altogether, the findings about molecular mechanism of cell aging may allow future development of novel approaches how to increase stability of the genome, which would lead to the extension of the cell life and minimize risk of malignant transformation.

The future research will need tools how to assess biomolecular interaction mechanisms in living cells. The state of the art fluorescence microscopy methods based on simultaneous detection of a whole plane in cell (Capoulade et al. 2011; Brazda et al. 2014) will be further developed in order to better understand protein diffusion processes and local dynamic equilibria inside cells leading to nonuniform spatial distributions of biomolecules, which are essential for processes such as nuclear organization and signaling in cell division, differentiation and migration; the processes that determine and contribute to genome homeostasis.

Using the above-mentioned very powerful tools the future studies aimed at unraveling the role of protein dynamics involved in the multi-subunit nucleoprotein complexes are expected. As a consequence, gene expressions and the regulation could be understood and controlled at the molecular level in living cells. Consequently, the knowledge of molecular mechanism controlling life-essential processes could be used in new therapeutic approaches targeting crucial genes or biomolecules in well described biological pathways.

5. References

Aguilar-Quesada R, Munoz-Gamez JA, Martin-Oliva D, Peralta-Leal A, Quiles-Perez R, Rodriguez-Vargas JM, Ruiz de Almodovar M, Conde C, Ruiz-Extremera A, Oliver FJ. 2007. Modulation of transcription by PARP-1: consequences in carcinogenesis and inflammation. *Current medicinal chemistry* **14**: 1179-1187.

Bhattacharyya A, Ear US, Koller BH, Weichselbaum RR, Bishop DK. 2000. The breast cancer susceptibility gene BRCA1 is required for subnuclear assembly of Rad51 and survival following treatment with the DNA cross-linking agent cisplatin. *J Biol Chem* **275**: 23899-23903.

Bilaud T, Koering CE, BinetBrasselet E, Ancelin K, Pollice A, Gasser SM, Gilson E. 1996. The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Res* **24**: 1294-1303.

Bjorndal MT, Fygenson DK. 2002. DNA melting in the presence of fluorescent intercalating oxazole yellow dyes measured with a gel-based assay. *Biopolymers* **65**: 40-44.

Bosi GJ, Feldman, D. R., Bajorin, D. E., et al. 2011. Cancer of the testis. in *Cancer: Principles and Practice of Oncology* (ed. LT DeVita VT Jr, Rosenberg SA), pp. 1280-1301. Lippincott Williams & Wilkins, Philadelphia.

Brabec V, Kasparkova J, Vrana O, Novakova O, Cox JW, Qu Y, Farrell N. 1999. DNA modifications by a novel bifunctional trinuclear platinum phase I anticancer agent. *Biochemistry* **38**: 6781-6790.

Brabec V, Leng M. 1993. DNA interstrand cross-links of trans-diamminedichloroplatinum(II) are preferentially formed between guanine and complementary cytosine residues. *Proceedings of the National Academy of Sciences of the United States of America* **90**: 5345-5349.

Brazda P, Krieger J, Daniel B, Jonas D, Szekeres T, Langowski J, Toth K, Nagy L, Vamosi G. 2014. Ligand binding shifts highly mobile retinoid X receptor to the chromatin-bound state in a coactivator-dependent manner, as revealed by single-cell imaging. *Mol Cell Biol* **34**: 1234-1245.

Bursova V, Kasparkova J, Hofr C, Brabec V. 2005. Effects of monofunctional adducts of platinum(II) complexes on thermodynamic stability and energetics of DNA duplexes. *Biophysical journal* **88**: 1207-1214.

Capoulade J, Wachsmuth M, Hufnagel L, Knop M. 2011. Quantitative fluorescence imaging of protein diffusion and interaction in living cells. *Nature biotechnology* **29**: 835-839.

Chen H, Parkinson JA, Parsons S, Coxall RA, Gould RO, Sadler PJ. 2002. Organometallic ruthenium(II) diamine anticancer complexes: arene-nucleobase stacking and stereospecific hydrogen-bonding in guanine adducts. *Journal of the American Chemical Society* **124**: 3064-3082.

Cho DS, Yang W, Lee JT, Shiekhattar R, Murray JM, Nishikura K. 2003. Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J Biol Chem* **278**: 17093-17102.

Cohen GL, Ledner JA, Bauer WR, Ushay HM, Caravana C, Lippard SJ. 1980. Sequence Dependent Binding of Cis-Dichlorodiammineplatinum(Ii) to DNA. *J Am Chem Soc* **102**: 2487-2488.

Corda Y, Job C, Anin MF, Leng M, Job D. 1993. Spectrum of DNA--platinum adduct recognition by prokaryotic and eukaryotic DNA-dependent RNA polymerases. *Biochemistry* **32**: 8582-8588.

Coste F, Malinge JM, Serre L, Shepard W, Roth M, Leng M, Zelwer C. 1999. Crystal structure of a double-stranded DNA containing a cisplatin interstrand cross-link at 1.63 A resolution: hydration at the platinated site. *Nucleic acids research* **27**: 1837-1846.

Court R, Chapman L, Fairall L, Rhodes D. 2005. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. *EMBO Rep* **6**: 39-45.

de Lange T. 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes & development* **19**: 2100-2110.

deLange T. 2010. How shelterin solves the telomere end-protection problem. *Cold Spring Harbor symposia on quantitative biology* **75**: 167-177.

Denchi EL, de Lange T. 2007. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448**: 1068-1071.

Dreyfuss G, Kim VN, Kataoka N. 2002. Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* **3**: 195-205.

Dvorackova M, Rossignol P, Shaw PJ, Koroleva OA, Doonan JH, Fajkus J. 2010. AtTRB1, a telomeric DNA-binding protein from Arabidopsis, is concentrated in the nucleolus and shows highly dynamic association with chromatin. *The Plant journal : for cell and molecular biology* **61**: 637-649.

Eastman A, Jennerwein MM, Nagel DL. 1988. Characterization of bifunctional adducts produced in DNA by trans-diamminedichloroplatinum(II). *Chemico-biological interactions* **67**: 71-80.

Fairall L, Chapman L, Moss H, de Lange T, Rhodes D. 2001. Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Molecular cell* **8**: 351-361.

Fanizzi FP, Intini FP, Maresca L, Natile G, Quaranta R. 1987. Biological-Activity of Platinum Complexes Containing Chiral Centers on the Nitrogen or Carbon-Atoms of a Chelate Diamine Ring. *Inorg Chim a-Bioinor* **137**: 45-51.

Farrell N. 2000. Polynuclear charged platinum compounds as a new calss of anticancer agents-toward a new paradigm. in *Platinum-based Drugs in Cancer Therapy* (ed. NPaK Farrell, L. R.), pp. 321-338. Humana Press, Totowa, NJ.

Fichtinger-Schepman AM, van der Veer JL, den Hartog JH, Lohman PH, Reedijk J. 1985. Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry* **24**: 707-713.

Gelasco A, Lippard SJ. 1998. NMR solution structure of a DNA dodecamer duplex containing a cis-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry* **37**: 9230-9239.

Ghersi D, Singh M. 2014. Interaction-based discovery of functionally important genes in cancers. *Nucleic Acids Res* **42**: e18.

Gifford G, Paul J, Vasey PA, Kaye SB, Brown R. 2004. The acquisition of hMLH1 methylation in plasma DNA after chemotherapy predicts poor survival for ovarian cancer patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* **10**: 4420-4426.

Hanahan D, Weinberg RA. 2000. The hallmarks of cancer. *Cell* **100**: 57-70.

Hanaoka S, Nagadoi A, Nishimura Y. 2005. Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities. *Protein Sci* **14**: 119-130.

Hanaoka S, Nagadoi A, Yoshimura S, Aimoto S, Li B, de Lange T, Nishimura Y. 2001. NMR structure of the hRap1 Myb motif reveals a canonical three-helix bundle lacking the positive surface charge typical of Myb DNA-binding domains. *J Mol Biol* **312**: 167-175.

Hobor F, Pergoli R, Kubicek K, Hrossova D, Bacikova V, Zimmermann M, Pasulka J, Hofr C, Vanacova S, Stefl R. 2011. Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (NAB) 3 protein. *J Biol Chem* **286**: 3645-3657.

Hofr C, Brabec V. 2001. Thermal and thermodynamic properties of duplex DNA containing site-specific interstrand cross-link of antitumor cisplatin or its clinically ineffective trans isomer. *The Journal of biological chemistry* **276**: 9655-9661.

Hofr C, 2005. Thermal stability and energetics of 15-mer DNA duplex interstrand crosslinked by trans-diamminedichloroplatinum(II). *Biopolymers* **77**: 222-229.

Hofr C, Farrell N, Brabec V. 2001. Thermodynamic properties of duplex DNA containing a site-specific d(GpG) intrastrand crosslink formed by an antitumor dinuclear platinum complex. *Nucleic acids research* **29**: 2034-2040.

Hofr C, Sultesova P, Zimmermann M, Mozgova I, Prochazkova Schrumpfova P, Wimmerova M, Fajkus J. 2009. Single-Myb-histone proteins from Arabidopsis thaliana: a quantitative study of telomere-binding specificity and kinetics. *The Biochemical journal* **419**: 221-228, 222 p following 228.

Holford J, Raynaud F, Murrer BA, Grimaldi K, Hartley JA, Abrams M, Kelland LR. 1998. Chemical, biochemical and pharmacological activity of the novel sterically hindered platinum co-ordination complex, cis-[amminedichloro(2-methylpyridine)] platinum(II) (AMD473). *Anti-cancer drug design* **13**: 1-18.

Holzer AK, Samimi G, Katano K, Naerdemann W, Lin X, Safaei R, Howell SB. 2004. The copper influx transporter human copper transport protein 1 regulates the uptake of cisplatin in human ovarian carcinoma cells. *Molecular pharmacology* **66**: 817-823.

Houghtaling BR, Cuttonaro L, Chang W, Smith S. 2004. A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Current Biology* **14**: 1621-1631.

Huang H, Zhu L, Reid BR, Drobny GP, Hopkins PB. 1995. Solution structure of a cisplatin-induced DNA interstrand cross-link. *Science* **270**: 1842-1845.

Hwang MG, Cho MH. 2007. Arabidopsis thaliana telomeric DNA-binding protein 1 is required for telomere length homeostasis and its Myb-extension domain stabilizes plant telomeric DNA binding. *Nucleic Acids Res* **35**: 1333-1342.

Ishida S, Lee J, Thiele DJ, Herskowitz I. 2002. Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc Natl Acad Sci U S A* **99**: 14298-14302.

Janoušková E, Nečasová I, Pavloušková J, Zimmermann M, Hluchý M, Marini V, Nováková M, Hofr C. 2014. Human Rap1 modulates TRF2 attraction to telomeric DNA.

Jung Y, Lippard SJ. 2007. Direct cellular responses to platinum-induced DNA damage. *Chemical reviews* **107**: 1387-1407.

Kabir S, Sfeir A, de Lange T. 2010. Taking apart Rap1: an adaptor protein with telomeric and non-telomeric functions. *Cell Cycle* **9**: 4061-4067.

Kasparkova J, Delalande O, Stros M, Elizondo-Riojas MA, Vojtiskova M, Kozelka J, Brabec V. 2003. Recognition of DNA interstrand cross-link of antitumor cisplatin by HMGB1 protein. *Biochemistry* **42**: 1234-1244.

Kašpárková J, Mellish KJ, Qu Y, Brabec V, Farrell N. 1996. Site-Specific d(GpG) Intrastrand Cross-Links Formed by Dinuclear Platinum Complexes. Bending and NMR Studies†. *Biochemistry* **35**: 16705-16713.

Kelland L. 2007. The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer* **7**: 573-584.

Kelland LR, Mistry P, Abel G, Freidlos F, Loh SY, Roberts JJ, Harrap KR. 1992. Establishment and characterization of an in vitro model of acquired resistance to cisplatin in a human testicular nonseminomatous germ cell line. *Cancer research* **52**: 1710-1716.

Kim SH, Kaminker P, Campisi J. 1999. TIN2, a new regulator of telomere length in human cells. *Nat Genet* **23**: 405-412.

Koberle B, Masters JR, Hartley JA, Wood RD. 1999. Defective repair of cisplatin-induced DNA damage caused by reduced XPA protein in testicular germ cell tumours. *Curr Biol* **9**: 273-276.

Kramara J, Willcox S, Gunisova S, Kinsky S, Nosek J, Griffith JD, Tomaska L. 2010. Tay1 protein, a novel telomere binding factor from Yarrowia lipolytica. *J Biol Chem* **285**: 38078-38092.

Kraus WL, Lis JT. 2003. PARP goes transcription. *Cell* **113**: 677-683.

Kuchar M. 2006. Plant telomere-binding proteins. *Biol Plantarum* **50**: 1-7.

Kuchar M, Fajkus J. 2004. Interactions of putative telomere-binding proteins in Arabidopsis thaliana: identification of functional TRF2 homolog in plants. *FEBS Lett* **578**: 311-315.

Levi F, Misset JL, Brienza S, Adam R, Metzger G, Itzakhi M, Caussanel JP, Kunstlinger F, Lecouturier S, Descorps-Declere A et al. 1992. A chronopharmacologic phase II clinical trial with 5-fluorouracil, folinic acid, and oxaliplatin using an ambulatory multichannel programmable pump. High antitumor effectiveness against metastatic colorectal cancer. *Cancer* **69**: 893-900.

Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB. 1992. Telomere end-replication problem and cell aging. *Journal of molecular biology* **225**: 951-960.

Maas S, Patt S, Schrey M, Rich A. 2001. Underediting of glutamate receptor GluR-B mRNA in malignant gliomas. *Proc Natl Acad Sci U S A* **98**: 14687-14692.

Mackereth CD, Madl T, Bonnal S, Simon B, Zanier K, Gasch A, Rybin V, Valcarcel J, Sattler M. 2011. Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* **475**: 408-411.

Mackereth CD, Sattler M. 2012. Dynamics in multi-domain protein recognition of RNA. *Current opinion in structural biology* **22**: 287-296.

Maeda Y, Nunomura K, Ohtsubo E. 1990. Differential scanning calorimetric study of the effect of intercalators and other kinds of DNA-binding drugs on the stepwise melting of plasmid DNA. *J Mol Biol* **215**: 321-329.

Malina J, Hofr C, Maresca L, Natile G, Brabec V. 2000. DNA interactions of antitumor cisplatin analogs containing enantiomeric amine ligands. *Biophysical journal* **78**: 2008-2021.

Malinge J-M, Leng M. 2006. Interstrand Cross-Links in Cisplatin- or Transplatin-Modified DNA. in *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug* (ed. PB Lippert), pp. 159-180. Verlag Helvetica Chimica Acta.

Marian CO, Bordoli SJ, Goltz M, Santarella RA, Jackson LP, Danilevskaya O, Beckstette M, Meeley R, Bass HW. 2003. The maize Single myb histone 1 gene, Smh1, belongs to a novel gene family and encodes a protein that binds telomere DNA repeats in vitro. *Plant physiology* **133**: 1336-1350.

Martindale. 2011. *The Complete Drug Reference*. Pharmaceutical Press, London.

McKeage MJ. 2007. Satraplatin in hormone-refractory prostate cancer and other tumour types: pharmacological properties and clinical evaluation. *Drugs* **67**: 859-869.

McKeage MJ, Raynaud F, Ward J, Berry C, O'Dell D, Kelland LR, Murrer B, Santabarabara P, Harrap KR, Judson IR. 1997. Phase I and pharmacokinetic study of an oral platinum complex given daily for 5 days in patients with cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **15**: 2691-2700.

Messias AC, Sattler M. 2004. Structural basis of single-stranded RNA recognition. *Accounts of chemical research* **37**: 279-287.

Morgan AS, Sanderson PE, Borch RF, Tew KD, Niitsu Y, Takayama T, Von Hoff DD, Izbicka E, Mangold G, Paul C et al. 1998. Tumor efficacy and bone marrow-sparing properties of TER286, a cytotoxin activated by glutathione S-transferase. *Cancer research* **58**: 2568-2575.

Mozgova I, Schrumpfova PP, Hofr C, Fajkus J. 2008a. Functional characterization of domains in AtTRB1, a putative telomere-binding protein in Arabidopsis thaliana. *Phytochemistry* **69**: 1814-1819.

Novakova O, Chen H, Vrana O, Rodger A, Sadler PJ, Brabec V. 2003. DNA interactions of monofunctional organometallic ruthenium(II) antitumor complexes in cell-free media. *Biochemistry* **42**: 11544-11554.

Novakova O, Kasparkova J, Bursova V, Hofr C, Vojtiskova M, Chen H, Sadler PJ, Brabec V. 2005. Conformation of DNA modified by monofunctional Ru(II) arene complexes: recognition by DNA binding proteins and repair. Relationship to cytotoxicity. *Chemistry & biology* **12**: 121-129.

O'Connor MS, Safari A, Xin HW, Liu D, Songyang Z. 2006. A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 11874-11879.

Ohga T, Koike K, Ono M, Makino Y, Itagaki Y, Tanimoto M, Kuwano M, Kohno K. 1996. Role of the human Y box-binding protein YB-1 in cellular sensitivity to the DNA-damaging agents cisplatin, mitomycin C, and ultraviolet light. *Cancer Res* **56**: 4224-4228.

Ohndorf UM, Rould MA, He Q, Pabo CO, Lippard SJ. 1999. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature* **399**: 708-712.

Palm W, de Lange T. 2008. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**: 301-334.

Paquet F, Boudvillain M, Lancelot G, Leng M. 1999. NMR solution structure of a DNA dodecamer containing a transplatin interstrand GN7-CN3 cross-link. *Nucleic acids research* **27**: 4261-4268.

Peska V, Schrumpfova PP, Fajkus J. 2011. Using the Telobox to Search for Plant Telomere Binding Proteins. *Curr Protein Pept Sc* **12**: 75-83.

Pietras RJ, Fendly BM, Chazin VR, Pegram MD, Howell SB, Slamon DJ. 1994. Antibody to HER-2/neu receptor blocks DNA repair after cisplatin in human breast and ovarian cancer cells. *Oncogene* **9**: 1829-1838.

Pilch DS, Dunham SU, Jamieson ER, Lippard SJ, Breslauer KJ. 2000. DNA sequence context modulates the impact of a cisplatin 1,2-d(GpG) intrastrand cross-link on the conformational and thermodynamic properties of duplex DNA. *J Mol Biol* **296**: 803-812.

Plumb JA, Strathdee G, Sludden J, Kaye SB, Brown R. 2000. Reversal of drug resistance in human tumor xenografts by 2'-deoxy-5-azacytidine-induced demethylation of the hMLH1 gene promoter. *Cancer research* **60**: 6039-6044.

Poklar N, Pilch DS, Lippard SJ, Redding EA, Dunham SU, Breslauer KJ. 1996. Influence of cisplatin intrastrand crosslinking on the conformation, thermal stability, and energetics of a 20-mer DNA duplex. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 7606-7611.

Raynaud FI, Boxall FE, Goddard PM, Valenti M, Jones M, Murrer BA, Abrams M, Kelland LR. 1997. cis-Amminedichloro(2-methylpyridine) platinum(II) (AMD473), a novel sterically hindered platinum complex: in vivo activity, toxicology, and pharmacokinetics in mice. *Clinical cancer research : an official journal of the American Association for Cancer Research* **3**: 2063-2074.

Regad F, Lebas M, Lescure B. 1994. Interstitial Telomeric Repeats within the Arabidopsis-Thaliana Genome. *Journal of Molecular Biology* **239**: 163-169.

Sandler A, Gray R, Perry MC, Brahmer J, Schiller JH, Dowlati A, Lilenbaum R, Johnson DH. 2006. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *The New England journal of medicine* **355**: 2542-2550.

Schrumpfova P, Kuchar M, Fajkus J. 2005. Analysis of plant telomere-binding proteins. *Febs J* **272**: 77-77.

Schrumpfova P, Kuchar M, Mikova G, Skrisovska L, Kubicarova T, Fajkus J. 2004. Characterization of two Arabidopsis thaliana myb-like proteins showing affinity to telomeric DNA sequence. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **47**: 316-324.

Schrumpfova PP, Kuchar M, Palecek J, Fajkus J. 2008. Mapping of interaction domains of putative telomere-binding proteins AtTRB1 and AtPOT1b from Arabidopsis thaliana. *FEBS letters* **582**: 1400-1406.

Schrumpfova PP, Vychodilova I, Dvorackova M, Majerska J, Dokladal L, Schorova S, Fajkus J. 2014. Telomere repeat binding proteins are functional components of Arabidopsis telomeres and interact with telomerase. *Plant Journal* **77**: 770-781.

Sfeir A, de Lange T. 2012. Removal of Shelterin Reveals the Telomere End-Protection Problem. *Science* **336**: 593-597.

Shakirov EV, Surovtseva YV, Osbun N, Shippen DE. 2005. The Arabidopsis Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. *Mol Cell Biol* **25**: 7725-7733.

Sharp SY, O'Neill CF, Rogers P, Boxall FE, Kelland LR. 2002. Retention of activity by the new generation platinum agent AMD0473 in four human tumour cell lines possessing acquired resistance to oxaliplatin. *European journal of cancer* **38**: 2309-2315.

Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB et al. 2010. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell* **143**: 225-237.

Stefl R, Skrisovska L, Allain FH. 2005. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep* **6**: 33-38.

Stefl R, Xu M, Skrisovska L, Emeson RB, Allain FH. 2006. Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure* **14**: 345-355.

Sun D, Luo M, Jeong M, Rodriguez B, Xia Z, Hannah R, Wang H, Le T, Faull KF, Chen R et al. 2014. Epigenomic Profiling of Young and Aged HSCs Reveals Concerted Changes during Aging that Reinforce Self-Renewal. *Cell stem cell* **14**: 673-688.

Takahara PM, Rosenzweig AC, Frederick CA, Lippard SJ. 1995. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature* **377**: 649-652.

Todd RC. 2010. Structural and functional consequences of platinum anticancer drug binding to free and nucleosomal DNA. in *Dept of Chemistry*. Massachusetts Institute of Technology, Boston.

Tomas-Loba A, Flores I, Fernandez-Marcos PJ, Cayuela ML, Maraver A, Tejera A, Borras C, Matheu A, Klatt P, Flores JM et al. 2008. Telomerase reverse transcriptase delays aging in cancer-resistant mice. *Cell* **135**: 609-622.

Turner N, Tutt A, Ashworth A. 2005. Targeting the DNA repair defect of BRCA tumours. *Current opinion in pharmacology* **5**: 388-393.

Twelves MW, Ann B, Roy AJS, Christopher. 2006. *Oncology*. OUP Oxford.

Valente L, Nishikura K. 2007. RNA binding-independent dimerization of adenosine deaminases acting on RNA and dominant negative effects of nonfunctional subunits on dimer functions. *J Biol Chem* **282**: 16054-16061.

Visacka K, Hofr C, Willcox S, Necasova I, Pavlouskova J, Sepsiova R, Wimmerova M, Simonicova L, Nosek J, Fajkus J et al. 2012. Synergism of the two Myb domains of Tay1 protein results in high affinity binding to telomeres. *J Biol Chem* **287**: 32206-32215.

Wang D, Lippard SJ. 2005. Cellular processing of platinum anticancer drugs. *Nature reviews Drug discovery* **4**: 307-320.

Wheate NJ, Walker S, Craig GE, Oun R. 2010. The status of platinum anticancer drugs in the clinic and in clinical trials. *Dalton transactions* **39**: 8113-8127.

Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW. 1996. Telomerase activity in human germline and embryonic tissues and cells. *Developmental genetics* **18**: 173-179.

Wu Y, Bhattacharyya D, King CL, Baskerville-Abraham I, Huh SH, Boysen G, Swenberg JA, Temple B, Campbell SL, Chaney SG. 2007. Solution structures of a DNA dodecamer duplex with and without a cisplatin 1,2-d(GG) intrastrand cross-link: comparison with the same DNA duplex containing an oxaliplatin 1,2-d(GG) intrastrand cross-link. *Biochemistry* **46**: 6477-6487.

Ye JZ, Donigian JR, van Overbeek M, Loayza D, Luo Y, Krutchinsky AN, Chait BT, de Lange T. 2004a. TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J Biol Chem* **279**: 47264-47271.

Ye JZS, Hockemeyer D, Krutchinsky AN, Loayza D, Hooper SM, Chait BT, de Lange T. 2004b. POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Gene Dev* **18**: 1649-1654.

Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang HW et al. 2013. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* **504**: 168-171.

Zaludova R, Zakovska A, Kasparkova J, Balcarova Z, Kleinwachter V, Vrana O, Farrell N, Brabec V. 1997. DNA interactions of bifunctional dinuclear platinum(II) antitumor agents. *European journal of biochemistry / FEBS* **246**: 508-517.

Zhang CX, Chang PV, Lippard SJ. 2004. Identification of nuclear proteins that interact with platinum-modified DNA by photoaffinity labeling. *Journal of the American Chemical Society* **126**: 6536-6537.

Zhong Z, Shiue L, Kaplan S, de Lange T. 1992. A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol Cell Biol* **12**: 4834-4843.

Zimmermann M, Lottersberger F, Buonomo SB, Sfeir A, de Lange T. 2013. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science* **339**: 700-704.

Summary

This habilitation work summarizes research studies focused on nucleic acid interactions which are in close relation to the stability of the genome. The habilitation thesis text is divided into three main parts.

The first part is dedicated to the description of changes in structure and thermodynamics of DNA induced by the binding of new and clinically used anticancer drugs derived from metal compounds.

The second part describes interaction of telomeric proteins that selectively bind DNA on chromosomes ends. This part is focused on summarizing new results that contributed to understanding the mechanism of selective protein recognition of telomeric DNA. The conclusion from studies of binding of telomeric proteins from different species (human, yeast, plant) are presented.

In the final part, it is demonstrated how functional binding studies and molecular structure data are combined to obtain a more comprehensive and detailed view of mechanism of double-stranded RNA recognition.

This habilitation thesis summarizes the candidate's professional effort to reveal the molecular mechanism of anticancer therapeutic effects of drug binding to DNA and to understand roles of nucleic acids interactions in formations and functions of essential nucleoprotein complexes.

Abstrakt

Tato habilitační práce shrnuje výzkumné studie zaměřené na interakce nukleových kyselin, které jsou v úzkém vztahu ke stabilitě genomu. Text habilitační práce je rozdělen do tří hlavních částí.

První část se věnuje popisu změny struktury a termodynamiky DNA, které jsou způsobené vazbou nově navrhnutých a klinicky používaných protinádorových léčiv odvozených od komplexů těžkých kovů.

Druhá část popisuje interakce telomerových proteinů, které se selektivně vážou a chrání konce chromosomů. Tato část práce je zaměřena na shrnutí nových poznatků, které přispěly k pochopení mechanismu rozpoznávání telomerové DNA těmito proteiny. Závěry vyplývající ze studií mechanismu vazby telomerových proteinů původu lidského, rostlinného a kvasinkového jsou představeny.

V závěrečné části habilitační práce je ukázáno, jak jsou funkční vazebné studie a strukturní data kombinovány k detailnějšímu pochopení obrazu mechanismu rozpoznání dvoušroubovicové RNA.

Tato habilitační práce shrnuje odbornou snahu kandidáta o poznání a popis molekulárního mechanismu protinádorového léčebného účinků léčiv vázajících se na DNA a pochopení rolí jednotlivých interakcí nukleových kyselin, které jsou podstatné pro vznik a funkci životně důležitých nukleoproteinových komplexů.

List of publications included in this habilitation work

1. Malina J, **Hofr C**, Maresca L, Natile G and Brabec V. DNA interactions of antitumor cisplatin analogs containing enantiomeric amine ligands. *Biophysical Journal.* 2000. 78, 2008-2021.

2. Nováková O, **Hofr C**, Brabec V. Modification of natural, double-helical DNA by antitumor cis- and trans-[Cl2(Me2SO)4Ru] in cell-free media. *Biochem. Pharmacol.* 2000. 60:1761- 1771.

3. **Hofr C.** and V. Brabec. Thermal and thermodynamic properties of duplex DNA containing site-specific interstrand cross-link of antitumor cisplatin or its clinically ineffective trans isomer. *J.Biol. Chem.* 2001. 276: 9655-9661.

4. **Hofr C**, Farrell N, Brabec V. Thermodynamic properties of duplex DNA containing site-specific d(GpG) intrastrand cross-link of antitumor dinuclear platinum komplex *Nucleic Acids Research*. 2001. 29:2034-40.

5. Bursova V,Kasparkova J, **Hofr C**, Brabec V. Effects of Monofunctional Adducts of Platinum(II) Complexes on Thermodynamic Stability and Energetics of DNA Duplexes. *Biophysical Journal.* 2005. 88:1207-14.

6. Novakova O, Kasparkova J, Bursova V, **Hofr C**, Vojtiskova M, Chen H, Sadler PJ, Brabec V. Conformation of DNA modified by monofunctional Ru(II) arene complexes: recognition by DNA binding proteins and repair. Relationship to cytotoxicity. *Chem. Biol.* 2005. 12:121-9.

7. **Hofr C**, Brabec V. Thermal stability, energetics and melting cooperativity of DNA interstrand cross-linked by trans-diamminedichloroplatinum(II). *Biopolymers.* 2005 77: 222-9.

8. **Hofr C***, Mikrokalorimetrie biologicky významných molekul. Československý časopis pro fyziku, 2006. 56: 288-292.

9. Mozgová I, Schrumpfová PP, **Hofr C**, Fajkus J. Functional characterization of domains in AtTRB1, a putative telomere-binding protein in Arabidopsis thaliana. *Phytochemistry*. 2008;69:1814-9.

10. **Hofr C***, Šultesová P, Zimmermann M, Mozgová I, Procházková Schrumpfová P, Wimmerová M, Fajkus J. Single-Myb-histone proteins from Arabidopsis thaliana: a quantitative study of telomere-binding specificity and kinetics. *Biochem J.* 2009;419:221-8.

11. Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, **Hofr C**, Emeson RB, Allain FH. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell.* 2010;143:225-37.

12. Hobor F, Pergoli R, Kubicek K, Hrossova D, Bacikova V, Zimmermann M, Pasulka, J, **Hofr C**, Vanacova S, Stefl R. Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (Nab)3 protein. *J Biol Chem.* 2010; 286: 3645-3657.

13. Visacka K, **Hofr C**, Willcox S, Necasova I, Pavlouskova J, Sepsiova R, Wimmerova M, Simonicova S, Nosek J, Fajkus J, Griffith JD, and Tomaska L. Synergism of the two Myb domains of Tay1 protein results in high-affinity binding to telomeres. *J Biol Chem.* 2012; 287:32206-15.

14. Kubicek K, Cerna H, Holub P, Pasulka J, Hrossova D, Loehr F, **Hofr C**, Vanacova S, and Stefl R. A Phosphoserine-Proline Switch in RNAP II CTD Controls Binding of an RNA Processing Factor. *Genes and Development*. 2012; 26:1891-6.

15. Buchlovič M, Kříž Z, Hofr C, Potáček M. New PAH derivatives functionalized by cyclic nitrone framework: synthetic design, anti-proliferative activity and interaction with DNA. *Bioorg Med Chem.* 2013. 21:1078-81.

16. Janoušková E, Nečasová I, Pavloušková J, Zimmermann M, Hluchý M, Marini V, Nováková M, and **Hofr C*** Human Rap1 modulates TRF2 attraction to telomeric DNA. Under review in *Nucleic Acids Research*.

*corresponding author

DNA Interactions of Antitumor Cisplatin Analogs Containing Enantiomeric Amine Ligands

Jaroslav Malina,* Ctirad Hofr,* Luciana Maresca,† Giovanni Natile,† and Viktor Brabec*
*Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic, and †Dipartimento Farmaco-Chimico, University of Bari, I-70125 Bari, Italy

ABSTRACT Modifications of natural DNA and synthetic oligodeoxyribonucleotide duplexes in a cell-free medium by analogs of antitumor cisplatin containing enantiomeric amine ligands, such as *cis*-[$PtCl_2$(RR-DAB)] and *cis*-[$PtCl_2$(SS-DAB)] (DAB = 2,3-diaminobutane), were studied by various methods of molecular biophysics and biophysical chemistry. These methods include DNA binding studies by pulse polarography and atomic absorption spectrophotometry, mapping of DNA adducts using transcription assay, interstrand cross-linking assay using gel electrophoresis under denaturing conditions, differential scanning calorimetry, chemical probing, and bending and unwinding studies of the duplexes containing single, site-specific cross-link. The major differences resulting from the modification of DNA by the two enantiomers are the thermodynamical destabilization and conformational distortions induced in DNA by the 1,2-d(GpG) intrastrand cross-link. It has been suggested that these differences are associated with a different biological activity of the two enantiomers observed previously. In addition, the results of the present work are also consistent with the view that formation of hydrogen bonds between the carbonyl oxygen of the guanine residues and the "quasi equatorial" hydrogen of the *cis* amine in the 1,2-d(GpG) intrastrand cross-link plays an important role in determining the character of the distortion induced in DNA by this lesion.

INTRODUCTION

Since the discovery of its anticancer activity, several new analogs of cisplatin [*cis*-diamminedichloroplatinum(II), *cis*-$PtCl_2(NH_3)_2$] have been synthesized and tested for biological activity. Some of these compounds are now considered potent anticancer drugs (Pasini and Zunino, 1987; Bloemink and Reedijk, 1996; Reedijk, 1996; O'Dwyer et al., 1999). Although the precise mechanism of antitumor action of platinum drugs is not completely understood, they are known to target DNA primarily by forming bifunctional adducts (Pinto and Lippard, 1985; Johnson et al., 1989). The anticancer activity displayed by cisplatin and its analogs is usually attributed to a unique type of intrastrand d(GpG) adduct with platinum cross-linking N7 atoms of neighboring guanine residues of DNA. It has been also shown that carrier amine ligands of cisplatin analogs appear to modulate the antitumor properties of this class of drugs. The antitumor activity is usually lost or diminished if the primary or secondary amines on platinum are replaced by tertiary amines (Sundquist and Lippard, 1990).

There are many possible roles for the carrier ligand of the platinum antitumor compounds. Hydrogen bonding between DNA and the carrier ligand could affect the initial attack of DNA by the drug and the type of DNA cross-linking (intra or interstrand). It is also reasonable to expect that the direction (5′ or 3′) of closure of monofunctional DNA adducts (formed in the first step of their binding to DNA) into cross-links is affected by hydrogen bonding. In addition, the carrier ligand may also affect biodistribution, and recognition of DNA adducts by repair enzymes, regulatory and/or DNA-binding proteins.

The biological activity of platinum complexes with enantiomeric amine ligands such as *cis*-[$PtCl_2$(RR-DACH)] and *cis*-[$PtCl_2$(SS-DACH)] (DACH = 1,2-diaminocyclohexane) and other enantiomeric pairs has been intensively investigated (Kidani et al., 1978; Noji et al., 1981, 1983; Coluccia et al., 1986, 1991; Fanizzi et al., 1987; Pasini and Zunino, 1987; Giannini and Natile, 1991; Vickery et al., 1993; Fenton et al., 1997). For instance, the DACH carrier ligand has been shown to significantly affect the ability of platinum-DNA adducts to block essential processes such as replication and transcription (Page et al., 1990). Also importantly, *cis*-[$PtCl_2$(N-N)] complexes with N-N = DACH or 1,2-diaminopropane (DAB) having an S configuration at the asymmetric carbon atoms were markedly more mutagenic toward several strains in *Salmonella typhimurium* than their R isomers (Fanizzi et al., 1987). Hence, although the asymmetry in the amine ligand in these platinum complexes did not involve the coordinated nitrogen atom, but rather an adjacent carbon atom, a dependence of the biological activity on the configuration of the amine was observed.

The major DNA adduct of cisplatin and its analogs is an intrastrand d(GpG) cross-link (Sherman and Lippard, 1987; Bloemink and Reedijk, 1996). It has been demonstrated that this cross-link adopts an *anti*, *anti* head-to-head (HH) conformation in both single- and double-stranded DNA. However, it has been speculated in numerous reports that this conformer equilibrates with other forms that interconvert too rapidly for separate characterization by NMR spectros-

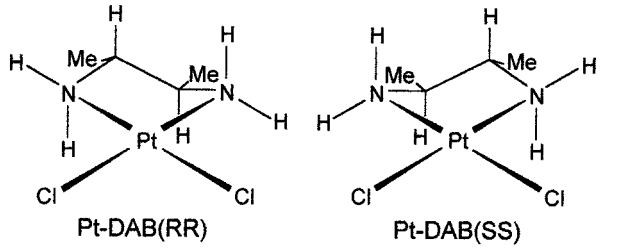
Received for publication 7 September 1999 and in final form 30 December 1999.

Address reprint requests to Dr. Viktor Brabec, Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, CZ-61265 Brno, Czech Republic. Tel.: 420-5-41517148; Fax: 420-5-41211293; E-mail: brabec@ibp.cz; URL: http://www.ibp.cz.

0006-3495/00/04/2008/14 $2.00

copy (Den Hartog et al., 1982; Neumann et al., 1984; Kline et al., 1989; Mukundan et al., 1991; Berners-Price et al., 1996, 1997; Van Boom et al., 1996). The dynamic character of the platinum-DNA adduct makes establishing a correlation between its stereochemistry and the configuration of the carrier ligand difficult.

In the present work modifications of DNA by *cis*-[$PtCl_2$(DAB)] enantiomers (Fig. 1) in cell-free media were investigated by using various techniques of molecular biophysics. The goal of these studies was to contribute to understanding how the chirality at the carbon atoms of the carrier ligand in cisplatin analogs can affect their biological activity. *cis*-[$PtCl_2$(DAB)] isomers were chosen for the studies described in the present paper as the representatives of platinum drugs with enantiomeric amine ligands because the effect of different chirality at the carbon atoms on the biological activity was most pronounced in the case of these compounds. The effect of the configuration of the carrier diamine in *cis*-[$PtCl_2$(DAB)] complexes with R, R and S, S configurations at the asymmetric carbons [these complexes are abbreviated as Pt-DAB(RR) and Pt-DAB(SS), respectively] was investigated. A schematic representation of the puckering of the chelate rings in Pt-DAB(RR) or Pt-DAB(SS) isomers is shown in Fig. 1.



d(TGGT)d(ACCA)(20-DSC): 5′-TCTCCTTCTGGTCTCTTCTC-3′
3′-AGAGGAAGACCAGAGAAGAG-5′

d(TGGT)d(ACCA)(20): 5′-CCTCTCCTTGGTCTCCTTCT-3′
3′- GAGAGGAACCAGAGGAAGAG-5′

d(TGGT)d(ACCA)(21): 5′-CCTCTCCTTGGTCTCCTTCTC-3′
3′- GAGAGGAACCAGAGGAAGAGG-5′

d(TGGT)d(ACCA)(22): 5′-CCTCTCCTTGGTCTCCTTCTCT-3′
3′- GAGAGGAACCAGAGGAAGAGAG-5′

d(TGGT)d(ACCA)(23): 5′-CCTCTCCTTGGTCTCCTTCTCTC-3′
3′- GAGAGGAACCAGAGGAAGAGAGG-5′

FIGURE 1 Structures of Pt-DAB(RR) and Pt-DAB(SS), and sequences of the synthetic oligodeoxyribonucleotides used in the present study with their abbreviations. The top and bottom strands of each pair are designated top and bottom, respectively, in the text. The bold letters in top strands indicate the location of the intrastrand cross-link after modification of the oligonucleotides by Pt-DAB complexes in the way described in the Experimental section.

MATERIALS AND METHODS

Starting materials

Cisplatin was synthesized and characterized in Lachema (Brno, Czech Republic). Pt-DAB(RR) and Pt-DAB(SS) complexes were prepared and characterized as described previously (Fanizzi et al., 1987). The stock solutions of the platinum complexes (5×10^{-4} M in 10 mM $NaClO_4$) were prepared in the dark at 25°C. Calf thymus (CT) DNA (42% G + C, mean molecular mass $\sim 2 \times 10^7$ Da) was also prepared and characterized as described previously (Brabec and Paleček, 1970, 1976). Plasmid pSP73KB [2455 bp (Lemaire et al., 1991)] was isolated according to standard procedures and banded twice in CsCl/EtBr equilibrium density gradients. The synthetic oligodeoxyribonucleotides were synthesized and purified as described previously (Brabec et al., 1992). Restriction endonucleases, T4 DNA ligase, Klenow fragment of DNA polymerase I, and T4 polynucleotide kinase were purchased from New England Biolabs (Beverly, MA). Riboprobe Gemini System II for transcription mapping containing SP6 and T7 RNA polymerases was purchased from Promega (Madison, WI). Ethidium bromide, agarose, acrylamide, bis(acrylamide), and NaCN were from Merck KgaA (Darmstadt, Germany). Dimethyl sulfate (DMS), $KMnO_4$, diethylpyrocarbonate (DEPC), KBr, and $KHSO_5$ were from Sigma, Prague. [γ-^{32}P]ATP and [α-^{32}P]dATP were from Amersham (Arlington Heights, IL).

Platination reactions

CT DNA and plasmid DNAs were incubated with the platinum complex in 10 mM $NaClO_4$ at 37°C for 48 h in the dark if not stated otherwise. The number of molecules of the platinum compound bound per nucleotide residue (r_b values) were determined by flameless atomic absorption spectrophotometry (FAAS) or by differential pulse polarography (DPP) (Kim et al., 1990). The oligonucleotides were allowed to react with the platinum compounds, and repurified as described previously (Brabec et al., 1992). Briefly, the oligonucleotides synthesized on an Applied Biosystems solid-phase synthesizer were purified by ion-exchange FPLC. The single-stranded oligonucleotides (the top strands in Fig. 1) were reacted in stoichiometric amounts with Pt-DAB(RR) or Pt-DAB(SS). The platinated oligonucleotides were purified by FPLC. It was verified by platinum FAAS and by the measurements of the optical density that the modified oligonucleotides contained one platinum atom. It was also verified using DMS footprinting of platinum on DNA (Lemaire et al., 1991; Brabec and Leng, 1993) that in the platinated top strands the N7 position of both neighboring guanines was not accessible for reaction with DMS. Briefly, platinated and nonmodified top strands (5′-end-labeled with ^{32}P) were reacted with DMS. DMS methylates the N7 position of guanine residues in DNA, producing alkali-labile sites (Maxam and Gilbert, 1980). However, if N7 is covalently bound to platinum, it cannot be methylated. The oligonucleotides were then treated with hot piperidine and analyzed by denaturing polyacrylamide gel electrophoresis. For the nonmodified oligonucleotides, shortened fragments due to the cleavage of the strand at the two methylated guanine residues were observed in the gel. However, no such bands were detected for the platinated oligonucleotides. These results indicate that one Pt-DAB molecule was coordinated to neighboring guanine residues, forming the 1,2-d(GpG) intrastrand cross-link. The platinated strands were allowed to anneal with nonplatinated complementary strands (the bottom strands in Fig. 1) in 50 mM NaCl plus 1 mM Tris · HCl with 0.1 mM EDTA, pH 7.4. FPLC purification and FAAS measurements were carried out on a Pharmacia Biotech FPLC System with a MonoQ HR 5/5 column and a Unicam 939 AA spectrometer equipped with a graphite furnace, respectively.

DNA transcription by RNA polymerases in vitro

Transcription of the (*NdeI*/*Hpa*I) restriction fragment of pSP73KB DNA with SP6 and T7 RNA polymerases and electrophoretic analysis of tran-

scripts was performed according to the protocols recommended by Promega [Promega Protocols and Applications, 43-46 (1989/90)] and previously described in detail (Lemaire et al., 1991; Brabec and Leng, 1993).

Interstrand cross-link assay

If not stated otherwise, Pt-DAB(RR) or Pt-DAB(SS) at varying concentrations were incubated with 2 μg pSP73KB DNA after it had been linearized by *Eco*RI. The platinated samples were precipitated by ethanol and the linear duplexes were then analyzed for DNA interstrand cross-links in the same way as described in several recent papers (Farrell et al., 1990; Lemaire et al., 1991; Brabec and Leng, 1993). The linear duplexes were first 3′-end-labeled by means of a Klenow fragment of DNA polymerase I and [α-^{32}P]dATP. The samples were deproteinized by phenol, precipitated by ethanol, and the pellet was dissolved in 18 μl of 30 mM NaOH with 1 mM EDTA, 6.6% sucrose, and 0.04% bromophenol blue. The amount of interstrand cross-links was analyzed by electrophoresis under denaturing conditions on alkaline agarose gel (1.5%). After the electrophoresis was completed, the intensities of the bands corresponding to single strands of DNA and interstrand cross-linked duplex were quantified by means of a Molecular Dynamics Phosphor Imager (Storm 860 system with ImageQuant software). The frequency of interstrand cross-links, F (the number of interstrand cross-links per adduct), was calculated as $F = XL/4910 \cdot r_b$ (pSP73KB plasmid contained 4910 nucleotide residues). XL is the number of interstrand cross-links per one molecule of the linearized DNA duplex which was calculated assuming Poisson distribution of the interstrand cross-links as $XL = -\ln A$, where A is the fraction of molecules running as a band corresponding to the non-cross-linked DNA (Farrell et al., 1990).

Differential scanning calorimetry

Excess heat capacity (ΔC_p) versus temperature profiles for the thermally induced transitions of d(TGGT)/d(ACCA)(20-DSC) duplex unmodified or containing a unique 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) were measured by using a VP-DSC Calorimeter (Microcal, Northampton, MA). In these experiments the heating rate was 60°C/h. Transition enthalpies (ΔH) and entropies (ΔS) were calculated from the areas under the experimental ΔC_p versus T and the derived $\Delta C_p/T$ versus T curves, respectively, by using the ORIGIN version 4.1 software (Microcal, Northampton, MA). The oligonucleotide duplexes at the concentration of 5 μM were dissolved in the buffer containing 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA. The samples were vacuum-degassed before the measurement. The formation of 1:1 complexes between the top and bottom strand of d(TGGT)/d(ACCA) nonmodified or containing the cross-link was verified by recording UV absorbance mixing curves at 25°C (Poklar et al., 1996). It was also verified in the same way as described in the previous paper (Poklar et al., 1996) that the melting transition of both the platinated and nonmodified duplexes were fully reversible.

Chemical modifications

The chemical probing of the conformation of the platinated oligonucleotide duplexes with the aid of NaCN was performed as described previously (Schwartz et al., 1990; Boudný et al., 1992). The top strand of the d(TGGT)/d(ACCA)(20) nonmodified or containing the Pt-DAB intrastrand cross-link was 5′-end-labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase before it was annealed with its complementary (bottom) nonlabeled strand. The oligonucleotide duplexes were treated with 0.2 M NaCN in 20 mM Tris · HCl, pH 8.3, and control nonmodified samples were run on a denaturing 24% polyacrylamide/8M urea gel.

The modifications by $KMnO_4$, DEPC, and $KBr/KHSO_5$ were also performed as described previously (Brabec et al., 1993; Bailly et al., 1994; Ross and Burrows, 1996; Bailly and Waring, 1997). The top or bottom strand of the d(TGGT)/d(ACCA)(20) was 5′-end-labeled with [γ-^{32}P]ATP before it was annealed with its complementary nonlabeled strands. In the case of the platinated oligonucleotides, platinum was removed after reaction of the DNA with the probe by incubation with 0.2 M NaCN (alkaline pH) at 45°C for 10 h in the dark.

Ligation and electrophoresis of oligonucleotides

Nonplatinated single strands (bottom strands in Fig. 1) were 5′-end-labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase. Then they were annealed with their phosphorylated complementary strands [nonplatinated or containing 1,2-d(GpG) intrastrand cross-links of Pt-DAB compounds]. Nonplatinated and intrastrand cross-link containing duplexes were allowed to react with T4 DNA ligase. The resulting samples along with ligated nonplatinated duplexes were subsequently examined on 8% native polyacrylamide [mono:bis(acrylamide) ratio = 29:1] electrophoresis gels. Other details of these experiments were as described in previously published papers (Koo et al., 1986; Bellon and Lippard, 1990).

RESULTS

DNA binding

Solutions of double-helical CT DNA at a concentration of 32 μg/ml were incubated with Pt-DAB(RR) or Pt-DAB(SS) at r_i values of 0.01 in 10 mM $NaClO_4$ at 37°C (r_i is defined as the molar ratio of free platinum complex to nucleotide phosphates at the onset of incubation with DNA). At various time intervals an aliquot of the reaction mixture was withdrawn and assayed by DPP for the amount of platinum bound to DNA (r_b) (Kim et al., 1990). The amount of platinum coordinated to DNA increased with time (not shown). After ~24 h, all molecules of Pt-DAB(RR) or Pt-DAB(SS) present in the reaction mixtures were coordinated to DNA [exhaustive dialysis of the samples of DNA treated with Pt-DAB(RR) or Pt-DAB(SS) against platinum-free background solution (10 mM $NaClO_4$) did not affect the amount of the platinum bound to DNA]. In these binding reactions both enantiomers coordinated to DNA with approximately the same rate, which indicates that isomerism in the non-leaving ligand of Pt-DAB compounds does not significantly affect the rate of the coordination of platinum moiety to natural double-helical DNA. The binding of Pt-DAB(RR) or Pt-DAB(SS) to CT DNA was also quantified in the following way. Aliquots of the reaction withdrawn at various time intervals were quickly cooled on an ice bath and then exhaustively dialyzed against 10 mM $NaClO_4$ at 4°C to remove free (unbound) platinum compound. The content of platinum in these samples was determined by FAAS. Results identical to those obtained using the DPP assay were obtained. Thus, DNA binding of Pt-DAB compounds resulted within <24 h in their complete coordination, which made it possible to prepare easily and precisely the samples of natural DNAs or their fragments modified by these compounds at a preselected r_b value.

In vitro transcription of DNA containing platinum adducts

In vitro RNA synthesis by RNA polymerases on DNA templates containing several types of bifunctional adducts of platinum complexes can be prematurely terminated at the level or in the proximity of adducts (Corda et al., 1991, 1992; Lemaire et al., 1991; Brabec and Leng, 1993; Brabec et al., 1994; Nováková et al., 1995; Žaludová et al., 1997). Importantly, monofunctional DNA adducts of several platinum complexes are unable to terminate RNA synthesis.

Cutting of pSP73KB DNA (Lemaire et al., 1991) by *Nde*I and *Hpa*I restriction endonucleases yielded a 212-bp fragment (a substantial part of its nucleotide sequence is shown in Fig. 2 *B*). This fragment contained convergent T7 and SP6 RNA polymerase promoters [in the upper and lower strands, respectively, close to its 3′-ends (Fig. 2 *B*)]. The experiments were carried out using this linear DNA fragment, modified by Pt-DAB(RR), Pt-DAB(SS), or cisplatin at $r_b = 0.01$, for RNA synthesis by T7 and SP6 RNA polymerases (Fig. 2 *A*, lanes RR, SS or cisDDP, respectively). RNA synthesis on the template modified by the platinum complexes yielded fragments of defined sizes, which indicates that RNA synthesis on these templates was prematurely terminated. The major stop sites produced by both Pt-DAB complexes were identical to those produced by cisplatin, and the corresponding bands produced by both enantiomers and cisplatin on the autoradiogram had similar intensity. The sequence analysis revealed that the major bands resulting from termination of RNA synthesis by the adducts of Pt-DAB(RR), Pt-DAB(SS), and cisplatin were identical and appeared at G sites and in a considerably less extent at A sites. These G and A sites were mostly contained in GG or AG sites, which are preferential DNA binding sites for untargeted cisplatin. Taken together, the results of the transcription mapping experiments suggest that base sequence selectivity of Pt-DAB(RR), Pt-DAB(SS), and cisplatin are similar.

Interstrand cross-linking

The experiments of the present work were carried out to compare the amounts of the interstrand cross-links formed by Pt-DAB(RR) or Pt-DAB(SS) in linear DNAs. We used in these experiments pSP73KB plasmid (2455 bp) modified by Pt-DAB complexes after it had been linearized by *Eco*RI (*Eco*RI cuts only once within pSP73KB plasmid). The samples were analyzed for the interstrand cross-links by agarose gel electrophoresis under denaturing conditions.

An electrophoretic method for precise and quantitative determination of interstrand cross-linking by platinum complexes in DNA was described previously (Farrell et al., 1990; Lemaire et al., 1991; Brabec and Leng, 1993). Upon electrophoresis under denaturing conditions, 3′-end-labeled strands of linearized pSP73KB plasmid containing no inter-

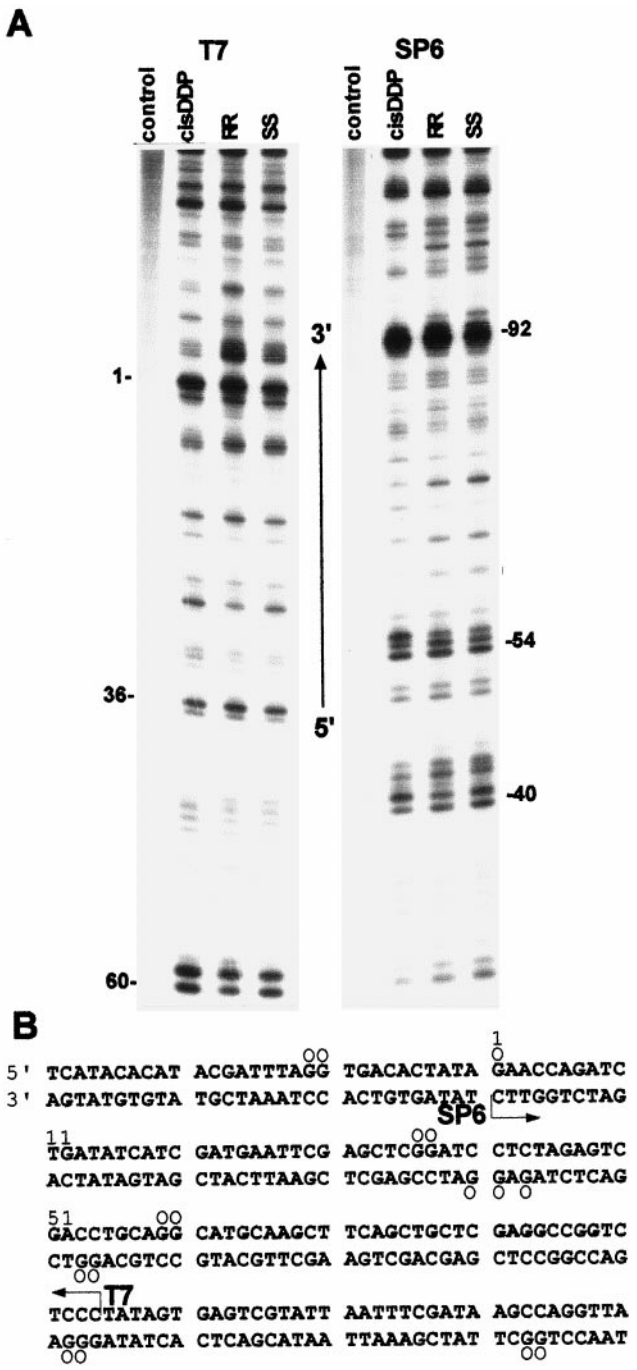


FIGURE 2 Inhibition of RNA synthesis by T7 (left) and SP6 (*right*) RNA polymerases on the *Nde*I/*Hpa*I fragment of pSP73KB plasmid modified by platinum complexes. (*A*) Autoradiogram of 6% polyacrylamide/8 M urea sequencing gel. *Lanes:* control, nonmodified template; cisDDP; RR; SS, the template modified by cisplatin; Pt-DAB(RR), Pt-DAB(SS) at $r_b = 0.01$, respectively. (*B*) Schematic diagram showing the portion of the nucleotide sequence of the template (upper) strand of the *Nde*I/*Hpa*I fragment used to monitor inhibition of RNA synthesis by platinum complexes. The arrows indicate the start of the T7 or SP6 RNA polymerases. (○), major stop signals (from *A*) for DNA modified by Pt-DAB(RR). The numbers correspond to the nucleotide numbering in the sequence map of the pSP73KB plasmid.

strand cross-links migrate as a 2455-nucleotide single strand, whereas the interstrand cross-linked strands migrate more slowly as a higher molecular mass species. The bands corresponding to more slowly migrating interstrand-cross-linked fragments were noticed if either Pt-DAB complex was used to modify DNA at r_b as low as 3×10^{-4} (Fig. 3 *A*). The intensity of the more slowly migrating band increased with the growing level of the modification. The radioactivity associated with the individual bands in each lane was measured to obtain estimates of the fraction of non-cross-linked or cross-linked DNA under each condition. The frequency of interstrand cross-links (the amount of interstrand cross-links per one molecule of Pt-DAB complex coordinated to DNA) was calculated using the Poisson distribution in combination with the r_b values and the fragment size (Farrell et al., 1990) (for details see also Materials and Methods).

As summarized in Fig. 3 *B*, both Pt-DAB complexes showed a relatively low but significant interstrand cross-linking efficiency in linear DNA (at $r_b = 0.001$, approximately 6%). There was no significant difference between the yields of DNA interstrand cross-linking by Pt-DAB enantiomers. Thus, these results indicate that the interstrand cross-links are only minor adducts in double-helical DNA modified by Pt-DAB complexes which are formed with a similar frequency like the same lesions of cisplatin. This observation is consistent with an idea that the spectrum of adducts produced on DNA by the two Pt-DAB compounds is similar for each enantiomer and similar to that reported for cisplatin (Brabec and Leng, 1993; Vrána et al., 1996) or for other analogs such as *cis*-[$PtCl_2$(DACH)] complexes (Boudný et al., 1992; Brabec, unpublished results).

Taken together, the results of the interstrand cross-linking assay and transcription mapping experiments are consistent with the idea that the replacement of NH_3 nonleaving ligands in cisplatin by the DAB carrier ligand in both enantiomeric forms (RR or SS) has not significantly altered base sequence selectivity of the parent platinum drug or the spectrum of its DNA adducts.

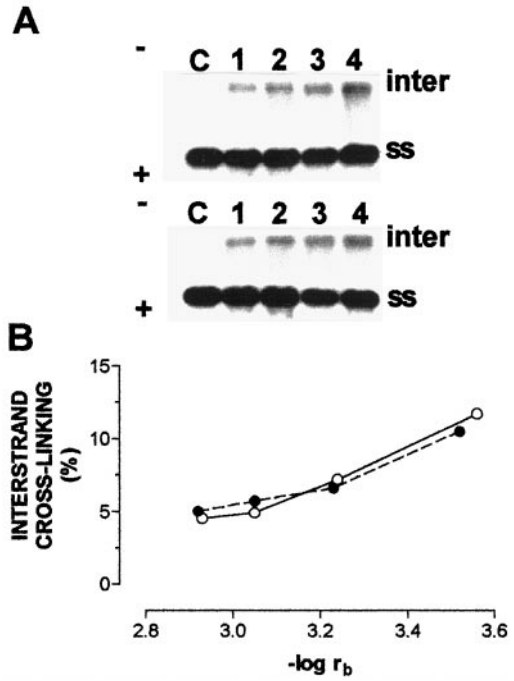


FIGURE 3 The formation of interstrand cross-links by Pt-DAB compounds in linearized pSP73KB plasmid. (*A*) Autoradiograms of a denaturing 1.5% agarose gel of linearized DNA, which was 3′-end-labeled; the interstrand cross-linked DNA appears as the top bands migrating on the gels more slowly than the single-stranded DNA (contained in the bottom bands); the plasmid linearized by *Eco*RI was incubated with Pt-DAB(RR) (*top*) or Pt-DAB(SS) (*bottom*) for 48 h at 37°C. *Lanes:* C, control, nonmodified DNA ($r_b = 0$); 1, $r_b = 3 \times 10^{-4}$; 2, $r_b = 6 \times 10^{-4}$; 3, $r_b = 9 \times 10^{-4}$; 4, $r_b = 1.2 \times 10^{-3}$. (*B*) Dependence on r_b of the percentage of interstrand cross-links per adduct [interstrand cross-linking (%)] formed by Pt-DAB(RR) (●) or Pt-DAB(SS) (○) in linearized DNA within 48 h. Data measured in triplicate varied on average ±3% from their mean.

Conformational changes produced in double-helical DNA by the site-specific d(GpG) intrastrand cross-link

The major DNA lesion of cisplatin and its simple analogs with different carrier amines is the 1,2-d(GpG) intrastrand adduct (Jennerwein et al., 1989; Page et al., 1990 and vide supra). The goal of our further work was to establish whether the steric structure of the non-leaving group of platinum DAB enantiomers could influence the distortions induced in DNA by the formation of the 1,2-d(GpG) intrastrand cross-link. We directed our further studies on establishing distortions and other biophysical properties of oligodeoxyribonucleotide duplexes containing a single, site-specific 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS).

Differential scanning calorimetry

A calorimetric technique was used to characterize the influence of the 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) on the thermal stability and energetics of the site-specifically platinated 20-mer DNA duplex. Such thermodynamic data can reveal how the platinum adduct influences duplex stability, a property that has been shown to play a significant role in the mechanism of antitumor activity of platinum drugs. Recently, calorimetric and spectroscopic techniques were used to characterize the influence of the 1,2-d(GpG) intrastrand cross-link on the thermal stability and energetics of a 20-mer DNA duplex site-specifically modified by cisplatin (Poklar et al., 1996). We expanded these studies on the oligodeoxyribonucleotide duplex containing unique 1,2-d(GpG) site-specific intrastrand adducts of the Pt-DAB(RR) or Pt-DAB(SS) complexes.

Fig. 4 shows DSC melting profiles (ΔC_p versus *T*) for the parent, nonmodified 20 bp duplex d(TGGT(/d(ACCA)(20-

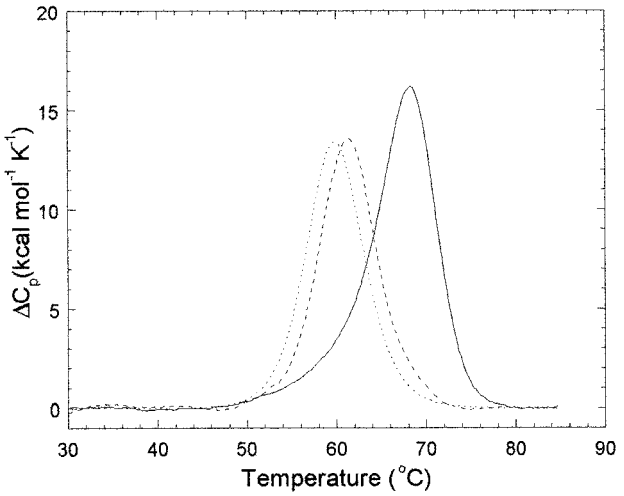


FIGURE 4 DSC thermograms for the d(TGGT)/d(ACCA)(20-DSC) duplex nonmodified (*solid curve*), containing the 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) (*dashed curve*), or Pt-DAB(SS) (*dotted curve*). The duplex concentrations were 5 μM, and the buffer conditions were 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA.

DSC) (*solid curve*) and the same duplex containing single 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) (*dashed curve*) or Pt-DAB(SS) (*dotted curve*). These curves were analyzed as described in Material and Methods to obtain the results listed in Table 1. Inspection of these thermodynamic parameters reveals a number of interesting features. First, cross-link formation of Pt-DAB(RR) and Pt-DAB(SS) reduced the duplex thermal stability by 6.9°C and 8.6°C, respectively. Second, cross-link formation by Pt-DAB(RR) and Pt-DAB(SS) resulted in a large increase of the enthalpy of duplex formation by 36 and 38 kcal/mol, respectively. In other words, the intrastrand cross-link of Pt-DAB enantiomers enthalpically destabilizes the duplex relative to their nonmodified counterpart. Third, cross-link formation by Pt-DAB(RR) and Pt-DAB(SS) resulted in a substantial increase in duplex transition entropy of 104 or 107 cal/K.mol ($T\Delta S$ = 31.0 or 31.9 kcal/mol at 25°C), respectively. In other words, the intrastrand cross-link of both DAB enantiomers entropically stabilizes the duplex. Thus, the 36 or 38 kcal/mol enthalpic destabilization of the duplex due to the cross-link of Pt-DAB(RR) or Pt-DAB(SS), respectively is partially, but not completely, compensated by the entropic cross-link-induced stabilization of the duplex of 31 or 32 kcal/mol at 25°C, respectively. The net result of these enthalpic and entropic effects is that 1,2-d(GpG) intrastrand cross-link formation by Pt-DAB(RR) or Pt-DAB(SS) at 25°C induces a decrease in duplex thermodynamic stability ($\Delta\Delta G_{25}$) of 5.0 or 7.6 kcal/mol, respectively, with this destabilization being enthalpic in origin. In this respect, the intrastrand cross-link of Pt-DAB(SS) was more effective than that of its RR counterpart.

Chemical probing of conformational distortions

Cyanide ions can rapidly remove cisplatin and its analogs from double-helical oligonucleotides containing 1,2-d(GpG) intrastrand cross-links of these platinum compounds. It has been shown (Schwartz et al., 1990; Boudný et al., 1992) that the kinetics of the reaction between cyanide ions and the d(GpG) intrastrand cross-link of cisplatin and its analogs is strongly dependent on the DNA conformation. The samples of d(TGGT)/d(ACCA) containing the intrastrand cross-link of Pt-DAB(RR) or Pt(DAB)(SS) were treated with a large excess of cyanide ions. At various times, aliquots were withdrawn and analyzed by gel electrophoresis under denaturing conditions (Schwartz et al., 1990; Boudný et al., 1992) (Fig. 5, *A* and *B*). As judged by the disappearance of the starting products (upper bands), cyanide ions were considerably less reactive with the double-stranded oligonucleotide containing the intrastrand cross-link of Pt-DAB(SS) as compared with that containing the same adduct of Pt-DAB(RR) (Fig. 5 *C*). It has been shown that the rate of removal of the bound platinum residues decreases when the distortion induced by the Pt-d(GpG) adduct is larger (Schwartz et al., 1990; Boudný et al., 1992). Thus, the results shown in Fig. 5 support the idea that the 1,2-d(GpG) intrastrand cross-link of Pt-DAB(SS) induces in

TABLE 1 Thermodynamic parameters for formation of the oligonucleotide duplex d(TGGT)/d(ACCA)(20-DSC) nonplatinated or containing the 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) and Pt-DAB(SS) determined by DSC

Duplex	T_m, °C	$\Delta H°$, kcal/mol duplex	$\Delta S°$, cal/K · mol duplex	$\Delta G°_{25}$,* kcal/mol duplex
no Pt	68.3	−151 ± 8	−446 ± 23	−18.0 ± 1.9
Pt-DAB(RR)	61.4	−115 ± 7	−342 ± 19	−13.0 ± 1.3
Pt-DAB(SS)	59.7	−113 ± 8	−339 ± 18	−10.3 ± 1.3

Calorimetric measurements were conducted as described in the text. The $\Delta H°$ and $\Delta S°$ values are averages derived from three independent experiments, with the indicated errors corresponding to the average deviation from the mean.
*$\Delta G°_{25}$ is the free energy of duplex formation at 25°C, as determined using the equation

$$\Delta G°_{25} = \Delta H° - (298.15)\Delta S° \quad (2)$$

The indicated uncertainties reflect the maximum possible errors in $\Delta G°_{25}$ that result from the corresponding uncertainties noted above in $\Delta H°$ and $\Delta S°$, as propagated through Eq. 2.

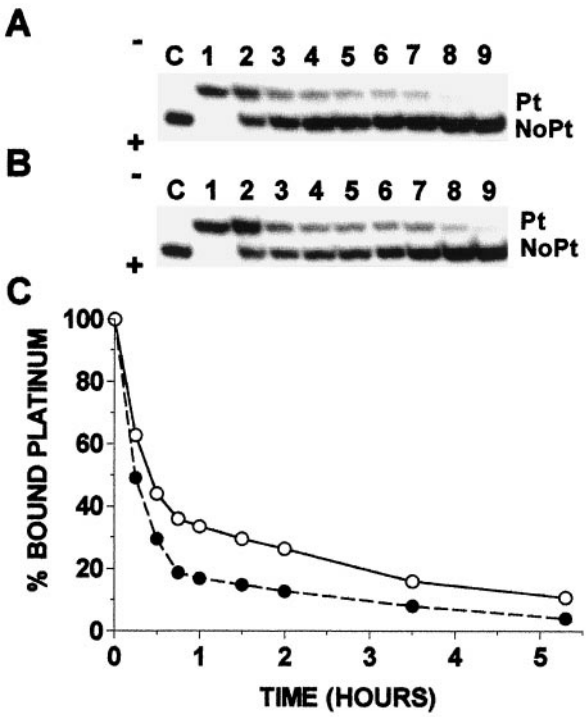


FIGURE 5 Autoradiograms of a denaturing 24% polyacrylamide gel of the products of the reaction between cyanide ions and the oligonucleotide duplex d(TGGT)/d(ACCA)(20) containing a single 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) (*A*) or Pt-DAB(SS) (*B*). The platinated samples at the concentration of 3×10^{-6} M were incubated at 37°C and in 0.2 M NaCN, 20 mM Tris · HCl adjusted at pH 8.3 by addition of HCl. At various times, the samples were precipitated with ethanol, washed three times with ethanol, and electrophoresed. *Lanes:* C, control, nonmodified duplex; 1–9 correspond to the following times (in hours) of incubation of the platinated duplex with NaCN: 1, 0.0; 2, 0.25; 3, 0.5; 4, 0.75; 5, 1.0; 6, 1.5; 7, 2; 8, 3.5; 9, 5.3. The top strand of the d(TGGT)/d(ACCA) was ^{32}P-end-labeled at the 5′ end. (*C*) The dependence of the amount of platinum coordinated to the oligonucleotide duplex d(TGGT)/d(ACCA)(20) containing single intrastrand cross-link of Pt-DAB(RR) (●) or Pt-DAB(SS) (○) on the time of incubation with NaCN. The experimental conditions were the same as in *A* and *B*.

DNA a larger conformational distortion than its RR counterpart. These results are also consistent with the DSC analysis (Fig. 4 and Table 1), which indicated a larger thermal and thermodynamic destabilization induced in DNA by the 1,2-d(GpG) intrastrand cross-link of Pt-DAB(SS) in comparison with the cross-link of Pt-DAB(RR).

To further characterize the distortion induced in DNA by intrastrand cross-links of Pt-DAB(RR) or Pt-DAB(SS), the d(TGGT)/d(ACCA)(20) containing the 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) was treated with several chemical agents that are used as tools for monitoring the existence of conformations other than canonical B-DNA. These agents include $KMnO_4$, DEPC, and bromine. They react preferentially with single-stranded DNA and distorted double-stranded DNA (Nielsen, 1990; Brabec et al., 1993; Bailly et al., 1994; Ross and Burrows, 1996; Bailly and Waring, 1997).

$KMnO_4$ is hyperreactive with thymine residues in single-stranded nucleic acids and in distorted DNA as compared to B-DNA (McCarthy et al., 1990; McCarthy and Rich, 1991; Bailly et al., 1994; Bailly and Waring, 1997). $KMnO_4$ reacted with no residue within the nonplatinated duplex (Fig. 6*A*, *left*, *lane* ds). All thymine residues were strongly reactive in the nonplatinated single-stranded top oligonucleotide (Fig. 6 *A*, *left*, *lane* ss). The duplex containing the intrastrand cross-link of Pt-DAB(RR) showed strong reactivity of the 5′ and 3′ thymine residues adjacent to the adduct (Fig. 6 *A*, *left*, *lane* RR). A weak but significant reactivity was also observed for the second 5′ thymine residue adjacent to the cross-link. The duplex containing the intrastrand cross-link of Pt-DAB(SS) showed strong reactivity of the two neighboring 5′ thymine residues in the top strand adjacent to the adduct, and a weak but significant reactivity was also observed for the 3′ thymine residue adjacent to the cross-link (Fig. 6 *A*, *left*, *lane* SS). No reactivity between $KMnO_4$ and the residues in the bottom strand of the cross-linked duplexes was apparent.

DEPC carbetoxylates purines at the N(7) position. It is hyperreactive with unpaired and distorted adenine residues in DNA and with left-handed Z-DNA (Herr, 1985; Johnston and Rich, 1985; Bailly et al., 1994; Bailly and Waring, 1997). Adenine and guanine residues within the nonplatinated single-stranded oligonucleotides (top and bottom) readily reacted with DEPC (shown for the bottom strand in Fig. 6 *A*, *center*, *lane* ss). No reactivity of adenine and guanine residues was observed within the nonplatinated double-stranded oligonucleotide (shown in Fig. 6 *A*, *center* for the bottom strand, *lane* ds). Within the double-stranded oligonucleotide containing the intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS), three base residues in the bottom strand became reactive (Fig. 6 *A*, *center*, *lanes* RR *and* SS): these are readily identified as the three adenine residues complementary to the reactive thymine residues of the top strand. Importantly, a strong reactivity with DEPC was only observed for adenine residues complementary to strongly reactive thymine residues, whereas adenine residues complementary to weakly reactive thymine residues also reacted with DEPC only weakly.

Bromination of cytosine and formation of piperidine-labile sites are observed when two simple salts, KBr and $KHSO_5$, are allowed to react with single-stranded or distorted double-stranded oligonucleotides (Ross and Burrows, 1996). The reaction proceeds via generation of Br_2 in situ, which reacts selectively with the 5,6 double bond to add Br and OH, respectively. H_2O is then eliminated to give 5-bromodeoxycytidine, which is susceptible to depyrimidination under basic conditions. All cytosine residues within the nonplatinated top and bottom strands of d(TGGT)d(ACCA) were readily reactive (shown for the bottom strand in Fig. 6 *A*, *right*, *lane* ss). No reactivity of these residues was ob-

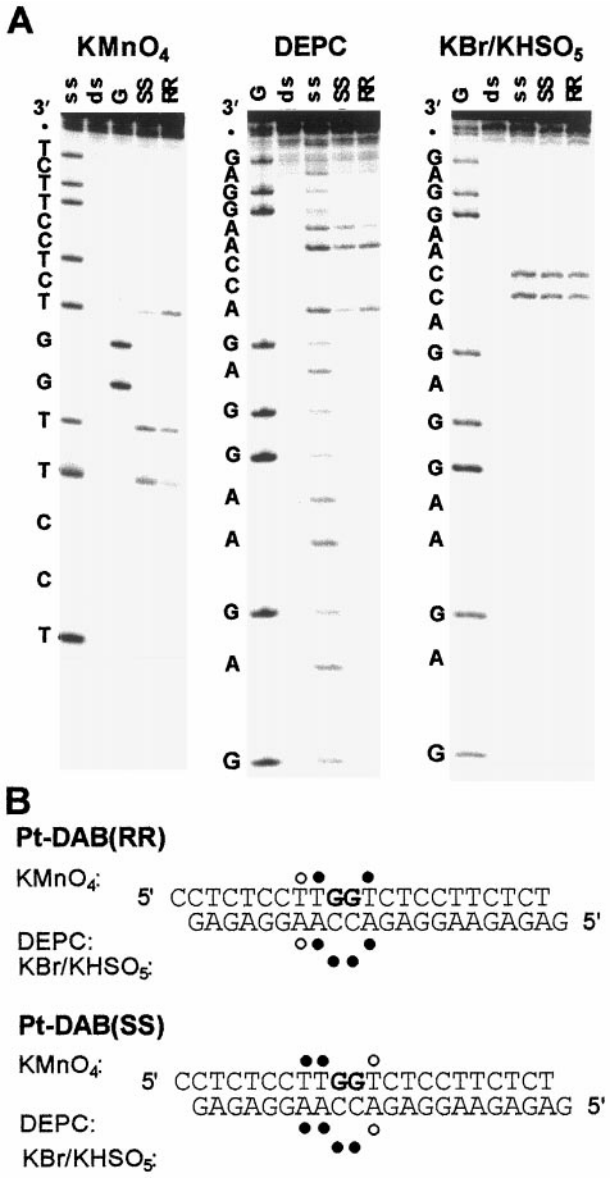


FIGURE 6 (*A*) Piperidine-induced specific strand cleavage at $KMnO_4$-modified (*left*), diethylpyrocarbonate-modified (*center*), and $KBr/KHSO_5$-modified (*right*) bases in nonplatinated and platinated d(TGGT)/d(ACCA)(20). The oligomers were 5′-end-labeled at their top strands in the case of the modification by $KMnO_4$ or at their bottom strands in the case of the modification by DEPC or $KBr/KHSO_5$. $KMnO_4$ (*left*): the lane ss is relative to the nonplatinated top strand. The lane ds is relative to the nonplatinated duplex. The lane G is a Maxam-Gilbert specific reaction for the nonplatinated duplex that had only top strand end-labeled. The lanes SS and RR are relative to the duplex containing Pt-DAB(SS) or Pt-DAB(RR) 1,2-d(GpG) intrastrand cross-link, respectively. DEPC (*center*): The lane G is a Maxam-Gilbert specific reaction for the nonplatinated duplex that had only bottom strand end-labeled. The lane ds is relative to the nonplatinated duplex. The lane ss is relative to the nonplatinated bottom strand. The lanes SS and RR are relative to the duplex containing Pt-DAB(SS) or Pt-DAB(RR) 1,2-d(GpG) intrastrand cross-link, respectively. $KBr/KHSO_5$ (*right*): The lane G is a Maxam-Gilbert specific reaction for the nonplatinated duplex. The lane ds is relative to the nonplatinated duplex. The lane ss is relative to the nonplatinated bottom strand that had only bottom strand end-labeled. The lanes SS and RR are relative to the duplex containing Pt-DAB(SS) or Pt-DAB(RR) 1,2-d(GpG) intrastrand cross-link, respectively, that had only bottom strand end-labeled. (*B*) Summary of the reactivity of chemical probes; ● and ○ designate strong or weak reactivity, respectively.

served within the nonplatinated duplex (shown for the bottom strand in Fig. 6 *A*, *right*, *lane* ds). Cytosine residues in the bottom strand complementary to the platinated guanine residues in the top strand containing the intrastrand cross-links of Pt-DAB(RR) or Pt-DAB(SS) were, however, strongly reactive.

The results obtained with chemical probes are summarized in Fig. 6 *B*. The pattern and degree of reactivity toward the chemical probes was different for the two Pt-DAB isomers indicating different character of the conformational distortion. In addition, the results suggest that the conformational distortion extends over at least five basepairs around the cross-link.

DNA unwinding and bending

Among the alterations of secondary and tertiary structure of DNA to which it may be subject, the role of intrinsic bending and unwinding of DNA is increasingly recognized as being potentially important in regulating replication and transcription functions through specific DNA-protein interactions. For cisplatin adducts, the structural details responsible for bending and subsequent protein recognition have recently been elucidated (Ohndorf et al., 1999; Zamble and Lippard, 1999). Given the recent advances in our understanding of the structural basis for the bending of DNA caused by cisplatin, it is of considerable interest to examine how the character of a carrier amine in the 1,2-d(GpG) intrastrand adduct affects conformational properties of DNA, such as bending and unwinding. In this work we further performed studies on the bending and unwinding induced by single, site-specific intrastrand cross-links of Pt-DAB enantiomers formed in oligodeoxyribonucleotide duplexes between neighboring guanine residues.

The top strands of the oligonucleotide duplexes examined in the present work were designed to contain only one high-affinity platinum binding site, the two adjacent guanine bases of the intrastrand cross-link (Fig. 1). All sequences were designed to leave a 1 nucleotide overhang at their 5′-ends in double-stranded form. These overhangs facilitate polymerization of the monomeric oligonucleotide duplexes by T4 DNA ligase in only one orientation, and maintain a constant interadduct distance throughout the resulting multimer. Autoradiograms of electrophoresis gels revealing resolution of the ligation products of 20–23-bp duplexes containing a unique 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) are shown in Fig. 7. A significant retardation was observed for the multimers of all platinated duplexes. Decreased gel electrophoretic mobility

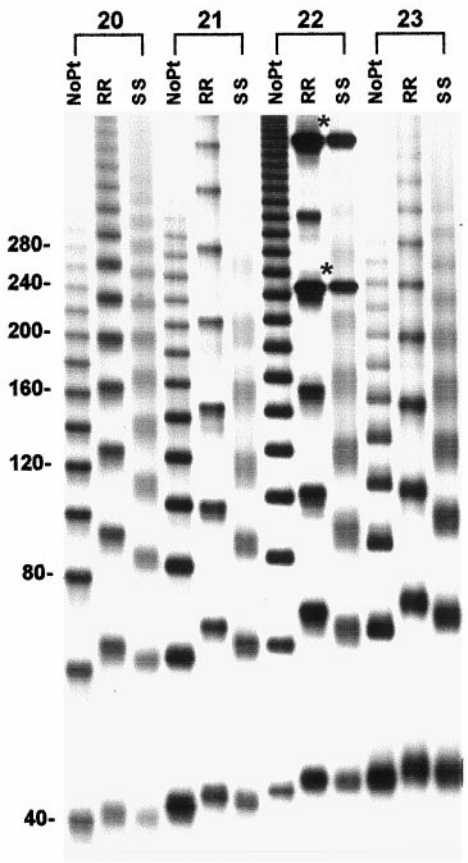


FIGURE 7 Autoradiograms of the ligation products of double-stranded oligonucleotides d(TGGT)d(ACCA)(20–23) containing a unique 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) separated on an 8% polyacrylamide gel (lanes RR and SS, respectively). Nonplatinated oligomers, lanes NoPt.

may result from a decrease in the DNA end-to-end distance (Koo and Crothers, 1988). Various platinum(II) complexes have been shown to form DNA adducts that decrease gel mobility of DNA fragments due to either stable curvature of the helix axis or increased isotropic flexibility (Rice et al., 1988; Bellon and Lippard, 1990; Leng, 1990; Brabec et al., 1993; Huang et al., 1995; Malinge et al., 1995). DNA multimers of identical length and number of stable bend units, but with differently phased bends, have different end-to-end distances. The DNA bends of a multimer must be, therefore, spaced evenly and phased with the DNA helical repeat in order to add constructively. Such constructively phased bends add in plane, yielding short end-to-end distances and the most retarded gel migration. In other words, gel electrophoresis of multimers of oligonucleotide duplexes, which only differ in length and contain a stable curvature induced by the same platinum adduct, should exhibit a phase effect, i.e., the maximum retardation should be observed for the multimers having the bends in phase with the helix screw. In contrast, the normal electrophoretic mobility should be observed for the multimers having the bends separated by a half-integral number of DNA turns. The *K* factor is defined as the ratio of calculated to actual length. The calculated length is based on a multimer's mobility, and is obtained from a calibration curve constructed from the mobilities of nonplatinated multimers. The variations of the *K* factor versus sequence length obtained for multimers of the duplexes 20–23 bp long and containing the unique 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) are shown in Fig. 8 *B*. Maximum retardation was observed for the 22-bp duplex containing the adduct of Pt-DAB(RR) (Fig. 8 *B, left*). This observation suggests that the natural 10.5-bp repeat of B-DNA and that of DNA perturbed by the Pt-DAB(RR) intrastrand cross-link are different as a consequence of DNA unwinding (Bellon et al., 1991). Similarly, in the case of the duplex containing the adduct of Pt-DAB(SS), maximum retardation was observed for the 21-bp duplex, but the 22-bp curve had only a slightly smaller slope, whereas the 20-bp curve differed more pronouncedly (Fig. 8 *B, right*). This asymmetry is also consistent with a significant DNA unwinding due to the formation of the cross-link by the SS enantiomer.

The exact helical repeat of the intrastrand cross-linked duplex and from it the unwinding angle were calculated by interpolation with the use of the *K* versus interadduct distance curve as described in the previous paper for intrastrand adducts of cisplatin (Bellon et al., 1991). The maximum of these curves constructed for the duplexes

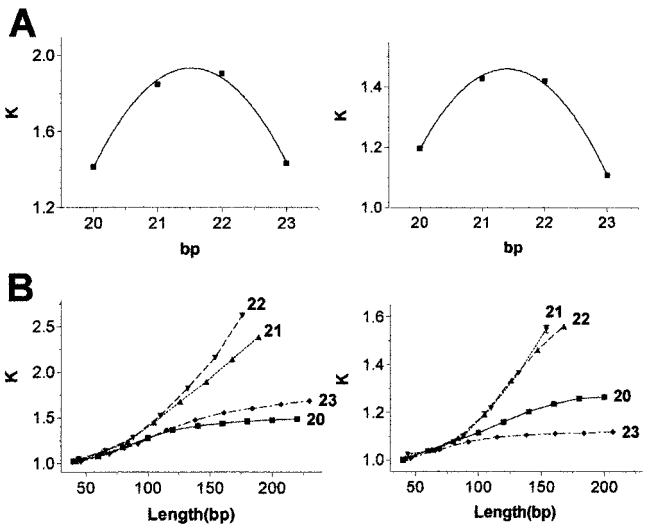


FIGURE 8 (*A*) Plots showing the relative mobility *K* versus interadduct distance in bp for the oligomers d(TGGT)/d(ACCA)(20–23) modified by Pt-DAB(RR) (*left*) or Pt-DAB(SS) (*right*) with a total length of 140 bp. The experimental points represent the average of three independent electrophoresis experiments. The curves represent the best fit of these experimental points to the equation $K = ad^2 + bd + c$ (Bellon et al., 1991). (*B*) Plots showing the relative mobility *K* versus sequence length curves for the oligomers d(TGGT)/d(ACCA)(20–23), denoted respectively as 20, 21, 22, and 23.

intrastrand cross-linked by Pt-DAB(RR) or Pt-DAB(SS) with a total length of 140 bp (Fig. 8 *B*) were determined to be 21.57 ± 0.01 or 21.44 ± 0.04 bp. Total sequence lengths other than 140 bp were examined and gave identical results. To convert the interadduct distance in basepairs corresponding to the curve maximum into a duplex unwinding angle in degrees, the value is compared with that of the helical repeat of B-DNA, which is 10.5 ± 0.05 bp (Wang, 1979; Rhodes and Klug, 1980). The difference between the helical repeat of B-DNA and the DNA containing intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) complex, therefore, is [(21.57 ± 0.01) − 2(10.5 ± 0.05)] = 0.57 ± 0.06 bp or [(21.44 ± 0.04) − 2(10.5 ± 0.05)] = 0.44 ± 0.09 bp, respectively. There are 360°/10.5 bp, so the DNA unwinding due to one intrastrand adduct of Pt-DAB(RR) or Pt-DAB(SS) is 20 ± 2° or 15 ± 3°, respectively. These unwinding angles are considerably greater than that found for the 1,2-d(GpG) intrastrand cross-link of cisplatin (13°) using the same experimental procedure (Bellon et al., 1991). One plausible explanation of this observation might be associated with an additional contribution to unwinding associated with interaction of the DAB moiety with the duplex upon covalent binding of platinum. In a similar way, large unwinding angles (~19°) produced by cisplatin tethered to intercalators were explained (Keck and Lippard, 1992).

The appreciation of the relationship between interadduct distance and phasing for self-ligated multimers composed of the identical number of monomeric duplexes (bend units) resulted in a bell-shaped pattern (Fig. 8 *B*) characteristic for bending (Rice et al., 1988; Bellon and Lippard, 1990; Leng, 1990; Brabec et al., 1993; Huang et al., 1995; Malinge et al., 1995). The quantitation of the bend angle of the intrastrand cross-links of Pt-DAB(RR) or Pt-DAB(SS) complexes was performed in the way described previously (Rice et al., 1988; Bellon and Lippard, 1990; Leng, 1990; Brabec et al., 1993; Huang et al., 1995; Malinge et al., 1995) utilizing the empirical equation

$$K - 1 = (9.6 \times 10^{-5}L^2 - 0.47)(\mathrm{RC})^2 \qquad (1)$$

where L represents the length of a particular oligomer with relative mobility K and RC the curvature relative to a DNA bending induced at the tract of six adenines (A_6 tract) (Rice et al., 1988). Application of Eq. 1 to the 132- or 154-bp multimers of the 22-bp oligomers containing the single intrastrand cross-link of Pt-DAB(RR) leads to curvatures of 0.87, relative to an A_6 tract. Similarly, the application of Eq. 1 to the 126- or 147-bp multimers of the 21-bp oligomers containing the single intrastrand cross-link of Pt-DAB(RR) leads to curvatures of 0.57, relative to an A_6 tract. The average bend angle per helix turn can be calculated by multiplying the relative curvature by the absolute value of an A_6 tract bend [20° (Bellon and Lippard, 1990; Koo et al., 1990)]. The results indicate that the bends induced by the intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) are ~35° or 24°, respectively. That these bands were oriented toward the major groove of DNA was verified in the same way as in the previously published paper (Huang et al., 1995). Other details of the calculations of the unwinding and bending angles are given in the previously published papers (Rice et al., 1988; Bellon and Lippard, 1990; Leng, 1990; Brabec et al., 1993; Huang et al., 1995; Malinge et al., 1995).

Also produced in ligations of monomers investigated in this work were separate bands arising from small DNA circles that migrate close to the top of the gel (see the bands marked by asterisk in Fig. 7 as example). The highest tendency to yield DNA circles was observed for the 22-bp multimers confirming a close match between the 22-bp sequence repeat and the helix screw (Ulanovsky et al., 1986; Rice et al., 1988).

DISCUSSION

The results of the present work (Figs. 2 and 3) are consistent with the view that the replacement of NH_3 nonleaving ligands in cisplatin by the DAB carrier ligand in both enantiomeric forms (RR or SS) changes neither the spectrum and frequency of DNA adducts nor the sequence selectivity of DNA binding of the parent drug. Thus, these features of DNA binding mode of Pt-DAB compounds are unlikely to be associated with the different biological activity of these platinum compounds. A possible explanation for the different biological activity of Pt-DAB enantiomers can be associated with the different conformational distortions induced in DNA by the adducts of these compounds and their different processing in the cell. To test this hypothesis, the experiments described in the present work were carried out.

Thermal and thermodynamical stability of duplexes containing single, site-specific 1,2-d(GpG) intrastrand adduct of either Pt-DAB enantiomer (this cross-link is the major DNA adduct of cisplatin and its direct analogs) and the resistance of this adduct to NaCN treatment demonstrate that the lesions formed by Pt-DAB(SS) were more effective at inducing overall destabilization of the duplex and global conformational alterations than those formed by Pt-DAB(RR). This result is consistent with the idea and supports the hypothesis that the enhancement of mutagenic activity of Pt-DAB(SS) compound is associated with an increase of the thermodynamical destabilization of the duplex and the character of the overall or global conformational alteration induced by this platinum compound in DNA.

The character of global conformational distortions and alterations of the overall stability of the double-helical DNA induced by its damage are determined by the sum of individual contributions from various features of the damage. Some of these individual features may result in stabilization

of the duplex, others may lower its stability. An important feature of the local conformational distortion induced by the 1,2-d(GpG) adduct of cisplatin and its analogs is bending of the duplex axis (Bellon and Lippard, 1990; Takahara et al., 1996; Gelasco and Lippard, 1998). The results of the present work indicate that bending and unwinding angles due to the Pt-DAB(SS) cross-link are smaller than those due to the cross-link of Pt-DAB(RR). However, the overall destabilization of the duplex due to the cross-link of Pt-DAB(SS) is greater than that due to the cross-link of Pt-DAB(RR). It was suggested recently (Poklar et al., 1996) that helical bending induced by the 1,2-d(GpG) intrastrand cross-link of cisplatin thermodynamically stabilized the duplex. A crude estimate has indicated that helical bending due to cisplatin-1,2-d(GpG) cross-link contributed ~6.4 kcal/mol toward stabilization of the global duplex structure. Thus, helical bending induced by the d(GpG) intrastrand cross-link has been suggested (Poklar et al., 1996) to partially compensate destabilization due to the formation of this adduct. With this qualification in mind, we suggest that one reason why the intrastrand cross-link of Pt-DAB(SS) globally destabilizes the DNA duplex more efficiently than the same cross-link of the RR enantiomer is also associated with a lower efficiency of the adduct of the SS enantiomer to contribute toward stabilization of the global duplex structure associated with the bending.

Another conformational parameter of the distortion induced by the formation of the 1,2-d(GpG) intrastrand cross-links of Pt-DAB compounds determined in the present work was unwinding of the double helix (lowering of the number of basepairs per a helical turn). The adduct of the RR enantiomer was slightly more effective in DNA unwinding than the cross-link of its SS counterpart. The energetics of DNA unwinding can be crudely estimated using the same approach as that used to calculate the free energy required to twist a DNA fragment 12 bp long containing a single 1,2-(GpG) intrastrand adduct of cisplatin about its helix axis (Bellon et al., 1991). The free energy of unwinding of only 0.29 kcal/mol was calculated assuming the unwinding angle 13° (Bellon et al., 1991). The same calculations were performed, assuming unwinding angles of 20° for the cross-link of Pt-DAB(RR), 15° for the cross-link of Pt-DAB(SS) (vide supra), and local twisting within the fragment 20 bp long [the fragment 20 bp long was taken for these calculations because energetics of the duplex of this length containing the single, site-specific intrastrand d(GpG) cross-link of Pt-DAB enantiomers was characterized in the present work (Fig. 4 and Table 1)]. These approximate calculations gave free energies of unwinding of only 0.42 and 0.23 kcal/mol, respectively. If this rough estimate is justified, then it seems reasonable to suggest that unwinding due to the intrastrand cross-link of Pt-DAB compounds contributes to the efficiency of these platinum complexes to affect the overall stability of the duplex only in a very small extent. More detailed proposals as to the exact nature of the effect of unwinding induced in DNA by the 1,2-d(GpG) intrastrand cross-link of Pt-DAB compounds must await the results of further experiments.

The structural perturbation caused by d(GpG) intrastrand cross-links of cisplatin has been subjected to numerous NMR investigations, which were recently reviewed (Ano et al., 1999). In these adducts, the polynucleotide chain confines the guanine to a head-to-head (HH) arrangement (Sherman and Lippard, 1987; Bloemink and Reedijk, 1996). The preferential orientations of the guanines in the HH conformation are those with one guanine close to perpendicular to the coordination plane (Ø = 100–110°) and the other rather tilted and forming a hydrogen bond between its carbonyl oxygen atom and the NH_3 group in *cis* position (Ø = 50–60°) (Kozelka et al., 1992). The tilting will be greater for hydrogen bond formation between the carbonyl oxygen of the guanine residue and a "quasi equatorial" hydrogen of the *cis* amine (Grabner et al., 1998). Molecular models indicate that one "quasi equatorial" amino proton of the Pt-DAB(SS)-[d(GpG)] cross-link is close to O(6)-5′ (carbonyl oxygen of the 5′-guanosine) and that one "quasi equatorial" amino proton of Pt-DAB(RR)-[d(GpG)] is close to O(6)-3′ (carbonyl oxygen of the 3′-guanosine). Consequently, the formation of the Pt-DAB(SS)-d(GpG) cross-link is expected to give greater tilt of the guanine residue on the 5′ side, and therefore the greater distortion on the 5′ side of the cross-link. In contrast, formation of the Pt-DAB(SS)-d(GpG) cross-link should result in the greater tilt for the guanine residue on the 3′ side, so that the distortion should be greater on the 3′ side of the cross-link.

These assumptions are in good agreement with the results of the experiments in which structural changes in DNA induced by the single, site-specific 1,2-d(GpG) intrastrand cross-link of Pt-DAB(RR) or Pt-DAB(SS) were investigated by studying the effect of this cross-link on the reactivity of $KMnO_4$ and DEPC toward DNA. $KMnO_4$ and DEPC are complementary probes capable of revealing the location of AT basepairs, the secondary structure of which has been perturbed by the cross-link. Importantly, these chemical probes do not represent measures of basepair disruption as they can proceed even if the basepairing is maintained (Bailly et al., 1994). Formation of adducts by these probes requires out-of-plane attack by the electrophile so that they will be sterically hindered by stacking of neighboring basepairs. Thus, $KMnO_4$ and DEPC are essentially probes of base stacking (Bailly et al., 1994) so that they are particularly suitable for proving distortions induced by the cross-links of Pt-DAB enantiomers predicted above. If the reactivity of the two chemical probes with the duplexes containing the cross-links of Pt-DAB(RR) or Pt-DAB(SS) is compared (Fig. 6), the duplex containing the cross-link of Pt-DAB(RR) shows considerably stronger reactivity of the AT basepair whose thymine residue is adjacent to the adduct on its 3′ side. In contrast, the duplex containing the intrastrand adduct of Pt-DAB(SS) shows

stronger reactivity of the two probes of the AT basepair on the other side of the adduct (particularly the second AT basepair 5′ to the adduct). Thus, the results obtained with the aid of these chemical probes highlight the importance of hydrogen bond formation between the carbonyl oxygen of the guanine residue and a "quasi equatorial" hydrogen of the *cis* amine as discussed above.

In conclusion, the excellent agreement between the previsions and the results of the present work demonstrates the potentiality of the techniques of molecular biophysics in highlighting very fine structural modifications, such as those promoted on DNA by enantiomeric Pt-DAB compounds. Importantly, the configuration of the asymmetric carbons in these complexes dictates the conformation of the chelate ring (*δ-gauche* and *λ-gauche* for the SS and RR isomer, respectively) and thus the "axial" or "equatorial" disposition of the hydrogen atoms on the coordinated nitrogen atoms. Also importantly, formation of hydrogen bonds between the carbonyl oxygen of the guanine residue and the "quasi equatorial" hydrogen of the *cis* amine determines propagation of the distortion of double-helical structure either on the 3′ or on the 5′ side of the cross-linked bases, depending on the configuration of the carrier ligand.

This work was supported by the Grant Agency of the Czech Republic (Grant 305/99/0695), the Grant Agency of the Academy of Sciences of the Czech Republic (Grant A5004702), and the Ministero dell'Universita' e della Ricerca Scientifica e Tecnologica (Cofinanziamento MURST) and the University of Bari. J.M. and C.H. are supported by doctoral fellowships from the Faculty of Sciences, Masaryk University, Brno. The research of V.B. was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute. This research is also a part of the European Cooperation in the field of Scientific and Technical Research network (projects COST D8/0009/97 and D8/0012/97) and of the Italian-Czech cooperation supported by the Italian Ministry for Foreign Affairs.

REFERENCES

Ano, S. O., Z. Kuklenyik, and L. G. Marzilli. 1999. Structure and dynamics of Pt anticancer drug adduct from nucleotides to oligonucleotides as revealed by NMR methods. *In* Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug. B. Lippert, editor. VHCA, WILEY-VCH, Zürich, Weinheim. 247–291.

Bailly, C., D. Gentle, F. Hamy, M. Purcell, and M. J. Waring. 1994. Localized chemical reactivity in DNA associated with the sequence-specific bisintercalation of echinomycin. *Biochem. J.* 300:165–173.

Bailly, C., and M. Waring. 1997. Diethylpyrocarbonate and osmium tetroxide as probes for drug-induced changes in DNA conformation in vitro. *In* Drug-DNA Interaction Protocols. K. R. Fox, editor. Humana Press Inc., Totowa, NJ. 51–79.

Bellon, S. F., J. H. Coleman, and S. J. Lippard. 1991. DNA unwinding produced by site-specific intrastrand cross-links of the antitumor drug *cis*-diamminedichloroplatinum(II). *Biochemistry.* 30:8026–8035.

Bellon, S. F., and S. J. Lippard. 1990. Bending studies of DNA site-specifically modified by cisplatin, *trans*-diamminedichloroplatinum(II) and *cis*-$Pt(NH_3)_2$(N3-Cytosine)Cl^+. *Biophys. Chem.* 35:179–188.

Berners-Price, S. J., K. J. Barnham, U. Frey, and P. J. Sadler. 1996. Kinetic analysis of the stepwise platination of single- and double-stranded GG oligonucleotides with cisplatin and *cis*-$[PtCl(H_2O)(NH_3)_2]^+$. *Chem. Eur. J.* 2:1283–1291.

Berners-Price, S. J., A. Corazza, Z. J. Guo, K. J. Barnham, P. J. Sadler, Y. Ohyama, M. Leng, and D. Locker. 1997. Structural transitions of a GG-platinated DNA duplex induced by pH, temperature and box A of high-mobility-group protein I. *Eur. J. Biochem.* 243:782–791.

Bloemink, M. J., and J. Reedijk. 1996. Cisplatin and derived anticancer drugs: Mechanism and current status of DNA binding. *In* Metal Ions in Biological Systems. A. Sigel, and H. Sigel, editors. Marcel Dekker, Inc., New York, Basel, Hong Kong. 641–686.

Boudný, V., O. Vrána, F. Gaucheron, V. Kleinwächter, M. Leng, and V. Brabec. 1992. Biophysical analysis of DNA modified by 1,2-diaminocyclohexane platinum(II) complexes. *Nucleic Acids Res.* 20: 267–272.

Brabec, V., V. Boudný, and Z. Balcarová. 1994. Monofunctional adducts of platinum(II) produce in DNA a sequence-dependent local denaturation. *Biochemistry.* 32:1316–1322.

Brabec, V., and M. Leng. 1993. DNA interstrand cross-links of *trans*-diamminedichloroplatinum(II) are preferentially formed between guanine and complementary cytosine residues. *Proc. Natl. Acad. Sci. USA.* 90:5345–5349.

Brabec, V., and E. Paleček. 1970. The influence of salts and pH on polarographic currents produced by denatured DNA. *Biophysik.* 6:290–300.

Brabec, V., and E. Paleček. 1976. Interaction of nucleic acids with electrically charged surfaces. II. Conformational changes in double-helical polynucleotides. *Biophys. Chem.* 4:76–92.

Brabec, V., J. Reedijk, and M. Leng. 1992. Sequence-dependent distortions induced in DNA by monofunctional platinum(II) binding. *Biochemistry.* 31:12397–12402.

Brabec, V., M. Šíp, and M. Leng. 1993. DNA conformational distortion produced by site-specific interstrand cross-link of *trans*-diamminedichloroplatinum(II). *Biochemistry.* 32:11676–11681.

Coluccia, M., M. Correale, D. Giordano, M. A. Mariggio, Moscelli S., F. P. Fanizzi, G. Natile, and L. Maresca. 1986. Mutagenic activity of some platinum complexes with monodentate and bidentate amines. *Inorg. Chim. Acta.* 123:225–229.

Coluccia, M., F. P. Fanizzi, G. Giannini, D. Giordano, F. P. Intini, G. Lacidogna, F. Loseto, M. A. Mariggio, A. Nassi, and G. Natile. 1991. Synthesis, mutagenicity, binding to pBR 322 DNA and antitumor activity of platinum(II) complexes with ethambutol. *Anticancer Res.* 11: 281–288.

Corda, Y., M. F. Anin, M. Leng, and D. Job. 1992. RNA polymerases react differently at d(ApG) and d(GpG) adducts in DNA modified by *cis*-diamminedichloroplatinum(II). *Biochemistry.* 31:1904–1908.

Corda, Y., C. Job, M. F. Anin, M. Leng, and D. Job. 1991. Transcription by eucaryotic and procaryotic RNA polymerases of DNA modified at a d(GG) or a d(AG) site by the antitumor drug *cis*-diamminedichloroplatinum(II). *Biochemistry.* 30:222–230.

Den Hartog, J. H. J., C. Altona, J. C. Chottard, J. P. Girault, J. Y. Lallemand, F. A. A. M. Leeuw, A. T. M. Marcelis, and J. Reedijk. 1982. Conformational analysis of the adduct *cis*-$[Pt(NH_3)_2\{d(GpG\}]^+$ in aqueous solution. A high field (500–300 MHz) nuclear magnetic resonance investigation. *Nucleic Acids Res.* 10:4715–4730.

Fanizzi, F. P., F. P. Intini, L. Maresca, G. Natile, R. Quaranata, M. Coluccia, L. Di Bari, D. Giordano, and M. A. Mariggió. 1987. Biological activity of platinum complexes containing chiral centers on the nitrogen or carbon atoms of a chelate diamine ring. *Inorg. Chim. Acta.* 137: 45–51.

Farrell, N., Y. Qu, L. Feng, and B. Van Houten. 1990. Comparison of chemical reactivity, cytotoxicity, interstrand cross-linking and DNA sequence specificity of bis(platinum) complexes containing monodentate or bidentate coordination spheres with their monomeric analogs. *Biochemistry.* 29:9522–9531.

Fenton, R. R., W. J. Easdale, H. M. Er, S. M. OMara, M. J. McKeage, P. J. Russell, and T. W. Hambley. 1997. Preparation, DNA binding, and in vitro cytotoxicity of a pair of enantiomeric platinum(II) complexes, [(R)- and (S)-3-aminohexahydroazepine]dichloro-platinum(II). Crystal structure of the S enantiomer. *J. Med. Chem.* 40:1090–1098.

Gelasco, A., and S. J. Lippard. 1998. NMR solution structure of a DNA dodecamer duplex containing a *cis*-diammineplatinum(II) d(GpG) in-

trastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry.* 37:9230–9239.

Giannini, G., and G. Natile. 1991. Steric constraints inside the metal-coordination sphere as revealed by diastereotopic splitting of methylene protons. *Inorg. Chem.* 30:2853–2855.

Grabner, S., J. Plavec, N. Bukovec, D. Di Leo, R. Cini, and G. Natile. 1998. Synthesis and structural characterization of platinum(II)-acyclovir complexes. *J. Chem. Soc. Dalton Trans.*1447–1451.

Herr, W. 1985. Diethyl pyrocarbonate: a chemical probe for secondary structure in negatively supercoiled DNA. *Proc. Natl. Acad. Sci. USA.* 82:8009–8013.

Huang, H. F., L. M. Zhu, B. R. Reid, G. P. Drobny, and P. B. Hopkins. 1995. Solution structure of a cisplatin-induced DNA interstrand cross-link. *Science.* 270:1842–1845.

Jennerwein, M. M., A. Eastman, and A. Khokhar. 1989. Characterization of adducts produced in DNA by isomeric 1,2-diaminocyclohexaneplatinum(II) complexes. *Chem.-Biol. Interactions.* 70:39–49.

Johnson, N. P., J.-L. Butour, G. Villani, F. L. Wimmer, M. Defais, V. Pierson, and V. Brabec. 1989. Metal antitumor compounds: the mechanism of action of platinum complexes. *Prog. Clin. Biochem. Med.* 10:1–24.

Johnston, B. H., and A. Rich. 1985. Chemical probes of DNA conformation: detection of Z-DNA at nucleotide resolution. *Cell.* 42: 713–724.

Keck, M. V., and S. J. Lippard. 1992. Unwinding of supercoiled DNA by platinum ethidium and related complexes. *J. Am. Chem. Soc.* 114: 3386–3390.

Kidani, Y., K. Inagaki, M. Iigo, A. Hoshi, and K. Kuretani. 1978. Antitumor activity of 1,2-diamminocyclohexane-platinum complexes against Sarcoma 180 ascites form. *J. Med. Chem.* 21:1315–1318.

Kim, S. D., O. Vrána, V. Kleinwächter, K. Niki, and V. Brabec. 1990. Polarographic determination of subnanogram quantities of free platinum in reaction mixture with DNA. *Anal. Lett.* 23:1505–1518.

Kline, T. P., L. G. Marzilli, D. Live, and G. Zon. 1989. Investigations of platinum amine induced distortions in single- and double-stranded oligodeoxyribonucleotides. *J. Am. Chem. Soc.* 111:7057–7067.

Koo, H. S., and D. M. Crothers. 1988. Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc. Natl. Acad. Sci. USA.* 85:1763–1767.

Koo, H. S., J. Drak, J. A. Rice, and D. M. Crothers. 1990. Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry.* 29:4227–4234.

Koo, H. S., H. M. Wu, and D. M. Crothers. 1986. DNA bending at adenine · thymine tracts. *Nature.* 320:501–506.

Kozelka, J., M. H. Fouchet, and J. C. Chottard. 1992. H8 chemical shifts in oligonucleotides cross-linked at a GpG sequence by *cis*-$Pt(NH_3)_2^{2+}$: a clue to the adduct structure. *Eur. J. Biochem.* 205:895–906.

Lemaire, M. A., A. Schwartz, A. R. Rahmouni, and M. Leng. 1991. Interstrand cross-links are preferentially formed at the d(GC) sites in the reaction between *cis*-diamminedichloroplatinum(II) and DNA. *Proc. Natl. Acad. Sci. USA.* 88:1982–1985.

Leng, M. 1990. DNA bending induced by covalently bound drugs—gel electrophoresis and chemical probe studies. *Biophys. Chem.* 35: 155–163.

Malinge, J.-M., C. Perez, and M. Leng. 1995. Base sequence-independent distortions induced by interstrand cross-links in *cis*-diamminedichloroplatinum(II)-modified DNA. *Nucleic Acids Res.* 22:3834–3839.

Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499–560.

McCarthy, J. G., and A. Rich. 1991. Detection of an unusual distortion in A-tract DNA using KMnO4: effect of temperature and distamycin on altered conformation. *Nucleic Acids Res.* 19:3421–3429.

McCarthy, J. G., L. D. Williams, and A. Rich. 1990. Chemical reactivity of potassium permanganate and diethyl pyrocarbonate with B-DNA: specific reactivity with short A-tracts. *Biochemistry.* 29:6071–6081.

Mukundan, S., Jr., Y. Xu, G. Zon, and L. G. Marzilli. 1991. Heteronuclear ^{13}C-^{1}H NMR investigation of the effects on an oligodeoxyribonucleotide of intrastrand cross-linking by Pt anticancer drug. A large shift of C3′ accompanies an S to N conformational change. *J. Am. Chem. Soc.* 113:3021–3027.

Neumann, J. M., S. Tran-Dinh, J. P. Girault, J. C. Chottard, T. Huynh-Dinh, and J. Igolen. 1984. DNA fragment conformations. *A* 1-NMR conformational analysis of the d(G-G)-chelated platinum-oligonucleotide d(A-T-G-G)cisPt. *Eur. J. Biochem.* 141:465–472.

Nielsen, P. E. 1990. Chemical and photochemical probing of DNA complexes. *J. Mol. Recognition.* 3:1–24.

Noji, M., S. Motoyama, T. Tashiro, and Y. Kidani. 1983. Synthesis and antitumor activity of Pt(II) complexes containing 2,3-diaminopropanol isomers. *Chem. Pharm. Bull. Tokyo.* 31:1469–1473.

Noji, M., K. Okamato, and Y. Kidani. 1981. Relation of conformation to antitumor activity of platinum(II) complexes of 1,2-cyclohexanediamine and 2-(aminoethyl)cyclohexamine isomers against leukemia P388. *J. Med. Chem.* 24:508–515.

Nováková, O., O. Vrána, V. I. Kiseleva, and V. Brabec. 1995. DNA interactions of antitumor platinum(IV) complexes. *Eur. J. Biochem.* 228:616–624.

O'Dwyer, P. J., J. P. Stevenson, and S. W. Johnson. 1999. Clinical status of cisplatin, carboplatin, and other platinum-based antitumor drugs. *In* Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug. B. Lippert, editor. VHCA, WILEY-VCH, Zürich, Weinheim. 31–72.

Ohndorf, U.-M., M. A. Rould, Q. He, C. O. Pabo, and S. J. Lippard. 1999. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature.* 399:708–712.

Page, J. D., I. Husain, A. Sancar, and S. G. Chaney. 1990. Effect of the diaminocyclohexane carrier ligand on platinum adduct formation, repair, and lethality. *Biochemistry.* 29:1016–1024.

Pasini, A., and F. Zunino. 1987. New cisplatin analogs—on the way to better antitumor agents. *Angew. Chem. Int. Ed.* 26:615–624.

Pinto, A. L., and S. J. Lippard. 1985. Binding of the antitumor drug *cis*-diamminedichloroplatinum(II) (cisplatin) to DNA. *Biochim. Biophys. Acta.* 780:167–180.

Poklar, N., D. S. Pilch, S. J. Lippard, E. A. Redding, S. U. Dunham, and K. J. Breslauer. 1996. Influence of cisplatin intrastrand crosslinking on the conformation, thermal stability, and energetics of a 20-mer DNA duplex. *Proc. Natl. Acad. Sci. USA.* 93:7606–7611.

Reedijk, J. 1996. Improved understanding in platinum antitumour chemistry. *Chem. Commun.* 801–806.

Rhodes, D., and A. Klug. 1980. Helical periodicity of DNA determined by enzyme digestion. *Nature.* 286:573–578.

Rice, J. A., D. M. Crothers, A. L. Pinto, and S. J. Lippard. 1988. The major adduct of the antitumor drug *cis*-diamminedichloroplatinum(II) with DNA bends the duplex by 40° toward the major groove. *Proc. Natl. Acad. Sci. USA.* 85:4158–4161.

Ross, S. A., and C. J. Burrows. 1996. Cytosine-specific chemical probing of DNA using bromide and monoperoxysulfate. *Nucleic Acids Res.* 24:5062–5063.

Schwartz, A., M. Sip, and M. Leng. 1990. Sodium cyanide: a chemical probe of the conformation of DNA modified by the antitumor drug *cis*-diamminedichloroplatinum(II). *J. Am. Chem. Soc.* 112:3673–3674.

Sherman, S. E., and S. J. Lippard. 1987. Structural aspects of platinum anticancer drug interactions with DNA. *Chem. Rev.* 87:1153–1181.

Sundquist, W. I., and S. J. Lippard. 1990. The coordination chemistry of platinum anticancer drugs and related compounds with DNA. *Coord. Chem. Rev.* 100:293–322.

Takahara, P. M., C. A. Frederick, and S. J. Lippard. 1996. Crystal structure of the anticancer drug cisplatin bound to duplex DNA. *J. Am. Chem. Soc.* 118:12309–12321.

Ulanovsky, L., M. Bodner, E. N. Trifonov, and M. Choder. 1986. Curved DNA: design, synthesis, and circularization. *Proc. Natl. Acad. Sci. USA.* 83:862–866.

Van Boom, S. S. G. E., D. Z. Yang, J. Reedijk, G. A. Van der Marel, and A. H. J. Wang. 1996. Structural effect of intra-strand cisplatin-crosslink on palindromic DNA sequences. *J. Biomol. Struct. Dyn.* 13:989–998.

Vickery, K., A. M. Bonin, R. R. Fenton, S. O'Mara, P. J. Russell, L. K. Webster, and T. W. J. Hambley. 1993. Preparation, characterization,

cytotoxicity, and mutagenicity of a pair of enantiomeric platinum(II) complexes with the potential to bind enantioselectively to DNA. *J. Med. Chem.* 36:3663–3668.

Vrána, O., V. Boudný, and V. Brabec. 1996. Superhelical torsion controls DNA interstrand cross-linking by antitumor *cis*-diamminedichloroplatinum(II). *Nucleic Acids Res.* 24:3918–3925.

Wang, J. C. 1979. Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci. USA.* 76:200–203.

Žaludová, R., A. Žákovská, J. Kašpárková, Z. Balcarová, V. Kleinwächter, O. Vrána, N. Farrell, and V. Brabec. 1997. DNA interactions of bifunctional dinuclear platinum(II) antitumor agents. *Eur. J. Biochem.* 246: 508–517.

Zamble, D. B., and S. J. Lippard. 1999. The response of cellular proteins to cisplatin-damaged DNA. *In* Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug. B. Lippert, editor. VHCA, WILEY-VCH, Zürich, Weinheim. 73–110.

Biochemical Pharmacology, Vol. 60, pp. 1761–1771, 2000.

ISSN 0006-2952/00/$–see front matter
PII S0006-2952(00)00495-0

Modification of Natural, Double-Helical DNA by Antitumor *cis-* and *trans-*[$Cl_2(Me_2SO_4)_4Ru$] in Cell-Free Media

*Olga Nováková, Ctirad Hofr and Viktor Brabec**

INSTITUTE OF BIOPHYSICS, ACADEMY OF SCIENCES OF THE CZECH REPUBLIC, CZ-61265 BRNO, CZECH REPUBLIC

ABSTRACT. Modifications of natural DNA in cell-free media by the antitumor ruthenium compounds *cis-* and *trans-*[$Cl_2(Me_2SO_4)_4Ru$] were studied by various biochemical and biophysical methods. These methods included: binding studies by means of flameless atomic absorption spectrophotometry, mapping of DNA adducts by means of transcription assay, use of ethidium bromide as a fluorescent probe of DNA adducts of metal complexes, an interstrand cross-linking assay employing gel electrophoresis under denaturing conditions, measurements of DNA unwinding by gel electrophoresis, differential pulse polarographic analysis of DNA conformation, and analysis of liquid crystalline dispersions of DNA by circular dichroism. The results indicated that both ruthenium compounds irreversibly coordinated to DNA; the rate of binding of the *cis* isomer was considerably lower than that of the *trans* isomer. The DNA-binding mode of *trans-*[$Cl_2(Me_2SO_4)_4Ru$] included formation of bifunctional adducts such as intrastrand cross-links between neighboring purine residues and a small amount (~1%) of interstrand cross-links. *cis-*[$Cl_2(Me_2SO_4)_4Ru$] formed mainly monofunctional lesions on natural DNA. Both ruthenium isomers induced conformational alterations of non-denaturational character in DNA, the *trans* compound being more effective. In addition, DNA adducts of *trans-*[$Cl_2(Me_2SO_4)_4Ru$] were capable of inhibiting RNA synthesis by DNA-dependent RNA polymerases, while the adducts of the *cis* isomer were not. Thus, several features of the DNA-binding mode of *trans-*[$Cl_2(Me_2SO_4)_4Ru$] were similar to those of antitumor *cis*-diamminedichloroplatinum (II), which may be relevant to the biological effects of this antitumor ruthenium drug. On the other hand, the different DNA-binding mode of *cis-*[$Cl_2(Me_2SO_4)_4Ru$] was consistent with its less pronounced biological effects. BIOCHEM PHARMACOL **60**;12:1761–1771, 2000. © 2000 Elsevier Science Inc.

KEY WORDS. DNA adducts; DNA conformation; transcription; antitumor; cross-link; liquid crystals

In recent years, metal-based antitumor drugs have played a relevant role in antiblastic chemotherapy [1–4]. Cisplatin† in particular is regarded as one of the most effective anticancer drugs used in the clinic. In spite of the great efficacy of cisplatin against several human tumors, this drug displays limited activity against some of the most common tumors, such as colon and breast cancers [4, 5]. In addition, a variety of adverse effects and acquired resistance are observed in patients receiving cisplatin chemotherapy. The great success of cisplatin on the one hand and these limitations on the other have initiated efforts to develop new metal-based agents that will display improved therapeutic properties.

Broadening the spectrum of antitumor drugs depends on understanding existing agents, with a view toward developing new modes of attack. After the discovery of cisplatin, the new platinum antitumor drugs so far introduced in the clinic were *cis*-diammine-1,1-cyclobutanedicarboxylatoplatinum(II) (carboplatin) and (*trans-R,R*)1,2-diamminocyclohexaneoxalatoplatinum(II) (oxaliplatin) [5–7]. Carboplatin and oxaliplatin differ from cisplatin only in the more inert leaving group. Hence, as these new compounds are direct structural analogues of cisplatin that exhibit reactivity qualitatively similar to that of cisplatin [8, 9], it is not surprising that they also induce similar biological consequences [5–7].

One approach in the search for new, metal-based anticancer agents that would exhibit antitumor activity markedly different from that of cisplatin and its direct analogues is to examine complexes that would contain another transition metal. Possible advantages in using transition metal ions other than platinum may involve additional coordination sites, alterations in ligand affinity and substitution kinetics, changes in oxidation state, and photodynamic approaches to therapy [3]. In the design of these new drugs, ruthenium complexes have raised great interest [1, 3, 10].

* Corresponding author: Dr. Viktor Brabec, Institute of Biophysics, Academy of Sciences of the Czech Republic, CZ-61265 Brno, Czech Republic. Tel. +420 5 415 17 148; FAX +420 5 412 11 293; E-mail: brabec@ibp.cz

† *Abbreviations:* cisplatin, *cis*-diamminedichloroplatinum(II); Me_2SO_4, dimethyl sulfoxide; [Cl(dien)Pt]Cl, chlorodiethylenetriamineplatinum(II) chloride; CT, calf thymus; EtBr, ethidium bromide; PEG, poly(ethyleneglycol); r_b, the number of molecules of the ruthenium compound bound per nucleotide residue; FAAS, flameless atomic absorption spectrophotometry; r_i, the molar ratio of free ruthenium complex to nucleotide phosphates at the onset of incubation with DNA; TE buffer, 10 mM Tris · HCl plus 1 mM EDTA, pH 7.2; and DPP, differential pulse polarography.

Received 21 February 2000; accepted 17 April 2000.

The antitumor activity of platinum and other metal-based drugs is frequently related to their binding to DNA [3, 11]. These drugs form adducts on DNA that block DNA and RNA synthesis and induce programmed cell death. Thus, intracellular interactions with these adducts are likely to be of importance in explaining their biological activity or at least some of its features. Dimethyl sulfoxide complexes of both Ru^{II} and Ru^{III} constitute a relatively new group of anticancer compounds [3, 12, 13]. For instance, these complexes exhibit antiblastic activity comparable to cisplatin at equitoxic dosage in animal models of metastasizing tumors, but with less severe side effects and prolonged host survival times [12, 14, 15]. In addition, these ruthenium compounds inhibit DNA replication, exhibit mutagenic activity, and induce the SOS repair mechanisms, which is consistent with the DNA binding of these compounds *in vivo* [16, 17].

A small series of complexes whose parent compounds are *cis-* and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ constitute one class of dimethyl sulfoxide Ru^{II} compounds. The examination of their effect on primary tumor and on metastasis development has revealed antimetastatic activities superior to the effects on primary tumor growth [12, 14, 18]. The initial studies were performed with *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ because of its similarity to cisplatin [19]. However, the comparisons between the antitumor effects of *cis-* and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ revealed the superiority of the latter [15]. *cis-* and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ contain two chlorides in the octahedral structure [14, 20]. In *cis*-$[Cl_2(Me_2SO_4)_4Ru]$, the three Me_2SO_4 molecules are S-bound in a facial configuration and the fourth is O-bonded. In *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ all the Me_2SO_4s are S-bound. When dissolved in water, the *cis* isomer immediately undergoes loss of the O-bonded dimethyl sulfoxide ligand, whereas the *trans* compound rapidly loses two yielding *cis*-diaqua species. Both hydrolyzed isomers then undergo slow reversible chloride dissociation, forming cationic compounds. After this step, the *trans* compound contains three reactive groups, the *cis* isomer only two [14, 20]. In addition, the three remaining Me_2SO_4 ligands in the *cis* isomer represent a considerable steric hindrance, which makes the *cis*-aqua species relatively inert in contrast to the *trans* isomer. Importantly, this difference correlates with a higher potency of the *trans* isomer to act as the antitumor agent [21].

Both *cis-* and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ bind to DNA in cell-free media [14, 21]. Whereas DNA modifications by platinum antitumor drugs have been studied systematically and described in detail [22], the modifications of high-molecular-mass natural DNA by *cis-* and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ have been much less explored [3]. Some early studies based on the analysis of circular dichroism (CD) spectra of DNA suggested that coordination of the *cis* isomer to DNA does not significantly alter the conformation of B-DNA [14, 23]. The *trans* isomer binds to DNA more rapidly, with some changes in the CD spectra indicating conformational alterations [14]. Studies of the interactions of both isomers with monomeric constituents of nucleic acids and very short single-stranded oligonucleotides (di- and tetranucleotides) have revealed that both isomers may have some preference for bifunctional binding to neighboring guanine residues at their N7 atoms, with the *trans* isomer being more effective [24–27]. This suggestion has been indirectly corroborated by the restriction enzyme analysis of plasmid DNA modified by these ruthenium compounds [21]. These studies showed that DNA modified by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ was protected most effectively from cutting by *Bam*HI restriction endonuclease, which contains two neighboring guanine residues in its recognition sequence.

In order to address further fundamental questions about the DNA-binding modes of ruthenium antitumor compounds containing Me_2SO_4 ligands, the experiments described in the present paper were carried out. More specifically, the interactions of polymeric natural DNAs with *cis-* and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ in cell-free media were investigated by various methods of molecular biophysics, the aim being to contribute to the understanding of the differences in the biological effects of these isomers.

MATERIALS AND METHODS

Starting Materials

cis- and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ were synthesized, recrystallized, and characterized as previously reported [20]. Cisplatin and [Cl(dien)Pt]Cl were synthesized and characterized in Lachema. Stock solutions of the ruthenium and platinum complexes (5×10^{-4} M in 10 mM $NaClO_4$) were prepared in the dark at 25°. CT DNA (42% G + C, mean molecular mass *ca.* 20,000 kDa) was also prepared and characterized as described previously [28, 29]. CT DNA used for the formation of liquid crystalline dispersions was depolymerized by ultrasound, so that its average molecular mass was ~500 kDa (determined by agarose gel electrophoresis). Plasmids pSP73 (2464 bp) and pSP73KB (2455 bp [30]) were isolated according to standard procedures and banded twice in CsCl/EtBr equilibrium density gradients. Restriction endonucleases were purchased from New England Biolabs. Klenow fragment of DNA polymerase I was from Boehringer Mannheim Biochemica. Riboprobe Gemini System II for transcription mapping containing T7 and SP6 RNA polymerase was purchased from Promega. EtBr and agarose were from Merck KgaA. PEG (molecular mass 4600 kDa) was purchased from Sigma. The radioactive products were from Amersham.

Metallation Reactions

CT or plasmid DNAs were incubated with the ruthenium or platinum complex in 10 mM $NaClO_4$ at 37° in the dark if not stated otherwise. After 48 hr, the samples of plasmid DNA were precipitated by ethanol and redissolved in the medium required for subsequent biochemical or biophysical analysis, whereas the samples of CT DNA were exhaus-

tively dialyzed against such a medium. An aliquot of these samples was used to determine the r_b value by FAAS.

The dependencies of the r_b values on time for the ruthenium complexes tested in the present work were obtained in the following way. Solutions of CT DNA at the concentration of *ca.* 0.1 mg/mL were incubated with the ruthenium complex at an r_i value of 0.1. At the various time intervals, aliquots were withdrawn and the reaction was stopped by adding 1/10 of a volume of 1.5 M NaCl, including quick cooling to −20°. DNA was then precipitated by ethanol. The amount of the free ruthenium complexes (not bound to DNA) in supernatant was determined by FAAS. The amount of ruthenium bound to DNA (r_b) was calculated by subtracting the amount of ruthenium remaining in solution (determined by FAAS) from the total amount of ruthenium present in the reaction.

DNA Transcription by RNA Polymerase In Vitro

Transcription of the (*Nde*I/*Hpa*I) restriction fragment of pSP73KB DNA with DNA-dependent T7 and SP6 RNA polymerases and electrophoretic analysis of the transcripts were performed according to the protocols recommended by Promega (Promega Protocols and Applications, 43–46 [1989/90]) and previously described in detail [30, 31].

Interstrand Cross-Link Assay

If not stated otherwise, *cis-* and *trans-*$[Cl_2(Me_2SO_4)_4Ru]$ at varying concentrations were incubated with 2 μg of pSP73 DNA either in negatively supercoiled form or after it had been linearized by *Eco*RI. The modified samples were precipitated by ethanol, and the circular DNA already modified by ruthenium complex was subsequently linearized by *Eco*RI. The linear duplexes were then analyzed for DNA interstrand cross-links in the same manner as described in several recent papers [31, 32]. The linear duplexes were first 3′-end labeled by means of Klenow fragment of DNA polymerase I and [α-^{32}P]dATP. The samples were deproteinized by phenol, precipitated by ethanol, and the pellet dissolved in 18 μL of 30 mM NaOH with 1 mM EDTA, 6.6% sucrose, and 0.04% bromophenol blue. The amount of interstrand cross-links was analyzed by electrophoresis under denaturing conditions on alkaline agarose gel (1%). After the electrophoresis was completed, the intensities of the bands corresponding to single strands of DNA and interstrand cross-linked duplex were quantified by means of a Molecular Dynamics PhosphorImager (Storm 860 system with ImageQuant software). The frequency of interstrand cross-links, F (the number of interstrand cross-links per adduct), was calculated as $F = \frac{XL}{4928 \cdot r_b}$ (pSP73 plasmid contained 4928 nucleotide residues). XL is the number of interstrand cross-links per one molecule of the linearized DNA duplex which was calculated assuming Poisson distribution of the interstrand cross-links as XL = −ln A, where A is the fraction of molecules running as a band corresponding to the non-cross-linked DNA [32].

Fluorescence Measurements

Fluorescence measurements of DNA modified by ruthenium complexes in the presence of EtBr were performed at an excitation wavelength of 546 nm, and the emitted fluorescence was analyzed at 590 nm. The fluorescence intensity was measured at 25° in 0.4 M NaCl to avoid secondary binding of EtBr to DNA [33, 34]. The concentrations were 0.01 mg/mL for DNA and 0.04 mg/mL for EtBr, which corresponded to the saturation of all intercalation sites of EtBr in DNA [33, 34]. These measurements were performed on a Shimadzu RF 40 spectrofluorophotometer using a 1-cm quartz cell.

Differential Pulse Polarography

DPP curves of DNA were measured after non-modified DNA or DNA modified by the ruthenium complex was redissolved in the medium of 0.3 M ammonium formate plus 0.05 M phosphate (Na_2HPO_4/NaH_2PO_4) buffer, pH 6.8 [35]. DPP curves were recorded with the aid of an EG&C PARC Electrochemical Analyzer, Model 384B at 25° using the following apparatus settings: voltage scan rate 2 mV/sec, pulse amplitude of 5 mV, drop time of 1.0 sec. The potentials are against the saturated calomel reference electrode.

Unwinding of Negatively Supercoiled DNA

Unwinding of closed circular supercoiled pSP73 plasmid DNA was assayed by an agarose gel mobility shift assay [36]. The unwinding angle Φ, induced per platinum–DNA adduct, was calculated upon the determination of the r_b value at which the complete transformation of the supercoiled to relaxed form of the plasmid was attained. Samples of pSP73 plasmid were incubated with *cis-* or *trans-*$[Cl_2(Me_2SO_4)_4Ru]$ for 48 hr, precipitated by ethanol, and redissolved in TAE buffer (0.04 M Tris–acetate + 1 mM EDTA, pH 7.0). An aliquot of the precipitated sample was subjected to electrophoresis on 1% agarose gels running at 25° in the dark with TAE buffer with a voltage set at 30 V. The gels were then stained with EtBr, followed by photography on Polaroid 667 film with transilluminator. The other aliquot was used for the determination of r_b values by FAAS.

Liquid Crystalline Dispersions of DNA

Liquid crystalline dispersions of DNA modified by the metal complexes were formed by mixing DNA and PEG solutions as described earlier [37, 38]. Briefly, 1 mL of non-modified DNA or DNA modified by the metal complex dissolved in 0.01 M $NaClO_4$ at the concentration of

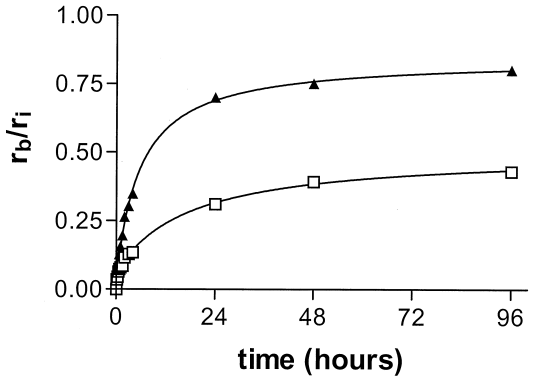


FIG. 1. Kinetics of the binding of *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ (□) and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ (▲) to CT DNA. Medium: 10 mM $NaClO_4$ at 37°. The binding was determined by FAAS after precipitation by ethanol. The concentration of DNA was 0.1 mg/mL, $r_i = 0.1$. Data measured in triplicate varied on average ±3% from their mean. For other details, see the text.

0.06 mg/mL was mixed vigorously with 3 mL PEG at the concentration of 200 mg/mL (also dissolved in 0.3 M $NaClO_4$) for 1 hr.

Other Methods

Absorption spectra were measured with a Beckmann DU-8 spectrophotometer. FAAS measurements were carried out on a Unicam 939 AA spectrometer with a graphite furnace. For FAAS analysis, DNA was precipitated with ethanol and dissolved in 0.1 M HCl. CD spectra were recorded at 25° using a JASCO spectropolarimeter, Model J720.

RESULTS

DNA Binding

Solutions of double-helical CT DNA at a concentration of 0.1 mg/mL were incubated with *cis-* and *trans-*$[Cl_2(Me_2SO_4)_4Ru]$ at r_i values of 0.1 at 37° in three media: 10 mM $NaClO_4$, TE buffer, or 1 mM phosphate buffer (NaH_2PO_4/Na_2HPO_4, pH 7.0). At various time intervals, an aliquot of the reaction mixture was withdrawn; the reaction was then terminated by adjusting NaCl concentration to 0.15 M followed by quick cooling to −20° and precipitating DNA by ethanol. The supernatant was assayed by FAAS for the amount of free ruthenium not bound to DNA. The amount of ruthenium bound to DNA, r_b, was calculated by subtracting the amount of ruthenium remaining in solution from the total amount of ruthenium present in the reaction. Figure 1 shows a plot of r_b against the time of DNA incubation with *cis-* or *trans-*$[Cl_2(Me_2SO_4)_4Ru]$ in 10 mM $NaClO_4$. The amount of ruthenium coordinated to DNA increased with time. After 48 hr, the binding reached its maximum values, which corresponded to approximately 40 or 75% of the molecules of *cis-* or *trans-*$[Cl_2(Me_2SO_4)_4Ru]$, respectively, present in the reaction mixtures. Importantly, both ruthenium compounds reacted with DNA at a similar rate, with r_i values in the range of 0.001–0.1. Thus, the rate of binding to natural polymeric DNA of both ruthenium compounds tested in the present work was lower than that of cisplatin [39]. These results also confirmed that DNA binding of the *cis-*$[Cl_2(Me_2SO_4)_4Ru]$ is considerably less effective than that of its *trans* isomer.

The binding studies described above were performed in the medium of 10 mM $NaClO_4$, which is used as a standard medium if DNA interactions of antitumor platinum complexes are investigated in cell-free media. For comparative purposes, the binding of the two ruthenium compounds to DNA was also investigated in TE or phosphate buffers, i.e. in the media used in some previous studies of DNA modifications by ruthenium complexes. The components of these buffers are potential ligands of transition metal complexes that compete with DNA for metal [40]. This fact can significantly affect the rate of the binding of ruthenium complexes to DNA. DNA binding of *trans-*$[Cl_2(Me_2SO_4)_4Ru]$ measured after 24 hr in TE or phosphate buffers was decreased from 70% in 10 mM $NaClO_4$ to 20 or 50%, respectively. Similarly, DNA binding of the *cis* isomer under the same experimental conditions was decreased from 30% to 5 or 10%, respectively. Identical results were obtained when the binding of both isomers was measured in the media containing TE or phosphate buffers at ten times lower concentrations. On the other hand, if DNA was first modified by *cis-* or *trans-*$[Cl_2(Me_2SO_4)_4Ru]$ in 10 mM $NaClO_4$ to an r_b value in the range of 0.001–0.05, the level of the modification (r_b) remained unchanged even after DNA modified by ruthenium complex was transferred into the media in which their rate of DNA binding is noticeably decreased (*vide supra*). This observation was also consistent with the results of the following experiment. CT DNA modified by the ruthenium complex in 10 mM $NaClO_4$ was precipitated by ethanol, redissolved in the medium of TE or phosphate buffer (containing no ruthenium complex), further incubated at 37° for an additional 48 hr, again precipitated by ethanol, and the content of ruthenium in these samples (r_b) determined by FAAS. No changes in r_b values were noticed.

The binding experiments of the present work indicate that modification reactions resulted in the irreversible coordination of molecules of the ruthenium complexes to polymeric double-helical DNA. In addition, these results made it possible to prepare easily and precisely the samples of DNA modified by the ruthenium complexes at a preselected value of r_b. Importantly, it is also reasonable to assume that the levels of DNA modifications by the ruthenium complexes reached in 10 mM $NaClO_4$ remained unchanged after the sample of DNA modified by the ruthenium complex was transferred into the medium required for subsequent biophysical or biochemical analyses. Thus, the samples of DNA modified by the ruthenium complexes and analyzed further by biophysical or biochemical methods were prepared in 10 mM $NaClO_4$ at 37°. After 48 hr of the reaction of DNA with the complex, the samples were precipitated in ethanol, dissolved in the

medium necessary for a particular analysis, and the r_b value in an aliquot of this sample checked by FAAS. In this way, the analyses described in the present paper were performed in the absence of unbound (free) ruthenium complex.

Transcription Mapping of DNA Adducts

In vitro RNA synthesis by RNA polymerases on DNA templates containing several types of bifunctional adducts of metal complexes can be prematurely terminated at the level or in the proximity of adducts [30, 31, 41, 42]. Importantly, monofunctional DNA adducts of several platinum complexes are unable to terminate RNA synthesis.

Cutting of pSP73KB DNA [30, 31] by *Nde*I and *Hpa*I restriction endonucleases yielded a 212-bp fragment (a substantial part of its nucleotide sequence is shown in Fig. 2B). This fragment contained convergent T7 and SP6 RNA polymerase promoters (in the upper and lower strands, respectively, close to its 3′-ends [Fig. 2B]). The experiments were carried out using this linear DNA fragment, modified by *cis*-$[Cl_2(Me_2SO_4)_4Ru]$, its *trans* isomer, or cisplatin at $r_b = 0.01$, for RNA synthesis by T7 and SP6 RNA polymerases (Fig. 2A, lanes cisRu, transRu, or cisPt, respectively). RNA synthesis on the template modified by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ or cisplatin yielded fragments of defined sizes (Fig. 2A, lanes transRu and cisPt), which indicates that RNA synthesis on these templates was prematurely terminated. The major stop sites produced by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ were identical to those produced by cisplatin, i.e. mainly appearing at G sites and to a considerably lesser extent at A sites. These G and A sites were mostly contained in GG or AG sites, which are preferential DNA-binding sites of cisplatin. Interestingly, the bands produced by the *trans* ruthenium complex had a lower intensity than those yielded by cisplatin, which suggests that *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ also produces DNA adducts incapable of terminating RNA synthesis and that the amount of these DNA adducts (presumably monofunctional lesions) was considerably higher than that formed in the case of cisplatin. Taken together, the results of these mapping experiments suggest that *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ can also form, on double-helical DNA, an amount of bidentate adducts similar to DNA cross-links formed by cisplatin.

DNA adducts of *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ produced only faint bands corresponding to stop sites (Fig. 2A, lane cisRu). This observation is consistent with capability of *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ to form a considerably smaller amount of adducts on DNA that would be similar to those formed by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ or cisplatin.

Characterization of DNA Adducts by EtBr Fluorescence

EtBr as a fluorescent probe has been used to characterize perturbations induced in DNA by adducts of several platinum compounds [33, 43–45]. Double-helical CT DNA was modified by cisplatin, monofunctional [Cl(dien)Pt]Cl, and

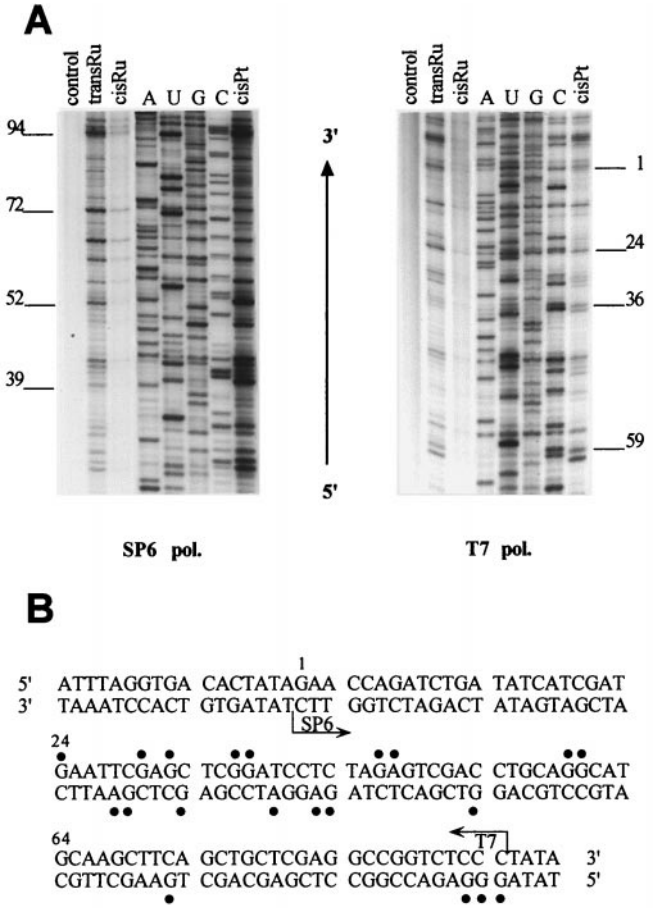


FIG. 2. Inhibition of RNA synthesis by SP6 (left) and T7 (right) RNA polymerases on the *Nde*I/*Hpa*I fragment of pSP73KB plasmid modified by ruthenium and platinum complexes. (A) Autoradiograms of 6% polyacrylamide/8 M urea sequencing gels. Lanes: control, non-modified template; cisPt, cisRu, and transRu, the templates modified by cisplatin, *cis*-$[Cl_2(Me_2SO_4)_4Ru]$, and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ at $r_b = 0.01$, respectively. (B) Schematic diagram showing the portion of the nucleotide sequence of the template (upper) strand of the *Nde*I/*Hpa*I fragment used to monitor inhibition of RNA synthesis by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$. The arrows indicate the start of the T7 or SP6 RNA polymerases. (●), major stop signals (from Fig. 2A, lanes transRu). The numbers correspond to the nucleotide numbering in the sequence map of pSP73KB plasmid.

cis- or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$. The levels of the modification corresponded to the values of r_b in the range of 0–0.1. Modification of DNA by all-metal complexes resulted in a decrease in EtBr fluorescence as compared with the control DNA–EtBr complex (Fig. 3). The decrease caused by the adducts of *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ complexes was similar to that induced by the DNA adducts of cisplatin at equivalent r_b albeit somewhat smaller. On the other hand, modification of DNA by *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ resulted in a pronouncedly smaller decrease in EtBr fluorescence as compared with that due to the modification by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ or cisplatin, but one that was still larger than that due to the modification by monofunctional [Cl(dien)Pt]Cl.

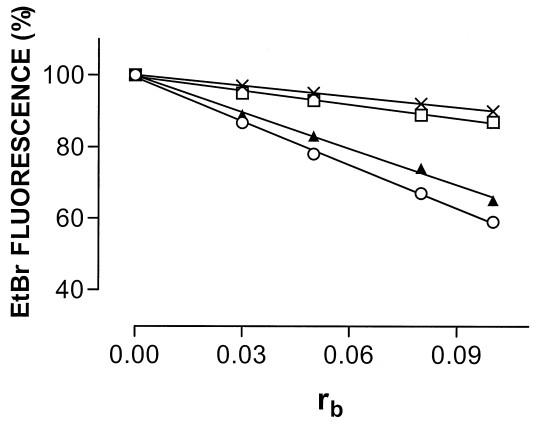


FIG. 3. Dependencies of ethidium bromide fluorescence on r_b for CT DNA modified by various metal complexes in 10 mM $NaClO_4$ at 37° for 48 hr. [Cl(dien)Pt]Cl (x), cisplatin (○), *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ (□), and *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ (▲). Data measured in triplicate varied on average ±2% from their mean.

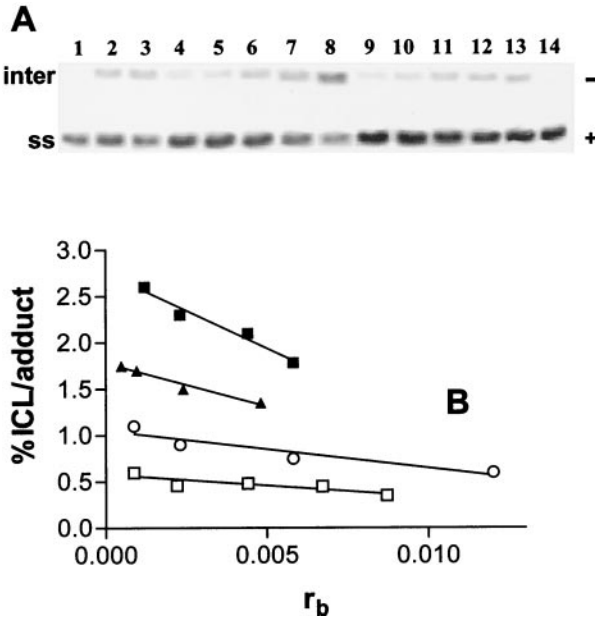


FIG. 4. The formation of interstrand cross-links by ruthenium complexes in negatively supercoiled and linearized pSP73 plasmid. (A) Autoradiogram of a denaturing 1% agarose gel of linearized DNA which was 3′-end labeled; the interstrand cross-linked DNA appears as the top bands migrating on the gels more slowly than the single-stranded (ss) DNA (contained in the bottom bands). The plasmid linearized by *Eco*RI was incubated with cisplatin (lanes 2 and 3), *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ (lanes 4–8), or *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ (lanes 9–13) for 48 hr at 37°; r_b values: 0 (control, non-modified DNA), lanes 1 and 14; 0.0005, lane 2; 0.001, lane 3; 0.0048, lane 4; 0.00096, lane 5; 0.0024, lane 6; 0.0048, lane 7; 0.0096, lane 8; 0.00087, lane 9; 0.0022, lane 10; 0.0044, lane 11; 0.0067, lane 12; 0.0087, lane 13. (B) Dependence on r_b of the percentage of interstrand cross-links (ICL) per adduct (interstrand cross-linking [%]) formed by *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ (▲, □) or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ (■, ○) in supercoiled (closed symbols) or linearized (open symbols) DNA within 48 hr. Data measured in triplicate varied on average ±3% from their mean.

Interstrand Cross-linking

Transcription mapping experiments and EtBr fluorescence analysis (Figs. 2 and 3) were consistent with the capability of ruthenium complexes to form on polymeric DNA bidentate adducts, but these studies could not distinguish whether these adducts were intrastrand or interstrand cross-links. Therefore, further experiments were carried out to compare the amounts of the interstrand cross-links formed by *cis*- or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ in negatively supercoiled or linear DNA. In these experiments, we used pSP73 plasmid (2464 bp, native supercoil density $\sigma = -0.063$), which was modified by *cis*- or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ complexes in two ways. One series of samples was prepared by modifying the plasmid only after it had been linearized by *Eco*RI (*Eco*RI cuts only once within pSP73 plasmid). The samples of the other series were prepared by a modification of the supercoiled plasmid by ruthenium complexes, and only after the modification reaction was completed was the plasmid linearized by *Eco*RI. Thus, we prepared two types of linear DNA molecules of the same length and nucleotide sequence modified by ruthenium complexes, which could be differently affected by this drug if different DNA topology were to play a role during the modification reaction. The two samples were analyzed for interstrand cross-links by agarose gel electrophoresis under denaturing conditions in an attempt to reveal these differences. It was also verified by FAAS whether the amount of ruthenium complexes coordinated to the base residues in DNA was independent of DNA topology during the modification reaction over a broad range of r_i values (0.001–0.1) and at any reaction time.

An electrophoretic method for precise and quantitative determination of interstrand cross-linking by metal complexes in DNA has been described previously [31, 32]. Upon electrophoresis under denaturing conditions, 3′-end labeled strands of linearized pSP73 plasmid containing no interstrand cross-links migrate as a 2464-nucleotide single strand, whereas the interstrand cross-linked strands migrate more slowly as a higher-molecular-mass species. The bands corresponding to more slowly migrating interstrand cross-linked fragments were observed if the *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ complex was used to modify DNA in both linearized and supercoiled forms at r_b as low as 5×10^{-3} (Fig. 4A). The intensity of the more slowly migrating band increased with the growing level of the modification. The radioactivity associated with the individual bands in each lane was measured to obtain estimates of the fraction of non-cross-linked or cross-linked DNA under each condition. The frequency of interstrand cross-links (the amount of interstrand cross-links per one molecule of *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ complex coordinated to DNA) was calculated using the Poisson distribution in combination with the r_b values and the fragment size [32] (for more details, see Materials and Methods).

As summarized in Fig. 4B, *trans*-$[Cl_2(Me_2SO_4)_4Ru]$

showed a relatively low interstrand cross-linking efficiency in both linear and negatively supercoiled DNA, i.e. approximately 6 times less than formed by cisplatin under identical conditions [31, 46]. Importantly, *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ forms an even smaller amount of interstrand cross-links (roughly half), which is consistent with a tendency of the *cis* isomer to preferentially form on double-helical DNA monofunctional adducts. These results indicate that the interstrand cross-links are only highly minor adducts in double-helical DNA modified by $[Cl_2(Me_2SO_4)_4Ru]$ complexes.

Differential Pulse Polarography

DPP analysis readily and with great sensitivity distinguishes between non-denaturational and denaturational conformational alterations induced in DNA by various physical or chemical agents [47, 48]. This analysis is based on the observation that intact double-helical DNA is polarographically inactive, because its reduction sites are involved in hydrogen bonds and are unable to make contact with the working electrode in a manner suitable for electron transfer. Electroreduction of adenine or cytosine residues present in distorted but still double-stranded (non-denatured) regions of DNA is responsible for the appearance of the small DPP peak II (Fig. 5A, curve 1). Base residues in these distorted regions become more accessible for electroreduction at the mercury electrode and can yield a small polarographic current. On the other hand, the appearance of a more negative peak III on DPP curves of DNA indicates the presence of single-stranded, denatured regions in the DNA molecule, in which hydrogen bonds between complementary bases have been broken [47, 48]. Differences in the adsorption properties of double-helical and denatured DNA at the mercury electrode have been suggested to give rise to the different reduction potentials that are observed for the two DNA conformations. Importantly, less than 1% denatured material in the excess of double-helical DNA can be determined by DPP [49].

DPP has already been used to analyze DNA modified by various physical or chemical agents, including platinum compounds with different clinical efficacy [50, 51]. It has been found that DNA globally modified by antitumor cisplatin or its analogues at r_b values up to 0.05 yields the DPP peak II, indicating that these antitumor drugs induce non-denaturational conformational changes in DNA [32, 35]. In contrast, the more negative DPP peak III is observed on DPP curves of DNA globally modified by clinically ineffective transplatin and other inactive platinum(II) complexes, indicating that the clinically ineffective platinum complexes induce denaturational conformational alterations in DNA [50, 51].

DPP analysis also sheds considerable light on the conformational basis for DNA binding of *cis*- or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$. The modification of CT DNA by these ruthenium complexes at r_b of 0.005–0.02 resulted in an increase in DPP peak II with a growing level of modification (shown for the *trans* isomer in Fig. 5A). The more negative peak III was not detected even on the DPP curves recorded for DNA modified at the highest r_b value used in our experiments (0.02). It could be argued that the absence of peak III on the DPP curves recorded for the samples of DNA modified at relatively high r_b values (~0.02) could be due to an increase in the slope of the part of the DPP curve corresponding to the background electrolyte discharge (Fig. 5A, curve 4). The observation that peak III was not buried under the background electrolyte discharge curve was

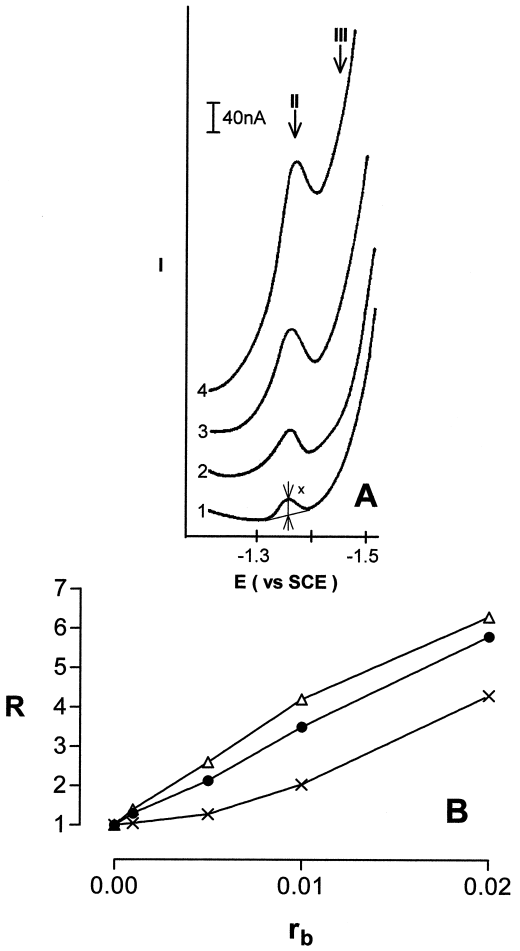


FIG. 5. Differential pulse polarographic analysis of CT DNA modified by ruthenium complexes. DNA at a concentration of 0.4 mg/mL in 0.3 M ammonium formate with 0.01 M phosphate buffer, pH 6.8. (A) DPP curves; DNA modified by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$; r_b values: 0 (control, non-modified DNA), curve 1; 0.005, curve 2; 0.01, curve 3; and 0.02, curve 4. The arrows in Fig. 5A marked by II and III indicate potentials E (against saturated calomel electrode [SCE]) at which native or denatured DNA samples yielded DPP peaks II or III, respectively (see the text). The way in which the height of the DPP peak (x) was measured in the present work is shown in Fig. 5A, curve 1. (B) Dependence of the relative height of DPP peak II, R, yielded by DNA modified by the ruthenium complexes on r_b: (x—x), *cis*-$[Cl_2(Me_2SO_4)_4Ru]$; (●——●), *trans*-$[Cl_2(Me_2SO_4)_4Ru]$; and (△——△), cisplatin. The value of R was calculated as the ratio of the peak height yielded by the modified DNA over the peak height yielded by the control (non-modified) DNA.

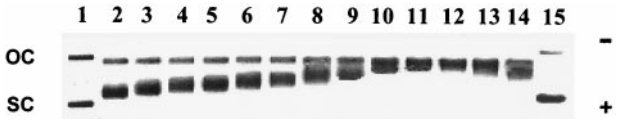


FIG. 6. Unwinding of negatively supercoiled pSP73 plasmid DNA by ruthenium complexes. The plasmid was incubated with *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ for 48 hr at 37° and precipitated by ethanol. r_b values: 0 (control, non-modified DNA), lanes 1 and 15; 0.06, lane 2; 0.07, lane 3; 0.08, lane 4; 0.09, lane 5; 0.1, lane 6; 0.11, lane 7; 0.12, lane 8; 0.13, lane 9; 0.14, lane 10; 0.15, lane 11; 0.16, lane 12; 0.2, lane 13; and 0.26, lane 14. The top bands (oc) correspond to the form of nicked plasmid and the bottom bands (sc) to the closed, negatively supercoiled plasmid.

verified using the sample of DNA modified by *cis-* or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ at $r_b = 0.02$ to which 0.8% thermally denatured CT DNA was added. This sample yielded a small, more negative peak III on the DDP curve (recorded under conditions specified for curve 4 in Fig. 5A) (not shown). Thus, the absence of the peak III on the DPP curves of DNA modified by *cis-* or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ suggests that these ruthenium complexes, as with antitumor cisplatin and other antitumor analogues of this drug, induce non-denaturational conformational distortions in DNA at relatively low levels of the global modification ($r_b \leq 0.02$). The DNA-binding mode of *cis-* or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ is also in profound contrast to the modification of DNA by clinically ineffective transplatin [50, 51]. In addition, the relative increase in peak II due to the global modification by *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ was considerably smaller at the same level of DNA modification (r_b) (Fig. 5B) than the increase in peak II due to the modification by *trans*-$[Cl_2(Me_2SO_4)_4Ru]$. This finding supports the view that non-denaturational distortions of DNA due to the global binding of *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ are more extensive than those due to the binding of *cis*-$[Cl_2(Me_2SO_4)_4Ru]$.

Unwinding

Electrophoresis in native agarose gel was used to quantify the unwinding induced in pSP73 plasmid by the ruthenium complexes by monitoring the degree of supercoiling (Fig. 6). A compound that unwinds DNA duplex reduces the number of supercoils in closed, negatively supercoiled DNA so that the negative superhelical density of closed circular DNA decreases. This decrease upon binding of unwinding agents causes a decrease in the rate of migration through agarose gel, which makes it possible for the unwinding to be observed and quantified [36]. Figure 6 shows an electrophoresis gel in which increasing amounts of *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ were bound to a mixture of relaxed and negatively supercoiled pSP73 DNA. Interestingly, *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ accelerated the mobility of the relaxed form in a similar fashion to cisplatin, whose bifunctional binding to DNA shortens and condenses the DNA helix [52, 53]. In contrast, *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ affected the mobility of the relaxed form markedly less (not shown), implying that its binding results in shortening or condensation of DNA to a very small extent. The unwinding angle is given by $\Phi = 18\ \sigma/r_b(c)$, where σ is the superhelical density and $r_b(c)$ is the value of r_b at which the supercoiled and relaxed forms co-migrate [36]. Under the present experimental conditions, σ was calculated to be -0.063 on the basis of the data for cisplatin, for which the $r_b(c)$ was determined in this study and $\Phi = 13°$ was assumed. Using this approach, the DNA unwinding angle of $7 \pm 1°$ was determined for *trans*-$[Cl_2(Me_2SO_4)_4Ru]$. On the contrary, the co-migration of the relaxed and negatively supercoiled DNAs was not observed even at such a high level of modification by *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ as is that corresponding to the r_b value of 0.25, indicating a negligible efficiency of the *cis* isomer to unwind DNA.

Liquid Crystals

The cholesteric liquid crystalline dispersions of DNA, which can simulate the principal properties of DNA molecules within cells, such as their spatial ordering in condensed and packed state, have been used as model systems *in vivo* [38, 54]. This model system was used in the present work to provide information on the possible consequences of the binding of *cis-* or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ on the formation and stability of the condensed DNA. In the presence of PEG water-containing salt solutions, linear double-stranded DNA is condensed, forming helically twisted liquid crystalline dispersions of DNA molecules (the left-handed helicoidal structure of cholesteric phase from the right-handed DNA molecules). The peculiar properties of this phase caused by anisotropic orientations can be analyzed by measuring CD spectra. The occurrence of the cholesteric liquid crystalline form of DNA is accompanied by the origin of the intense negative CD band at *ca.* 275 nm (Fig. 7) [55–57]. The decrease in the amplitude of this band in the CD spectra of liquid crystalline dispersions of DNA modified by antitumor platinum complexes is associated with the disappearance of the helical twist of the liquid crystalline microphase due to alterations of DNA secondary structure, such as disturbances in the stacking interactions of bases [37, 55, 58]. Interestingly, clinically ineffective platinum complexes, such as *trans* isomer of cisplatin or monofunctional platinum(II) compounds, only negligibly or weakly affected the CD spectra of liquid crystalline dispersions of DNA. Thus, the CD spectra of liquid crystalline microphases of DNA modified by platinum compounds are very sensitive to the status of the antitumor activity of platinum compounds and have been suggested [37] for preliminary testing of the antitumor activity of novel metal-based drugs.

We recorded CD spectra of the liquid crystalline dispersions of sonicated CT DNA modified by *cis-* or *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ at a wide range of r_b values in the presence of PEG (shown for the *trans* isomer in Fig. 7A), and the data were compared with those obtained for DNA modified by cisplatin and monofunctional [Cl(dien)Pt]Cl. In the CD spectrum of the liquid crystalline dispersions of

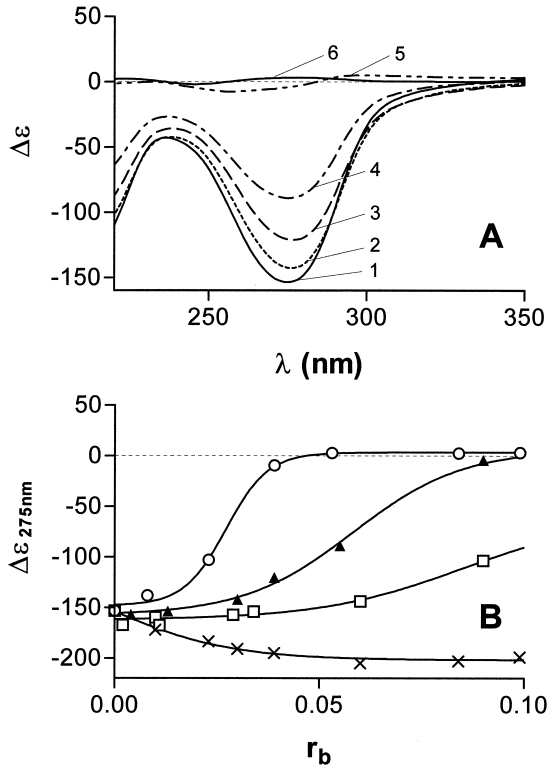


FIG. 7. (A) Circular dichroism spectra of liquid crystalline dispersions formed from non-modified CT DNA and CT DNA modified by *trans*-[$Cl_2(Me_2SO_4)_4Ru$] in the presence of 150 mg of PEG/mL. r_b values: 0 (control, non-modified DNA), curve 1; 0.03, curve 2; 0.039, curve 3; 0.055, curve 4; 0.098, curve 5; and 0.12, curve 6. (B) The dependence on r_b of the relative amplitude of the negative circular dichroism band at ~275 nm, $\Delta\epsilon_{275nm}$, of liquid crystalline dispersions formed from CT DNA modified by [Cl(dien)Pt]Cl (x), cisplatin (○), *cis*-[$Cl_2(Me_2SO_4)_4Ru$] (□), and *trans*-[$Cl_2(Me_2SO_4)_4Ru$] (▲) in the presence of 150 mg PEG/mL. For other details, see the text.

the control, non-modified DNA, an intense negative band at around 275 nm was observed. The amplitude of this band decreased as a consequence of the increasing amount of *cis*- or *trans*-[$Cl_2(Me_2SO_4)_4Ru$] bound to DNA (Fig. 7B). At the level of binding of *trans*-[$Cl_2(Me_2SO_4)_4Ru$] corresponding to $r_b \geq 0.1$, the negative band disappeared. It is also evident from Fig. 7B that *cis*-[$Cl_2(Me_2SO_4)_4Ru$] was markedly less efficient in decreasing the amplitude of the negative CD band at 275 nm than its *trans* isomer.

The results shown in Fig. 7B also indicate that *cis*-[$Cl_2(Me_2SO_4)_4Ru$] only disturbs DNA liquid crystals weakly, but more than monofunctional [Cl(dien)Pt]Cl. On the other hand, although *trans*-[$Cl_2(Me_2SO_4)_4Ru$] disturbs these crystals markedly more than its *cis* isomer, its efficiency is in this respect lower than that of cisplatin. Hence, these results are consistent with the view that disturbance of the secondary structure of DNA by *trans*-[$Cl_2(Me_2SO_4)_4Ru$] has some features similar to those induced by cisplatin, which might be relevant to the antitumor activity observed for this ruthenium compound earlier. On the other hand, *cis*-[$Cl_2(Me_2SO_4)_4Ru$] could be rather grouped with clinically less effective or ineffective metal-based compounds that alter conformation of DNA in a way only resulting in a negligible or weak disturbance of its liquid crystalline dispersions.

DISCUSSION

The present work demonstrates that antitumor *cis*- and *trans*-[$Cl_2(Me_2SO_4)_4Ru$] irreversibly coordinate to the residues in high-molecular-mass DNA. The rate of DNA binding of *cis*-[$Cl_2(Me_2SO_4)_4Ru$] is considerably lower than that of the *trans* isomer. The lower reactivity of the *cis* isomer is very likely associated with the presence of three bulky Me_2SO_4 ligands after this compound is dissolved in water, in contrast to only two Me_2SO_4 groups remaining in the dissolved molecules of the *trans*-isomer. Importantly, the reactivity of these metal-based compounds toward DNA is strongly affected by potential ligands of ruthenium complexes that may be present in the reaction medium, for instance as components of the reaction buffers.

Some details about the nature of the adducts formed by *cis*- and *trans*-[$Cl_2(Me_2SO_4)_4Ru$] on natural, double-helical DNA have emerged from the present work, such as transcription mapping of these adducts, the analysis of DNA by the EtBr fluorescent probe, and the DNA interstrand cross-linking assay. The results of these investigations are consistent with the view that *trans*-[$Cl_2(Me_2SO_4)_4Ru$], like antitumor cisplatin, forms a significant amount of intrastrand cross-links in DNA between neighboring purine residues, but their amount was considerably smaller than that formed by the platinum drug. Interestingly, *trans*-[$Cl_2(Me_2SO_4)_4Ru$] can also form DNA interstrand cross-links, a previously unobserved phenomenon, although their amount is very low (~1%). *cis*-[$Cl_2(Me_2SO_4)_4Ru$] preferentially forms on DNA monofunctional lesions.

One of the important features of the DNA-binding mode of metal-based antitumor drugs relevant to their biological effects is the conformational alteration induced by the formation of the adduct. The assays based on differential pulse polarographic analysis, gel electrophoresis of negatively supercoiled and nicked plasmids, and the formation of liquid crystalline dispersions of DNA revealed that both ruthenium complexes are capable of inducing local conformational alterations of a non-denaturational character. The extent of these alterations induced by the *trans* isomer was noticeably higher than that induced by the *cis* isomer. Thus, the global character of conformational changes induced in DNA by the two ruthenium isomers is similar to that induced by antitumor cisplatin and different from distortions induced in DNA by clinically ineffective transplatin or monofunctional platinum complexes such as [Cl(dien)Pt]Cl or [$Cl(NH_3)_3Pt$]Cl.

The results of these studies also demonstrate that bifunctional DNA adducts of *trans*-[$Cl_2(Me_2SO_4)_4Ru$] inhibit RNA synthesis and specifically block RNA polymerases at the sites of the modified guanine and, to a lesser extent, adenine residues. Importantly, DNA adducts of therapeutically less effective *cis*-[$Cl_2(Me_2SO_4)_4Ru$] are unable to

inhibit RNA synthesis, very likely because this ruthenium compound binds DNA in a monodentate fashion, a binding mode expected to result in antitumor activity only in rare cases.

Taken together, several features of the DNA-binding mode of *trans*-$[Cl_2(Me_2SO_4)_4Ru]$ described in the present work are similar to those of antitumor cisplatin, which may be relevant to the biological effects of this antitumor ruthenium drug. DNA binding of *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ mainly results in the formation of monofunctional adducts, and the extent of conformational distortions induced in DNA is considerably smaller in comparison with distortions induced by its *trans* isomer. This observation is consistent with the less pronounced biological effects of *cis*-$[Cl_2(Me_2SO_4)_4Ru]$ in comparison with its *trans* isomer. Whatever the detailed mechanism, however, the ruthenium compounds are interesting from a mechanistic point of view and, therefore worthy of additional testing. Further studies should reveal the extent to which ruthenium complexes hold promise as clinically useful antitumor compounds.

This work was supported by the Grant Agency of the Czech Republic (Grant nos. 305/99/0695 and 204/97/P028). V.B. was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute. This research was also a part of the European Cooperation in the field of Scientific and Technical Research network (project COST D8/0017/97).

References

1. Keppler BK, Metal complexes as anticancer agents. The future role of inorganic chemistry in cancer therapy. *New J Chem* **14:** 389–403, 1990.
2. Kopf-Maier P, Complexes of metals other than platinum as antitumor agents. *Eur J Clin Pharmacol* **47:** 1–16, 1994.
3. Clarke MJ, Zhu F and Frasca DR, Non-platinum chemotherapeutic metallopharmaceuticals. *Chem Rev* **99:** 2511–2533, 1999.
4. Wong E and Giandomenico CM, Current status of platinum-based antitumor drugs. *Chem Rev* **99:** 2451–2466, 1999.
5. O'Dwyer PJ, Stevenson JP and Johnson SW, Clinical status of cisplatin, carboplatin, and other platinum-based antitumor drugs. In: *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Ed. Lippert B), pp. 31–72. VHCA, WILEY-VCH, Zürich, Weinheim, 1999.
6. Lokich J and Anderson N, Carboplatin versus cisplatin in solid tumors: An analysis of the literature. *Ann Oncol* **9:** 13–21, 1998.
7. Raymond E, Chaney SG, Taamma A and Cvitkovic E, Oxaliplatin: A review of preclinical and clinical studies. *Ann Oncol* **9:** 1053–1071, 1998.
8. Blommaert FA, van Dijk-Knijnenburg HC, Dijt FJ, den Engelse L, Baan RA, Berends F and Fichtinger-Schepman AMJ, Formation of DNA adducts by the anticancer drug carboplatin: Different nucleotide sequence preferences *in vitro* and in cells. *Biochemistry* **34:** 8474–8480, 1995.
9. Woynarowski JM, Chapman WG, Napier C, Herzig MC and Juniewicz P, Sequence- and region-specificity of oxaliplatin adducts in naked and cellular DNA. *Mol Pharmacol* **54:** 770–777, 1998.
10. Sava G, Clerici K, Capozzi I, Cocchietto M, Gagliardi R, Alessio E, Mestroni G and Perbellini A, Reduction of lung metastasis by ImH[*trans*-$RuCl_4$(DMSO)Im]: Mechanism of the selective action investigated on mouse tumors. *Anticancer Drugs* **10:** 129–138, 1999.
11. Johnson NP, Butour JL, Villani G, Wimmer FL, Defais M, Pierson V and Brabec V, Metal antitumor compounds: The mechanism of action of platinum complexes. *Prog Clin Biochem Med* **10:** 1–24, 1989.
12. Sava G, Pacor S, Bregant F, Ceschia V and Mestroni G, Metal complexes of ruthenium: Antineoplastic properties and perspectives. *Anticancer Drugs* **1:** 99–108, 1990.
13. Mestroni G, Alessio E, Sava G, Pacor S, Coluccia M and Boccarelli A, Water-soluble ruthenium(III)–dimethyl sulfoxide complexes: Chemical behaviour and pharmaceutical properties. *Metal-Based Drugs* **1:** 41–63, 1994.
14. Mestroni G, Alessio E, Calligaris M, Attia WM, Quadrifoglio F, Cauci S, Sava G, Zorzet S, Pacor S, Monti-Bragadin C, Tamaro M and Dolzani L, Chemical, biological and antitumor properties of ruthenium(II) complexes with dimethylsulfoxide. *Prog Clin Biochem Med* **10:** 71–87, 1989.
15. Sava G, Pacor S, Zorzet S, Alessio E and Mestroni G, Antitumour properties of dimethylsulfoxide ruthenium(II) complexes in the Lewis lung carcinoma system. *Pharmacol Res* **21:** 617–628, 1989.
16. Monti-Bragadin C, Tamaro M and Banfi E, Mutagenic activity of platinum and ruthenium complexes. *Chem Biol Interact* **11:** 469–472, 1975.
17. Monti-Bragadin C, Ramani L, Samer L, Mestroni G and Zassinovich G, Effect of *cis*-dichlorodiammineplatinum(II) and related transition metal complexes on *Escherichia coli*. *Antimicrob Agents Chemother* **7:** 825–827, 1975.
18. Coluccia M, Sava G, Loseto F, Nassi A, Boccarelli A, Giordano D, Alessio E and Mestroni G, Anti-leukemic action of $RuCl_2(DMSO)_4$ isomers and prevention of brain involvement on P388 leukemia and on P388/DDP subline. *Eur J Cancer* **29A:** 1873–1879, 1993.
19. Sava G, Zorzet S, Giraldi T, Mestroni G and Zassinovich G, Antineoplastic activity and toxicity of an organometallic complex of ruthenium(II) in comparison with *cis*-DDP in mice bearing solid malignant neoplasms. *Eur J Cancer Clin Oncol* **20:** 841–847, 1984.
20. Alessio E, Mestroni G, Nardin G, Attia WM, Calligaris M, Sava G and Zorzet S, *cis*- and *trans*-Dihalotetrakis (dimethyl sulfoxide) ruthenium(II) complexes ($RuX_2(DMSO)_4$; X = Cl, Br): Synthesis, structure, and antitumor activity. *Inorg Chem* **27:** 4099–4106, 1988.
21. Loseto F, Alessio E, Mestroni G, Lacidogna G, Nassi A, Giordano D and Coluccia M, Interaction of $RuCl_2$(dimethylsulphoxide)$_4$ isomers with DNA. *Anticancer Res* **11:** 1549–1553, 1991.
22. Jamieson ER and Lippard SJ, Structure, recognition, and processing of cisplatin–DNA adducts. *Chem Rev* **99:** 2467–2498, 1999.
23. Cauci S, Alessio E, Mestroni G and Quadrifoglio F, Reaction of *cis*-RuII$(DMSO)_4Cl_2$ with DNA and with some of its bases in aqueous solution. *Inorg Chim Acta - Bioinorg Chem* **137:** 19–24, 1987.
24. Alessio E, Xu Y, Cauci S, Mestroni G, Quadrifoglio F, Viglino P and Marzilli LG, Novel diastereomers with opposite chirality at ruthenium formed by N7, a-PO4 chelation of 5′-dGMP to the antimetastatic agent *trans*-$RuCl_2(DMSO)_4$: NMR and CD evidence. *J Am Chem Soc* **111:** 7068–7071, 1989.
25. Cauci S, Viglino P, Esposito G and Quadrifoglio F, Reaction of the octahedral antitumor complex *trans*-$RuCl_2(DMSO)_4$ with 2′deoxyguanosine. *J Inorg Biochem* **43:** 739–751, 1991.
26. Esposito G, Cauci S, Fogolari F, Alessio E, Scocchi M, Quadrifoglio F and Viglino P, NMR structural characterization of the reaction product between d(GpG) and the

octahedral antitumor complex *trans*-$RuCl_2(DMSO)_4$. *Biochemistry* **31:** 7094–7103, 1992.
27. Anagnostopoulou A, Moldrheim E, Katsaros N and Sletten E, Interaction of *cis*- and *trans*-$RuCl_2(DMSO)_4$ with the nucleotides GpA, d(GpA), ApG, d(ApG) and d(CCTGGTCC): High-field NMR characterization of the reaction products. *J Biol Inorg Chem* **4:** 199–208, 1999.
28. Brabec V and Paleček E, The influence of salts and pH on polarographic currents produced by denatured DNA. *Biophysik* **6:** 290–300, 1970.
29. Brabec V and Paleček E, Interaction of nucleic acids with electrically charged surfaces. II. Conformational changes in double-helical polynucleotides. *Biophys Chem* **4:** 76–92, 1976.
30. Lemaire MA, Schwartz A, Rahmouni AR and Leng M, Interstrand cross-links are preferentially formed at the d(GC) sites in the reaction between *cis*-diamminedichloroplatinum(II) and DNA. *Proc Natl Acad Sci USA* **88:** 1982–1985, 1991.
31. Brabec V and Leng M, DNA interstrand cross-links of *trans*-diamminedichloroplatinum(II) are preferentially formed between guanine and complementary cytosine residues. *Proc Natl Acad Sci USA* **90:** 5345–5349, 1993.
32. Farrell N, Qu Y, Feng L and Van Houten B, Comparison of chemical reactivity, cytotoxicity, interstrand cross-linking and DNA sequence specificity of bis(platinum) complexes containing monodentate or bidentate coordination spheres with their monomeric analogues. *Biochemistry* **29:** 9522–9531, 1990.
33. Butour JL and Macquet JP, Differentiation of DNA–platinum complexes by fluorescence. The use of an intercalating dye as a probe. *Eur J Biochem* **78:** 455–463, 1977.
34. Butour JL, Alvinerie P, Souchard JP, Colson P, Houssier C and Johnson NP, Effect of the amine nonleaving group on the structure and stability of DNA complexes with *cis*-[Pt(R-$NH_2)_2(NO_3)_2$]. *Eur J Biochem* **202:** 975–980, 1991.
35. Brabec V, Reedijk J and Leng M, Sequence-dependent distortions induced in DNA by monofunctional platinum(II) binding. *Biochemistry* **31:** 12397–12402, 1992.
36. Keck MV and Lippard SJ, Unwinding of supercoiled DNA by platinum ethidium and related complexes. *J Am Chem Soc* **114:** 3386–3390, 1992.
37. Akimenko N, Cheltsov P, Balcarová Z, Kleinwächter V and Yevdokimov Y, A study of interactions of platinum(II) compounds with DNA by means of CD spectra of solutions and liquid crystalline microphases of DNA. *Gen Physiol Biophys* **4:** 597–608, 1985.
38. Yevdokimov YM, Skuridin SG and Salyanov VI, The liquid crystalline phases of double-stranded nucleic acids *in vitro* and *in vivo*. *Liquid Cryst* **3:** 1443–1459, 1988.
39. Bancroft DP, Lepre CA and Lippard SJ, Pt-195 NMR kinetic and mechanistic studies of *cis*-diamminedichloroplatinum and *trans*-diamminedichloroplatinum(II) binding to DNA. *J Am Chem Soc* **112:** 6860–6871, 1990.
40. Prenzler PD and McFadyen WD, Reactions of cisplatin and the *cis*-diamminediaqua platinum(II) cation with Tris and Hepes. *J Inorg Biochem* **68:** 279–282, 1997.
41. Žaludová R, Žákovská A, Kašpárková J, Balcarová Z, Kleinwächter V, Vrána O, Farrell N and Brabec V, DNA interactions of bifunctional dinuclear platinum(II) antitumor agents. *Eur J Biochem* **246:** 508–517, 1997.
42. Kašpárková J, Nováková O, Vrána O, Farrell N and Brabec V, Effect of geometric isomerism in dinuclear platinum antitumor complexes on DNA interstrand cross-linking. *Biochemistry* **38:** 10997–11005, 1999.
43. Žaludová R, Žakovská A, Kašpárková J, Balcarová Z, Vrána O, Coluccia M, Natile G and Brabec V, DNA modifications by antitumor *trans*-[$PtCl_2$(E-iminoether)$_2$]. *Mol Pharmacol* **52:** 354–361, 1997.
44. Žákovská A, Nováková O, Balcarová Z, Bierbach U, Farrell N and Brabec V, DNA interactions of antitumor *trans*-[$PtCl_2(NH_3)$(quinoline)]. *Eur J Biochem* **254:** 547–557, 1998.
45. Brabec V, Kasparkova J, Vrana O, Novakova O, Cox JW, Qu Y and Farrell N, DNA modifications by a novel bifunctional trinuclear platinum Phase I anticancer agent. *Biochemistry* **38:** 6781–6790, 1999.
46. Vrána O, Boudný V and Brabec V, Superhelical torsion controls DNA interstrand cross-linking by antitumor *cis*-diamminedichloroplatinum(II). *Nucleic Acids Res* **24:** 3918–3925, 1996.
47. Paleček E, Modern polarographic (voltammetric) techniques in biochemistry and molecular biology. Part II. Analysis of macromolecules. In: *Topics in Bioelectrochemistry and Bioenergetics* (Ed. Milazzo G), Vol. 5, pp. 65–155. J Wiley, New York, 1983.
48. Brabec V, Vetterl V and Vrána O, Electroanalysis of biomacromolecules. In: *Bioelectrochemistry: Principles and Practice* (Eds. Brabec V and Milazzo G), Vol. 3, pp. 287–359. Birkhäuser Verlag, Basel, 1996.
49. Vrána O and Brabec V, Electrochemical analysis of antitumour platinum drugs and their complexes with DNA. *Bioelectrochem Bioenerg* **19:** 145–160, 1988.
50. Vrána O, Brabec V and Kleinwächter V, Polarographic studies on the conformation of some platinum complexes: Relations to anti-tumour activity. *Anticancer Drug Des* **1:** 95–109, 1986.
51. Brabec V, Kleinwächter V, Butour JL and Johnson NP, Biophysical studies of the modification of DNA by antitumour platinum coordination complexes. *Biophys Chem* **35:** 129–141, 1990.
52. Cohen GL, Bauer WR, Barton JK and Lippard SJ, Binding of *cis*- and *trans*-dichlorodiammineplatinum(II) to DNA: Evidence for unwinding and shortening of the double helix. *Science* **203:** 1014–1016, 1979.
53. Scovell WM and Collart F, Unwinding of supercoiled DNA by *cis*- and *trans*-diamminedichloroplatinum(II): Influence of the torsional strain on DNA unwinding. *Nucleic Acids Res* **13:** 2881–2895, 1985.
54. Livolant F, Levelut AM and Benoit JP, The highly concentrated liquid-crystalline phase of DNA is columnar hexagonal. *Nature* **339:** 724–726, 1989.
55. Akimenko N, Kleinwächter V and Yevdokimov Y, Liquid crystalline microphases of DNA molecules complexed with compounds of platinum(II). *FEBS Lett* **156:** 58–62, 1983.
56. Belyakov VA, Orlov VP, Semenov SV, Skuridin SG and Yevdokimov YM, Comparison of calculated and observed CD spectra of liquid crystalline dispersions formed from double-stranded DNA and from DNA complexed with coloured compounds. *Liquid Cryst* **20:** 777–784, 1996.
57. Livolant F and Maestre M, Circular dichroism microscopy of compact forms of DNA and chromatin *in vivo* and *in vitro*: Cholesteric liquid-crystalline phases of DNA and single dinoflagellate nuclei. *Biochemistry* **27:** 3056–3068, 1988.
58. Yevdokimov YM, Skuridin SG, Salyanov VI, Damaschun G, Damaschun H, Misselwitz R and Kleinwächter V, Effect of platinum (II) chemotherapeutic agents on properties of DNA liquid crystals. *Biophys Chem* **35:** 143–153, 1990.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 13, Issue of March 30, pp. 9655–9661, 2001

Printed in U.S.A.

Thermal and Thermodynamic Properties of Duplex DNA Containing Site-specific Interstrand Cross-link of Antitumor Cisplatin or Its Clinically Ineffective Trans Isomer*

Received for publication, November 9, 2000
Published, JBC Papers in Press, December 4, 2000, DOI 10.1074/jbc.M010205200

Ctirad Hofr‡ and Viktor Brabec§

From the Institute of Biophysics, Academy of Sciences of the Czech Republic, CZ-61265 Brno, Czech Republic

The effect of the single, site-specific interstrand cross-link formed by cisplatin or transplatin on the thermal stability and energetics of a 20-base pair DNA duplex is reported. The cross-linked or unplatinated 20-base pair duplexes were investigated with the aid of differential scanning calorimetry, temperature-dependent UV absorption, and circular dichroism. The cross-link of both platinum isomers increases the thermal stability of the modified duplexes by changing the molecularity of denaturation. The structural perturbation resulting from the interstrand cross-link of cisplatin increases entropy of the duplex and in this way entropically stabilizes the duplex. This entropic cross-link-induced stabilization of the duplex is partially but not completely compensated by the enthalpic destabilization of the duplex. The net result of these enthalpic and entropic effects is that the structural perturbation resulting from the formation of the interstrand cross-link by cisplatin induces a decrease in duplex thermodynamic stability, with this destabilization being enthalpic in origin. By contrast, the interstrand cross-link of transplatin is enthalpically almost neutral with the cross-link-induced destabilization entirely entropic in origin. These differences are consistent with distinct conformational distortions induced by the interstrand cross-links of the two isomers. Importantly, for the duplex cross-linked by cisplatin relative to that cross-linked by transplatin, the compensating enthalpic and entropic effects almost completely offset the difference in cross-link-induced energetic destabilization. It has been proposed that the results of the present work further support the view that the impact of the interstrand cross-links of cisplatin and transplatin on DNA is different for each and might also be associated with the distinctly different antitumor effects of these platinum compounds.

The thermal and thermodynamic stability of DNA play an important role in many biological processes. In addition, agents of biological significance that modify DNA may also affect its thermal and thermodynamic stability, which may be associated with the mechanism underlying biological activity of such agents. Thus, the studies of thermal and thermodynamic stability of DNA modified by various agents are of great interest.

It is well established that platinum coordination complexes exhibit antitumor effects (1–4). The success of platinum complexes in killing tumor cells results from their ability to form on DNA various types of covalent adducts that are capable of terminating DNA synthesis (5, 6) and the cellular processes triggered by the presence of those adducts on DNA (7). The first platinum complex introduced in the clinic is *cis*-diamminedichloroplatinum(II) (cisplatin)[1] (1). Although the antitumor effects of cisplatin were discovered more than 30 years ago, the mechanism of its antitumor activity has not yet been fully understood. It has been shown (8, 9) that this bifunctional platinum complex mainly forms intrastrand cross-links on DNA between neighboring purine residues (~90%). Other minor adducts are intrastrand cross-links between two purine nucleotides separated by one or more nucleotides; few adducts remain monofunctional. Importantly, cisplatin also forms interstrand cross-links (~6% in cell-free media in linearized plasmid DNA (10, 11)). Transplatin (the *trans* isomer of cisplatin) is clinically ineffective, so that both isomers have been used frequently in studies of the structure-pharmacological activity relationship of platinum complexes. Transplatin-DNA adducts are also intrastrand cross-links but between nonadjacent nucleotides (12). Transplatin also forms in DNA interstrand cross-links (~12%) (10), and a relatively large portion of the adducts remains monofunctional even after long periods of DNA modification (13).

It has been postulated (6, 14) that the antitumor properties of cisplatin are mediated by damaged DNA-binding proteins (for instance those containing a HMG (high mobility group) domain). It has been also shown (15) that the recognition of adducts formed on DNA by platinum complexes is dependent on the extent of thermal or thermodynamic destabilization imposed on the duplex by the adduct. The increase of the thermodynamic destabilization results in the reduced recognition and binding of HMG domain proteins to platinated DNA. In addition, the thermal stability of DNA modified by various platinum compounds, which differ in their antitumor effects, has been also studied. These studies have revealed (16–20) that the important factors influencing the thermal stability of platinated DNA are also interstrand cross-links, which contribute to the global stabilization of DNA.

Despite great effort devoted to the understanding of how

* This research was supported by the Grant Agency of the Czech Republic (Grants 305/99/0695 and 301/00/0556), the Grant Agency of the Academy of Sciences of the Czech Republic (Grant A5004702), and the Internal Grant Agency of the Ministry of Health of the Czech Republic (Grants NL6058-3/2000 and NL6069-3/2000). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by a doctoral fellowship from the Faculty of Sciences, Masaryk University, Brno, Czech Republic.

§ To whom correspondence should be addressed: Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-61265 Brno, Czech Republic. Tel.: 420-5-41517148; Fax: 420-5-41240499; E-mail: brabec@ibp.cz.

[1] The abbreviations used are: cisplatin, *cis*-diamminedichloroplatinum(II); transplatin, the *trans* isomer of cisplatin; DSC, differential scanning calorimetry; bp, base pair(s); FPLC, fast protein liquid chromatography; FAAS, flameless atomic absorption spectrophotometry.

cisplatin modifies DNA and how these modifications are associated with the antitumor effects of cisplatin, the relative efficacy of its intrastrand and interstrand cross-links is unknown. Whereas the thermal and thermodynamic properties of DNA duplexes containing intrastrand cross-link of cisplatin or its analogues have already been studied in detail (15, 21, 22), no attention has been paid to thermal stability and energetics of DNA interstrand cross-links of platinum drugs. It is so despite the fact that DNA interstrand cross-links of platinum complexes could play a very significant role in the biological activity of these compounds because these covalent cross-links preventing separation of the two strands of DNA could block DNA replication markedly more efficiently than intrastrand adducts (23). In addition, it has been also suggested that nucleotide excision repair may reduce antitumor effects of platinum complexes. Thus, the fact that nucleotide excision repair of interstrand cross-links in general is much more difficult than that of intrastrand cross-links (24–26) may also serve to emphasize the importance of interstrand cross-links of platinum complexes for their antitumor effects even when these cross-links are only minor lesions. Here we examine the effect of the single, site-specific interstrand cross-link formed by cisplatin or transplatin on the thermal stability and energetics of a 20-base pair (bp) DNA duplex. The differential scanning calorimetric (DSC), temperature-dependent UV absorption, and circular dichroism (CD) properties of the platinated or unplatinated 20-bp duplex were investigated.

EXPERIMENTAL PROCEDURES

Chemicals—Cisplatin and transplatin were from Sigma. The synthetic oligodeoxyribonucleotides d(TGCT) and d(AGCA) (Fig. 1) were purchased from IDT, Inc. (Coralville, IA) and purified as described previously (10, 27, 28). Molar extinction coefficients for the single-stranded oligonucleotides were determined by phosphate analysis (21). The following extinction coefficients at 260 nm and 25 °C were obtained: 148,000 for unmodified d(TGCT) and 189,000 for unmodified d(AGCA). Isothermal mixing experiments (21) using unmodified d(TGCT) and d(AGCA) strands revealed 1:1 stoichiometries for both complexes, a ratio consistent with duplex formation. Dimethyl sulfate was from Sigma.

Platinations of Oligonucleotides—The single-stranded oligonucleotide d(TGCT) (the *top strand* in Fig. 1) at a concentration of 125 μM was reacted with a monoaquamonochloro derivative of cisplatin or transplatin generated by allowing these complexes to react with 0.9 molar equivalent of $AgNO_3$ at an input platinum to strand molar ratio of 3:1 or 3.9: 1, respectively, in 10 mM $NaClO_4$ (pH 5.2) at 37 °C. The mixture with cisplatin was incubated for 13 min and the mixture with transplatin for 15 min. Then, the NaCl concentration was adjusted to 0.1 M, and the platinated oligonucleotides were again purified by fast protein liquid chromatography (FPLC). Using platinum flameless atomic absorption spectrophotometry (FAAS) and measurements of the optical density, it was verified that the modified oligonucleotide contained one platinum atom. It was also verified using dimethyl sulfate footprinting of platinum on DNA (10) that in the platinated top strands the N7 position of the central G was not accessible for reaction with dimethyl sulfate, which implies that this G residue was platinated. The platinated strands were allowed to anneal with unplatinated complementary strand d(AGCA) in 0.4 M NaCl (pH 7.4) at 25 °C for 24 h, precipitated by ethanol, dissolved in 0.1 M $NaClO_4$ and incubated for 48 h in the dark at 37 °C. The resulting products were still purified by FPLC in an alkaline gradient. Using this denaturing gradient, non-interstrand cross-linked strands were eluted as 20-nucleotide single strands, whereas the interstrand cross-linked strands were eluted later in a single peak as a higher molecular mass species. This single peak was only collected so that the samples of the interstrand cross-linked duplexes contained no single-stranded molecules. Alternatively, the duplexes containing the interstrand cross-links were separated on a 12% polyacrylamide, 8 M urea denaturing gel, and the single bands corresponding to interstrand cross-linked duplexes were cut off from the gel, eluted, precipitated by ethanol, and dissolved in a solution consisting of 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA. Both procedures of the purification of interstrand cross-linked duplexes provided products of which subsequent analysis (see below) gave identical results. The yields of these interstrand cross-linking reactions were ~15 and 30% for cisplatin and transplatin, respectively. The duplexes were still further analyzed for platinum content by FAAS. Additional quantitation of cross-linked duplex by UV absorption spectrophotometry was used to ascertain that 1:1 adducts (one Pt/duplex) had formed. The sites involved in interstrand cross-links were deduced in the same way as described earlier (10, 27–30), *i.e.* mainly from Maxam-Gilbert footprinting experiments. It was verified in this way that the interstrand cross-link of cisplatin was formed between guanine residues in neighboring base pairs in the 5′-GC·5′-GC central sequence, whereas the interstrand cross-link of transplatin was formed between central guanine residue in the top strand of the d(TGCT)·d(AGCA) duplex and its complementary cytosine residue. FPLC purification and FAAS measurements were carried out on an Amersham Biotech FPLC system with MonoQ HR 5/5 column and a Unicam 939 AA spectrometer equipped with a graphite furnace, respectively. The concentration of the purified and characterized duplexes containing the interstrand cross-link of cisplatin and transplatin was further estimated by determining the platinum concentration by means of FAAS. Other details can be found in previously published papers (10, 28, 29).

Differential Scanning Calorimetry—Excess heat capacity (ΔC_p) *versus* temperature profiles for the thermally induced transitions of d(TGCT)·d(AGCA) duplex, unmodified or containing a unique interstrand cross-link of cisplatin or transplatin, were measured using a VP-DSC calorimeter (Microcal, Northampton, MA). In these experiments, the heating rate was 60 °C/h and a maximum temperature was 95 °C. Enthalpies (ΔH_{cal}) and entropies (ΔS_{cal}) of duplex formation were calculated from the areas under the experimental ΔC_p *versus* T and the derived $\Delta C_p/T$ *versus* T curves, respectively, using ORIGIN version 5.0 software (Microcal). The free energy of duplex formation at 25 °C (ΔG_{25}) was calculated using the standard thermodynamic relationship given in Equation 1 and the corresponding values of ΔH and ΔS.

$$\Delta G_{25} = \Delta H - (298.15)\Delta S \qquad \text{(Eq. 1)}$$

The oligonucleotide duplexes at the concentration of 10 μM were dialyzed against the buffer containing 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA. The samples were vacuum-degassed before the measurement. It was also verified in the same way as described in the previous paper (21) that the melting transition of both the platinated and unmodified duplexes were fully reversible.

UV Absorption Spectrophotometry—UV absorbance measurements were conducted on a Beckman DU-7400 spectrophotometer equipped with a thermoelectrically controlled cell holder and quartz cells with a path length of 1 cm. Absorbance *versus* temperature profiles were measured at 260 nm. The temperature was raised using linear heating rate of 1.0 °C/min. For each optically detected transition, the melting temperature (T_m) was determined as described previously (16). The DNA solutions ranged from 0.2 to 10 μM in duplex and contained 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA.

Circular Dichroism Spectrophotometry—CD spectra were recorded using a Jasco J-720 spectropolarimeter equipped with a thermoelectrically controlled cell holder. The cell path length was 1 cm. Isothermal CD spectra were recorded from 220 to 320 nm in 1-nm increments with an averaging time of 5 s. The DNA concentration was 6 μM in duplex, and buffer conditions were 10 mM sodium cacoldylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA.

RESULTS

Differential scanning calorimetry measurements were conducted to characterize the thermally induced denaturation of the 20-bp duplex with the specific goal of elucidating the thermal and thermodynamic consequences of modifying and constraining DNA via a single, site-specific interstrand cross-link of antitumor cisplatin or its clinically ineffective *trans* isomer (Fig. 1). The cross-links were formed in the center of these duplexes between the nucleotide residues preferentially involved in these adducts when high molecular mass DNA is globally modified by cisplatin or transplatin, *i.e.* between guanine residues in neighboring base pairs in the sequence 5′-GC·5′-GC in the case of cisplatin cross-link (30) or between guanine and complementary cytosine of the same duplex in the case of the cross-link formed by transplatin (10). The results of

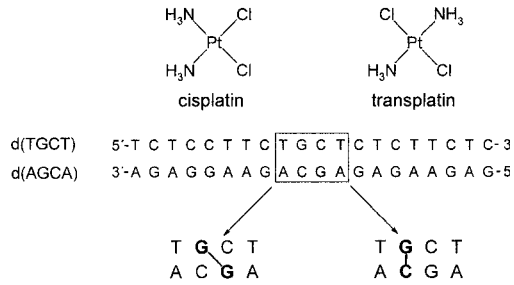


FIG. 1. **Structures of cisplatin and transplatin along with the base sequence of the synthetic oligodeoxyribonucleotide duplex used in the present study with their abbreviations.** The top and bottom strands in the pair of oligonucleotides are designated *top* and *bottom*, respectively, in the text. The central part of the duplex containing the base pairs interstrand cross-linked by cisplatin or transplatin is framed and also shown separately with the manifestation of the cross-linking. The *bold letters* in this central part indicate the location of the interstrand cross-link after modification of the oligonucleotide by cisplatin or transplatin as described under "Experimental Procedures."

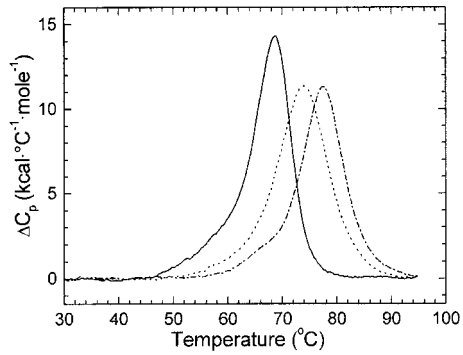


FIG. 2. **DSC thermograms for the d(TGCT)·d(AGCA) unplatinated duplex (*solid line*) containing a single, site-specific interstrand cross-link of cisplatin (*dot-dash line*) or transplatin (*dashed line*).** The duplex concentration was 10.0 μM, and the buffer conditions were 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA.

these studies are shown in Fig. 2 with the associated data listed in Table I. Denaturation (heating) and renaturation (cooling) curves for the unmodified and the platinated duplex were superimposable, which is consistent with the reversibility of this melting equilibrium. Thus, meaningful thermodynamic data from our calorimetric and spectrophotometric measurements described below could be obtained. In addition, the pre- and post-base lines coincide for both unmodified and platinated duplexes, which suggests no differential heat capacity change resulting from the presence of the cross-link. Comparing the calorimetrically determined melting temperatures (T_m) for the cross-linked duplexes and for the unconstrained (unplatinated) duplex reveals that formation of either cross-link results in a substantial increase in the thermal stability of the duplex (ΔT_m (defined as the difference between T_m values of the cross-linked and unplatined duplex) was +8.7 or +5.0 °C for the cross-link of cisplatin or transplatin, respectively).

The unconstrained (unplatinated) duplex denatures in a bimolecular reaction to form two single strands. As a consequence, melting of the unplatinated duplex was dependent on the overall oligonucleotide concentration. For instance, increasing the duplex concentration from 0.2 to 10.0 μM increased the T_m of the unplatinated duplex from 61.5 to 68.8 °C. In contrast, the duplexes containing the interstrand cross-link of cisplatin or transplatin melted in a concentration-independent manner to a single-stranded state, which is consistent with the expectation that the molecularity had been reduced from bimolecular to monomolecular. Thus, the observed ΔT_m differences could also result from the change in molecularity. To support this assumption, we performed a "correction" for the concentration dependence of the T_m in the bimolecular unplatinated duplex following the general procedure outlined by Marky and Breslauer (31). Using this approach and our calorimetrically determined enthalpy (Table I), we estimated a reduced concentration-independent T_m of our unplatinated duplex as 88.1 °C. This "reduced" T_m value of the unplatinated duplex is significantly different from the T_m values of the cross-linked duplexes. Hence, the overall impact of the single interstrand cross-link of cisplatin or transplatin should not be associated only with the change in molecularity of the duplex system but another mechanism affecting the thermal stability of the duplex also has to be involved (32).

The transition entropy for a bimolecular complex depends on strand concentration. To eliminate the effect of different molecularities of the unplatinated and interstrand cross-linked oligomer systems, we also performed a correction for this concentration dependence again using the general procedure outlined by Marky and Breslauer (31) to calculate a reduced entropy (ΔS^*) (Table I) from the observed ΔS_{cal} values. A comparison of the calorimetrically obtained enthalpy (ΔH_{obs}) and reduced entropy (ΔS^*) values for unplatinated duplex with those measured for the duplex containing the interstrand cross-link of cisplatin revealed a significant decrease of the change of enthalpy of duplex formation by 18 kcal/mol and an increase in the reduced duplex entropy ($\Delta\Delta S^* = 37$ cal/K · mol so that $T\Delta\Delta S^* = 11$ kcal/mol at 25 °C). In other words, the structural perturbation resulting from the interstrand cross-link of cisplatin increases entropy of the duplex and in this way entropically stabilizes the duplex. Thus, the 18 kcal/mol enthalpic destabilization of the duplex resulting from the cross-link of cisplatin is partially, but not completely, compensated by the entropic cross-link induced stabilization of the duplex of 11 kcal/mol at 25 °C. The net result of these enthalpic and entropic effects is that the structural perturbation resulting from formation of the interstrand cross-link by cisplatin induces a decrease in duplex thermodynamic stability ($\Delta\Delta G_{25}^*$) of +7.0 kcal/mol with this destabilization being enthalpic in origin.

By contrast, the cross-link formed by transplatin does not considerably alter the enthalpic stability of the duplex, whereas entropically destabilizing the host duplex by 15 cal/K · mol ($\Delta\Delta S^* = -15$ cal/K · mol so that $T\Delta\Delta S^* = -4.5$ kcal/mol at 25 °C). In other words, the interstrand cross-link of transplatin is enthalpically neutral with the cross-link-induced destabilization entirely entropic in origin. These results illustrate that the isomerization of diamminedichloroplatinum(II) can modulate the magnitude of the destabilization induced by the interstrand cross-link as well as the relative enthalpic and entropic contributions to this destabilization. The net result of the enthalpic and entropic effects noted above is that formation of the interstrand cross-link by transplatin at 25 °C reduces the thermodynamic stability (ΔG_{25}^*) of the duplex by 6.5 kcal/mol. Thus, formation of the interstrand cross-links of cisplatin and transplatin has an almost identical impact on the reduction in thermodynamic stability of the duplex (ΔG_{25}^*), although the origins of these destabilizations (ΔH, ΔS^*) are different.

From a comparison of the model-dependent van't Hoff and the model-independent calorimetric transition enthalpies, it is possible in principle to conclude whether a transition occurs in a two-state (all-or-none) manner with no significant thermodynamic contribution from intermediate states (33, 34). Table II lists the directly measured, model-independent calorimetric transition enthalpies (ΔH_{cal}) and the indirectly derived, model-dependent van't Hoff transition enthalpies (ΔH_{vH}). The ΔH_{vH} values were obtained by analyzing the shapes of each calorimetric curve using the approach described earlier (31). A com-

TABLE I

Calorimetrically derived thermodynamic parameters for the formation of the 20-mer duplex d(TGCT)·d(AGCA), unplatinated or containing a single, site-specific interstrand cross-link of cisplatin or transplatin

The T_m, ΔH_{cal} and ΔS_{cal} values are averages derived from three independent experiments.

	Unplatinated duplex	Duplex cross-linked by cisplatin[a]	Duplex cross-linked by transplatin[b]
T_m (°C)[c]	68.8	77.5	73.8
ΔT_m (°C)[d]		8.7	5.0
ΔH_{cal} (kcal/mol duplex)[e]	−148	−130	−146
$\Delta\Delta H_{cal}$ (kcal/mol duplex)[f]		18	2
ΔS_{cal} (cal/K·mol duplex)[e]	−434	−370	−422
$\Delta\Delta S_{cal}$ (cal/K·mol duplex)[f]		64	12
ΔS^* (cal/K·mol duplex)[g]	−407		
$\Delta\Delta S^*$ (cal/K·mol duplex)[f]		37	−15
ΔG_{25} (kcal/mol duplex)[e]	−18.6	−19.7	−20.2
$\Delta\Delta G_{25}$ (kcal/mol duplex)[f]		−1.1	−1.6
$\Delta G_{25}{}^*$ (kcal/mol duplex)[h]	−26.7		
$\Delta\Delta G_{25}{}^*$ (kcal/mol duplex)[f]		7.0	6.5

[a] The interstrand cross-link of cisplatin was formed between guanine residues contained in the adjacent base pairs in the central sequence 5′-GC·5′-GC.

[b] The interstrand cross-link of transplatin was formed between central guanine residue in the top strand and its complementary cytosine residue.

[c] T_m denotes the temperature corresponding to the maximum in the DSC melting profile. The indicated T_m values were derived from the DSC profiles shown in Fig. 2; the error is ±0.3 °C.

[d] $\Delta T_m = T_{m\ \text{platinated}} - T_{m\ \text{unplatinated}}$.

[e] ΔH_{cal}, ΔS_{cal}, and ΔG_{25} denote, respectively, the observed enthalpy, entropy, and free energy (at 25 °C) of duplex formation; the error is ±3%.

[f] $\Delta\Delta(H_{cal}, S_{cal}, G_{25}) = \Delta(H_{cal}, S_{cal}, G_{25},)_{platinated} - \Delta(H_{cal}, S_{cal}, G_{25})_{unplatinated}$; $\Delta\Delta(S^*, G_{25}{}^*) = \Delta(S^*, G_{25}{}^*)_{platinated} - \Delta(S, G_{25})_{nonmodified}$.

[g] ΔS^* denotes reduced entropy calculated from the observed ΔS_{cal} values (for details, see the text).

[h] $\Delta G_{25}{}^*$ denotes reduced free energy (at 25 °C) of duplex formation calculated from the observed ΔH_{cal} values and the reduced entropy ΔS^* values (for details, see the text).

TABLE II

Calorimetric (model-independent) and van't Hoff enthalpies for formation of the 20-mer duplexes d(TGGT)·d(ACCA), unplatinated or containing a single, site specific interstrand cross-link of cisplatin or transplatin

	ΔH_{cal}[a]	ΔH_{vH}[b]	$\Delta H_{vH}/\Delta H_{cal}$
	kcal/mol duplex	*kcal/mol duplex*	
Unplatinated duplex	−148	−149	1.01
Duplex cross-linked by cisplatin	−130	−96	0.74
Duplex cross-linked by transplatin	−146	−78	0.53

[a] Model-independent calorimetric enthalpies (ΔH_{cal}) were derived from DSC experiments as described in the text; the error is ±3%.

[b] van't Hoff enthalpies (ΔH_{vH}) values were obtained by analyzing the shapes of each calorimetric curve using the approach described in Ref. 31; the error is ±3%.

parison of the ΔH_{vH} and ΔH_{cal} data listed in Table II reveals that for the unplatinated duplex the van't Hoff value is identical within the experimental uncertainty to the corresponding model-independent calorimetric value. This result is consistent with an all-or-none, two-state melting behavior of the unplatinated duplex. On the other hand, van't Hoff values were considerably smaller than the corresponding model-independent calorimetric data for the interstrand cross-linked duplexes (Table II). This disparity demonstrates that both interstrand cross-links alter the ability of the duplex to propagate those interactions required for cooperative melting and that the efficiency of the cross-links of cisplatin and transplatin to alter this ability is different.

The interpretation of our data described below is also based on the assumption that all thermodynamic parameters for formation of the unmodified and platinated duplexes are ascribed to differences in the initial duplex states. This implies that the final single-stranded states should be thermodynamically equivalent at the elevated temperatures at which they are formed. The CD signals of the high-temperature denatured state (recorded at 95 °C) of the unplatined duplex and those cross-linked by either cisplatin or transplatin agree within the noise of the measurement, whereas significant differences exist in the intensity of the CD signals of the native states (Fig. 3). These results are consistent with the local perturbations in the

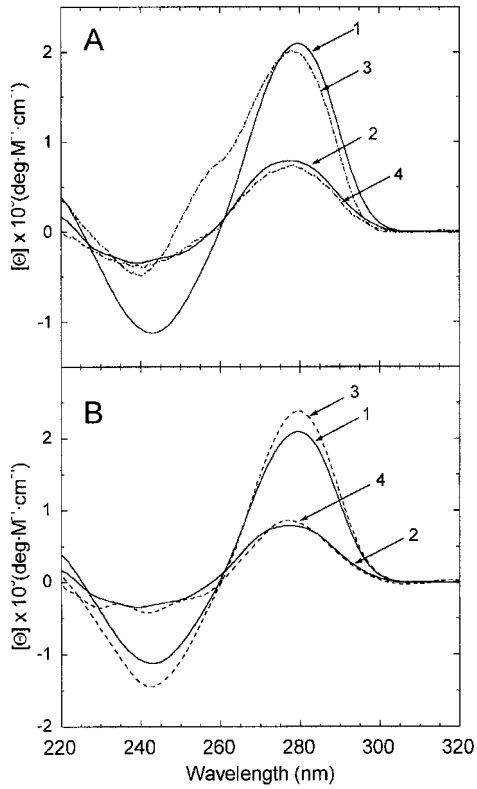


FIG. 3. **CD spectra for the d(TGCT)·d(AGCA) duplex unplatinated or containing a single, site-specific interstrand cross-link of cisplatin (*A*) or transplatin (*B*) recorded at 25 or 95 °C.** The duplex concentration was 6 μM, and the buffer conditions were 10 mM sodium cacodylate (pH 7.2), 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA. *Curves: 1* and *2*, unplatinated duplex at 25 and 95 °C, respectively; *3* and *4*, the cross-linked duplex at 25 and 95 °C, respectively.

native duplex state in agreement with the structural studies performed with the duplexes containing an interstrand cross-link of cisplatin or transplatin (23, 28, 35–37). The similarity of the CD signals of the denatured states of unplatinated and cross-linked duplexes may reflect a similar degree of base unstacking in these duplexes at the elevated temperatures, al-

though such a conclusion may well exceed the information content of the CD measurement.

DISCUSSION

The CD spectrum of the duplex containing single, site-specific interstrand cross-link of cisplatin recorded at 25 °C (Fig. 3*A*) confirms that this lesion considerably alters the global geometry of the parent duplex. It has been shown (23, 35, 36) that cisplatin interstrand cross-link, which is preferentially formed between opposite guanines in the 5′-GC·5′-GC sequence (30), induces several irregularities in the cross-linked base pairs and their immediate adjacent pairs in a base sequence-independent manner (38). The cross-linked deoxyriboguanosine residues are not paired with hydrogen bonds to the complementary deoxyribocytidines, which are located outside the duplex and not stacked with other aromatic rings. All other base residues are paired, but distortion extends over at least four base pairs at the site of the cross-link (38). In addition, the *cis*-diammineplatinum(II) bridge resides in the minor groove (23, 35, 36) and the double helix is locally reversed to a left-handed, Z-DNA-like form. The change of the helix sense and the extrusion of deoxyribocytidine residues (complementary to the platinated deoxyriboguanosine residues) from the duplex results in the helix unwinding by ~80° relative to B-DNA (35), which is very likely responsible for a marked reduction of the amplitude of the negative CD band at around 240 nm observed at 25 °C (Fig. 3*A*). The interstrand cross-link of cisplatin also induces the bending of ~40° of the helix axis at the cross-linked site toward the minor groove (23, 35, 36).

The CD spectrum at 25 °C of the duplex d(TGCT)·d(AGCA) is affected by the site-specific interstrand cross-link of transplatin (formed preferentially between complementary guanine and cytosine (10)) much less than with cisplatin (Fig. 3*B*). Consistent with this CD behavior is the observation that the conformational alterations induced by the interstrand cross-link of transplatin (28, 37) are much less severe than those induced by the cross-link of cisplatin. The platinated deoxyriboguanosine residue in the cross-link of transplatin adopts a *syn* conformation. In addition, the duplex is slightly distorted on both sides of the cross-link, but all bases are still paired and hydrogen-bonded. The cross-link of transplatin unwinds the double helix by ~12° and induces a slight, flexible bending of ~20° of its axis toward minor groove (28).

The distinctly different structural features of the interstrand cross-links of cisplatin and transplatin are also reflected by their different thermal and thermodynamic properties (Table I). Importantly, the increase of the thermal stability of the d(TGCT)·d(AGCA) duplex resulting from the interstrand cross-link of cisplatin or transplatin is because of the change in the molecularity of the oligomer system. If the change observed in T_m is due entirely to the molecularity of the system, then one might expect to observe changes only in entropy. This is observed in the case of the formation of the interstrand cross-link of transplatin. On the other hand, significant changes in enthalpy are observed in addition to entropy changes as a consequence of the formation of the interstrand cross-link of cisplatin. Thus, whereas the structural perturbation resulting from the formation of the latter cross-link induces a decrease in duplex thermodynamic stability ($\Delta\Delta G_{25}$*) of +7.0 kcal/mol, with this destabilization being enthalpic in origin, the structural distortion associated with the formation of transplatin cross-link induces a very similar decrease in duplex thermodynamic stability ($\Delta\Delta G_{25}$*) of +6.5 kcal/mol, but this destabilization is entropic in origin. In other words, for the duplex containing the interstrand cross-link of cisplatin relative to the duplex containing the interstrand cross-link of transplatin, the compensating enthalpic and entropic effects almost completely offset the difference in cross-link-induced energetic destabilization. The observation that the transplatin interstrand cross-link is nearly enthalpically neutral in contrast to the same lesion formed by cisplatin is consistent with a relatively small conformational distortion induced by the cross-link of transplatin in comparison with a markedly more severe distortion induced by the cross-link of cisplatin.

We also attempted to rationalize the enthalpic destabilizing effect of the interstrand cross-link of cisplatin ($\Delta\Delta H_{cal}$ = 18 kcal/mol (Table I)) in terms of the cross-link-induced structural perturbations in the host duplex. Conformational changes in DNA induced by the single, site-specific interstrand cross-link of cisplatin have been investigated already (see above) by various techniques. The view that the conformational distortion induced by the interstrand cross-link of cisplatin is not only localized to the platinated base pairs is supported by the results of DNase I footprinting of cisplatin-modified DNA (39). The approach, based on using a chemical nuclease such as 1,10-phenantroline-copper (38), even permitted us to determine the approximate number of base pairs, the secondary structure of which had been perturbed by the cross-link. Importantly, this chemical nuclease does not represent measures of base pair disruption, as they can proceed even if the base pairing is distorted in a nondenaturational manner (40). The analysis of the duplex containing the site-specific interstrand cross-link of cisplatin has revealed that the interstrand cross-link of cisplatin results in distortions between ~4 to 5 base pairs at the d(GC)·d(GC) site involved in the cross-link (38). Such a conformational alteration should be enthalpically unfavorable, which is consistent with observations of the present work. A rough estimation of an upper limit for this cross-link-induced perturbation, based on the enthalpic cost resulting from disruption of 4–5 relevant nearest-neighbor stacking interactions, yields $\Delta\Delta H$ value of 25–32.5 kcal/mol (41). Our experimental calorimetric data (see Table I) reflect a cross-link-induced $\Delta\Delta H_{cal}$ value of 18 kcal/mol.

An attempt to account for the differences between the predicted upper limit $\Delta\Delta H$ value of 25–32.5 kcal/mol and the measured calorimetric value of 18 kcal/mol may invoke at least some energetic consequences of the formation of the interstrand cross-link of cisplatin. An important feature of the structure of this adduct is that cytosine residues complementary to the platinated guanines are no longer hydrogen-bonded and are extrahelical (see above). On the other hand, the unpaired platinated guanine residues are stacked with the adjacent base pairs, which contributes to the stabilization of the duplex. Similarly, it has been proposed that the hydration of the interstrand cross-link (42) and bending induced by the adducts of cisplatin (21) thermodynamically stabilize the duplex. Although the stabilization resulting from hydration cannot be quantified because of the limited thermodynamic database on DNA hydration, at least a very crude estimate can be obtained in the case of helical bending. It was shown that the bending due to the formation of the 1,2-d(GpG) intrastrand cross-link contributed roughly 6.4 kcal/mol toward stabilization of the global duplex structure (21) so that it seems reasonable to assume that the contribution of the bending induced by the interstrand cross-link of cisplatin is at least that derived for its intrastrand cross-link. Another important conformational parameter of the distortion induced by the formation of the interstrand cross-links of cisplatin is the unwinding of the double helix (lowering of the number of base pairs per helical turn). The energetics of the destabilizing effect of DNA unwinding can be estimated crudely using the same approach as used to calculate the free energy required to twist a 12-bp-long DNA fragment containing a single 1,2-d(GpG) intrastrand adduct of

cisplatin about its helix axis (43). The free energy of unwinding of only 0.29 kcal/mol was calculated assuming the unwinding angle is 13° (43). The same calculations were performed, assuming unwinding angles of 80° for the interstrand cross-link of cisplatin (see above) and local twisting within the 20-bp-long fragment (the 20-bp-long fragment was taken for these calculations because the energetics of the duplex d(TGCT)·d(AGCA) of this length containing the single, site-specific interstrand cross-link of cisplatin was characterized in the present work (Fig. 2 and Table I)). These approximate calculations give a rough estimate of the free energy of unwinding of 6.7 kcal/mol. If this rough estimate is justified, then it seems reasonable to suggest that the destabilization of the duplex because of its unwinding as a consequence of the formation of the interstrand cross-link of cisplatin compensates at least partially for the stabilizing effect of bending. Furthermore, a local reversal of the double helix to a left-handed form at the site of the cross-link undoubtedly also contributes to the observed energetic impact of the interstrand cross-link of cisplatin. In general, a prediction of the energetic consequences of conformational changes induced by the interstrand cross-link of cisplatin is difficult because of the limited current knowledge on the thermodynamic consequences of distortions and transitions on DNA duplexes. Despite these possible microscopic interpretations of our macroscopic data, the calorimetric results reported here reveal that the formation of the interstrand cross-link of cisplatin induces on one hand a substantial thermal stabilization associated mainly with the change in the molecularity of the system and on the other hand thermodynamic destabilization of the host duplex that is enthalpic in origin. The interstrand cross-link of transplatin also induces thermal stabilization associated mainly with the change in the molecularity of the system, but it is enthalpically neutral so that the thermodynamic destabilization of the duplex induced by this transplatin adduct is entirely entropic in origin.

From the ratio of the model-dependent van't Hoff and the model-independent calorimetric transition enthalpies (Table II), one can define the fraction of a duplex that undergoes transition as a single thermodynamic unity (31, 33, 34). Thus, for the duplex containing an interstrand cross-link of cisplatin this ratio is 0.74. This value indicates that the largest size of the unit in the host duplex, which melts in the all-or-none manner, should involve 74% of the duplex. We speculate that this value reflects the existence of a severe local distortion of the duplex at the central cross-link of cisplatin (23, 35, 36, 38) and that the base pairs in this distorted segment melt cooperatively and more easily than the rest of the duplex. Hence, one intermediate state of the cisplatin-cross-linked duplex during its thermal melting could involve a short, denatured central structure arising from the segment ~4–5 base pairs long consisting of the two base pairs involved in the cross-link and approximately two or three base pairs flanking the cross-linked base pairs. This size of the denatured central structure, which represents 20–25% of the duplex examined in the present work, is deduced from the observation that severe distortion induced by cisplatin cross-link extends over approximately four or five base pairs at the site of the adduct (38). The remaining ~15 or 16 base pairs of the duplex d(TGCT)·d(AGCA) represent 75 or 80% of its size; this value is very close to the 74% found for the largest size of the unit melting cooperatively in a two-state process in the host duplex on the basis of the ratio of the model-dependent van't Hoff and the model-independent calorimetric transition enthalpies (Table II). The two-state melting of the duplex containing two marginal segments of a similar length and GC content, which are separated by the central denatured region (4–5 base pairs long) around the cross-link,

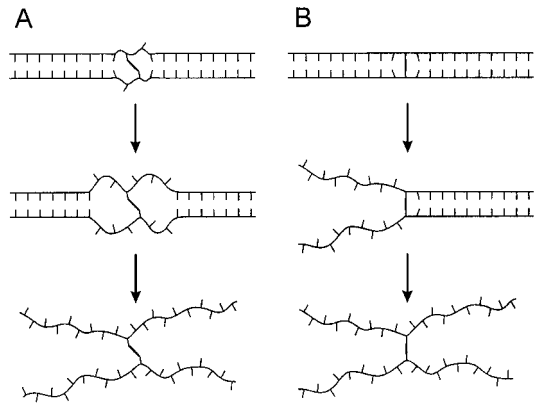


FIG. 4. **Schematic representation of thermal melting of the duplexes containing central, site-specific interstrand cross-link of cisplatin (*A*) or transplatin (*B*).**

deserves additional discussion. The two-state melting was observed even in the case of an immobile DNA junction composed of the four isolated octameric duplex arms in which not all arms had the same melting temperature (44). For the duplex examined in the present work consisting of the two marginal segments separated by the central denatured region, because of two different sets of circumstances the two-state melting can take place in a manner analogous to that proposed for the melting behavior of the DNA junction structure (44). In one case, the two marginal duplex parts, seven or eight base pairs long, separated by the central denatured region around the cross-link may fortuitously possess the same T_m. When this situation prevails, the remainder of the duplex consisting of the two marginal segments will melt in an apparent two-state manner without requiring any cooperative communication across the central denatured part around the cross-link. A second alternative case might truly reflect an all-or-none melting event. When this case prevails, molecular communication should occur between the marginal parts of the duplex across the denatured central part in a manner that results in a cooperative melting event. Our data do not allow us to differentiate between these two alternatives.

On the other hand, the $\Delta H_{vH}/\Delta H_{cal}$ ratio found for the duplex containing an interstrand cross-link of transplatin was 0.55. This value is consistent with this duplex containing two units each melting cooperatively, the size of which is close to one-half of the duplex. The conformational distortion in the central part of the duplex around the cross-link of transplatin is much less severe than that induced by the cross-link of cisplatin. As a consequence of this markedly less pronounced distortion, the base pairs in the short segment of the duplex around the cross-link of transplatin would not melt independently as a separated cooperative unit as in the case of the distorted segment around the cross-link of cisplatin. The schematic representation of the melting of the duplexes, which are interstrand cross-linked by cisplatin or transplatin, proposed on the basis of the results of the present work is shown in Fig. 4.

In the aggregate, our results reveal and characterize the profound but different impacts that interstrand cross-linking by antitumor cisplatin and clinically ineffective transplatin can have on DNA stability and melting behavior. Such assessments are important in a range of applications including those aimed at understanding the molecular mechanisms underlying the biological effects of bifunctional agents that modify DNA. Moreover, the results of the present work further support the view that the impact of the interstrand cross-links of cisplatin and transplatin on DNA is different, which might be also associated with distinctly different antitumor effects of these platinum compounds.

REFERENCES

1. Rosenberg, B. (1999) in *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., ed) pp. 3–30, Wiley-VCH, Zürich
2. O'Dwyer, P. J., Stevenson, J. P., and Johnson, S. W. (1999) in *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., ed) pp. 31–72, Wiley-VCH, Zürich
3. Kelland, L. R., Sharp, S. Y., ONeill, C. F., Raynaud, F. I., Beale, P. J., and Judson, I. R. (1999) *J. Inorg. Biochem.* **77,** 111–115
4. Farrell, N., Qu, Y., Bierbach, U., Valsecchi, M., and Menta, E. (1999) in *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., ed) pp. 479–496, Wiley-VCH, Zürich
5. Comess, K. M., Burstyn, J. N., Essigmann, J. M., and Lippard, S. J. (1992) *Biochemistry* **31,** 3975–3990
6. Jamieson, E. R., and Lippard, S. J. (1999) *Chem. Rev.* **99,** 2467–2498
7. Eastman, A. (1999) in *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., ed) pp. 111–134, Wiley-VCH, Zürich
8. Eastman, A. (1987) *Pharmacol. Ther.* **34,** 155–166
9. Fichtinger-Schepman, A. M. J., Van der Veer, J. L., Den, H.artog, J. H. J., Lohman, P. H. M., and Reedijk, J. (1985) *Biochemistry* **24,** 707–713
10. Brabec, V., and Leng, M. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 5345–5349
11. Vrana, O., Boudny, V., and Brabec, V. (1996) *Nucleic Acids Res.* **24,** 3918–3925
12. Eastman, A., Jennerwein, M. M., and Nagel, D. L. (1988) *Chem. Biol. Interact.* **67,** 71–80
13. Eastman, A., and Barry, M. A. (1987) *Biochemistry* **26,** 3303–3307
14. Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) *Nature* **399,** 708–712
15. Pilch, D. S., Dunham, S. U., Jamieson, E. R., Lippard, S. J., and Breslauer, K. J. (2000) *J. Mol. Biol.* **296,** 803–812
16. Zaludova, R., Kleinwächter, V., and Brabec, V. (1996) *Biophys. Chem.* **60,** 135–142
17. Zaludova, R., Zakovska, A., Kasparkova, J., Balcarova, Z., Kleinwächter, V., Vrána, O., Farrell, N., and Brabec, V. (1997) *Eur. J. Biochem.* **246,** 508–517
18. Zaludova, R., Zakovska, A., Kasparkova, J., Balcarova, Z., Vrana, O., Coluccia, M., Natile, G., and Brabec, V. (1997) *Mol. Pharmacol.* **52,** 354–361
19. Zakovska, A., Novakova, O., Balcarova, Z., Bierbach, U., Farrell, N., and Brabec, V. (1998) *Eur. J. Biochem.* **254,** 547–557
20. Brabec, V., Kasparkova, J., Vrana, O., Novakova, O., Cox, J. W., Qu, Y., and Farrell, N. (1999) *Biochemistry* **38,** 6781–6790
21. Poklar, N., Pilch, D. S., Lippard, S. J., Redding, E. A., Dunham, S. U., and Breslauer, K. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 7606–7611
22. Malina, J., Hofr, C., Maresca, L., Natile, G., and Brabec, V. (2000) *Biophys. J.* **78,** 2008–2021
23. Brabec, V. (2000) in *Platinum Based Drugs in Cancer Therapy* (Kelland, L. R., and Farrell, N. P., eds) pp. 37–61, Humana Press Inc., Totowa, NJ
24. Zamble, D. B., Mu, D., Reardon, J. T., Sancar, A., and Lippard, S. J. (1996) *Biochemistry* **35,** 10004–10013
25. McA'Nulty, M. M., and Lippard, S. J. (1996) *Mutat. Res.* **362,** 75–86
26. Zamble, D. B., and Lippard, S. J. (1995) *Trends Biochem. Sci.* **20,** 435–439
27. Brabec, V., Reedijk, J., and Leng, M. (1992) *Biochemistry* **31,** 12397–12402
28. Brabec, V., Sip, M., and Leng, M. (1993) *Biochemistry* **32,** 11676–11681
29. Brabec, V., Neplechova, K., Kasparkova, J., and Farrell, N. (2000) *J. Biol. Inorg. Chem.* **5,** 364–368
30. Lemaire, M. A., Schwartz, A., Rahmouni, A. R., and Leng, M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 1982–1985
31. Marky, L. A., and Breslauer, K. J. (1987) *Bioploymers* **26,** 1601–1620
32. Osborne, S. E., Völker, J., Stevens, S. Y., Breslauer, K. J., and Glick, G. D. (1996) *J. Am. Chem. Soc.* **118,** 11993–12003
33. Breslauer, K. J. (1995) *Methods Enzymol.* **259,** 221–242
34. Pilch, D. S., Plum, G. E., and Breslauer, K. J. (1995) *Curr. Opin. Struct. Biol.* **5,** 334–342
35. Huang, H. F., Zhu, L. M., Reid, B. R., Drobny, G. P., and Hopkins, P. B. (1995) *Science* **270,** 1842–1845
36. Coste, F., Malinge, J.-M., Serre, L., Shepard, W., Roth, M., Leng, M., and Zelwer, C. (1999) *Nucleic Acids Res.* **27,** 1837–1846
37. Paquet, F., Boudvillain, M., Lancelot, G., and Leng, M. (1999) *Nucleic Acids Res.* **27,** 4261–4268
38. Malinge, J. M., Perez, C., and Leng, M. (1994) *Nucleic Acids Res.* **22,** 3834–3839
39. Schwartz, A., and Leng, M. (1994) *J. Mol. Biol.* **236,** 969–974
40. Sigman, D. S., and Chen, C. H. B. (1990) *Annu. Rev. Biochem.* **59,** 207–236
41. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 3746–3750
42. Malinge, J. M., Giraud-Panis, M. J., and Leng, M. (1999) *J. Inorg. Biochem.* **77,** 23–29
43. Bellon, S. F., Coleman, J. H., and Lippard, S. J. (1991) *Biochemistry* **30,** 8026–8035
44. Marky, L. A., Kallenbach, N. R., McDonough, K. A., Seeman, N. C., and Breslauer, K. J. (1987) *Biopolymers* **26,** 1621–1634

Chemistry & Biology, Vol. 12, 121–129, January, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.chembiol.2004.11.008

Conformation of DNA Modified by Monofunctional Ru(II) Arene Complexes: Recognition by DNA Binding Proteins and Repair. Relationship to Cytotoxicity

Olga Novakova,[1] Jana Kasparkova,[1] Vendula Bursova,[1] Ctirad Hofr,[1] Marie Vojtiskova,[1] Haimei Chen,[2] Peter J. Sadler,[2,*] and Viktor Brabec[1,*]

[1]Institute of Biophysics
Academy of Sciences of the Czech Republic
Kralovopolska 135
CZ-61265 Brno
Czech Republic

[2]School of Chemistry
University of Edinburgh
West Mains Road
Edinburgh, EH9 3JJ
United Kingdom

Summary

We analyzed DNA duplexes modified at central guanine residues by monofunctional Ru(II) arene complexes [(η^6-arene)Ru(II)(en)(Cl)]$^+$ (arene = tetrahydroanthracene or *p*-cymene, Ru-THA or Ru-CYM, respectively). These two complexes were chosen as representatives of two different classes of Ru(II) arene compounds for which initial studies revealed different binding modes: one that may involve DNA intercalation (tricyclic-ring Ru-THA) and the other (mono-ring Ru-CYM) that may not. Ru-THA is ~20 times more toxic to cancer cells than Ru-CYM. The adducts of Ru-THA and Ru-CYM have contrasting effects on the conformation, thermodynamic stability, and polymerization of DNA in vitro. In addition, the adducts of Ru-CYM are removed from DNA more efficiently than those of Ru-THA. Interestingly, the mammalian nucleotide excision repair system has low efficiency for excision of ruthenium adducts compared to cisplatin intrastrand crosslinks.

Introduction

There is much current interest in the potential of ruthenium complexes as new metal-based antitumor drugs [1, 2]. Although the pharmacological target for antitumor ruthenium compounds has not been unequivocally identified, several ruthenium(III) compounds have been found to inhibit DNA replication, exhibit mutagenic activity, induce the SOS repair mechanism, bind to nuclear DNA, and reduce RNA synthesis, which is consistent with DNA binding of these compounds in vivo [2]. Thus, DNA interactions of antitumor ruthenium agents are of potential importance. Organometallic ruthenium(II) arene complexes of the type [(η^6-arene)Ru(II)(en)Cl][PF_6] (en = ethylenediamine) constitute a relatively new group of anticancer compounds [3, 4]. These monodentate complexes appear to be novel anticancer agents with a mechanism of action different from those of the ruthenium(III) complexes (ImH)[*trans*-$Ru(III)Cl_4$Im(Me_2SO)] (Im = imidazole, NAMI-A) and (IndH)[*trans*-$RuCl_4(Ind)_2$] (Ind = indazole, KP1019), which are currently in clinical trials [5]. The (η^6-arene)Ru(II) π bonds in the monofunctional [(η^6-arene)Ru(II)(en)(Cl)]$^+$ complexes are inert toward hydrolysis, but the chloride ligand is readily lost, and the complex is transformed into the corresponding more reactive, aquated species [6]. It has also been shown that in cell-free media ethylenediamine Ru(II) arene compounds, in which arene = biphenyl, dihydroanthracene, tetrahydroanthracene, *p*-cymene, or benzene, bind preferentially to guanine residues in natural double-helical DNA. In addition, DNA binding of the complexes containing biphenyl, dihydroanthracene, or tetrahydroanthracene ligands can involve combined coordination to G N7 and noncovalent, hydrophobic interactions between the arene ligand and DNA, which may include arene intercalation and minor groove binding [7, 8]. In contrast, the single hydrocarbon rings in the *p*-cymene and benzene ruthenium complexes cannot interact with double-helical DNA by intercalation [8]. Interestingly, adducts of the complex containing the *p*-cymene ligand, which has methyl and isopropyl substituents, distort the conformation and thermally destabilize double-helical DNA distinctly more than the adducts of the tricyclic-ring Ru(II) arene compounds. It has been suggested that the different character of conformational alterations induced in DNA as a consequence of its global modification, and the resulting thermal destabilization, may affect differently further "downstream" effects of damaged DNA [8] and consequently may result in different biological effects of this new class of metal-based antitumor compounds.

To achieve a rational design of novel antitumor Ru(II) arene compounds capable of circumventing inherent or acquired resistance to metal-based drugs already used in the clinic, it is important to understand in detail the differences in DNA binding properties of these new ruthenium complexes and their possible relationship to cytotoxicities in different tumor cell lines. This may provide grounds for establishing new structure-pharmacological activity relationships for this class of metal-based complexes. In this work, we have considered the activity of two Ru(II) arene complexes from the [(η^6-arene)Ru(II)(en)(Cl)]$^+$ family (arene = tetrahydroanthracene and *p*-cymene, Ru-THA and Ru-CYM, respectively, Figure 1A) in two tumor cell lines. These two complexes were chosen as representatives of two different classes of Ru(II) arene compounds for which initial studies of global modification of natural DNA revealed [8] different binding modes: one that may involve DNA intercalation (tricyclic-ring Ru-THA) and the other (mono-ring Ru-CYM) that cannot interact with double-helical DNA by intercalation. We compare the cytotoxicity data with those for DNA binding obtained previously [8], new data obtained in the present work relating to conformational distortions induced by single, site-specific monofunctional adducts of the Ru(II) arene complexes in short oligodeoxyribonucleotide duplexes, and the recognition of these DNA adducts by specific proteins and their repair, i.e., the most important factors that

*Correspondence: pjs01@staffmail.ed.ac.uk; brabec@ibp.cz

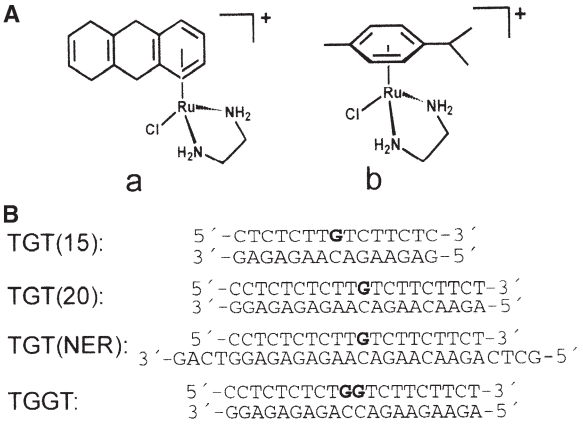


Figure 1. Structures of the Ruthenium Complexes and the Nucleotide Sequences of the Synthetic Oligodeoxyribonucleotide Duplexes with the Abbreviations Used in the Present Study

(A) Structures. (a) Ru-THA; (b) Ru-CYM.

(B) Sequences. The top and bottom strands of the pair of oligonucleotides are designated "top" and "bottom," respectively. The boldface letter in the top strand of the duplexes indicates the ruthenated residues.

modulate the antitumor effects of platinum antitumor drugs already used in the clinic.

Results

Chemical Probes of DNA Conformation

We demonstrated in our preceding paper [8] that Ru(II) arene compounds bind preferentially to guanine residues in natural double-helical DNA forming monofunctional adducts. In order to obtain information on how these adducts affect DNA conformation, the oligonucleotide duplexes containing a site-specific monofunctional adduct of Ru-THA or Ru-CYM (Figure 1A), [(η^6-tetrahydroanthracene)Ru(II)(en)(Cl)]$^+$ and [(η^6-*p*-cymene)Ru(II)(en)(Cl)]$^+$, respectively, at the G residue were further analyzed by chemical probes of DNA conformation. The ruthenated duplex TGT(20) (Figure 1B) was treated with several chemical agents that are used as tools for monitoring the existence of conformations other than canonical B DNA. These agents include $KMnO_4$ and diethyl pyrocarbonate (DEPC). They react preferentially with single-stranded DNA and distorted double-stranded DNA [9, 10]. We used for this analysis exactly the same methodology as described in detail in our recent papers in which we studied DNA adducts of various antitumor platinum drugs [9, 10], and, therefore, these experiments are described in more detail in the Supplemental Data (see the Supplemental Data available with this article online). The results demonstrated in Figure 2 indicate that the distortion induced by the nonintercalating Ru-CYM extended over at least 7 base pairs (bp), whereas the distortion induced by Ru-THA was less extensive.

Isothermal Titration Calorimetry

A calorimetric technique was employed to characterize the influence of the monofunctional adduct of Ru-THA

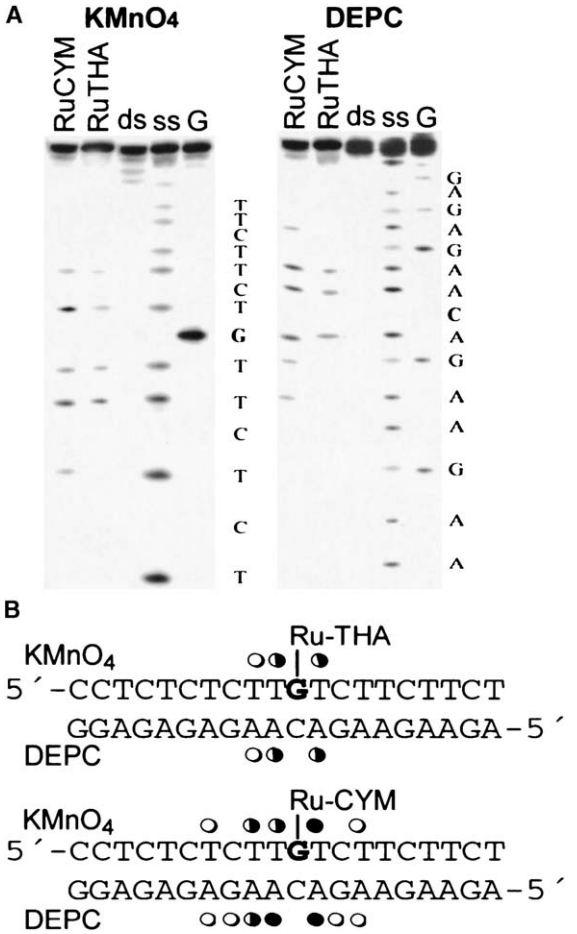


Figure 2. Chemical Probes of DNA Conformation

(A) Piperidine-induced specific strand cleavage at $KMnO_4$-modified (lanes marked with $KMnO_4$) and DEPC-modified (lanes marked with DEPC) bases in the nonmodified TGT(20) duplex or that containing the single, monofunctional adduct at the central G in the top strand. The oligomer was 5′-end labeled at its top ($KMnO_4$) or bottom (DEPC) strand. Lanes: ss, the nonmodified single strand; ds, the nonmodified duplex; RuTHA, the duplex containing an adduct of Ru-THA; RuCYM, the duplex containing an adduct of Ru-CYM; G, a Maxam-Gilbert-specific reaction for the unplatinated duplex. The boldface letters in the sequences indicate the ruthenated G and complementary C residues.

(B) Summary of the reactivity of chemical probes of DNA conformation. Upper panel, adduct of Ru-THA; lower panel, adduct of Ru-CYM. Filled circle, strong reactivity; half-filled circle, medium reactivity; open circle, weak reactivity. The boldface letter in the top strand of the duplexes indicates the ruthenated G residue.

and Ru-CYM on the thermal stability and energetics of the site-specifically ruthenated 15 bp DNA duplex. Such thermodynamic data can reveal how the ruthenium adduct influences duplex stability, a property that has been shown to play a significant role in cellular processes such as recognition of DNA damage by DNA binding proteins and repair of this damage, i.e., the processes that may modulate potency of antitumor drugs, including metal-based cytostatics. Recently, differential scanning calorimetry (DSC) was employed to characterize the influence of different crosslinks of platinum

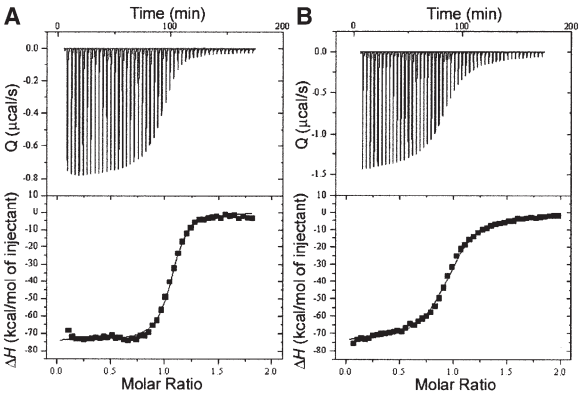


Figure 3. Isothermal Titration Calorimetry

(A and B) ITC binding isotherm for the association of the top strand of the 15 bp TGT duplex containing a single, monofunctional adduct of (A) Ru-THA or (B) Ru-CYM with the complementary (nonmodified) strand (bottom strand of the duplex TGT(15)) at 25°C in 10 mM phosphate buffer (pH 7.0) containing 50 mM NaCl. The upper panels show total heat released upon injecting 5 μl aliquots of the 50 μM bottom strand into a 1.4 ml reaction cell containing the 5 μM top strand. The lower panels show the resultant binding isotherm (full squares) obtained by integrating the peak areas of each injection. The continuous line represents the nonlinear least squares fit of the affinity (*K*), enthalpy (Δ*H*), and stoichiometry (*n*) to a single-site binding model. For other details, see the text.

antitumor drugs on the thermal stability and energetics of 15–20 bp DNA duplexes site-specifically modified by these drugs [11–13]. We decided to expand these studies to oligodeoxyribonucleotide duplexes containing unique monofunctional adducts of the Ru-THA or Ru-CYM complexes. DSC makes it possible to measure excess heat capacity versus temperature profiles for the thermally induced transitions of nonmodified DNA duplexes and those containing a unique adduct of the metal-based drug. Such thermograms are usually recorded with the heating rate of 60°C/hr, and after reaching the maximum temperature of 95°C, the samples are cooled at the same rate to the starting temperature of 25°C [11–13]. This implies that the duplexes containing the unique adduct are exposed to higher temperatures for a relatively long period of time. Therefore, we verified first the stability of the Ru(II) arene adducts at various temperatures and found that the adducts formed in the TGT(15) duplex by Ru-THA and Ru-CYM are stable for more than 2 hr only at temperatures lower than 50°C. Hence, it is apparent that DSC cannot be used to analyze the duplexes containing the adducts of these two Ru(II) arene complexes. A suitable alternative is isothermal titration calorimetry (ITC), which makes it possible to study the thermodynamic parameters of the duplex formation from its two complementary single strands over a range of temperatures including those at which the DNA adducts of Ru(II) arene compounds were stable for a long period of time [14].

Figure 3 shows ITC profiles of duplex formation from the nonmodified bottom strand of the duplex TGT(15) titrated into the complementary nonmodified top strand of the duplex TGT(15) or the same strand containing a single monofunctional adduct of Ru-THA or Ru-CYM at 25°C. It was verified that the melting temperature of the duplexes was significantly higher than this temperature and that the adducts of Ru-THA and Ru-CYM formed in the TGT(15) duplex or in its single-stranded top strand were stable for more than 24 hr. The ITC profiles were analyzed as described in the Experimental Procedures to obtain the results listed in Table 1. Inspection of these thermodynamic parameters reveals that the exothermic formation of the single monofunctional adduct in the duplex TGT(15) by Ru-THA or Ru-CYM resulted in a large decrease of the change in the enthalpy of duplex formation by 4.4 and 7.4 kcal/mol, respectively. In other words, the monofunctional adduct of Ru-THA or Ru-CYM enthalpically destabilizes the duplex relative to its nonmodified counterpart. In addition, the formation of the monofunctional adducts by Ru-THA or Ru-CYM resulted in a substantial increase in the entropy of the duplex TGT(15) of 11.6 or 18.2 cal/K.mol ($T\Delta\Delta S$ = 3.6 and 5.4 kcal/mol at 25°C), respectively. In other words, the monofunctional adduct of Ru-THA or Ru-CYM increases the entropy of the ruthenated duplexes and, in this way, entropically stabilizes the duplex. Thus, the 4.4 or 7.4 kcal/mol enthalpic destabilization of the TGT(15) duplex due to the monofunctional adduct of Ru-THA or Ru-CYM is partially, but not completely, compensated by the entropic stabilization of the duplex induced by these adducts of 3.5 or 5.4 kcal/mol at 25°C, respectively. The net result of these enthalpic and entropic effects is that the formation of the monofunctional adducts of Ru-THA and Ru-CYM with the duplex TGT(15) at 25°C induces a decrease in duplex thermodynamic stability ($\Delta\Delta G_{25}$) of 0.8 or 2.0 kcal/mol, respectively, with this destabilization being enthalpic in origin. In this respect, the monofunctional adduct of Ru-CYM was considerably more effective than that of Ru-THA.

Probing by DNA Polymerase

It has been demonstrated that various DNA secondary structures have significant effects on the processivity of a number of prokaryotic, eukaryotic, and viral DNA polymerases. Interestingly, with DNA templates containing site-specifically placed adducts of various platinum compounds, a number of prokaryotic and eukaryotic DNA polymerases were blocked but could also traverse through the adducts depending on their character and conformational alterations induced in DNA. Inhibition of prokaryotic DNA-dependent RNA polymerase by the adducts on DNA globally modified by Ru arene compounds, including Ru-THA and Ry-CYM, has already been demonstrated in in vitro transcription mapping experiments [8]. Interestingly, monofunctional adducts of cisplatin or transplatin and those of the monodentate compounds such as chlorodiethylenetriamineplatinum(II) chloride ([PtCl(dien)]Cl) or $[PtCl(NH_3)_3]Cl$ terminate DNA synthesis by DNA polymerases in vitro markedly less efficiently than crosslinks of platinum complexes [15, 16]. It is, therefore, interesting to examine whether a DNA polymerase, which processes DNA substrates containing monofunctional adducts of Ru-THA or Ru-CYM, can reveal potential specific features of conformational alterations imposed on DNA by the monofunctional adducts of these two Ru(II) arene compounds.

Table 1. Calorimetrically Derived Thermodynamic Parameters for the Formation of the 15 bp Nonmodified Duplexes or Those Containing Single, Site-Specific Monofunctional Adducts of Ru-THA or Ru-CYM at 25°C

Duplex	ΔH^a (kcal/mol)	ΔS^a (cal/K.mol)	$\Delta G_{25}{}^a$ (kcal/mol)	K^b (M^{-1})	N^b
TGT(15)	–79.3	–228	–11.3	1.98×10^8	1.08
TGT(15)·Ru-THA	–74.9	–216	–10.5	3.8×10^7	1.06
TGT(15)·Ru-CYM	–71.9	–210	–9.3	6×10^6	0.94

The ΔH and ΔS values are averages derived from three independent experiments.
[a] ΔH, ΔS, and ΔG_{25} denote, respectively, the enthalpy, entropy, and free energy (at 25°C) of duplex formation.
[b] K and n denote, respectively, association constant and binding stoichiometry for strand association.

We constructed 8-mer/23-mer primer/template nonmodified duplexes or those containing the monofunctional adduct of Ru-THA, Ru-CYM or [PtCl(dien)]Cl in the central TGT sequence (for sequences, see Figure 4). The first eight nucleotides on the 3′ terminus of the 23-mer template strand were complementary to the nucleotides of the 8-mer primer, and the guanine involved in the monofunctional adduct of Ru-THA, Ru-CYM, or [PtCl(dien)]Cl on the template strand was located at the 13th position from the 3′ terminus (Figure 4). After annealing the 8-mer primer to the 3′ terminus of the nonmodified or metallated template strand, positioning the 3′ end of the primer five bases before the adduct in the template strand, we examined DNA polymerization through the single, monofunctional adducts of Ru-THA, Ru-CYM, or [PtCl(dien)]Cl by a Klenow fragment of DNA polymerase I (KF) in the presence of all four deoxyribonucleoside 5′ triphosphates. The reaction was stopped at various time intervals, and the products were analyzed by using a sequencing 24% polyacrylamide (PAA)/8 M urea gel (shown for the monofunctional adducts of Ru-THA and Ru-CYM in Figure 4). Polymerization with the 23-mer template containing the adduct of Ru-THA proceeded rapidly up to the nucleotide preceding and at the sites opposite the adduct, such that the 12 and 13 nucleotide products accumulated to a significant extent (shown in Figure 4, lanes 6–10). The larger DNA intermediates were not observed to a considerable extent, whereas no intermediate products were seen with the 23-mer control template or the template containing the monofunctional adduct of [PtCl(dien)]Cl as the full-length products were formed (shown in Figure 4 for control template, lanes 1–5). The full-length products were also noticed with the 23-mer template containing the adduct of Ru-THA (Figure 4, lanes 6–10). This result demonstrates that the monofunctional adduct of Ru-THA effectively inhibits DNA synthesis, but translesion synthesis may occur. Under the same experimental conditions, DNA polymerization by KF with the template containing the monofunctional adduct of Ru-CYM proceeded up to the nucleotide preceding the site opposite the ruthenated G involved in the adduct and to the following nucleotide residue (Figure 4, lanes 11–15). There was no accumulation of shorter intermediates, but larger DNA intermediates (corresponding to 14 and 15 nucleotide products) and the full-length products were noticed. The amount of the full-length products increased with reaction time, but with a somewhat lower rate compared to the polymerization with the template containing the adduct of Ru-THA.

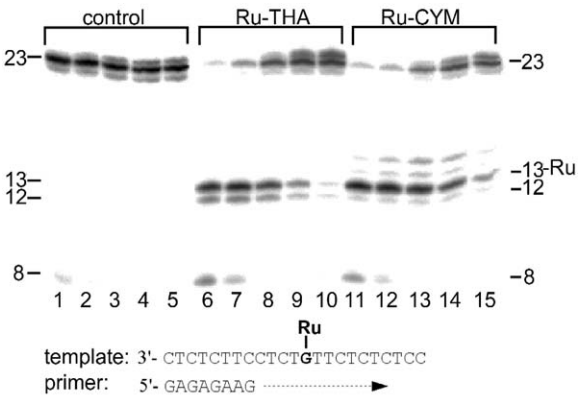


Figure 4. Primer Extension Activity of Klenow Fragment of DNA Polymerase I

Experiments were conducted by incubating 8-mer/23-mer primer/template duplex for various times (lanes 1–5), or the template containing a monofunctional adduct of Ru-THA (lanes 6–10) or of Ru-CYM (lanes 11–15). Timings were as follows: 1 min, lanes 1, 6, and 11; 3 min, lanes 2, 7, and 12; 15 min, lanes 3, 8, and 13; 30 min, lanes 4, 9, and 14; 60 min, lanes 5, 10, and 15. The pause sites opposite the ruthenated guanine and the preceding residues are marked 13 and 12, respectively (the site opposite the ruthenated residue is still marked "Ru"). The nucleotide sequences of the template and the primer are shown beneath the gels.

Repair

Figure 5A illustrates an experiment that measures DNA repair synthesis by a repair-proficient HeLa cell-free extract (CFE) in pUC19 plasmid modified at $r_b = 0.05$ by Ru-THA or Ru-CYM, and for comparative purposes, also by cisplatin. Repair activity was monitored by measuring the amount of incorporated radiolabeled nucleotide. A similar amount of undamaged pBR322 of a slightly different size is included in the reactions to show the background incorporation into undamaged plasmid. This background incorporation was subtracted from that found for metallated pUC19 plasmid. Considerably different levels of damage-induced DNA repair synthesis were detected in the plasmid modified by Ru-THA, Ru-CYM, and cisplatin (Figures 5A and 5B). The level of the synthesis detected in the plasmid modified by Ru-THA was ~6 times lower than that in the plasmid modified by Ru-CYM.

DNA repair synthesis can be due to various DNA repair mechanisms. Bulky, helix-distorting DNA adducts, such as those generated by various chemotherapeutics, including cisplatin, are removed from DNA by nucleotide excision repair (NER), which is an important

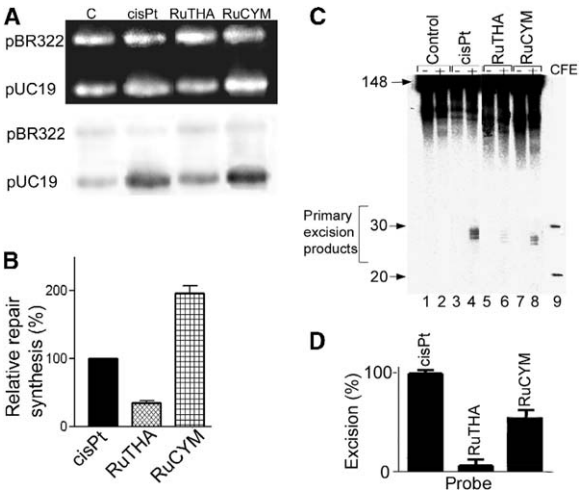


Figure 5. Repair DNA Synthesis and Nucleotide Excision Repair

(A and B) In vitro repair synthesis assay of the extract prepared from the repair-proficient HeLa cell line. Repair synthesis used as substrates nonmodified pBR322 plasmid and nonmodified pUC19 plasmid (lane C) or pUC19 plasmid modified at r_b = 0.05 by cisplatin, Ru-THA, or Ru-CYM (lanes cisPt, RuTHA, and RuCYM, respectively). (A) Results of a typical experiment. The top panel is a photograph of the EtBr-stained gel, and the bottom panel is the autoradiogram of the gel showing incorporation of [α-^{32}P]dCMP. (B) Incorporation of dCMP into nonmodified, platinated, or ruthenated plasmids. For all quantifications representing mean values of three separate experiments, incorporation of radioactive material is corrected for the relative DNA content in each band. The radioactivity associated with the incorporation of [α-^{32}P]dCMP into DNA modified by cisplatin was taken as 100%.

(C and D) Excision of the adducts of ruthenium complexes by rodent excinuclease. (C) The 148 bp substrates were incubated with CHO AA8 CFE and subsequently treated overnight with NaCN prior to analysis in 10% PAA/8 M urea denaturing gel; lanes: 1 and 2, control, nonmodified substrate; 3 and 4, the substrate containing the 1,2-GG intrastrand crosslink of cisplatin; 5 and 6, monofunctional adduct of Ru-THA; 7 and 8, the monofunctional adduct of Ru-CYM. Lanes 1, 3, 5, and 7, no extract added; lanes 2, 4, 6, and 8, the substrates were incubated with CHO AA8 CFE for 40 min at 30°C. Lane M, the 20- and 30-mer markers. (D) Quantitative analysis of removal of the adducts. The columns marked cisPt, RuTHA, and RuCYM represent 1,2-GG intrastrand crosslink of cisplatin, the monofunctional adduct of Ru-THA, and the monofunctional adduct of Ru-CYM, respectively. The radioactivity associated with the fragments excised from the duplex containing the 1,2-GG intrastrand CL of cisplatin was taken as 100%. Data are the average of two independent experiments performed under the same conditions.

component of the mechanism underlying the biological effects of these agents. Efficient removal of crosslinks formed in DNA by platinum antitumor compounds has already been reported for various NER systems, including human and rodent excinucleases [17–20]. The result presented in Figure 5C, lane 4 is consistent with these reports. The major excision fragment contains 28 nucleotides, and other primary excision fragments are 24–29 nucleotides in length [17–20]. In contrast, the monofunctional adducts of Ru-THA and Ru-CYM were also excised by both human and rodent excinucleases (shown for rodent excinuclease in Figures 5C and 5D), although with a noticeably lower efficiency than the major intrastrand crosslink of cisplatin; the adduct of Ru-CYM was excised slightly more than that of Ru-THA. Consistent with this observation were the results of the gel mobility shift assay analysis (Supplemental Figure S1) employing replication protein A (RPA) (which belongs to the initial damage-sensing factors of eukaryotic excision nuclease initiating repair) and DNA probes containing the adducts of Ru-THA, Ru-CYM, or cisplatin. The analysis performed with the substrate containing the major intrastrand crosslink of cisplatin revealed considerably higher binding than that performed with the substrate containing the adduct of Ru-THA or Ru-CYM (Supplemental Figure S1). In addition, a lower binding of RPA to the substrate containing the adduct of Ru-THA than to the substrate with the adduct of Ru-CYM was observed (Supplemental Figure S1). Thus, these results (described in detail in the Supplemental Data) corroborate the findings in Figures 5C and 5D and demonstrate the low efficiency of the mammalian NER systems employed in the present work to excise Ru(II) arene adducts and especially that of Ru-THA.

Cytotoxicity

The cytotoxic activity of the Ru(II) arene compounds tested in the present work was evaluated as described previously [3, 4] and has been determined in two cancer cell lines, A2780 and HT29. The compounds were incubated for 24 hr with the tumor cell lines [3, 4]. IC_{50} values (compound concentration that produces 50% of cell killing) of 0.4 and 10 μM were obtained in A2780 cells, and values of 3 and >100 μM were obtained in HT29 cells for Ru-THA and Ru-CYM, respectively (D.I. Jodrell and R. Aird, personal communication). Hence, the tricyclic-ring complex Ru-THA was considerably more potent than the Ru-CYM complex. Thus, the capability of the Ru(II) arene complex to intercalate DNA correlates with the noticeably enhanced activity of this class of Ru(II) arene compounds in several cancer cell lines.

Discussion

Our initial studies [8] suggested that Ru(II) arene compounds containing multi-ring biphenyl, dihydroanthracene, or tetrahydroanthracene ligands bind to DNA differently in comparison to the complexes containing single hydrocarbon rings, such as *p*-cymene or benzene. DNA binding of the multiring Ru(II) arene complexes involves not only coordination to G N7 together with C6O··(en) H–bonding, but also noncovalent, hydrophobic interactions between the arene ligand and DNA, which may include arene intercalation and minor groove binding. In contrast, the Ru arene compounds containing single hydrocarbon rings cannot interact with double-helical DNA by intercalation. Therefore, we tried first to find out whether this different DNA binding mode correlates with the cytotoxicity of the Ru(II) arene compounds in different tumor cell lines. Interestingly, the results of the previous [3, 4] and present work (vide supra) support the view that the presence of an arene ligand in these classes of ruthenium complexes that is capable of noncovalent, hydrophobic interaction with DNA (presumably intercalation) considerably enhances the cytotoxicity in a number of tumor cell lines.

It has been demonstrated that the biological activity of several transition metal-based complexes is modulated by the "downstream" effects of damaged DNA, such as recognition of damaged DNA by specific proteins and/or repair [21, 22]. For instance, recognition of DNA adducts of several antitumor metal-based drugs and removal of these adducts from DNA is dependent on the character of the distortion and thermodynamic destabilization induced in DNA by these adducts [11, 23]. A more detailed analysis of conformational distortions induced in DNA by Ru-THA and Ru-CYM carried out in the present work revealed substantial differences in the character of these distortions. Their analysis by chemical probes of DNA conformation demonstrated (Figure 2) that the distortion induced by the nonintercalating Ru-CYM extended over at least 7 bp, whereas the distortion induced by Ru-THA was less extensive. Consistent with this observation were the results of the ITC analysis (Figure 3 and Table 1). The association constants, *K*, for the formation of duplexes containing adducts of Ru-THA or Ru-CYM were 5 or 33 times lower, respectively, than the *K* value for the formation of the control (nonmodified) duplex (Table 1). Hence, the adducts of both Ru-THA and Ru-CYM thermodynamically destabilized DNA, with this destabilization being enthalpic in origin. The adduct of Ru-CYM destabilized DNA significantly more than the adduct of Ru-THA, whose DNA binding mode, additionally, involves noncovalent, hydrophobic interactions between the arene ligand and DNA, such as arene intercalation. Various intercalators thermodynamically stabilize DNA since they lengthen and unwind DNA, increasing the phosphate spacing along the helix axis [24, 25]. Hence, it is reasonable to suggest that the higher thermodynamic stability of DNA containing the adducts of Ru-THA observed in the present work is associated with this hydrophobic interaction.

Consistent with the different character of the adducts of Ru-THA and Ru-CYM and with the different impact of these adducts on DNA conformation and stability are also their different effects on primer extension activity of KF (Figure 4). The results of the present work suggest that the monofunctional adducts of Ru-THA and Ru-CYM efficiently inhibit DNA polymerization and in different ways. These studies demonstrate that the monofunctional adducts of Ru-THA and Ru-CYM constitute a fairly strong block to DNA synthesis catalyzed by KF; however, this block is not absolute, allowing translesion DNA synthesis with a limited efficiency. Hence, DNA polymerization appears to be inhibited by Ru(II) arene adducts markedly more strongly than by the adducts of simple monofunctional platinum(II) compounds. This provides a new dimension to the design of Ru(II) arene compounds for affecting processes in tumor cells, possibly including replication or DNA repair.

It has been suggested [21, 22] that HMG domain proteins play a role in sensitizing cells to cisplatin. It has been shown [26] that HMG domain proteins recognize and bind to DNA crosslinks formed by cisplatin. The details of how the binding of HMG domain proteins to cisplatin-modified DNA sensitize tumor cells to cisplatin are still not completely resolved, but possibilities such as shielding cisplatin-DNA adducts from repair or that these proteins could be recruited from their native transcriptional regulatory function have been suggested [21, 22] as clues for how these proteins are involved in the antitumor activity. In addition, an important structural motif recognized by HMG domain proteins on DNA modified by cisplatin is a directional bend of the helix axis toward the major groove [27]. No recognition of the DNA monofunctional adducts of Ru-THA or Ru-CYM by HMGB1 protein was observed in the present work (see the Supplemental Data). A plausible explanation of this observation may be that these adducts do not bend DNA, thus affording no structural motif recognized by HMG domain proteins. From these considerations, we could conclude that the mechanism of antitumor activity of Ru(II) arene compounds does not involve recognition of its DNA adducts by HMG domain proteins as a crucial step, in contrast to the proposals for cisplatin and its direct analogs [21, 22].

Another important feature of the mechanism underlying antitumor effects of DNA binding metal-based compounds is repair of their DNA adducts [21, 28]. A persistence of these DNA adducts may potentiate their antitumor effects in the cells sensitive to these compounds [21, 22, 29]. DNA repair synthesis was investigated in the present work by using the CFE from human tumor cells and DNA substrates randomly modified by the Ru(II) arene compounds (Figures 5A and 5B). Importantly, Ru-THA adducts induced a considerably lower level of repair synthesis than the adducts of Ru-CYM and also of cisplatin (Figure 5B), suggesting a less efficient removal from DNA and enhanced persistence of the adducts of more potent multi-ring and intercalating Ru(II) arene compounds in comparison with the adducts of less potent and nonintercalating Ru(II) arene compounds. Additionally, the level of DNA repair synthesis induced by the adducts of Ru-CYM was still markedly higher than that induced by cisplatin; this finding is consistent with the lower cytotoxicity of this ruthenium compound (vide supra).

There are several types of DNA repair, for instance base excision, NER, mismatch, and recombination repair. The assay based on the measurement of DNA repair synthesis may reflect the effectiveness of all of these repair mechanisms. Several reports have demonstrated [30–32] that NER is a major mechanism contributing to cisplatin resistance. The examination of excision of monofunctional adducts of Ru-THA and Ru-CYM has revealed that these adducts also can be removed from DNA by NER (Figures 5C and 5D), but considerably less efficiently than the adducts of cisplatin. This is in contrast to the results of DNA repair synthesis (Figures 5A and 5B) and implies a less significant role of NER in the mechanism underlying the antitumor effects of Ru(II) arene compounds than in the mechanism of cisplatin. In other words, the results of the present work indicate that the adducts of Ru(II) arene compounds are preferentially removed from DNA by repair mechanisms other than NER, which provides additional support for a mechanism of antitumor activity of Ru(II) arene compounds different from that of cisplatin. Nevertheless, the results of both repair assays employed in the present work (Figure 5) demonstrate clearly that the adducts of Ru-CYM, which distort and destabilize DNA more than the adducts of Ru-THA, are removed from

DNA more effectively, independent of the type of the repair mechanism.

Hence, the character and extent of DNA distortion induced in DNA by the adducts of Ru(II) arene complexes and resulting thermodynamic destabilization of DNA control the biological effects of this class of ruthenium complexes. The results of the present work afford further details, which allow for improving the structure-pharmacological relationship of Ru(II) arene compounds, and should provide a more rational basis for the design of new antitumor ruthenium drugs and chemotherapeutic strategies.

Significance

Organometallic ruthenium(II) arene complexes of the type [(η^6-arene)Ru(II)(en)Cl][PF_6] (en = ethylenediamine) constitute a new group of anticancer compounds. To achieve a rational design of novel antitumor Ru(II) arene compounds, it is important to understand the differences in DNA binding properties of these complexes and their possible relationship to cytotoxicities in different tumor cell lines. In this work, we studied the activity of two Ru(II) arene complexes from the [(η^6-arene)Ru(II)(en)(Cl)]$^+$ family (arene = tetrahydroanthracene and p-cymene, Ru-THA, and Ru-CYM, respectively) in two tumor cell lines, conformational distortions induced by monofunctional adducts of these complexes, and their recognition by DNA binding proteins and repair, i.e., the most important factors that modulate the antitumor effects of related platinum drugs. These two ruthenium complexes were chosen as representatives of two different classes of Ru(II) arene compounds that modify DNA differently: one that may interact with DNA by intercalation (tricyclic-ring Ru-THA), and the other (mono-ring Ru-CYM) that cannot.

The presence of the arene ligand in this class of ruthenium complexes capable of noncovalent, hydrophobic interaction with DNA considerably enhances cytotoxicity in several tumor cell lines. An analysis of DNA duplexes modified by Ru-THA and Ru-CYM revealed substantial differences in the impact of their monofunctional adducts on the conformation and thermodynamic stability of DNA and DNA polymerization in vitro. In addition, the adducts of Ru-CYM are removed from DNA more efficiently than those of Ru-THA. Interestingly, the adducts of Ru(II) arene compounds are preferentially removed from DNA by mechanisms other than nucleotide excision repair. This provides additional support for a mechanism underlying antitumor activity of Ru(II) arene compounds different from that of cisplatin. Hence, the character of DNA distortion induced in DNA by the adducts of Ru(II) arene complexes and the resulting thermodynamic destabilization of DNA control the biological effects of this class of ruthenium complexes.

Experimental Procedures

Starting Materials

The complexes Ru-THA and Ru-CYM (Figure 1A) were prepared by the methods described in detail previously [3, 7]. Cisplatin, glycogen, and dimethyl sulfate (DMS) were obtained from Sigma. [PtCl(dien)]Cl was kindly provided by G. Natile. The stock solutions of the ruthenium and platinum complexes at the concentration of 5×10^{-4} M in H_2O were prepared in the dark at 25°C. Plasmids pUC19 (2686 bp) and pBR322 (4363 bp) were isolated according to standard procedures. The synthetic oligodeoxyribonucleotides were purchased from VBC-Genomics (Vienna, Austria) and were purified as described previously [33]. Restriction endonucleases, T4 polynucleotide kinase, KF, and bovine serum albumin were purchased from New England Biolabs. A CFE was prepared from the HeLa S3 cell line as described [18]. This extract was kindly provided by J.T. Reardon and A. Sancar from the University of North Carolina. Acrylamide, agarose, bis(acrylamide), ethidium bromide (EtBr), urea, and NaCN were purchased from Merck KgaA. Creatine phosphokinase and creatine phosphate were purchased from ICN Biomedicals, Inc. The radioactive products were purchased from Amersham.

Metallation of Oligonucleotides

The single-stranded oligonucleotides (the top, pyrimidine-rich strands containing a single central G of the TGT(15), TGT(20), or TGT(NER) duplexes; Figure 1B) were reacted in stoichiometric amounts with either Ru-THA, Ru-CYM, or [PtCl(dien)]Cl. The ruthenated or platinated oligonucleotides were purified by ion-exchange fast protein liquid chromatography (FPLC). It was verified by ruthenium or platinum flameless atomic absorption spectrophotometry (FAAS) and by optical density measurements that the modified oligonucleotides contained one ruthenium or platinum atom. It was also verified by using DMS footprinting of ruthenium or platinum on DNA [34] that one molecule of ruthenium or platinum complex was coordinated to the N7 atom of the single G in the top strand of each duplex. FPLC purification and FAAS measurements were carried out on a Pharmacia Biotech FPLC System with a MonoQ HR 5/5 column and a Unicam 939 AA spectrometer equipped with a graphite furnace, respectively. The duplexes containing a single, central 1,2-GG intrastrand crosslink of cisplatin in the pyrimidine-rich top strand were prepared as described [12]. The nonmodified, ruthenated, or platinated duplexes used in the studies of recognition by RPA protein were purified by electrophoresis on native 15% PAA gels (mono:bis[acrylamide] ratio = 29:1).

Isothermal Titration Calorimetry

The standard isothermal titration calorimetry (ITC) buffer for these studies contained 50 mM NaCl with 10 mM phosphate buffer (Na_2HPO_4/NaH_2PO_4 [pH 7.0]). Sufficient quantities of ITC solutions were prepared to perform a set of titrations of the 50 μM solution of the bottom strand of the duplex TGT(15) (for its sequence, see Figure 1B) into the 5 μM solution of the top strand of nonmodified TGT(15) or that containing the single, site-specific monofunctional adduct of Ru-THA or Ru-CYM at 25°C. Molar extinction coefficients for the single-stranded oligonucleotides (related to the strands that were 15 nucleotides long) used in ITC experiments were determined by phosphate analysis [35]. The following extinction coefficients at 260 nm and 25°C were obtained: 108,000 and 128,000 $M^{-1}\cdot cm^{-1}$ for the upper and bottom strands of the nonmodified TGT(15) duplex, respectively; 113,000 and 111,000 $M^{-1}\cdot cm^{-1}$ for the upper strand of the TGT(15) containing the single monofunctional adduct of Ru-THA and Ru-CYM, respectively. Stock solutions of the strands for ITC studies were prepared in the ITC buffer and were exhaustively dialyzed against this buffer. It was verified that enthalpies of ITC injections of each individual oligomer into buffer, of buffer into buffer, and of excess oligomer into a solution of duplex were all the same as water into water injections, within error. From these data, it was concluded that effects of any solvent mismatching are negligible. Titrations were carried out on a VP-ITC instrument (MicroCal LLC, Northampton, MA). For each titration, the top strand of the nonmodified TGT(15) duplex or that containing the single, site-specific adduct of Ru-THA or RuCYM was loaded into the 1.4 ml sample cell, and the complementary oligomer (bottom strand of the duplex TGT(15)) was loaded into the 300 μl injection syringe. The stirring rate of the injection syringe was 400 rpm, and samples were equilibrated thermally prior to a titration until the baseline had leveled off and the rms noise was less than 0.015 $\mu cal\cdot s^{-1}$. A typical titration consisted of 50 injections of 5 μl each,

with 3 min between injections. Data from individual titrations were analyzed by using the Origin 5.0 software package (Origin, Northampton, MA) to extract the relevant thermodynamic parameters (the enthalpy change [ΔH], the entropy change [ΔS], the stoichiometry [n], and the equilibrium constant [K] for strand association).

Inhibition of DNA Polymerization

We investigated DNA polymerization using the templates site-specifically modified by Ru-THA or Ru-CYM by KF. The DNA polymerase I class of enzymes has served as the prototype for studies on structural and biochemical mechanisms of DNA replication [36, 37]. The 23-mer templates containing a single monofunctional adduct of Ru-THA or Ru-CYM were prepared in the same way as described above. The eight-mer DNA primer was complementary to the 3′ terminus of the 23-mer template. The DNA substrates were formed by annealing templates and 5′-end-labeled primers at a molar ratio of 3:1. All experiments were performed at 25°C in a volume of 50 μl in a buffer containing 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 50 μg/ml BSA, 25 μM dATP, 25 μM dCTP, 25 μM dGTP, 25 μM TTP and 0.5 U KF. Reactions were terminated by the addition of EDTA so that its resulting concentration was 20 μM and by heating at 100°C for 30 s. Products were resolved on a denaturing 24% PAA/8 M urea gel and then visualized and quantified by using the FUJIFILM bio-imaging analyzer and AIDA image analyzer software.

Repair Synthesis by Human Cell Extracts

Repair DNA synthesis of CFEs was assayed by using pUC19 and pBR322 plasmids. Each reaction of 50 μl contained 250 ng nonmodified pBR322 and 250 ng nonmodified or platinated pUC19; 2 mM ATP; 30 mM KCl; 0.5 mg/ml creatine phosphokinase (rabbit muscle); 20 mM of each dGMP, dCTP, and TTP; 8 mM dATP; 74 kBq [α-^{32}P]dAMP in the buffer composed of 40 mM HEPES-KOH (pH 7.5), 5 mM $MgCl_2$, 0.5 mM dithiothreitol, 22 mM creatine phosphate, 1.4 mg/ml bovine serum albumin, and 150 μg CFE. Reactions were incubated for 3 hr at 25°C and terminated by adding EDTA to a final concentration of 20 mM, SDS to 0.6%, and proteinase K to 250 μg/ml and then incubating for 30 min. The products were extracted with 1 volume 1:1 phenol:chloroform. The DNA was precipitated from the aqueous layer by the addition of 1/50 volume 5 M NaCl, 5 mg glycogen, and 2.5 volumes ethanol. After 20 min of incubation on dry ice and centrifugation at 12,000 × g for 30 min at 4°C, the pellet was washed with 0.5 ml 70% ethanol and dried in a vacuum centrifuge. DNA was finally linearized before electrophoresis on a 1% agarose gel containing 0.3 mg/ml EtBr. The basic principles of this assay are shown schematically in Supplemental Figure S2A.

Nucleotide Excision Assay

The 149 bp substrates containing a single monofunctional adduct of Ru-THA or Ru-CYM were assembled from three oligonucleotide duplexes. The central duplex was TGT(NER) duplex (shown in Figure 1B) to which two duplexes (arms) with random base pair sequences with overhangs partially overlapping those of the modified duplex were ligated (one to each side) by T4 DNA ligase. The top strand of the modified central duplexes were 5′-end labeled with ^{32}P before ligation. Substrates containing a single, central 1,2-GG intrastrand crosslink of cisplatin were prepared in a similar way to that described previously [38]. Full-length substrates (nonmodified, containing the monofunctional adduct of Ru-THA or Ru-CYM or the 1,2-intrastrand crosslink of cisplatin) were separated from unligated products on a denaturing 6% PAA gel, purified by electroelution, reannealed, and stored in annealing buffer (50 mM Tris-HCl [pH 7.9], 100 mM NaCl, 10 mM $MgCl_2$, and 1 mM dithiothreitol) at 20°C. In vitro repair of monofunctional adducts of Ru(II) arene complexes and of the 1,2-intrastrand crosslink of cisplatin was measured in an excision assay as described previously [38], with minor modifications. The reaction mixtures (25 μl) contained 10 fmol radiolabeled DNA, 50 μg CFE, 20 μM dATP, 20 μM dCTP, 20 μM dGTP, and 20 μM TTP in reaction buffer (23 mM HEPES [pH 7.9], 44 mM KCl, 4.8 mM $MgCl_2$, 0.16 mM EDTA, 0.52 mM dithiothreitol, 1.5 mM ATP, 5 μg bovine serum albumin, and 2.5% glycerol) and were incubated at 30°C for 40 min. DNA was deproteinized and precipitated by ethanol. Reaction products were treated overnight with 0.4 M NaCN (pH 10–11) at 45°C and precipitated by ethanol prior to resolution on the gels. The excision products were separated on denaturing 10% PAA gels and visualized by using the PhosphorImager. The basic principles of this assay are shown schematically in Supplemental Figure S2B.

Supplemental Data

A description of the experiments with chemical probes of DNA conformation, recognition by HMGB1 and RPA proteins, corresponding experimental procedures, and the basic principles of repair DNA synthesis and nucleotide excision repair assays are available at http://www.chembiol.com/cgi/content/full/12/1/121/DC1/.

Acknowledgments

This research was supported by the Grant Agency of the Czech Republic (Grants 305/02/1552), the Grant Agency of the Academy of Sciences of the Czech Republic (Grants B5004301), Wellcome Trust (Grant 073646/Z/03/Z), Edinburgh Technology Fund, and Oncosense Ltd. J.K. is the international research scholar of the Howard Hughes Medical Institute. We are grateful to Dr. Duncan Jodrell (University of Edinburgh) for providing cytotoxicity data and for stimulating discussions and comments on the manuscript. The authors acknowledge that their participation in the European Commission Cooperation in the Field of Scientific and Technical Research Chemistry Action D20 enabled them to exchange regularly the most recent ideas in the field of ruthenium anticancer drugs with several European colleagues.

Received: October 7, 2004
Revised: November 3, 2004
Accepted: November 5, 2004
Published: January 21, 2005

References

1. Keppler, B.K., Lipponer, K.-G., Stenzel, B., and Kratz, F. (1993). New tumor-inhibiting ruthenium complexes. In Metal Complexes in Cancer Chemotherapy, B Keppler, ed. (Weinheim, NY: VCH Verlagsgesellschaft, VCH Publishers), pp. 187–220.
2. Clarke, M.J. (2003). Ruthenium metallopharmaceuticals. Coord. Chem. Rev. *236*, 209–233.
3. Morris, R.E., Aird, R.E., Murdoch, P.D., Chen, H.M., Cummings, J., Hughes, N.D., Parsons, S., Parkin, A., Boyd, G., Jodrell, D.I., et al. (2001). Inhibition of cancer cell growth by ruthenium(II) arene complexes. J. Med. Chem. *44*, 3616–3621.
4. Aird, R., Cummings, J., Ritchie, A., Muir, M., Morris, R., Chen, H., Sadler, P., and Jodrell, D. (2002). In vitro and in vivo activity and cross resistance profiles of novel ruthenium (II) organometallic arene complexes in human ovarian cancer. Br. J. Cancer *86*, 1652–1657.
5. Sava, G., and Bergamo, A. (2000). Ruthenium-based compounds and tumour growth control (Review). Int. J. Oncol. *17*, 353–365.
6. Wang, F., Chen, H., Parsons, S., Oswald, I.D.H., Davidson, J.E., and Sadler, P.J. (2003). Kinetics of aquation and anation of ruthenium(II) arene anticancer complexes, acidity and X-ray structures of aqua adducts. Chem. Eur. J *9*, 5810–5820.
7. Chen, H.M., Parkinson, J.A., Parsons, S., Coxall, R.A., Gould, R.O., and Sadler, P.J. (2002). Organometallic ruthenium(II) diamine anticancer complexes: arene-nucleobase stacking and stereospecific hydrogen-bonding in guanine adducts. J. Am. Chem. Soc. *124*, 3064–3082.
8. Novakova, O., Chen, H., Vrana, O., Rodger, A., Sadler, P.J., and Brabec, V. (2003). DNA interactions of monofunctional organometallic ruthenium(II) antitumor complexes in cell-free media. Biochemistry *42*, 11544–11554.
9. Nielsen, P.E. (1990). Chemical and photochemical probing of DNA complexes. J. Mol. Recognit. *3*, 1–24.
10. Brabec, V., Sip, M., and Leng, M. (1993). DNA conformational

distortion produced by site-specific interstrand cross-link of trans-diamminedichloroplatinum(II). Biochemistry *32*, 11676–11681.

11. Pilch, D.S., Dunham, S.U., Jamieson, E.R., Lippard, S.J., and Breslauer, K.J. (2000). DNA sequence context modulates the impact of a cisplatin 1,2-d(GpG) intrastrand cross-link an the conformational and thermodynamic properties of duplex DNA. J. Mol. Biol. *296*, 803–812.
12. Hofr, C., Farrell, N., and Brabec, V. (2001). Thermodynamic properties of duplex DNA containing a site-specific d(GpG) intrastrand crosslink formed by an antitumor dinuclear platinum complex. Nucleic Acids Res. *29*, 2034–2040.
13. Malina, J., Hofr, C., Maresca, L., Natile, G., and Brabec, V. (2000). DNA interactions of antitumor cisplatin analogs containing enantiomeric amine ligands. Biophys. J. *78*, 2008–2021.
14. Holbrook, J.A., Capp, M.W., Saecker, R.M., and Record, M.T. (1999). Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. Biochemistry *38*, 8409–8422.
15. Heiger-Bernays, W.J., Essigmann, J.M., and Lippard, S.J. (1990). Effect of the antitumor drug cis-diamminedichloroplatinum(II) and related platinum complexes on eukaryotic DNA replication. Biochemistry *29*, 8461–8466.
16. Kasparkova, J., Novakova, O., Farrell, N., and Brabec, V. (2003). DNA binding by antitumor trans-[$PtCl_2(NH_3)$(thiazole)]. Protein recognition and nucleotide excision repair of monofunctional adducts. Biochemistry *42*, 792–800.
17. Zamble, D.B., Mu, D., Reardon, J.T., Sancar, A., and Lippard, S.J. (1996). Repair of cisplatin-DNA adducts by the mammalian excision nuclease. Biochemistry *35*, 10004–10013.
18. Reardon, J.T., Vaisman, A., Chaney, S.G., and Sancar, A. (1999). Efficient nucleotide excision repair of cisplatin, oxaliplatin, and bis-aceto-ammine-dichloro-cyclohexylamine-platinum(IV) (JM216) platinum intrastrand DNA diadducts. Cancer Res. *59*, 3968–3971.
19. Kasparkova, J., Novakova, O., Marini, V., Najajreh, Y., Gibson, D., Perez, J.-M., and Brabec, V. (2003). Activation of trans geometry in bifunctional mononuclear platinum complexes by a piperidine ligand: mechanistic studies on antitumor action. J. Biol. Chem. *278*, 47516–47525.
20. Zehnulova, J., Kasparkova, J., Farrell, N., and Brabec, V. (2001). Conformation, recognition by high mobility group domain proteins, and nucleotide excision repair of DNA intrastrand crosslinks of novel antitumor trinuclear platinum complex BBR3464. J. Biol. Chem. *276*, 22191–22199.
21. Cohen, S.M., and Lippard, S.J. (2001). Cisplatin: from DNA damage to cancer chemotherapy. In Progress in Nucleic Acid Research and Molecular Biology, Volume 67, K Moldave, ed. (San Diego: Academic Press, Inc.), pp. 93–130.
22. Brabec, V. (2002). DNA modifications by antitumor platinum and ruthenium compounds: their recognition and repair. In In Progress in Nucleic Acid Research and Molecular Biology, Volume 71, K Moldave, ed. (San Diego: Academic Press, Inc.), pp. 1–68.
23. Plum, G.E., Gelfand, C.A., and Breslauer, K.J. (1999). Physicochemical approaches to structural elucidation. Effects of 3,N4-ethenodeoxycytidine on duplex stability and energetics. In Exocyclic DNA Adducts in Mutagenesis and Carcinogenesis, Publication No. *150*, B. Singer, and H. Bartsch, eds. (Lyon: International Agency for Research on Cancer), pp. 169–177.
24. Maeda, Y., Nunomura, K., and Ohtsubo, E. (1990). Differential scanning calorimetric study of the effect of intercalators and other kinds of DNA-binding drugs on the stepwise melting of plasmid DNA. J. Mol. Biol. *215*, 321–329.
25. Bjorndal, M.T., and Fygenson, D.K. (2002). DNA melting in the presence of fluorescent intercalating oxazole yellow dyes measured with a gel-based assay. Biopolymers *65*, 40–44.
26. Jamieson, E.R., and Lippard, S.J. (1999). Structure, recognition, and processing of cisplatin-DNA adducts. Chem. Rev. *99*, 2467–2498.
27. Ohndorf, U.M., Rould, M.A., He, Q., Pabo, C.O., and Lippard, S.J. (1999). Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. Nature *399*, 708–712.
28. Brabec, V., and Kasparkova, J. (2002). Molecular aspects of resistance to antitumor platinum drugs. Drug Resist. Updat. *5*, 147–161.
29. Kasparkova, J., Zehnulova, J., Farrell, N., and Brabec, V. (2002). DNA interstrand cross-links of the novel antitumor trinuclear platinum complex BBR3464. Conformation, recognition by high mobility group domain proteins, and nucleotide excision repair. J. Biol. Chem. *277*, 48076–48086.
30. Furuta, T., Ueda, T., Aune, G., Sarasin, A., Kraemer, K.H., and Pommier, Y. (2002). Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. Cancer Res. *62*, 4899–4902.
31. Wang, D., Hara, R., Singh, G., Sancar, A., and Lippard, S.J. (2003). Nucleotide excision repair from site-specifically platinum-modified nucleosomes. Biochemistry *42*, 6747–6753.
32. Selvakumaran, M., Pisarcik, D.A., Bao, R., Yeung, A.T., and Hamilton, T.C. (2003). Enhanced cisplatin cytotoxicity by disturbing the nucleotide excision repair pathway in ovarian cancer cell lines. Cancer Res. *63*, 1311–1316.
33. Brabec, V., Reedijk, J., and Leng, M. (1992). Sequence-dependent distortions induced in DNA by monofunctional platinum(II) binding. Biochemistry *31*, 12397–12402.
34. Brabec, V., and Leng, M. (1993). DNA interstrand cross-links of trans-diamminedichloroplatinum(II) are preferentially formed between guanine and complementary cytosine residues. Proc. Natl. Acad. Sci. USA *90*, 5345–5349.
35. Murphy, J.H., and Trapane, T.L. (1996). Concentration and extinction coefficient determination for oligonucleotides and analogs using a general phosphate analysis. Anal. Biochem. *240*, 273–282.
36. Lam, W.C., Van der Schans, E.J.C., Sowers, L.C., and Millar, D.P. (1999). Interaction of DNA polymerase I (Klenow fragment) with DNA substrates containing extrahelical bases: implications for proofreading of frameshift errors during DNA synthesis. Biochemistry *38*, 2661–2668.
37. Patel, P.H., Suzuki, M., Adman, E., Shinkai, A., and Loeb, L.A. (2001). Prokaryotic DNA polymerase I: evolution, structure, and 'base flipping' mechanism for nucleotide selection. J. Mol. Biol. *308*, 823–837.
38. Malina, J., Kasparkova, J., Natile, G., and Brabec, V. (2002). Recognition of major DNA adducts of enantiomeric cisplatin analogs by HMG box proteins and nucleotide excision repair of these adducts. Chem. Biol. *9*, 629–638.

Ctirad Hofr
Viktor Brabec
Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-61265 Brno, Czech Republic

Thermal Stability and Energetics of 15-mer DNA Duplex Interstrand Crosslinked by *trans*-Diamminedichloroplatinum(II)

Received 11 October 2004; revised 17 November 2004; accepted 23 November 2004
Published online 21 January 2005 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20216

Abstract: *The effect of the location of the interstrand cross-link formed by* trans-*diamminedichloroplatinum(II) (transplatin) on the thermal stability and energetics of 15-mer DNA duplex has been investigated. The duplex containing single, site-specific cross-link, thermodynamically equivalent model structures (hairpins) and nonmodified duplexes were characterized by differential scanning calorimetry, temperature-dependent uv absorption, and circular dichroism. The results demonstrate that the formation of the interstrand cross-link of transplatin does not affect pronouncedly thermodynamic stability of DNA: the cross-link induces no marked changes not only in enthalpy, but also in "reduced" (concentration independent) monomolecular transition entropy. These results are consistent with the previous observations that interstrand cross-links of transplatin structurally perturb DNA only to a relatively small extent. On the other hand, constraining the duplex with the interstrand cross-link of transplatin results in a significant increase in thermal stability that is primarily due to entropic effects: the cross-link reduces the molecularity of the oligomer system from bimolecular to monomolecular. Importantly, the position of the interstrand cross-link within the duplex modulates cooperativity of the melting transition of the duplex and consequently its thermal stability.* © 2005 Wiley Periodicals, Inc. Biopolymers 77: 222–229, 2005

Keywords: *DNA; interstrand cross-link; transplatin; calorimetry; van't Hoff enthalpy; circular dichroism*

INTRODUCTION

Reaction of cellular DNA with environmental and chemotherapeutic agents can give rise to a variety of lesions, including interstrand cross-links (CLs). Because these lesions can prevent DNA strand separation interfering with critical cellular events, such as DNA transcription and replication, they represent a serious impediment to cell survival. Thus, the mechanism of the action of interstrand cross-linking agents and the design of new types of cross-linking compounds have been extensively investigated.

The interstrand CLs not only prevent separation of the strands of DNA, but also may alter its conformation. Both of these factors may affect in a fundamental way processing of these lesions in cells. Efforts to reveal the details of the mechanisms by which interstrand cross-links are processed have been hampered by the difficulty to distinguish between the effects of covalent linking of two complementary strands of DNA itself and resulting conformational alterations.

Correspondence to: V. Brabec; email: brabec@ibp.cz

In order to contribute to understanding the biological effects of interstrand CLs, we have analyzed in our previous work[1] the short oligodeoxyribonucleotide duplex containing single and central, site-specific interstrand CL of transplatin [*trans*-diamminedichloroplatinum(II)] by differential scanning calorimetry (DSC) and circular dichrosim (CD) spectropolarimetry. This CL is preferentially formed in DNA between complementary cytosine and guanine residues[2,3] and, in contrast to the interstrand CL formed in DNA by its *cis* isomer, affects DNA conformation to a very small extent.[4,5] Thus, the interstrand CL of this simple and purely inorganic compound (it is composed of only 11 atoms, 6 of which are hydrogens) is a suitable model to investigate the effect of interstrand cross-linking not markedly affected by the conformational alterations imposed on DNA by this lesion.

The interstrand CL of transplatin increased the thermal stability of the modified duplexes by changing the molecularity of denaturation. Other results were interpreted to mean that the interstrand CL of transplatin was proposed to be enthalpically almost neutral. The CL-induced thermodynamic destabilization was found to be entirely entropic in origin.

The conclusions of our previous work[1] were based on the extraction of thermodynamic data from equilibrium transition curves on short duplexes of different molecularity. The unconstrained (nonplatinated) duplex denatured in a bimolecular reaction to form two single strands whereas the interstrand crosslinked duplex denatured in monomolecular reaction since the crosslinked strands could not separate. To dissect the effect of different molecularities of the nonplatinated and interstrand crosslinked oligomer systems from other mechanisms that could affect thermal stability of the crosslinked duplexes, we performed in our previous work[1] a correction for the concentration dependence using a theoretical approach.[6]

In order to understand the effects of interstrand cross-linking by transplatin on thermal stability and energetics of DNA more deeply and to verify our previous conclusions,[1] we used in the present work an empirical approach to eliminate the effect of different molecularities of the nonplatinated and interstrand crosslinked oligomer systems. We used DSC to measure thermal transition of the duplexes (nonplatinated or containing a site-specific interstrand CL of transplatin) and corresponding monomolecular structures: hairpin duplexes. These duplexes also contained a short, single-stranded loop preventing a change of molecularity of the system if the noncrosslinked duplexes were thermally denatured. In addition, when analyzing short, interstrand crosslinked duplexes, the results might be also affected by the distance of the CL from the ends of the duplex. This is because the single interstrand CL divides the duplex into two fractions separated by the CL that may melt independently and cooperatively. These fractions may differ in the length depending on the position of the CL within the duplex, which may affect the thermodynamic data associated with the formation of the interstrand CL. As this aspect of DNA cross-linking has not been examined yet, we also examined in the present work a short duplex containing single, asymmetrically positioned interstrand CL of transplatin and compared these results with those shown in the previous report[1] describing the results obtained with the CL of transplatin positioned centrally.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were purchased from IDT, Inc. (Coralville, IA). Molar extinction coefficients of the oligonucleotides were determined by using standard phosphate analysis.[7] The following molar extinction coefficients at 260 nm and 25°C (in units of mole strand·liter^{-1}·cm^{-1}) were obtained: 114,000 for upper strand S1; 139,000 for strand S2; 272,000 for hairpin SH. Isothermal mixing experiments[8] using strands S1 and S2 revealed 1:1 stoichiometries for both nonmodified duplexes, a ratio consistent with duplex formation. Dimethyl sulfate (DMS) used for platinum footprinting on DNA was from Sigma.

Platination of Oligonucleotides

The oligonucleotides were modified by transplatin, which was purchased from Sigma. The single-stranded oligonucleotide S1 (Figure 1) at a concentration of 150 μM was reacted with a monoaquamonochloro derivative of transplatin generated by allowing this complex to react with 0.9 *M* equivalent of $AgNO_3$ at an input of platinum to strand molar ratio of 3.9:1 in deionized water at 37°C. After incubating the mixture for 15 min, the concentration of NaCl was adjusted to 0.1 *M*, and the platinated oligonucleotides were purified by ion-exchange high-pressure liquid chromatography (HPLC) with linear 0.2 to 0.6 *M* NaCl gradient in 50 mM Tris (pH 7.4). It was verified by platinum flameless atomic absorption spectrophotometry (FAAS) and by the measurements of the optical density that the modified oligonucleotide S1 contained one or two platinum atoms, respectively. It was also verified using DMS footprinting of platinum on DNA[2] that, in the platinated strand S1, the N7 position of the G residue was not accessible for reaction with DMS, which implies that this G residue was platinated. The platinated strand was allowed to anneal with nonplatinated complementary strand in 0.4 *M* NaCl (pH 7.4) at 25°C for 24 h, precipitated by ethanol, dissolved in 0.1 *M* Na-

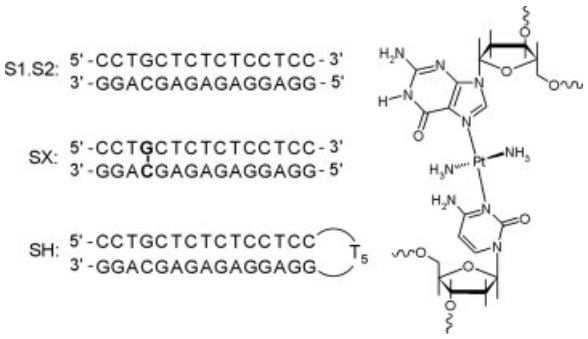


FIGURE 1 Structures of the synthetic oligodeoxyribonucleotides used in the present study together with their abbreviations. The bold letters connected by the line indicate the location of the interstrand CL of transplatin. The structure of the interstrand CL formed by transplatin between complementary guanine and cytosine residues is shown on the right.

ClO_4, and incubated for 18 h in the dark at 37°C. The resulting products were purified by HPLC in an alkaline gradient. Using this denaturing gradient, noninterstrand crosslinked strands were eluted as a 15-nucleotide single strand, whereas the interstrand crosslinked strand was eluted later in a single peak as a higher-molecular-mass species. The crosslinked oligomer (0.1 A_{260} unit) was cleaved by using PI nuclease and then treated with alkaline phosphatase (5 units) in the buffer containing 100 mM Tris (pH 9) and 15 mM $MgCl_2$. It was verified by gel electrophoresis that digestion to nucleosides was complete. The resulting mixture of nucleosides was analyzed by reversed phase HPLC with uv detection and with a gradient of acetonitrile in ammonium acetate at 40°C.[9] The duplex SX was still analyzed for platinum contents by FAAS. Additional quantitation of the crosslinked duplex by uv absorption spectrophotometry was used to ascertain that the ratio of platinum adduct per duplex SX was 1. The position of the platinated site in the bottom strand of this duplex was determined by hydroxyl radical footprinting[10] and revealed that the cytosine residues complementary to the platinated G residues in the top strand were the sites involved in the interstrand CL.

HPLC purification was carried out on a Merck-Hitachi Lachrom HPLC system with an anion exchange Nucleogel SAX 1000–8 column (Macherey Nagel) or reverse phase LC-318 column (Supelco). FAAS measurement was carried out on a Unicam 939 AA spectrometer equipped with a graphite furnace.

Differential Scanning Calorimetry

Excess heat capacity ($\Delta C_{p,xs}$) versus temperature profiles for the thermally induced transitions of nonmodified duplexes or duplexes containing interstrand CL of transplatin were measured by using VP Differential Scanning Calorimeter (Microcal, Northampton, MA). In the DSC experiments, the concentration of the duplexes was 15 μM, the heating rate was 1°C/min, and a maximum temperature was 100°C. After reaching the maximum temperature the samples were cooled at the same rate to the starting temperature of 25°C. In the present paper, $\Delta C_{p,xs}$ is defined as excess heat capacity that is baseline subtracted and concentration normalized.[11] The reference scans were subtracted from the sample scans to obtain $\Delta C_{p,xs}$ versus temperature profiles. Enthalpies (ΔH) and entropies (ΔS) of duplex formation were calculated from the areas under the experimental $\Delta C_{p,xs}$ versus T and the derived $\Delta C_{p,xs}/T$ versus T curves, respectively, by using ORIGIN version 5.0 software (Microcal, Northampton, MA). In order to minimize possible error arising from dependence of enthalpy change on temperature in the case of nonmodified duplexes and to simulate physiological conditions, the free energy of duplex formation was determined for the temperature of 37°C. Free energy for each duplex or hairpin transition (ΔG_{37}) was calculated by using the standard thermodynamic relationship given in Eq. (1) and the corresponding values of ΔH and ΔS.

$$\Delta G_{37} = \Delta H - (310.15)\Delta S \quad (1)$$

The oligonucleotide duplexes were dialyzed against the buffer containing 10 mM sodium cacodylate (pH 7.2), 0.1 M NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA. The samples were vacuum degassed before the measurement. It was verified by comparing repeated heating scans of samples that the melting transition of both the platinated and nonmodified duplexes was fully reversible.

UV Absorption Spectrophotometry

UV absorbance measurements were conducted on a Beckman DU-7400 spectrophotometer equipped with a thermoelectrically controlled cell holder and quartz cells with the pathlength of 1 or 0.1 cm. Absorbance versus temperature profiles were measured at 260 nm. The temperature was raised using a linear heating rate of 1.0°C/min. For each optically detected transition, the melting temperature (t_m) was determined as previously described.[6] The final absorbance versus temperature profile of each duplex (Figure 2b) was determined as an average of six independent measurements. The DNA solutions contained 10 mM sodium cacodylate (pH 7.2), 0.1 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA.

Circular Dichroism (CD) Spectrophotometry

CD spectra were recorded using a Jasco J-720 spectropolarimeter equipped with a thermoelectrically controlled cell holder. The cell pathlength was 1 cm. Isothermal CD spectra were recorded from 220 to 320 nm in 1-nm increments with an averaging time of 5 s. The DNA concentration was 7 μM in duplex and buffer conditions were 10 mM sodium cacodylate (pH 7.2), 0.1 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA.

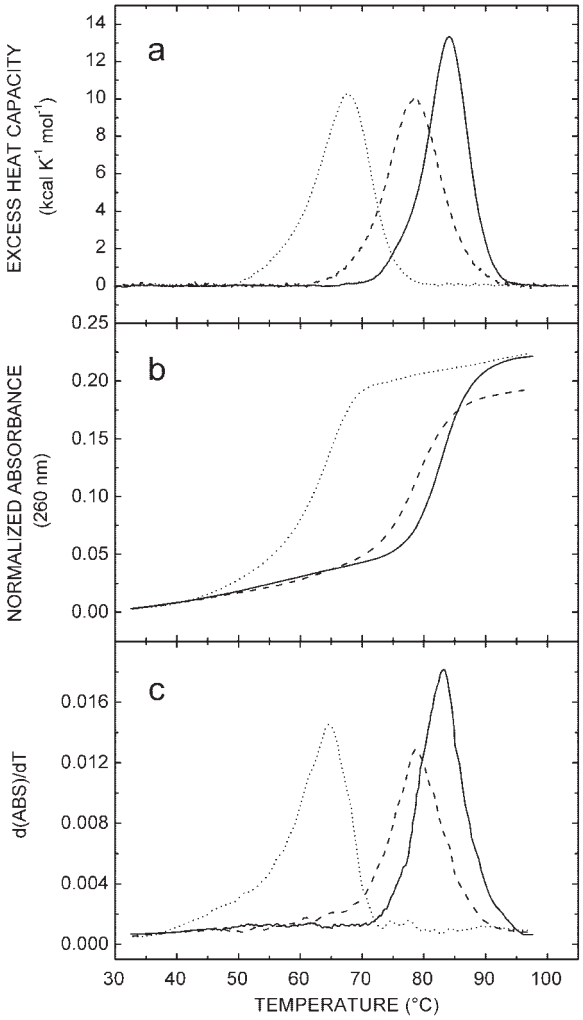


FIGURE 2 Differential scanning calorimetry and temperature-dependent uv spectroscopy data. (a) DSC thermograms for the nonmodified duplexes S1.S2 (dot); the crosslinked duplexes SX (dash), and hairpin SH (solid). The duplex concentration was 15 μM, and the buffer conditions were 10 mM sodium cacodylate (pH 7.2), 0.1 M NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA. (b) uv melting profiles at 260 nm for the nonmodified duplexes, the hairpin, and the crosslinked duplexes at concentration 3.5 μM. The melting curves were normalized for clarity. (c) Derivation of uv melting curves. Lines corresponding to duplexes and hairpin in b and c are same as described in a.

RESULTS

Melting of Duplexes with Free Ends

DSC and temperature-dependent uv absorption measurements were conducted to characterize the thermally induced melting of the 15-bp DNA duplexes with the specific goal of elucidating the thermal and thermodynamic consequences of constraining DNA via a site-specific interstrand cross-linking by transplatin. One CL between complementary cytosine and guanine residues was asymmetrically positioned in the duplex S1.S2 forming the duplex SX (the platinated G.C pair was separated from the ends of the duplex SX by 3 and 11 bp). The results of the calorimetric studies along with uv temperature denaturations of the crosslinked duplex SX are shown in Figure 2 with associated data listed in Table I. Denaturation (heating) curves and renaturation (cooling) curves for the nonmodified and platinated duplexes were superimposable (not shown), which is consistent with the reversibility of the melting equilibrium. Comparison of the t_m values of the crosslinked and corresponding nonconstrained (nonplatinated) duplexes determined calorimetrically revealed that a single (asymmetrical) transplatin CL resulted in a marked increase in thermal stability of the duplex S1.S2. The value of Δt_m (defined as the difference between t_m values of the crosslinked and nonplatinated duplex) was 11.1°C. In addition, the hyperchromism associated with the melting of the duplexes was reduced as a consequence of the cross-linking (Figure 2b).

The nonconstrained short duplexes denature in a bimolecular reaction to form two single strands. As a consequence, melting of the nonmodified duplexes S1.S2 was dependent on the overall oligonucleotide concentration. For instance, if the concentration of the nonplatinated duplex S1.S2 increased from 3.5 to 15 μM, the value of t_m of the duplex increased from 64.6°C (Figure 2c) to 67.7°C (Figure 2a). In contrast, the interstrand crosslinked duplexes SX melted in a concentration-independent manner to a denaturated state, which is consistent with monomolecular behavior of the oligonucleotide system. The molecularity of the transition of platinated oligonucleotide system

Table I Thermodynamic Parameters Obtained from Calorimetric Experiments[a]

Duplex	t_m[b]	ΔH[c]	ΔS[c]
S1.S2	67.7	−119	−352
SX	78.8	−116	−329
SH	82.1	−118	−331

[a] The t_m, ΔH, and ΔS values are averages derived from three independent measurements.

[b] t_m denotes the temperature in °C corresponding to the maximum in the DSC melting profiles shown in Figure 2 and error is ±0.2°C.

[c] ΔH and ΔS denote, respectively, the transition enthalpy in kcal/mol and transition entropy in cal/(K mol) of duplex or hairpin formation; the relative error is ±3%.

was one, because the covalent interstrand CL did not allow separation of the strands. Thus, the observed Δt_m values are apparently affected by the fact that the original bimolecular oligomer system became monomolecular as a consequence of the cross-linking.

Melting of Hairpin Duplexes

Upon introduction of a single transplatin CL, the change in t_m can result not only from the change in the molecularity of the system, but also from a different mechanism of melting transition. We tried to dissect these two origins of the observed crosslinked induced shift in thermal stability. If the observed change in t_m is due mainly to the molecularity of the system, then one might expect mainly to observe changes in entropy. In accord with this assumption, we observe, as a consequence of the formation of the interstrand CL of transplatin in the S1.S2 duplex, only small changes in enthalpy in addition to entropy changes (Table I). To estimate how many of the observed Δt_m differences result from the change in molecularity, we used, in our recent work[1] (when we analyzed the 20-mer duplex containing the central interstrand CL of transplatin), "theoretical" corrections.[6] In the present work, we applied another ("experimental") approach based on examination of thermal transition of the hairpin SH (Figure 1). The stem duplex of the hairpin structure SH contained two complementary nucleotide sequences identical to those of the top and bottom strands of the duplex S1.S2, which were still linked by a short single-stranded loop composed of five thymine residues. The thymine residues in this short loop behave as denatured single strands and should contribute little to the transition enthalpy of an adjacent stem duplex.[12] This assumption is fully justified since the values of the transition enthalpy for the nonmodified duplex S1.S2 and the hairpin SH were very close so that the differences in the thermodynamic stability of these nonplatinated structures S1.S2 and SH were only due to the differences in the changes in the transition entropy (Table I). Thus, the effect of the T5 loop was only to prevent a separation of the two complementary sequences of the hairpin structure SH during its thermal melting and had no effect on the conformation of its double-stranded stem, which had a nucleotide sequence identical to that of the duplex S1.S2. Hence, it is reasonable to expect that the hairpin SH melts in the same way as the duplex S1.S2, but in a monomolecular, concentration-independent reaction. In this way, if Δt_m is calculated as the difference between the values of t_m for the interstrand crosslinked duplex SX and the nonplatinated hairpin SH, one "corrects" for the concentration dependence of the t_m of the duplex S1.S2 and obtains the change in t_m corresponding to effects other than the change in molecularity of the system.

Examination of the thermal melting of the nonmodified hairpin SH revealed that its melting temperature was 82.1°C. Hence, this value can be taken in the first approximation as a "reduced" concentration-independent melting temperature for the nonmodified duplex S1.S2. This "reduced" t_m value is quite close to the t_m value observed for the (mono)crosslinked counterpart SX, which suggests that the overall impact of a single asymmetrically positioned transplatin CL is mainly due to the change in molecularity of the oligomer system.

The transition entropy for a nonmodified duplex S1.S2, which melts in a bimolecular manner, also depends on the strand concentration. To correct for this concentration dependence, we used again the value of transition entropy obtained for the hairpin SH. Comparison of the values of transition entropy for the duplex SX (crosslinked by transplatin) and for the hairpin SH revealed that the transition entropy for the duplex SX was by only 2 cal/(K mol) higher. In other words, formation of the CL covalently linked complementary strands of the duplex S1.S2 and in this way entropically increased its stability only negligibly. The formation of the CL by transplatin also slightly decreased transition enthalpy and consequently decreased stability of the duplex by 3 $kcal \cdot mol^{-1}$. Hence, the interstrand CLs of transplatin enthalpically destabilized the duplex and this small destabilization was almost completely compensated by a CL-induced entropic stabilization of the duplex.

The net result of enthalpic and entropic effects is that formation of the interstrand CL in the duplex S1.S2 by transplatin at 37°C reduced its overall stability ΔG_{37} by 2.4 kcal/mol in comparison with stability of the hairpin SH as a model of a comparable monomolecular oligomer system ($\Delta G_{37} = \Delta H - 310.15\ \Delta S_{red}$, where ΔS_{red} is ΔS of the hairpin SH). Thus, formation of the interstrand CL in DNA by transplatin slightly decreased the thermodynamic stability of DNA and this decrease was found to be enthalpic in origin.

Van't Hoff Transition Enthalpy

Further insight into the nature of a thermal transition of DNA duplexes interstrand crosslinked by transplatin provides a quantitative comparison of the model-independent calorimetric and model-dependent van't Hoff transition enthalpy determined independently. This comparison allows one to conclude whether the transition proceeds in a two state (all-or-none) manner

Table II Calorimetric (Model-Independent) and van't Hoff Enthalpies for Transition of the Oligomer Systems Nonplatinated or Containing a Single, Site-Specific Interstrand Cross-Link of Transplatin

Duplex	ΔH_{cal}[a]	ΔH_{vH}[b]	$\Delta H_{vH}/\Delta H_{cal}$
S1.S2	−119	−117	0.98
SX	−116	−87	0.75
SH	−118	−118	1.00
SX_{20}[c]	−146	−78	0.53

[a] Model-independent calorimetric enthalpies ΔH_{cal} are in kcal · mol^{-1}, the values were derived from DSC experiments as described in the text, and the relative error is ±3%.

[b] van't Hoff enthalpies ΔH_{vH} are in kcal · mol^{-1} and the values were obtained by analyzing the shapes of each calorimetric curve using the approach described previously,[6] and the relative error is ±5%.

[c] Data obtained for the 20-bp duplex containing a single, central interstrand CL were taken from Hoff and Brabec.[1]

or whether the transition includes intermediate states that are responsible for broadening the thermograms.[6,13] The broadening of equilibrium melting curves is associated with a reduced van't Hoff transition enthalpy (ΔH_{vH}) relative to the calorimetric value (ΔH_{cal}). The ratio of the ΔH_{vH} and the ΔH_{cal} provides a measure of the fraction of the structure that melts as a single thermodynamic entity, i.e., may afford the size of the largest cooperative unit.[6] Table II lists the directly measured, model-independent calorimetric transition enthalpies and the indirectly derived, model-dependent van't Hoff transition enthalpies. The ΔH_{vH} values were obtained by analyzing the shapes of each calorimetric curve using the approach described earlier.[6] A comparison of the ΔH_{vH} and ΔH_{cal} data listed in Table II reveals that the van't Hoff values for the nonplatinated duplexes are identical within the experimental uncertainty to the corresponding model-independent calorimetric values. This result is consistent with all-or-none, two-state melting behavior of the nonconstrained duplex. On the other hand, the van't Hoff value was considerably smaller than the corresponding model-independent calorimetric data for the duplex interstrand crosslinked by transplatin (Table II). This disparity demonstrates that denaturation of the duplexes interstrand crosslinked by transplatin does not proceed via a two-state process or, in other words, that this cross-linking alters the ability of the duplexes to propagate the interactions required for cooperative melting.

Circular Dichroism

The CD spectra of the duplex SX containing a single, asymmetrically positioned CL and its nonmodified counterpart at 37 and 95°C were also measured to study the macroscopic helical geometry of nonmodified and crosslinked duplexes and to monitor difference in overall structure between these duplexes. At 37°C the perturbation of B-conformation of DNA due to the formation of the interstrand CL results in the red shift of the positive band of DNA at 274 nm and a decrease in intensity of the negative band at 238 nm (not shown). On the other hand, the CD signals of the high-temperature denatured states (recorded at 95°C) of the nonmodified duplex and that containing one CL of transplatin agree within the noise of the measurement, which implies that the final denatured states are not structurally and thermodynamically markedly different. Thus, validity of the assumption that the final single-stranded states of the unmodified and platinated duplexes were thermodynamically equivalent at the elevated temperatures was also verified for the duplexes investigated in the present work. In aggregate, meaningful thermodynamic data from our calorimetric measurements described in the present work could be obtained.

DISCUSSION

Interstrand Cross-Links of Transplatin Do Not Affect Thermodynamic Stability of DNA

The major goal of this work was to identify new aspects of the thermodynamic impact of interstrand cross-linking by transplatin on DNA. Increase of thermal stability of the host DNA duplex containing a single interstrand CL of transplatin located in the central part of a 20-bp duplex due to the change in molecularity of the oligomer system has been already reported in our recent study.[1] In addition, the previous work has also demonstrated that the interstrand CL of transplatin appears nearly enthalpically neutral. The results reported in the present work describing the effect of asymmetrically positioned CL are consistent with these findings. However, the results of our previous work were also interpreted to mean that the CL of transplatin entropically destabilized the host 20-bp duplex by 15 cal/(K·mol). The small enthalpic effect was overcome by entropic effects; the extent of which was estimated on the base of the theoretical correction of entropy change for monomolecular transition. On aggregate, the net result of the previous analysis was that the formation of interstrand CL of transplatin induced a decrease in duplex thermodynamic stability with this destabilization being entropic in origin. On the other hand, the results of the present work dem-

onstrate that the formation of the interstrand CL of transplatin induces no marked changes not only in enthalpy, but also in "experimentally reduced" (corrected) concentration-independent transition entropy (Table I). Thus, the results of the present work are consistent with the view and support the hypothesis that interstrand CLs of transplatin structurally perturb DNA in the way that does not lead to the pronounced thermodynamic destabilization. On the other hand, constraining the duplex with the interstrand CL of transplatin results in a significant increase in thermal stability that is primarily due to the changes in the molecularity of the oligomer system from bimolecular to monomolecular.

Only a relatively slight effect of the interstrand CL of transplatin (Table I) on the thermodynamic properties of DNA duplexes correlates with existing structural data.[4,5] The basic characteristics of the distortions induced in DNA by interstrand CLs of transplatin involve only a weak distortion on both sides of the crosslinked base pair. The crosslinked bases remain paired and hydrogen bonded. The CL induces only a slight flexible bending of its axis toward the minor groove and local unwinding by approximately 12°. The deoxyriboguanosine residue involved in the CL adopts *syn* conformation that produces rearrangement to the Hoogsteen type base pairing between NH_2 of the complementary cytosine residue and oxygen of the guanine residue. Hence, the results of the present work also suggest that following the "experimental" approach based on the use of hairpin duplexes to correct for the change in molecularity as a consequence of the formation of interstrand CL in DNA may afford results that better reflect the properties of the interstrand crosslinked duplexes than the theoretical calculations used in our previous work.[1]

The Position of the Interstrand Cross-Link within the Duplex Modulates Its Thermal Stability

From the ratio of the model-dependent van't Hoff and the model-independent calorimetric transition enthalpies (Table II), one can define the fraction of a duplex that undergoes transition as a single thermodynamic unity.[1,6,13,14] The $\Delta H_{vH}/\Delta H_{cal}$ ratio found for the 15-bp duplex SX containing one asymmetric interstrand CL of transplatin was 0.75. This value indicates that the length of the largest cooperative unit, which melts in an all-or-none manner, involves 75% of the duplex SX, i.e., its part that is 11 bp long. This length corresponds very accurately to the length of the longer part of the duplex SX, if it is divided into two parts by the interstrand CL as shown in Figure 1. Consistent with this is the observation that the values of t_m and the model-dependent van't Hoff transition enthalpies ΔH_{vH} for transition of the hairpin CH and (mono)crosslinked duplex SX are close (Table II). This indicates that one important factor affecting the melting of the whole interstrand crosslinked duplex is the value of t_m of the largest part of the interstrand crosslinked duplex, which melts cooperatively in a concentration- independent monomolecular manner. Importantly, if the single interstrand CL of transplatin was formed approximately in the middle of the 20-bp duplex, the $\Delta H_{vH}/\Delta H_{cal}$ ratio was close to 0.5 (see data in Table II for the duplex SX_{20}).[1] This result has been interpreted to mean that this duplex also contained two units melting cooperatively, the size of which was close to one-half of the duplex. The schematic representation of the melting of the duplexes interstrand crosslinked by transplatin proposed on the basis of the results of the present work is shown in Figure 3.

CONCLUSION

Irrespective of these possible microscopic interpretations of our macroscopic data, the calorimetric and spectroscopic results reported here reveal that, in short oligomer duplexes, the position of the interstrand CL affects the length of the unit, which melts cooperatively. In addition, the formation of the interstrand CL by transplatin induces a substantial thermal stabilization of the short host duplexes, which is mainly associated with the reduction in molecularity of oligomer system. The bimolecular melting transition of short duplexes, such as S1.S2 (Figure 1) formed by hybridization of two short complementary single strands does not resemble the local monomolecular melting within a much longer double-helical molecule of natural nucleic acids. The melting of short oligonucleotide duplexes yields two single strands because the ends are free. On the other hand, the local melting within a long molecule of natural nucleic acid yields an interior "denaturation bubble," the ends of which are constrained. Hence, the results most relevant to the effects of interstrand cross-linking by transplatin on thermal and thermodynamic stability of nucleic acids in vivo might be those that were corrected for the concentration dependence of melting temperature and transition entropy. These corrected (reduced) values are quite close to the values observed for the corresponding nonmodified hairpin duplexes in accord with only subtle conformational distortions induced by transplatin interstrand CLs.[4,5]

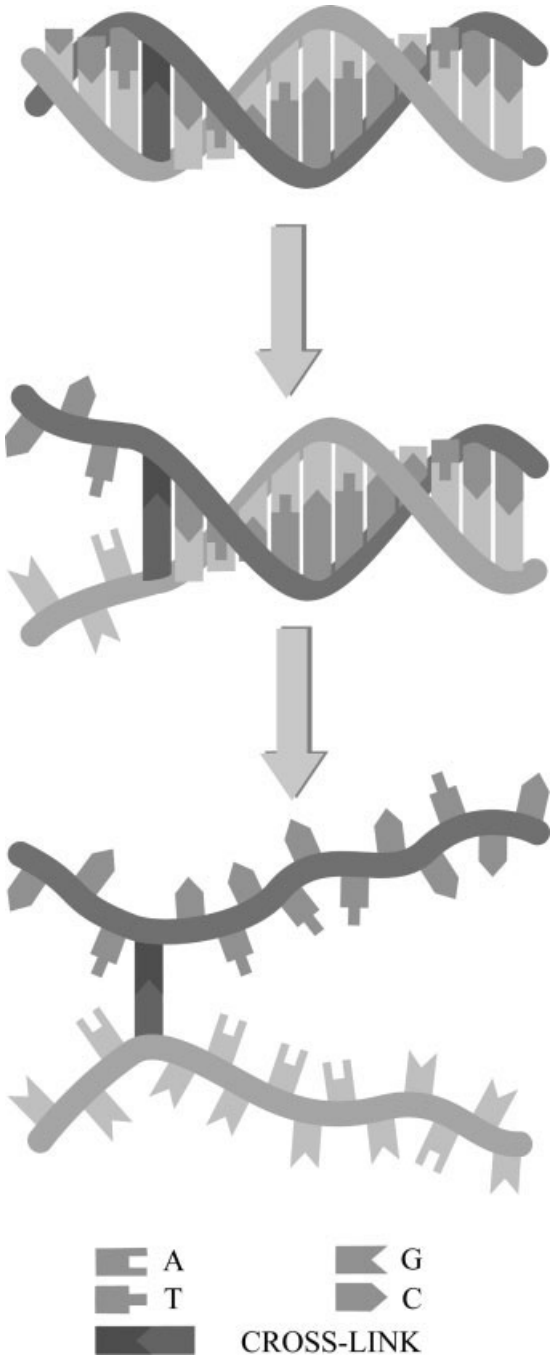


FIGURE 3 Schematic representation of thermal melting of the duplexes containing one assymetric interstrand CL of transplatin. The cross-link divides the duplex into two parts of different length, each melting cooperatively and independently on each other.

This research was supported by the Grant Agency of the Czech Republic (Grants 202/01/D110 and 305/02/1552). The research of CH was also supported by the Ministry of Education, Youth, and Sports of the Czech Republic (Project No. 1K03010). The authors acknowledge that their participation in the EC COST Chemistry Action D20 enabled them to exchange regularly the most recent ideas in the field of platinum anticancer drugs with several European colleagues.

REFERENCES

1. Hofr, C.; Brabec, V. J Biol Chem 2001, 276, 9655–9661.
2. Brabec, V.; Leng, M. Proc Natl Acad Sci U S A 1993, 90, 5345–5349.
3. Brabec, V. In Platinum-Based Drugs in Cancer Therapy; Kelland, L. R.; Farrell, N. P., Eds.; Humana Press Inc: Totowa, NJ, 2000; pp 37–61.
4. Brabec, V.; Sip, M.; Leng, M. Biochemistry 1993, 32, 11676–11681.
5. Paquet, F.; Boudvillain, M.; Lancelot, G.; Leng, M. Nucleic Acids Res 1999, 27, 4261–4268.
6. Marky, L. A.; Breslauer, K. J. Bioploymers 1987, 26, 1601–1620.
7. Murphy, J. H.; Trapane, T. L. Anal Biochem 1996, 240, 273–282.
8. Poklar, N.; Pilch, D. S.; Lippard, S. J.; Redding, E. A.; Dunham, S. U.; Breslauer, K. J. Proc Natl Acad Sci and USA 1996, 93, 7606–7611.
9. Lepre, C. A.; Chassot, L.; Costello, C. E.; Lippard, S. J. Biochemistry 1990, 29, 811–823.
10. Kasparkova, J.; Novakova, O.; Marini, V.; Najajreh, Y.; Gibson, D.; Perez, J.-M.; Brabec, V. J Biol Chem 2003, 278, 47516–47525.
11. Leharne, S. A.; Chowdhry, B. Z. In Biocalorimetry: Applications of Calorimetry in the Biological Sciences; Ladbury, J. E.; Chowdhry, B. Z., Eds.; J. Wiley & Sons: Chichester, UK, 1998; pp 157–182.
12. Erie, D.; Sinha, N.; Olson, W.; Jones, R.; Breslauer, K. Biochemistry 1987, 26, 7150–7159.
13. Pilch, D. S.; Plum, G. E.; Breslauer, K. J. Curr Opin Struct Biol 1995, 5, 334–342.
14. Breslauer, K. J. Methods Enzymol 1995, 259, 221–242.

Reviewing Editor: Kenneth J. Breslauer

MIKROKALORIMETRIE BIOLOGICKY VÝZNAMNÝCH MOLEKUL

Ctirad Hofr, *Oddělení funkční genomiky a proteomiky, Přírodovědecká fakulta MU, Kamenice 5, 625 00 Brno*

Díky technologickému pokroku došlo v posledním desetiletí ke zvýšení citlivosti mikrokalorimetrických technik na úroveň, která umožňuje jejich rutinní použití pro získávání komplexního popisu termodynamického chování biologicky významných molekul v roztoku. Mikrokalorimetrické metody jsou schopny přímo sledovat interakci makromolekul a popisovat jejich energetickou stabilitu. Tento článek vysvětluje princip izotermální titrační kalorimetrie a diferenční skenovací kalorimetrie a zabývá se určováním základních termodynamických parametrů z mikrokalorimetrických záznamů. Přímé stanovení kinetických a energetických parametrů vazby molekul umožňuje využívání mikrokalorimetrických přístupů při farmakokinetických studiích. Na příkladech jsou demonstrovány možnosti použití termodynamické analýzy při in-situ testování nových typů léčiv a je přiblíženo využití termodynamických dat získaných analýzou interakce stávajících léčiv pro cílené navrhování léčiv nových.

HISTORIE A POKROKY KALORIMETRIE

Od počátku studia přírodních zákonitostí byl vědecký zájem zaměřen na sledování procesů tepelné výměny u živých organismů. Při jednom z prvních kalorimetrických experimentů bylo do tepelně izolované nádoby uzavřeno morče. Byl sledován poměr výdeje tepla těla morčete a množství uvolněného oxidu uhličitého. Bylo zjištěno, že změřená hodnota byla podobná poměru určenému pro spalování uhlí. To společně s dalšími experimenty prokázalo, že dýchání je zvláštní formou pomalého spalování. Tento pokus provedli v tzv. ledovém kalorimetru Athoine Lavoaiser a Pierre-Simon Laplace v roce 1872. Je to první známý experiment, při kterém byla změřena tepelná výměna u biologického systému, a je považován také za první v rámci nové vědní disciplíny zabývající se stanovením tepelné výměny u biologických systémů, která byla později pojmenována biokalorimetrie.

Technický pokrok postupně posouval hranice citlivosti detekce od množství tepla produkovaného celými organismy až na současnou úroveň několika mikrojoulů, kdy je možno sledovat tepelné změny doprovázející vzájemné interakce makromolekul. Zvýšení citlivosti kalorimetrů bylo umožněno nejen díky technologickému rozvoji, ale zejména díky nově navrženému systému detekce tepelné výměny [1]. Technické vylepšení instrumentace a rozvoj separačních technik, zejména kapalinové chromatografie, umožnily začít rutinně měřit energetické charakteristiky biologicky významných makromolekul, zejména proteinů a nukleových kyselin. První kalorimetrická měření přispěla k potvrzení a rozvinutí teorie termodynamického chování makromolekul v roztocích [2]. Byla rovněž podrobně zpracována teorie tepelně indukovaných strukturních přechodů nukleových kyselin [3]. Rozšíření teoretických základů vysvětlujících termodynamické chování makromolekul umožnilo relativně rychlou aplikaci kalorimetrie ve vědních disciplínách zabývajících se interakcí makromolekul. Dokladem výrazného zvýšení popularity mikrokalorimetrie a možností její aplikace v širokém spektru vědeckých disciplín je nárůst počtu vědeckých prací, které obsahují kalorimetrická měření. S rostoucím zájmem o kalorimetrii souvisí mimo jiné také relativně vysoká komerční úspěšnost firem, které se zabývají vývojem a výrobou přístrojů pro mikrokalorimetrická měření [4], [5].

TERMODYNAMICKÁ CHARAKTERIZACE INTERAKCE MAKROMOLEKUL

Při vzájemné vazbě makromolekul se obvykle uplatňují nekovalentní interakce, jejichž energetické příspěvky jsou relativně malé, ale díky jejich velkému množství a často kooperativnímu působení mají výrazný vliv na způsob a rychlost vzájemné vazby makromolekul.

Jestliže uvažujeme vazbu molekuly M a jiné molekuly L, kterou můžeme pro zjednodušení nazývat ligand, může být reakce vyjádřena:

$$M + L \Leftrightarrow M \bullet L. \qquad (1)$$

Vazebná konstanta K_v příslušná reakci je pak dána poměrem molární koncentrace vznikajícího kom-

plexu [M•L] a součinu molárních koncentrací jednotlivých reagujících molekul [M] a [L]:

$$K_v = \frac{[M \bullet L]}{[M][L]}. \tag{2}$$

Volná energie je pro vazbu M a L vyjádřena za standardních podmínek

$$\Delta G^0 = -RT \ln K_v, \tag{3}$$

kde R je molární plynová konstanta a T je absolutní teplota v kelvinech. Vazba makromolekul M a L je možná, pouze je-li změna volné energie ΔG po asociaci makromolekul L a M negativní.

Z následující rovnice plyne vztah mezi volnou energií, entalpií a entropií systému molekul:

$$\Delta G = \Delta H - T\Delta S. \tag{4}$$

Jak je patrno z rovnice (4), entalpický a entropický příspěvek působí na celkovou změnu volné energie ΔG opačně. Zjednodušeně je možno říct, že vazebná entalpie ΔH je z molekulárního hlediska teplo spojené se vznikem, zánikem a deformací chemických vazeb. Vazebná entalpie ΔH obvykle popisuje změny v počtu a typu vodíkových můstků. Hodnota ΔH je negativní, jestliže během vzájemné vazby molekul dochází k celkovému zvýšení množství a zesílení vodíkových můstků; ΔH je naopak pozitivní, jestliže dojde k zániku nebo zeslabení vodíkových můstků.

Pro měření termodynamických veličin makromolekulárních interakcí je výhodné jejich vyjádření pomocí tepelné kapacity, kterou lze experimentálně určovat. Tepelná kapacita C_p udává teplo, které je nutno dodat za konstantního tlaku soustavě, aby se ohřála o jeden stupeň.

Vazebná entalpie H při ději za stálého tlaku je zapsána pomocí tepelné kapacity

$$\Delta H = \int_{T_{poc}}^{T_{kon}} C_p \mathrm{d}T. \tag{5}$$

Změna entropie ΔS při dané teplotě T je spojena se změnou v uspořádanosti systému molekul. Změna entropie je pozitivní, jestliže při vzájemné vazbě makromolekul dojde k uvolnění molekul vody vázaných na povrchu makromolekul z prostoru vazebného místa do volného roztoku. Tím se zvýší neuspořádanost systému molekul a dojde ke zvýšení celkové entropie. Negativní změna entropie většinou provází změny, při kterých dochází k omezení možných konformačních stavů makromolekul.

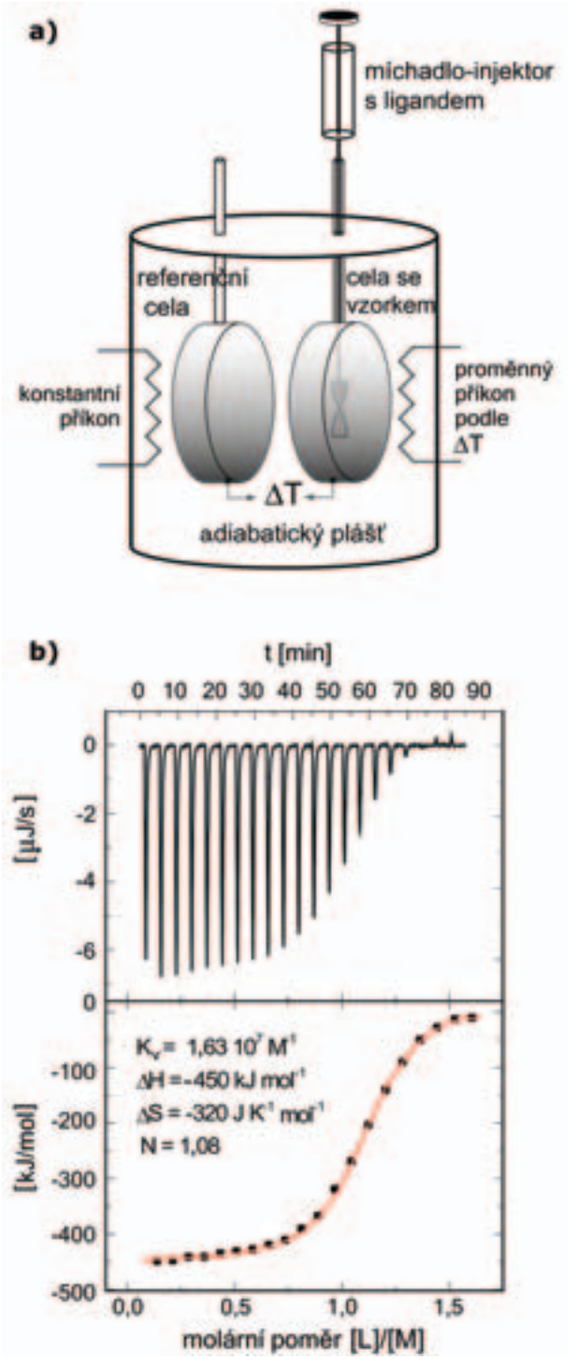
$$\Delta S = \int_{T_{poc}}^{T_{kon}} \frac{C_p}{T} \mathrm{d}T. \tag{6}$$

Jestliže po vzájemné vazbě makromolekul dojde k pozitivní změně entropie ΔS, která je doprovázena negativní změnou tepelné kapacity ΔC_p, je to známka toho, že se při vazbě uplatnily hydrofobní interakce. V tomto případě se při vazbě makromolekul do volného roztoku uvolňují molekuly vody, které byly původně uspořádány do vrstev kolem později vzájemně reagujících hydrofobních částí makromolekul. To má za následek snížení schopnosti celého systému molekul přijímat teplo, což se projeví celkovým snížením tepelné kapacity.

IZOTERMÁLNÍ TITRAČNÍ KALORIMETRIE

Izotermální titrační kalorimetrie (ITC z angl. *Isothermal Titration Calorimetry*) přímo měří změnu vazebné entalpie při interakci molekul v roztoku. Uspořádání měřicí části izotermálního titračního kalorimetru ukazuje obr. 1, část a. Referenční a vzorková měřicí cela mají tvar mince. Plní se kapilárami, přičemž referenční cela je zpravidla plněna pouze samotným pufrem. Objem cel je řádově jeden mililitr. Teplota obou cel je během celého měření udržována konstantní. Při měření je do reakční cely s roztokem makromolekul postupně, za stálého míchání, přidáván roztok vázající se molekuly – ligandu. Obvykle je koncentrace ligandu přibližně desetkrát vyšší než koncentrace molekuly v reakční cele. Celkový objem ligandu v injektoru je 100–250 µl. Při interakci makromolekul s ligandem dochází k tepelným efektům, které se projeví rozdílem teploty referenční cely a cely se vzorkem. Rozdíl teplot je detekován speciálně tvarovaným termočlánkem umístěným v prostoru mezi celami. Termočlánek je spojen s okruhem, který při známé tepelné kapacitě cely vypočte dle vztahu (5) tepelný rozdíl a zajistí změnu přísunu elektrické energie pro vyrovnání teplot obou cel. Při exotermické reakci, kdy se při vazbě molekul uvolňuje teplo, je přísun energie snížen, naopak v případě endotermické reakce se teplo spotřebovává a přísun energie je zvýšen. Změna dodávané energie v závislosti na čase je zaznamenávána a je vztažena k okamžité koncentraci jednotlivých složek v reakční cele. Výsledkem měření je závislost změny tepelné kapacity na čase. Následně je provedena korekce na tepelné vlivy spojené s ředěním ligandu a mícháním, což se provádí odečtením referenční křivky získané při měření za totožných podmínek, ale bez přítomnosti makromolekul v roztoku. Integrací je následně vypočtena křivka závislosti vazebné entalpie na molárním poměru ligandu a makromolekuly v cele. Nelineární analýza výsledné křivky umožňuje na základě jednoho měření určit vazebnou entalpii ΔH, vazebnou konstantu K_v a stechiometrii reakce N.

Výsledná sigmoidální křivka (obr. 1, část b) je tvořena body, které udávají vazebnou entalpii příslušnou ke každému přídavku ligandu do reakce. Hodnota signálu je největší na začátku titrace, kdy je

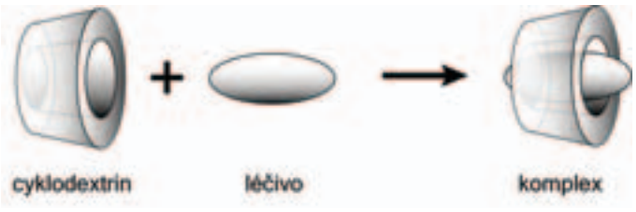


1/ (a) Schéma izotermálního titračního kalorimetru. (b) V horní části je záznam tepla dodávaného do reakční cely při titračním experimentu, kdy byl do roztoku krátkého jednořetězcového fragmentu DNA postupně přidáván komplementární řetěz. Docházelo k hybridizaci – vzniku dvoušroubovice DNA, při kterém se uvolňuje teplo, takže bylo nutno snížit dodávanou tepelnou energii po každém přidání komplementárního řetězce. Dolní panel ukazuje záznam, který vznikl integrací jednotlivých signálů. Ze záznamu lze po nelineárním proložení odečíst vazebnou entalpii ΔH, vazebnou konstantu K_V a stechiometrii reakce N.

velký nadbytek makromolekul v cele a zpravidla určuje hodnotu vazebné entalpie ΔH ligandu k makromolekule. Postupně v průběhu titrace dochází k obsazování vazebných míst ligandem a odpovídajícímu snižování tepelného signálu až do stavu, kdy dojde k nasycení vazebných míst a v roztoku se nachází volný ligand. Z tvaru titrační křivky lze určit vazebnou konstantu K_v a ze vztahu (3) následně změnu volné energie ΔG spojené se vzájemnou vazbou molekul. Poloha inflexního bodu křivky udává stechiometrický poměr N ligandu a makromolekuly při vzájemné vazbě. Právě schopnost určit přímo stechiometrii vazby makromolekul činí z izotermální titrační kalorimetre jedinečný nástroj pro studium interakce multimerních molekulových komplexů, kdy je stanovení vzájemného poměru reagujících molekul často složité.

ITC PŘI NAVRHOVÁNÍ NOVÝCH LÉČIV

Izotermální titrační kalorimetrie našla své uplatnění při navrhování léčiv. Limitujícím faktorem při vývoji nových léčiv je schopnost léčiva dosáhnout místa svého působení, dále zajištění postupného uvolňování léčiva a častý požadavek na rozpustnost léčiva ve vodě. Chemická povaha léčiva někdy neumožňuje optimální splnění všech výše uvedených podmínek. Jednou z možností jak zlepšit farmakokinetické vlastnosti léčiva je navázat léčivo na molekulu speciálně navrženého nosiče. Po podání komplexu nosiče a léčiva pak může dojít k cílenému transportu na místo účinku, kde postupně dochází k uvolňování léčiva. Příkladem takových nosičových molekul jsou cyklodextriny a jejich deriváty (obr. 2). Cyklodextriny jsou v přírodě se vyskytující cyklické oligosacharidy. Uvolňování léčiva je řízeno vazebnou afinitou léčiva a cyklodextrinu. Při hledání vhodné nosičové molekuly cyklodextrinu se provádí měření vazebné charakteristiky za použití ITC. Na základě změřených hodnot vazebných charakteristik se zvolí nejvhodnější kandidát pro nosičovou cyklodextrinovou molekulu daného léčiva. Cyklodextrinové nosiče jsou v současnosti testovány při vývoji nových léčiv proti malárii, rakovině tlustého střeva a různým druhům zánětů [6]–[8].



2/ Cyklodextriny mohou sloužit jako nosičové molekuly pro různé typy léčiv.

DIFERENČNÍ SKENOVACÍ KALORIMETRIE

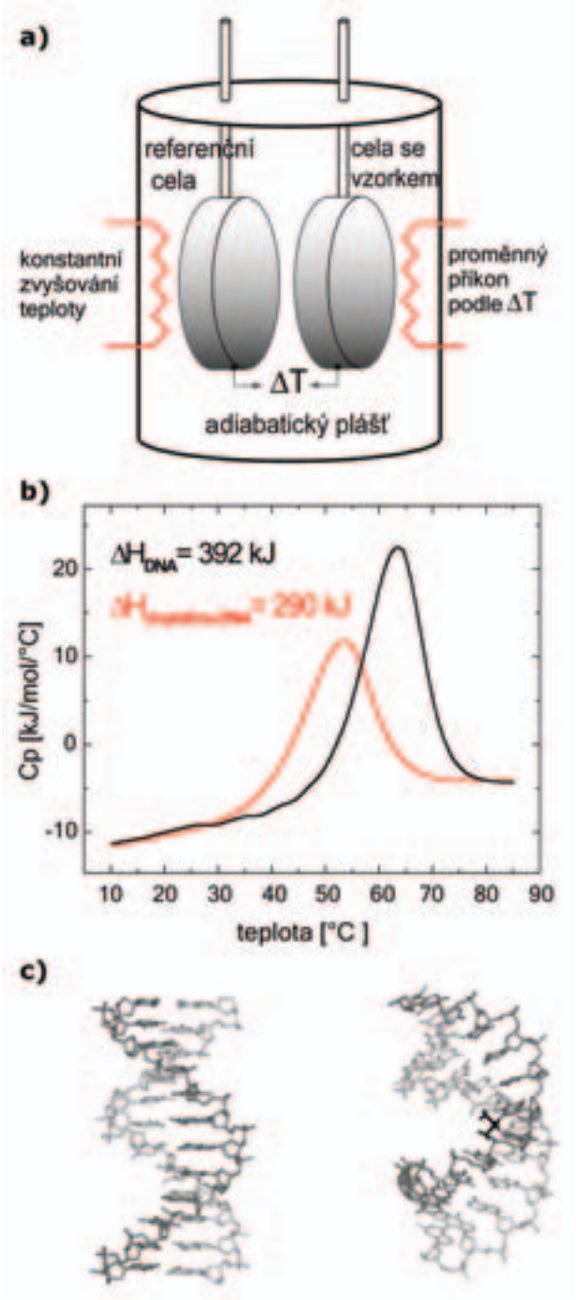
Další kalorimetrická metoda, která umožňuje sledovat změny termodynamických parametrů molekul po vzájemné vazbě, je diferenční skenovací kalorimetrie (DSC z angl. *Differential Scanning Calorimetry*). Jedno měření za použití DSC umožňuje určit termodynamické parametry stability makromolekul z množství tepla, které je uvolněno nebo spotřebováno při tepelně indukovaných konformačních přechodech (např. tání DNA nebo denaturaci proteinů). Z rozdílu termodynamických charakteristik pro makromolekulu s navázaným ligandem a nemodifikovanou makromolekulu je pak určen termodynamický vliv vazby ligandu na energetickou stabilitu makromolekuly. Nejdříve se stanoví hodnoty termodynamických parametrů pro denaturaci nemodifikované makromolekuly, které se následně odečtou od hodnot parametrů pro denaturaci makromolekuly s navázaným ligandem, jak je ukázáno následujícím vztahem.

$$\Delta H_{naz} = \Delta H_{M\bullet L} - \Delta H_M . \qquad (7)$$

Při měření je kontinuálně zahřívána cela se vzorkem a cela referenční naplněná pouze pufrem (obr. 3, část a). Teplota se obvykle mění v rozsahu 0-100 °C s gradientem 1 °C/min. Jestliže dojde v cele se vzorkem k denaturaci neboli tání dvoušroubovice DNA, kdy se část dodaného tepla spotřebuje pro narušení nekovalentních vazeb a oddělení komplementárních řetězců DNA, bude teplota cely se vzorkem nižší než teplota cely referenční. Z rozdílu teplot mezi oběma celami, který je zaznamenáván termočlánkem, je určena dodatečná energie, která je potřebná pro vyrovnání teplot obou cel. Tato dodatečná energie je zaznamenávána. Výsledkem měření je termogram, který udává závislost tepelné kapacity C_p roztoku makromolekuly na teplotě. Integrací termogramu lze přímo určit změnu entalpie H příslušnou tání DNA podle (5) a po integraci závislosti C_p/T na teplotě ΔS podle (6). Hodnota volné energie ΔG je vypočtena podle vztahu (4). Teplota odpovídající maximu termogramu je teplota tání T_m, která charakterizuje teplotní stabilitu DNA. Při teplotě tání T_m je právě polovina molekul v denaturovaném stavu.

VYUŽITÍ DSC PŘI HLEDÁNÍ NOVÝCH LÉČIV

Diferenční skenovací kalorimetrie byla úspěšně využita při popisu termodynamických změn vyvolaných kovalentní vazbou protinádorově účinných koordinačních sloučenin kovů na DNA. Protinádorový účinek komplexů platiny je obecně spojován se strukturními a energetickými změnami DNA, které jsou jejich vazbou vyvolávány. Klinicky



3/ (a) Schéma diferenčního skenovacího kalorimetru. (b) Normalizované záznamy měření tání krátkého fragmentu DNA bez modifikace a s navázanou protinádorově účinnou cisplatinou. Hodnoty změny entalpie ΔH ukazují, že vytvoření komplexu cisplatina-DNA výrazně snížilo hodnotu entalpie pro tání DNA, což odpovídá narušení vodíkových vazeb přibližně v rozsahu pěti párů bází v okolí navázané cisplatiny. (c) Strukturní změny dvoušroubovice DNA po vytvoření komplexu cisplatina-DNA

používané protinádorové komplexy, jako je cisplatina, jsou však účinné proti relativně úzkému spektru nádorů. Léčba s sebou nese nežádoucí vedlejší

účinky a po čase dochází ke vzniku rezistence. Pak je nutno zvyšovat dávku podávaných léčiv. Tyto skutečnosti podněcují hledání nových protinádorových farmak, která by umožnila účinnější léčbu s menšími vedlejšími účinky. Při cíleném hledání nových léčiv může být s výhodou využito poznatků získaných podrobným popisem změn struktury a termodynamické stability DNA, které jsou vyvolány protinádorově účinnými komplexy a jejich neúčinnými izomery [9], [10]. Změny ve struktuře a termodynamické stabilitě DNA jsou signálem pro specifické proteiny, které rozpoznávají modifikovanou DNA a mohou se podílet na mechanismu protinádorového účinku. Na základě těchto poznatků jsou pak navrhovány nové typy sloučenin, které mohou být účinné proti doposud neléčitelným typům nádorů, a léčba těmito novými farmaky nebude tolik zatěžovat organismus pacienta.

VÝHLEDY DO BUDOUCNA

Aby mohlo dojít k širšímu používání kalorimetrie, bude nutno dále zvyšovat její citlivost. Při extenzivním hledání nových léčiv vyvstává požadavek analyzovat v krátkém čase velké množství vzorků. Nedávno se na trhu objevily DSC a ITC přístroje, které jsou schopny měřit vzorky automaticky [5]. Vyvíjejí se kalorimetry na bázi čipů, které umožní zvýšit počet vzorků měřených při jedné analýze a zároveň snížit jejich spotřebu [11]. Mikrokalorimetrie je metodou, která má velký potenciál pro monitorování mezimolekulových interakcí bez nutnosti speciální vizualizace reagujících molekul. Jako jedna z mála experimentálních metod umožňuje přímou a rychlou energetickou charakterizaci molekul. Nové poznatky o termodynamickém chování molekul mohou být pak s výhodou využity při následné počítačové simulaci, což dovolí rychlejší a efektivnější cílené navrhování nových typů farmak, která budou mít žádané farmakokinetické vlastnosti. Tyto skutečnosti činí mikrokalorimetrické přístupy velmi atraktivní pro využití v moderní strategii vývoje nových farmak. Jestliže dojde k naplnění očekávání vkládaných do vývoje nových typů kalorimetrů, význam mikrokalorimetrie při hledání nových léčiv bude dále vzrůstat.

Literatura

[1] V. Plotnikov, J. M. Brandts, L. L. Lin, J. F. Brandts, Analytical Biochemistry **250**, 237 (1997).
[2] T. M. J. Record, C. F. Anderson, T. M. Lohman, *Quarterly Reviews of Biophysics II* **2**, 104 (1978).
[3] L. A. Marky, K. J. Breslauer, Biopolymers **26**, 1601 (1987).
[4] Calorimetric Sciences Corp: [06-2006], URL: http://www.calscorp.com.
[5] Microcal: [06-2006], URL:http://www.microcalorimetry.com.
[6] S. A.Charman, C. S. Perry, F. C. K. Chiu, K. A. Mcintos, R. J. Prankerd, W. N. Charman, Journal of Pharmaceutical Sciences **95**, 256 (2005).
[7] T. Schluep, J. Cheng, K. T. Khin, M. E. Davis, Cancer Chemotherapy and Pharmacology. **57**, 654 (2006).
[8] K. Uekama, Chem. Pharm. Bull. **52**, 900 (2004).
[9] C. Hofr, V. Brabec, J. Biol. Chem. **276**, 9655 (2001).
[10] C. Hofr, V. Brabec, Biopolymers **77**, 222 (2005).
[11] F. E. Torres a kol., Proc. Natl. Acad. Sci. U. S. A. **101**, 9517 (2004).

Phytochemistry 69 (2008) 1814–1819

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Functional characterization of domains in AtTRB1, a putative telomere-binding protein in *Arabidopsis thaliana*

Iva Mozgová [a], Petra Procházková Schrumpfová [a], Ctirad Hofr [a], Jiří Fajkus [a,b,*]

[a] *Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic*
[b] *Laboratory of DNA-Molecular Complexes, Institute of Biophysics, Czech Academy of Sciences, Královopolská 135, CZ-61265 Brno, Czech Republic*

ARTICLE INFO

Article history:
Received 10 March 2008
Received in revised form 24 March 2008
Available online 12 May 2008

Keywords:
Plant
Arabidopsis thaliana
Telomere
Single-Myb-Histone protein
AtTRB1
DNA-protein interaction
EMSA
PFO-ELFO

ABSTRACT

Telomeres are nucleoprotein structures ensuring the stability of eukaryotic chromosome ends. Two protein families, TRFL (*TFL-L*ike) and SMH (*S*ingle-*M*yb-*H*istone), containing a specific *telobox* motif in their Myb domain, have been identified as potential candidates involved in a functional nucleoprotein structure analogous to human "shelterin" at plant telomeres. We analyze the DNA-protein interaction of the full-length and truncated variants of AtTRB1, a SMH-family member with a typical structure: N-terminal Myb domain, central H1/5 domain and C-terminal coiled-coil. We show that preferential interaction of AtTRB1 with double-stranded telomeric DNA is mediated by the Myb domain, while the H1/5 domain is involved in non-specific DNA-protein interaction and in the multimerization of AtTRB1.

1. Introduction

Telomeres are nucleoprotein structures which form the ends of linear chromosomes, ensuring their stability and complete replication. The telomeric DNA is composed of minisatellite tandem repeats, for example $(TTTAGGG)_n$ in the majority of plants or $(TTAGGG)_n$ in vertebrates, which form a terminal 3′-single-stranded G-rich overhang (Moyzis et al., 1988; Richards and Ausubel, 1988). Telomeric DNA is complexed with histones and non-histone proteins which mediate the flexibility of telomere structure (Griffith et al., 1999), regulate its accessibility for other proteins including telomerase, prevent recombination and serve as recognition marks of intact chromosomal termini to distinguish them from double-strand breaks (Bertuch and Lundblad, 2006).

Proteins bind telomere either by directly interacting with the double-stranded (ds) or single-stranded (ss) telomeric DNA or through protein–protein interactions with the DNA-binding proteins. In humans, the constitutive complex of six telomere-binding proteins has been termed "shelterin", encompassing the dsDNA-binding proteins TRF1 and TRF2, the ssDNA-binding protein POT1 and the proteins TIN2, TPP1 and hRap1 which are bound to the complex via protein–protein interactions (de Lange, 2005). Telomeric dsDNA-binding proteins show the presence of a single highly conserved motif, the three-helical Myb domain in which the third helix is defined by the presence of a conserved amino acid sequence termed a "*telobox*" (Bilaud et al., 1996), which is responsible for the recognition of telomeric dsDNA.

Numerous proteins have been suggested to participate in the plant analog of human shelterin. Based on the presence of a *telobox* within their Myb domain, a group of TBP-like (*T*elomere-*B*inding-*P*rotein-like) proteins has been found *in silico* in *Arabidopsis thaliana* (Yanhui et al., 2006). The plant Myb domain of the TBP-like proteins can be placed either at the C-terminus of the protein, as in TRFL (TRF-like) proteins (Karamysheva et al., 2004), or at the N-terminus, as in SMH (*S*ingle-*M*yb-*H*istone) proteins (Marian et al., 2003). In TRFL proteins, the Myb domain can either be extended by a fourth helix (TRFLI family) or the three-helical structure of the Myb domain is maintained (TRFLII) (Hwang and Cho, 2007; Hwang et al., 2005; Karamysheva et al., 2004; Sue et al., 2006).

SMH proteins are plant-specific and have been reported to preferentially bind plant telomeric dsDNA. Apart from the N-terminally localized three-helical Myb-domain, they contain a histone-like H1/5 domain in the central part of the protein and a coiled-coil domain at the C-terminus (Marian et al., 2003; Schrumpfova et al., 2004). The *A. thaliana SMH* family of genes encodes five proteins (AtTRB1-5), two of which (At TRB2 and At TRB3) have been characterized previously and are capable to bind preferentially to plant

* Corresponding author. Address: Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic. Tel.: +420 549494003; fax: +420 549492654.
E-mail address: fajkus@sci.muni.cz (J. Fajkus).

0031-9422/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2008.04.001

\telomeric dsDNA. Considering their unique structure, it has been suggested that binding of these proteins to dsDNA is mediated and affected not only by the Myb domain but possibly also by the H1/5 domain, whose role in DNA binding has however not been elucidated (Schrumpfova et al., 2004). AtTRB1 (At1g49950), a member of the SMH family, interacts with the telomere-associated protein AtPOT1b, one of the POT1-like proteins in *A. thaliana* (Kuchar and Fajkus, 2004), which in turn participates in telomere capping and maintenance of genome stability (Shakirov et al., 2005).

In this study, we analyze the mode of binding of AtTRB1 to DNA. We show that the high-affinity sequence-specific interaction between telomeric dsDNA and AtTRB1 is mediated predominantly by the Myb domain, although other parts of the protein also contribute to the interaction. Furthermore, we show that AtTRB1 is capable of forming multimers and that this interaction is mediated mainly by the H1/5 domain. We suggest that AtTRB1 shows the characteristics of a telomere-binding protein and may be a part of the protein complex safeguarding plant telomeres.

2. Results and discussion

2.1. AtTRB1 preferentially binds telomeric dsDNA

Full-length AtTRB1 (Fig. 1) was subjected to EMSA (Electrophoretic Mobility Shift Assay) using ss (AtTR3-G, AtTR3-C) or ds (AtTR3) trimers of a plant telomeric sequence and a non-telomeric oligonucleotide (NonTR-C, NonTR-G, NonTR). The telomeric oligonucleotide was used as the labeled probe while the non-telomeric oligonucleotide served as competitor DNA and *vice versa.* The results show that AtTRB1 is indeed a telomeric dsDNA-binding protein, similarly to proteins AtTRB2 and AtTRB3 which we have described previously (Schrumpfova et al., 2004), and shows no or very low-affinity to non-telomeric dsDNA (Fig. 2) or ssDNA of either telomeric or non-telomeric sequence (data not shown).

2.2. The Myb domain mediates the specific interaction between AtTRB1 and telomeric dsDNA

In order to determine which domain of AtTRB1 is primarily responsible for binding telomeric dsDNA, truncated variants of AtTRB1 (Myb domain only – *m*; Myb + H1/5 domains – *mh*; H1/5 domain only – *h*; and H1/5 + coiled-coil domains – *hc*) were used in EMSA (Fig. 1). All truncated variants of AtTRB1 showed a certain affinity towards telomeric dsDNA (Fig. 3); the level and character of the binding differed in each case. The Myb domain showed the strongest interaction with telomeric dsDNA of all the protein fragments tested; telomeric DNA was not released from the DNA-protein complex even upon addition of 100-fold excess of non-telomeric competitor (Fig. 3a). Two telomeric repetitions were adequate for the binding of Myb domain (data not shown). Unlike in TRFLI proteins, the Myb domain of SMH proteins does not contain

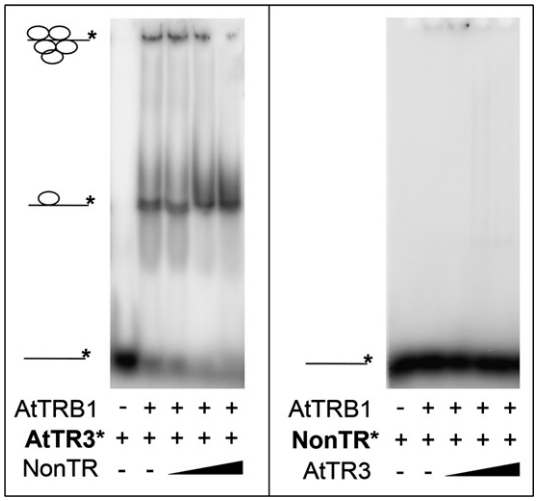
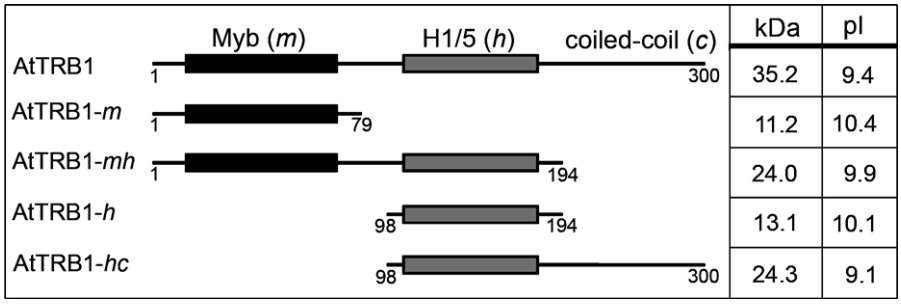


Fig. 2. EMSA of full-length AtTRB1 binding a radioactively-labeled ds telomeric (AtTR3∗) or non-telomeric (NonTR∗) oligonucleotide with unlabeled NonTR or AtTR3 respectively as competitor DNA. The concentration of unlabeled competitor increases from 1-, 20- to 100-fold the concentration of the labeled probe (as depicted by the triangle). DNA probe/protein concentration is 1/100.

the C-extension domain which has been proposed to condition the ability of the proteins to bind to telomeric DNA (Hwang and Cho, 2007; Karamysheva et al., 2004). However, recent solution of the structure of the C-extension-containing Myb domain of AtTRP1 has shown that the contact between DNA and AtTRP1 is mediated merely by amino acid residues of helix 1 (H1) and H3 and the loop connecting H3 and H4, while H4 itself stabilizes the Myb domain into a tetrahedral structure (Sue et al., 2006). The level of conservation of the amino acid sequence within the C-extension region thus may not be critical for binding of telomeric DNA, as the domain-stabilizing function may be provided by a different amino acid sequence.

The fragment *mh* bound telomeric dsDNA in a similar mode as fragment *m*, but formation of a high molecular weight complex, which did not migrate into the gel, was observed. This complex was partially disassembled by the addition of a 100-fold excess of non-telomeric competitor, which however did not release the protein from the complex with telomeric DNA (Fig. 3b). The H1/5 domain of AtTRB1 formed a high molecular weight complex with the telomeric DNA which was released from the complex by addition of non-telomeric competitor (Fig. 3c). Similarly, the fragment *hc* retarded the telomeric DNA probe but was partially loosened from the complex by a 100-fold excess of non-telomeric competitor (Fig. 3d). None of the fragments of AtTRB1 displayed affinity towards ssDNA or non-telomeric dsDNA with the exception of the H1/5 domain, which bound non-telomeric dsDNA in a similar manner as the telomeric dsDNA (data not shown). In this case, how-



	kDa	pI
AtTRB1	35.2	9.4
AtTRB1-*m*	11.2	10.4
AtTRB1-*mh*	24.0	9.9
AtTRB1-*h*	13.1	10.1
AtTRB1-*hc*	24.3	9.1

Fig. 1. Cloned regions of AtTRB1 schematic representation of the full-length AtTRB1 and its truncated variants. Numbers in diagrams denote amino acid residues. pI denotes isoelectric point as determined using pI/MW calculator at ExPASy Proteomics Server.

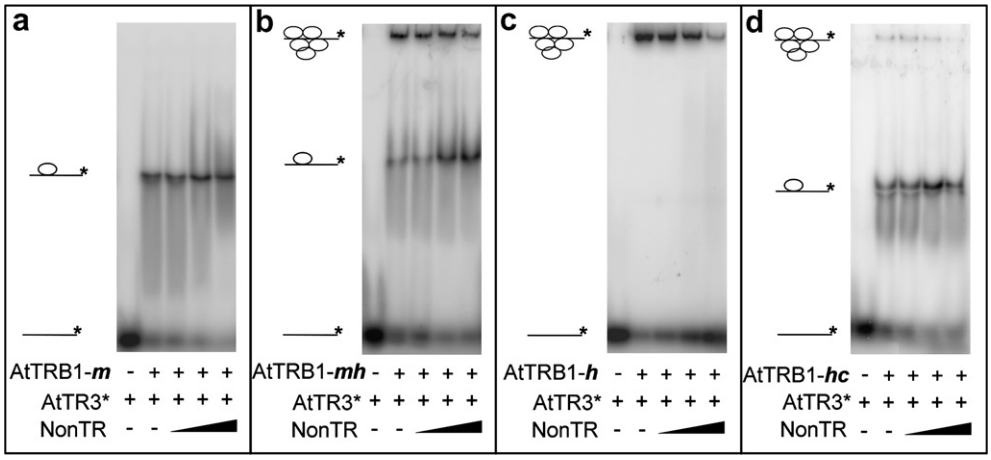


Fig. 3. EMSA of fragments of AtTRB1 binding a radioactively-labeled ds telomeric (AtTR3∗) oligonucleotide with unlabeled non-telomeric oligonucleotide (NonTR) as competitor DNA. The concentration of unlabeled competitor increases from 1-, 20- to 100-fold the concentration of the labeled probe (as depicted by the triangle). DNA probe/protein concentration is 1/100.

ever, telomeric dsDNA did not compete for the interaction as efficiently, and the non-telomeric probe was not released from the complex. This shows that the H1/5 domain interacts non-specifically with any DNA without preference for either telomeric or non-telomeric sequence (Ellen and van Holde, 2004) which, together with the high pI of the AtTRB1 fragments, suggests that electrostatic interactions take part in the interaction of the fragments of AtTRB1 with telomeric dsDNA. This may interfere with determining the sequence specificity of AtTRB1 binding to telomeric DNA. The specificity of the interaction is nevertheless demonstrated by the fact that the H1/5 domain alone as well as the H1/5 + coiled-coil domains are released from the DNA-protein complex in excess of non-telomeric DNA, while the Myb and Myb + H1/5 domains form stable complexes with telomeric DNA even when competed by non-telomeric DNA.

To determine the relative contribution of the Myb domain to the telomeric DNA binding affinity of AtTRB1, the amount of protein required for the formation of a DNA-protein complex which could be clearly and reproducibly detected in the gel was compared between fragments *m* and *hc*. The Myb domain formed a clear retarded band at a DNA/protein ratio 1/5, whereas fragment *hc* formed retarded bands only from a DNA/protein ratio above 1/50 (see Fig. 4).

Taken together, these data suggest that the Myb domain is the main region of the protein responsible for and promoting the binding of AtTRB1 to telomeric dsDNA, but also other regions have the potential to contribute to this interaction.

2.3. The Myb domain of AtTRB1 binds plant and human telomeric DNA with a similar affinity

To test the specificity of the Myb domain of AtTRB1 in recognition of related telomeric sequences, the binding to plant telomeric dsDNA (AtTR3) and human telomeric dsDNA (HuTR3) was compared when competed by the non-telomeric dsDNA (NonTR), AtTR3 or HuTR3. While the protein was not released from the plant telomeric probe by the non-telomeric DNA, the complex was equally dissociated by an excess of either human or plant telomeric DNA (Fig. 5). Similarly, the fragment covering the Myb domain together with the H1/5 domain (fragment *mh*) did not bind plant telomeric DNA stronger than human telomeric DNA (data not shown).

The non-selective binding of the Myb domain to either plant (TTTAGGG) or human (TTAGGG) telomeric sequence appears to be a general feature of the *A. thaliana* SMH proteins. Although the Myb domain has not been studied separately before, the bind-

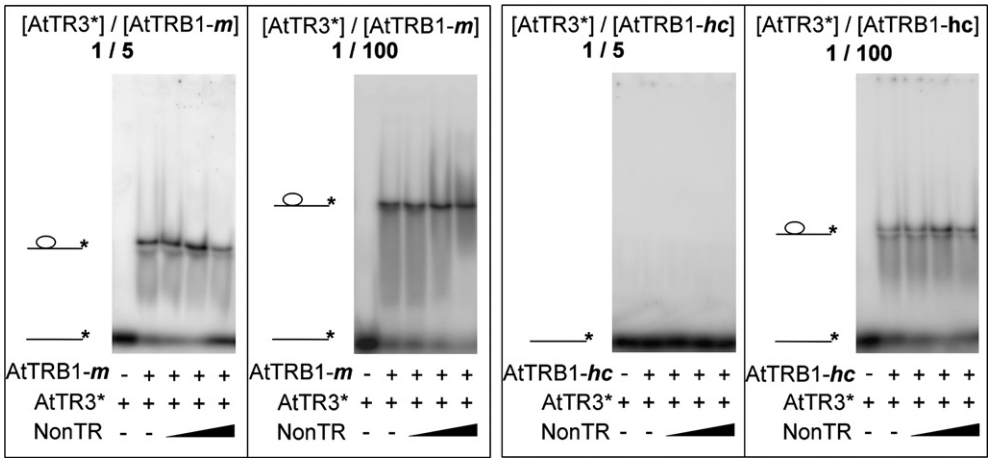


Fig. 4a. Comparison of the binding affinity of the Myb domain (fragment AtTRB1-*m*) and the H1/5 + coiled-coil domains (fragment AtTRB1-*hc*) to a radioactively-labeled ds telomeric (AtTR3∗) oligonucleotide with unlabeled non-telomeric oligonucleotide (NonTR) as competitor. The concentration of the protein exceeds the concentration of the labeled probe 5-fold or 100-fold.

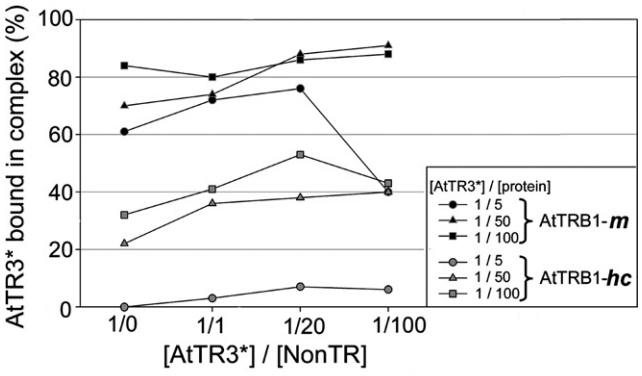


Fig. 4b. Percentage of labeled telomeric probe (AtTR3∗) bound in a DNA-protein complex by the Myb domain (AtTRB1-*m*) or the H1/5 + coiled-coil domains (AtTRB1-*hc*) with increasing concentration of unlabelled non-telomeric competitor (NonTR).

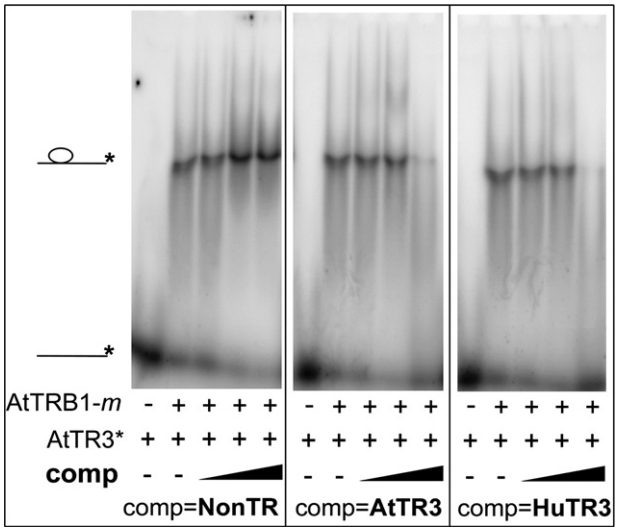


Fig. 5. Affinity of the Myb domain (AtTRB1-*m*) to labeled plant telomeric probe (AtTR3∗) in the presence of increasing amounts of unlabeled non-telomeric (NonTR), plant telomeric (AtTR) or human telomeric (HuTR) competitor. DNA probe/protein concentration is 1/100.

ing of two other full-length SMH proteins, AtTRB2 and AtTRB3, to plant telomeric oligonucleotide has been efficiently competed by human telomeric sequence as well (Schrumpfova et al., 2004). Although the proteins are not able to discriminate between related telomeric sequences of *A. thaliana* and human their stronger binding of telomeric than non-telomeric DNA suggests the following possibilities: (i) the core sequence recognized by the Myb domain without the C-extension is centered so that there is no difference between plant and human sequences (e.g. TTAGGGTT). Myb domains with a C-extension are however centered around the TTT sequence (Chen et al., 2001; Sue et al., 2006; Yang et al., 2003; Yu et al., 2000); (ii) At TRB proteins do not distinguish small changes in the sequence, as there has never been competition of so highly similar sequences throughout the evolution of these proteins in *A. thaliana*. In addition, the ability of AtTRB1 to bind typical plant and human telomeric motifs with a similar affinity could be important for an easier adaptation to a change in telomere sequence from a plant to a human-like motif, which has occurred during the evolution of plants of the order Asparagales (Sykorova et al., 2003). This also corresponds with the finding that proteins binding ss telomeric DNA in plants with human-type telomeres show affinity to both the ancestral and present telomeric DNA (Rotkova et al., 2004; Rotková et al., 2007).

2.4. AtTRB1 forms multimers under the conditions of PFO-PAGE (Perfluoro-octanoic acid PolyAcrylamide Gel Electrophoresis)

In order to investigate the ability of the AtTRB1 fragments to form self-dimers or multimers, PAGE with a weak detergent, perfluoro-octanoic acid (PFO), was used. This method can be used for detection and molecular mass determination of protein complexes since (in contrast to SDS-PAGE), PFO-PAGE preserves high-affinity protein–protein interactions (Ramjeesingh et al., 1999). The results confirm the strong tendency of the H1/5 domain to multimerize and the same holds true for all the fragments of AtTRB1 which contain the H1/5 domain. Dimerization has been proved to increase the efficiency of the binding of telomere-associated proteins to telomeric DNA (Bianchi et al., 1997, 1999; Fairall et al., 2001) and may also enhance the selectivity for binding longer tracts of telomeric sequence (Karamysheva et al., 2004). Candidate plant telomeric proteins containing the Myb domain on their C-terminus have also been shown to interact with the telomeric DNA as dimers. AtTRP1, a member of the TRFLI family (Karamysheva et al., 2004), has been shown to multimerize via its central domain and even the isolated Myb domain with C-extension has been shown to multimerize through the C-terminal residues (Sue et al., 2006). Myb domain of the rice RTBP also interacts with plant telomeric DNA in the form of a homodimer (Yu et al., 2000). A capability for multimerization therefore seems to characterize telomere dsDNA-binding proteins. In the case of AtTRB1, the N-terminal Myb domain itself does not form higher molecular weight complexes (Fig. 6), therefore the length of the DNA recognition sequence should be determined only by the number of DNA-protein interactions within one molecule of protein. This is in agreement with our finding that two telomeric repetitions are sufficient for the binding of the Myb domain. Similarly, Sue et al. have shown that the monomeric Myb domain of AtTRP1$^{464–560}$ interacts with 13 bp of DNA containing a single repeat of *A. thaliana* telomeric sequence (Sue et al., 2006).

3. Concluding remarks

AtTRB1 is a member of the plant-specific SMH family of *A. thaliana* proteins which are characterized by the presence of a *telobox*-containing Myb domain located at their N-terminus, a histone-like H1/5 domain in the centre and a coiled-coil domain at the C-terminus (Marian et al., 2003). Because the presence of a *telobox* is common to telomere dsDNA-binding proteins (Bilaud et al., 1996), the SMH proteins are potential candidates to cap and regulate the plant telomeres.

We show that AtTRB1 is indeed a telomeric dsDNA-binding protein, similarly as proteins AtTRB2 and AtTRB3 which we have described previously (Schrumpfova et al., 2004). The Myb domain of AtTRB1 preferentially binds telomeric over non-telomeric DNA. Two telomeric repeats are sufficient for the Myb domain to bind DNA. Homomultimer formation through the H1/5 domain has been shown by PFO-PAGE. This suggests a possible role for the H1/5 domain of AtTRB proteins in promoting either their self- or mutual interactions or their interaction with other partners. Compared to other dsDNA-binding proteins, the triple-domain AtTRB proteins have a potential for a remarkable multiplicity of modes of binding to telomeres: the Myb-domain primarily ensures direct sequence-specific binding of AtTRB1 to telomeric DNA, the H1/5 domain may enhance this binding by protein dimerization, while extensive protein aggregation is counteracted by competitive loading of the H1/5 domain to non-specific DNA (Fig. 7). Thus, paradoxically, DNA-

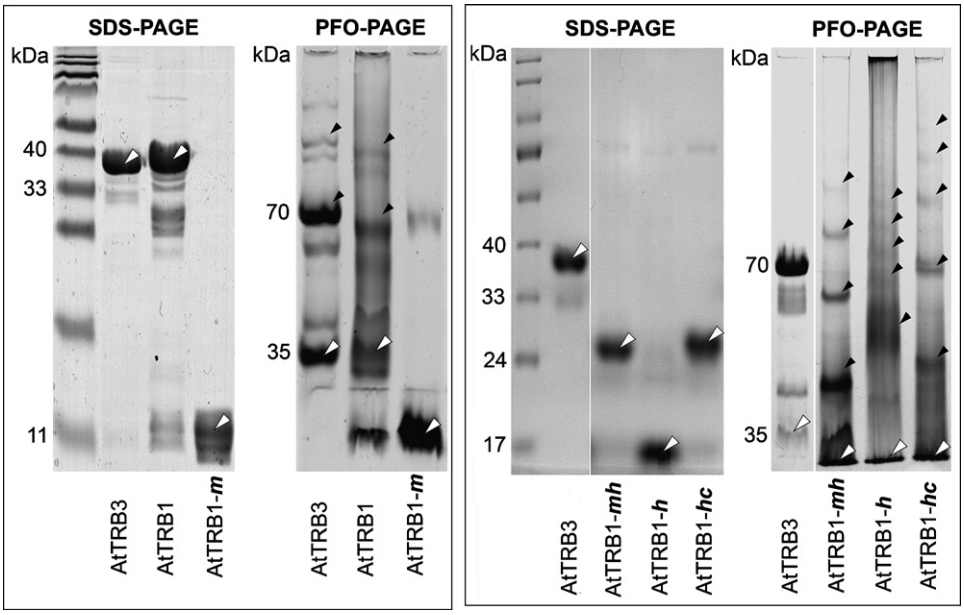


Fig. 6. Comparison of SDS-PAGE and PFO-PAGE of AtTRB1 and its fragments. White arrowheads point to monomers, black to homomultimers; the AtTRB3 monomer was used as an approximate marker of MW 35 kDa. Left panel: AtTRB1 and AtTRB1-*m* (15% SDS-PAGE gel, AA:BIS 37:1; 9% PFO-PAGE gel, AA:BIS 37:1); right panel: AtTRB1 – *mh, h, hc* (12.5% SDS-PAGE gel, AA:BIS 37:1; 8% PFO-PAGE gel, AA:BIS 37:1).

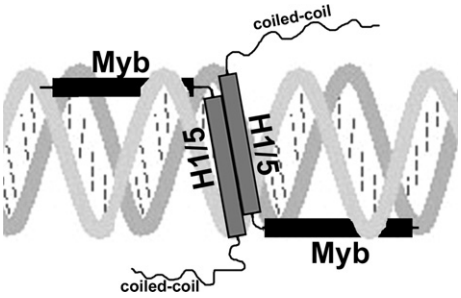


Fig. 7. Schematic view of binding of AtTRB1 to telomeric DNA. The Myb-domain primarily ensures direct sequence-specific binding of AtTRB1 to telomeric DNA, while the H1/5 domain may enhance this binding by protein dimerization and sequence-non-specific binding to DNA.

sequence-non-specific binding of the H1/5 domain promotes specific interaction of the Myb-domain with telomeric DNA, as illustrated in Figs. 2 and 4b. Moreover, the H1/5 domain of AtTRB proteins was recently shown to interact with the OB-fold domain of the AtPOT1b telomeric protein (Prochazkova Schrumpfova et al., 2008) suggesting yet another mode of AtTRB recruitment to telomeres. Interestingly, our results do not show any important role of the coiled-coil domain of AtTRB1 in the interactions studied, although this domain was originally suggested as responsible for protein-protein interactions (Marian et al., 2003). However, we cannot exclude that the coiled-coil domain participates in dimerization upon binding of the AtTRB1 to DNA, when the H1/5 domain becomes engaged in DNA-protein interaction, or in interaction with AtPOT1b. Our data thus reveal at least partial outlines of a plant analogy of the mammalian shelterin complex (de Lange, 2005).

We conclude that AtTRB1 is a telomeric-dsDNA-binding protein which is able to form multimers. The specific binding to telomeric DNA is primarily conducted by the Myb domain, while the multimerization and sequence-non-specific interactions with DNA are mediated by the H1/5 domain.

4. Experimental

4.1. Cloning, expression and purification of AtTRB1 and its fragments

The cDNA sequence of AtTRB1 (locus At1g49950) was obtained by RT-PCR from total RNA as described previously (Kuchar and Fajkus, 2004). In order to analyze the binding of AtTRB1 to telomeric DNA and to examine the contribution of its three individual major domains in the binding, AtTRB1 and its truncated variants (Fig. 1) were cloned into pET15b (Novagen) between the NdeI and BamHI sites using the following sets of primers:

AtTRB1 (full-length): A (5′-AGTATTCATATGGGTGCTCCTAAGCAGA-3′) + B (5′-CGGGATCCTCAGGCACGGATCATCATT-3′); AtTRB1-*m*: A + C (5′-GAGGATCCCTAAAGGGAGAACGTCCTTTT-3′); AtTRB1-*mh*: A + D (5′-TTGGATCCTCATCTCCTGTGGGAAGACAAGG-3′); AtTRB1-*h*: E (5′-AGTGTTCATATGGAAAATGTGGATGCGACCT-3′) + D, AtTRB1-*hc*: E + B.

Proteins were expressed by autoinduction in ZYM-5052 (Studier, 2005) in *E. coli* C41(DE3) (Miroux and Walker, 1996) grown for 3 h at 37 °C followed by 18 h at 20 °C. Expression of proteins of the correct molecular weight was confirmed by immunodetection. His-tagged proteins were purified by affinity FPLC (Fast Protein Liquid Chromatography) (Pharmacia ÄKTA purifier system) on BD TALON™ Superflow Resin (BD Biosciences) in 50 mM Na-phosphate buffer with 300 mM NaCl. A step gradient from 10 mM to 80 mM imidazole was used for elution. Protein concentration was estimated according to Bradford (1976).

4.2. Electrophoresis mobility shift assay (EMSA)

Oligonucleotides used in EMSA are described in Table 1. Telomeric oligonucleotides containing different numbers of telomere repeat units were adjusted to 28 nt by supplementing their telomeric part with a non-telomeric sequence (NonTR). Ds oligonucleotides were formed by annealing of complementary ss oligonucleotides. Oligonucleotides used as probes were end-labeled using [γ-^{32}P]ATP (MP Biomedicals) and polynucleotide kinase (NEB). 2.5–250 pmol of proteins (as indicated in Results)

Table 1
Oligonucleotides used in EMSA. The plant telomeric sequence is shown in bold

Name	Length (nt)	Sequence 5′-3′
AtTR2-G	28	**GGTTTAGGGTTTAG**AGTACCAGCCATGA
AtTR2-C	28	TCATGGCTGGTACT**CTAAACCCTAAACC**
AtTR3-G	28	**GGTTTAGGGTTTAGGGTTTAG**AGTACCA
AtTR3-C	28	TGGTACT**CTAAACCCTAAACCCTAAACC**
AtTR4-G	28	**GGTTTAGGGTTTAGGGTTTAGGGTTTAG**
AtTR4-C	28	**CTAAACCCTAAACCCTAAACCCTAAACC**
HuTR4-G	24	GGTTAGGGTTAGGGTTAGGGTTAG
HuTR4-C	24	CTAACCCTAACCCTAACCCTAACC
NonTR-G	28	CATCATGGCTGGTCATGGCTGGTACTAG
NonTR-C	28	CTAGTACCAGCCATGACCAGCCATGATG

were preincubated with 2.5, 50 or 250 pmol of competitor DNA for 10 min in 50 mM Na-phosphate buffer pH 8 + 200 mM NaCl; the labeled probe (2.5 pmol) was added, and the reaction was further incubated for 10 min at 25 °C. Total reaction volume was 20 μl. The reaction mixture was loaded onto an 8% w/v non-denaturing polyacrylamide gel (AA:BIS = 37:1, 0.25 × Tris-borate-EDTA (TBE) buffer). Electrophoresis was performed at 4 °C in 0.25 × TBE for 4 h at 120 V. Results were visualized on a STORM840 phosphoimager and evaluated using ImageQuant software (Molecular Dynamics). Protein-DNA interaction was quantified from the signal of unbound DNA probe and expressed as the percentage of labeled probe bound in a protein-DNA complex.

4.3. Polyacrylamide gel electrophoresis with perfluoro-octanoic acid (PFO-PAGE)

PFO-PAGE was performed according to Ramjeesingh et al. (1999) with a slight modification in the content of PFO in the running buffer (0.2% w/v compared to 0.5%). No DTT was used in the loading buffer. Ten micrograms of protein were used per lane. Electrophoresis in 8% or 9% gel (AA:BIS = 37:1) was carried out at 80 V for 3 h. Identical samples were simultaneously separated by SDS-PAGE to compare the denatured state of the proteins.

Acknowledgements

We thank Mrs. Jana Kapustová for excellent technical assistance and Professor Ronald Hancock, Laval University Cancer Res. Ctr., Quebec for critical reading of the manuscript. This work was supported by the Grant Agency of the Czech Republic (Project 204/08/H054 and 521/08/P452), Czech Ministry of Education (LC06004) and the institutional support (MSM0021622415 and AVOZ50040507).

References

Bertuch, A.A., Lundblad, V., 2006. The maintenance and masking of chromosome termini. Curr. Opin. Cell Biol. 18, 247–253.
Bianchi, A., Smith, S., Chong, L., Elias, P., de Lange, T., 1997. TRF1 is a dimer and bends telomeric DNA. EMBO J. 16, 1785–1794.
Bianchi, A., Stansel, R.M., Fairall, L., Griffith, J.D., Rhodes, D., de Lange, T., 1999. TRF1 binds a bipartite telomeric site with extreme spatial flexibility. EMBO J. 18, 5735–5744.
Bilaud, T., Koering, C.E., Binet-Brasselet, E., Ancelin, K., Pollice, A., Gasser, S.M., Gilson, E., 1996. The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. Nucl. Acids Res. 24, 1294–1303.
Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.
Chen, C.M., Wang, C.T., Ho, C.H., 2001. A plant gene encoding a Myb-like protein that binds telomeric GGTTAG repeats in vitro. J. Biol. Chem. 276, 16511–16519.
de Lange, T., 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev. 19, 2100–2110.
Ellen, T.P., van Holde, K.E., 2004. Linker histone interaction shows divalent character with both supercoiled and linear DNA. Biochemistry 43, 7867–7872.
Fairall, L., Chapman, L., Moss, H., de Lange, T., Rhodes, D., 2001. Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. Mol. Cell 8, 351–361.
Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., de Lange, T., 1999. Mammalian telomeres end in a large duplex loop. Cell 97, 503–514.
Hwang, M.G., Cho, M.H., 2007. Arabidopsis thaliana telomeric DNA-binding protein 1 is required for telomere length homeostasis and its Myb-extension domain stabilizes plant telomeric DNA binding. Nucl. Acids Res. 35, 1333–1342.
Hwang, M.G., Kim, K., Lee, W.-K., Cho, M.H., 2005. AtTBP2 and AtTRP2 in Arabidopsis encode proteins that bind plant telomeric DNA and induce DNA bending in vitro. Mol. Genet. Genomics 273, 66–75.
Karamysheva, Z.N., Surovtseva, Y.V., Vespa, L., Shakirov, E.V., Shippen, D.E., 2004. A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in Arabidopsis. J. Biol. Chem. 279, 47799–47807.
Kuchar, M., Fajkus, J., 2004. Interactions of putative telomere-binding proteins in *Arabidopsis thaliana*: identification of functional TRF2 homolog in plants. FEBS Lett. 578, 311–315.
Marian, C.O., Bordoli, S.J., Goltz, M., Santarella, R.A., Jackson, L.P., Danilevskaya, O., Beckstette, M., Meeley, R., Bass, H.W., 2003. The maize Single myb histone 1 gene, Smh1, belongs to a novel gene family and encodes a protein that binds telomere DNA repeats in vitro. Plant Physiol. 133, 1336–1350.
Miroux, B., Walker, J.E., 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Mol. Biol. 260, 289–298.
Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., Wu, J.R., 1988. A highly conserved repetitive DNA sequence, $(TTAGGG)_n$, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA 85, 6622–6626.
Prochazkova Schrumpfova, P., Kuchar, M., Palecek, J., Fajkus, J., 2008. Mapping of interaction domains of putative telomere-binding proteins AtTRB1 and AtPOT1b from *Arabidopsis thaliana*. FEBS Lett. 582, 1400–1406.
Ramjeesingh, M., Huan, L.J., Garami, E., Bear, C.E., 1999. Novel method for evaluation of the oligomeric structure of membrane proteins. Biochem. J. 342, 119–123.
Richards, E.J., Ausubel, F.M., 1988. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53, 127–136.
Rotkova, G., Sklenickova, M., Dvorackova, M., Sykorova, E., Leitch, A.R., Fajkus, J., 2004. An evolutionary change in telomere sequence motif within the plant section Asparagales had significance for telomere nucleoprotein complexes. Cytogenet. Genome Res. 107, 132–138.
Rotková, G., Sýkorová, E., Fajkus, J., 2007. Characterization of nucleoprotein complexes in plants with human-type telomere motifs. Plant Physiol. Biochem. 45, 716–721.
Shakirov, E.V., Surovtseva, Y.V., Osbun, N., Shippen, D.E., 2005. The Arabidopsis Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. Mol. Cell Biol. 25, 7725–7733.
Schrumpfova, P., Kuchar, M., Mikova, G., Skrisovska, L., Kubicarova, T., Fajkus, J., 2004. Characterization of two *Arabidopsis thaliana* myb-like proteins showing affinity to telomeric DNA sequence. Genome 47, 316–324.
Studier, F.W., 2005. Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. 41, 207–234.
Sue, S.-C., Hsiao, H.-H., Chung, B.C.P., Cheng, Y.-H., Hsueh, K.-L., Chen, C.M., Ho, C.H., Huang, T.-H., 2006. Solution structure of the *Arabidopsis thaliana* telomeric repeat-binding protein DNA binding domain: a new fold with an additional C-terminal helix. J. Mol. Biol. 356, 72–85.
Sykorova, E., Lim, K.Y., Kunicka, Z., Chase, M.W., Bennett, M.D., Fajkus, J., Leitch, A.R., 2003. Telomere variability in the monocotyledonous plant order Asparagales. Proc. Biol. Sci. 270, 1893–1904.
Yang, S.W., Kim, D.H., Lee, J.J., Chun, Y.J., Lee, J.-H., Kim, Y.J., Chung, I.K., Kim, W.T., 2003. Expression of the telomeric repeat binding factor gene NgTRF1 is closely coordinated with the cell division program in tobacco BY-2 suspension culture cells. J. Biol. Chem. 278, 21395–21407.
Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G., Li-Jia, Q., 2006. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. 60, 107–124.
Yu, E.Y., Kim, S.E., Kim, J.H., Ko, J.H., Cho, M.H., Chung, I.K., 2000. Sequence-specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1. J. Biol. Chem. 275, 24208–24214.

BJ www.biochemj.org

BJ Plant

Biochem. J. (2009) **419**, 221–228 (Printed in Great Britain) doi:10.1042/BJ20082195

online data

Single-Myb-histone proteins from *Arabidopsis thaliana*: a quantitative study of telomere-binding specificity and kinetics

Ctirad HOFR*[1], Pavla ŠULTESOVÁ*, Michal ZIMMERMANN*, Iva MOZGOVÁ*, Petra PROCHÁZKOVÁ SCHRUMPFOVÁ*, Michaela WIMMEROVÁ† and Jiří FAJKUS*‡[1]

*Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, CZ-62500 Brno, Czech Republic, †National Centre for Biomolecular Research and Department of Biochemistry, Faculty of Science, Masaryk University, CZ-61137 Brno, Czech Republic, and ‡Laboratory of DNA–Molecular Complexes, Institute of Biophysics, Czech Academy of Sciences, CZ-61265 Brno, Czech Republic

Proteins that bind telomeric DNA modulate the structure of chromosome ends and control telomere function and maintenance. It has been shown that AtTRB (*Arabidopsis thaliana* telomere-repeat-binding factor) proteins from the SMH (single-Myb-histone) family selectively bind double-stranded telomeric DNA and interact with the telomeric protein AtPOT1b (*A. thaliana* protection of telomeres 1b), which is involved in telomere capping. In the present study, we performed the first quantitative DNA-binding study of this plant-specific family of proteins. Interactions of full-length proteins AtTRB1 and AtTRB3 with telomeric DNA were analysed by electrophoretic mobility-shift assay, fluorescence anisotropy and surface plasmon resonance to reveal their binding stoichiometry and kinetics. Kinetic analyses at different salt conditions enabled us to estimate the electrostatic component of binding and explain different affinities of the two proteins to telomeric DNA. On the basis of available data, a putative model explaining the binding stoichiometry and the protein arrangement on telomeric DNA is presented.

Key words: *Arabidopsis thaliana*, fluorescence anisotropy, kinetics, single-Myb-histone protein (SMH protein), surface plasmon resonance, telomere protein–DNA interaction.

INTRODUCTION

Telomeres are nucleoprotein complexes consisting of repetitive DNA sequences, general chromatin proteins and telomere-specific proteins. Tandem repeats of telomeric DNA are short T- and G-rich sequences, such as d(GGGTTA) in humans and d(GGGTTTA) in the majority of plants.

Telomeres form protective capping structures at the ends of chromosomes [1]. These structures are essential for cell viability as they prevent chromosomes from unwanted end-to-end joining and recognition of chromosome tips as unrepaired double-strand breaks by the repair system of the cell. Changes in telomere structure and function induce chromosomal abnormalities and are directly connected with human aging and cancer [2].

Telomeres are usually maintained by telomerase, a ribonucleoprotein enzyme that adds telomeric repeats to the 3′-overhang of the G-rich DNA strand. The action of telomerase is regulated by its expression and by numerous proteins that control telomerase access to telomeres and organize telomeres into specific capping structures, such as telomeric loops that were observed in a number of organisms, including humans and plants [3,4].

Three DNA-binding proteins have been found to be responsible for specific recognition and direct interactions with the telomeric repeat sequence in humans. Two of them, TRF1 and TRF2 (where TRF is telomeric repeat-binding factor), described as negative regulators of telomere length [5], show substantial structural similarity and bind double-stranded telomeric DNA. The third protein, POT1 (protection of telomeres 1), binds the G-rich strand of telomeric DNA, participates in chromosome capping and is able to control telomere extension by telomerase, both positively and negatively [6,7]. The human TRFs and their homologues in other organisms possess a well conserved DNA-binding structural motif similar to the c-Myb-family of transcriptional activators [8]. The Myb domain of TRFs is C-terminally positioned and consists of three helices connected in a helix–turn–helix manner. The third helix contains a conserved amino acid sequence called a 'telobox', which has been shown to be important for recognition of telomeric double-stranded DNA [8].

Numerous TRF-like proteins have been identified in plants (reviewed in [9]), and, in a few cases, the influence of these proteins on telomere length homoeostasis has been demonstrated [10,11]. Interestingly, besides the TRF-like proteins, a plant-specific family of other telobox proteins has been described [12]. This group of proteins, termed the SMH (single-Myb-histone) family, is characterized by a triple-domain structure consisting of an N-terminal Myb domain, central globular histone H1/5 domain, and a C-terminal coiled-coil domain. In *Arabidopsis thaliana* (At), five SMH proteins were identified (AtTRB1–AtTRB5, where TRB is telomere-repeat-binding factor) [12], and three of them have been characterized [13,14]. These proteins show not only specific interactions with telomeric DNA, but also a number of protein–protein interactions functionally related to telomeres. In addition to their ability to form homodimers (similarly to TRFs), they can also form heterodimers and both homo- and heterotypic multimers [13–15] via their H1/5 histone domain. They also interact (using the same H1/5 domain) with one of the POT1 proteins in *A. thaliana*, AtPOT1b [15,16], which participates in telomere capping [17].

The emerging complexity of interactions of AtTRBs urges more detailed and quantitative studies of their DNA–protein and protein–protein interactions to reveal principles of their regulatory role. So far, only structural data for the Myb DNA-binding

Abbreviations used: At, *Arabidopsis thaliana*; EMSA, electrophoretic mobility-shift assay; FA, fluorescence anisotropy; LB, Luria–Bertani; POT1, protection of telomeres 1; RedX, Rhodamine Red-X; RT, reverse transcription; SMH, single-Myb-histone; SPR, surface plasmon resonance; TRB, telomere-repeat-binding factor; TRF, telomeric repeat-binding factor.

[1] Correspondence may be addressed to either of these authors (email hofr@sci.muni.cz or fajkus@sci.muni.cz).

domain are available [18]. Similarly, kinetic studies are limited to the interaction of a Myb-domain-bearing fragment with a short telomeric DNA oligonucleotide (13 bp) [18], and a non-equilibrium technique was used to describe binding kinetics of TRFs in rice [19]. In order to describe binding interactions more thoroughly, associations of the full-length proteins with telomeric DNA need to be evaluated.

The equilibrium binding kinetics of the full-length proteins can be studied by quantitative biophysical approaches. The binding of proteins to fluorescently labelled DNA may be monitored by FA (fluorescence anisotropy). This method gives well-resolved binding isotherms at different buffer conditions and therefore reliable kinetic and energetic parameters of binding. If the solution contains only free DNA molecules, FA is relatively low, owing to the fast rotational rearrangement of DNA molecules. After the binding of protein to DNA, a bulky slower-rotating protein–DNA complex is formed and the anisotropy is increased. Thus the anisotropy change of fluorescently labelled DNA duplexes, after each addition of protein into solution, describes the extent of protein–DNA binding [20,21].

In the present paper, we report a detailed study to reveal stoichiometry and kinetics of AtTRB1 and AtTRB3 binding to telomeric DNA. Proteins AtTRB1 and AtTRB3 have been chosen for these functional assays because they showed the highest structural stability within the AtTRB family of proteins. Interactions of full-length proteins with telomeric DNA are analysed by a combination of EMSA (electrophoretic mobility-shift assay) and quantitative biophysical methods employing FA and SPR (surface plasmon resonance). Kinetic analyses at different salt conditions enable us to estimate the electrostatic component of binding and explain different affinities of the two AtTRBs to telomeric and non-telomeric DNA. The kinetic measurements also contribute to the estimation of the length of double-stranded DNA for proper protein binding. On the basis of these data, a speculative model for binding stoichiometry and protein arrangement on telomeric DNA is presented.

EXPERIMENTAL

Cloning, expression and purification of AtTRB1 and AtTRB3

The cDNA sequence of AtTRB1 (locus At1g49950) was obtained by RT (reverse transcription)–PCR from total RNA as described previously [16]. AtTRB1 has been cloned into pET15b vector (Novagen) and expressed as a His-tagged fusion protein in *Escherichia coli* C41(DE3) cells [14]. The cells were grown on LB (Luria–Bertani) medium with ampicillin (100 μg/ml) at 37 °C overnight. The next day, cells were diluted 20-fold into ZYM 5052 complex autoinducing medium containing ampicilin [22]. The cells were incubated at 37 °C for 5 h. Then the temperature was set to 20 °C, and the incubation continued overnight.

The cDNA sequence of AtTRB3 (locus At3g49850) was obtained by RT–PCR from total RNA as described previously [13]. AtTRB3 has been cloned into pET30a(+) vector (Novagen) and expressed as a His-tagged fusion protein in *E. coli* BL21(DE3)pLysS cells. The cells were grown on LB medium with kanamycin (50 μg/ml) at 37 °C for 4 h. At a D_{600} of 0.6, the overexpression of AtTRB3 was induced by the addition of IPTG (isopropyl β-D-thiogalactoside) to a concentration of 1 mM. After lowering the incubation temperature to 25 °C, the growth continued for an additional 3 h.

The following extraction and purification steps were the same for both recombinant proteins. After harvesting by centrifugation at 8000 ***g*** for 8 min, the pellet was dissolved in buffer containing 50 mM sodium phosphate (pH 8.0) with 300 mM NaCl and 10 mM imidazole and was sonicated for 5 min. The sonicated cell extract was cleared by centrifugation at 14 000 rev./min for 1 h at 4 °C using a Beckman JA 14 rotor and subsequent filtration (0.45 μm filter). Affinity purification was performed on a column filled with a TALON® metal-affinity resin (BD Biosciences). Protein was eluted at 80 mM imidazole. The eluent was loaded on to a heparin HiTrap™ column (GE Healthcare). A concentration gradient of NaCl from 0.4 to 1 M NaCl was used for protein elution. The fractions containing pure protein were concentrated, and buffer-exchanged usually into 50 mM sodium phosphate (pH 7.5) with 100 mM NaCl by ultrafitration (Amicon 10K, Millipore) or by extensive dialysis. A typical yield was 1 mg of purified protein per litre of bacterial culture. The concentration of purified protein was determined using the Bradford assay [23].

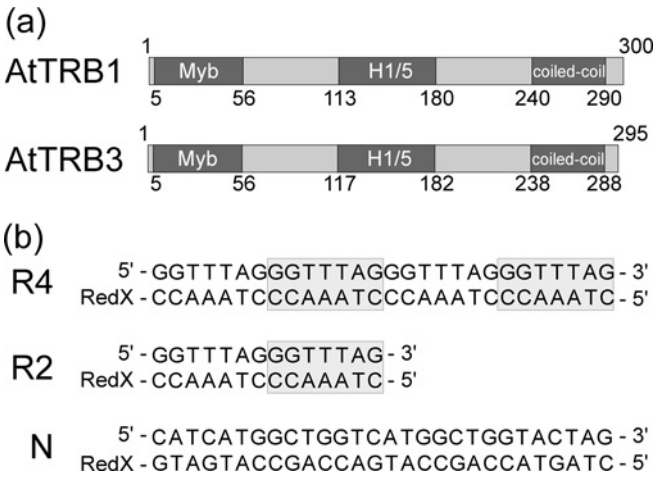


Figure 1 Proteins and oligonucleotide duplexes used for binding studies

(**a**) Organization of the AtTRB1 and AtTRB3 polypeptide chains. The localization of the Myb domain, histone-like H1/5 domain and coiled-coil domain is shown together with numbers denoting their positions in the sequence. (**b**) Base sequence of telomeric oligonucleotide duplex R4 and R2 along with non-telomeric duplex N. RedX denotes fluorescent label Rhodamine RedX. The nucleotides of putative Myb-domain-binding sites are shaded grey [18].

DNA substrates

Oligodeoxynucleotides were synthesized and HPLC-purified by Core Laboratory at Masaryk University. One of the strands in the duplexes was synthesized with the 3′-end C_6 aminoalkyl linker and labelled with RedX (Rhodamine Red-X) (Molecular Probes) using the protocol provided by the manufacturer. The duplexes comprising four and two telomeric repeats were denoted as R4 and R2 respectively. The DNA duplex with non-telomeric sequence was denoted as N. The molar absorption coefficients of the single strands were estimated with the employment of phosphate assay [24]. Molar absorption coefficients were 281 000 (RedX-labelled strand in R4), 278 000 (complementary strand in R4), 148 000 (RedX-labelled strand in R2), 140 000 (complementary strand in R2), 284 000 (RedX-labelled strand in N) and 265 000 $M^{-1} \cdot cm^{-1}$ (complementary strand in R2) for DNA oligonucleotides shown in Figure 1.

EMSA

Protein–DNA-binding reactions were performed in 10 μl volumes containing the same amount of labelled DNA duplex (30 pmol) and various concentrations of protein (0–180 pmol) in 50 mM sodium phosphate (pH 7.0) with 200 mM NaCl.

Reaction mixtures were incubated for 10 min at 25 °C. Protein–DNA complexes were resolved on horizontal 7.5 % (w/v) acrylamide/0.3 % bisacrylamide gels, as described in [25]. The electrophoresis proceeded at 1.5 V/cm for 30 min and for an additional 90 min at 3 V/cm. Gels were analysed with a LAS 3000 imaging system (Fujifilm). After the fluorescence imaging, Coomassie Blue staining of the gel was performed to reveal protein-containing bands in the gel.

Fluorescence anisotropy

Florescence anisotropy was measured on a FluoroMax-4 spectrofluorimeter (Horiba) with an L-format set up under control of an Origin-based FluorEssence software (version 2.1.6). Excitation and emission wavelengths were 572 and 591 nm respectively, with the same excitation and emission bandpath, 8 nm. The integration time was 3s. For each anisotropy value, five measurements were averaged. The titration experiments were carried out in a 10 mm × 4 mm quartz-glass cuvette with a magnetic bar stirrer. All measurements were conducted at 25 °C in 50 mM sodium phosphate buffer (pH 7.5) containing 100 mM NaCl if not stated otherwise. To 1500 μl of DNA solution (20 nM) in the buffer, protein solution was added stepwise. The decrease in DNA concentration during the titration was taken into account in the analysis of the data. A control titration of protein to RedX solution (without DNA) has been performed to confirm that there was no interaction between RedX and protein.

Dissociation constants of protein binding were evaluated by fitting of dilution-corrected binding isotherms using programs SigmaPlot 8 (Systat Software) and DynaFit3 (version 3.28) [26]. Analysis of the binding of protein to DNA duplexes was performed with the assumption of a non-co-operative binding mode. The association constants were calculated as reciprocal values of dissociation constants ($K_a = 1/K_d$). The association constants provided the free energies of association.

Electrostatic component of binding

In order to determine the contribution of electrostatic interactions upon binding of DNA with protein, the equilibrium binding constant was measured at different concentrations of NaCl (see Figure 4 and Table 2). The electrostatic component of binding originates from the formation of ion pairs between the cationic amino acid residues of the protein and the negatively charged DNA. The number of ion pairs formed upon protein–DNA binding and corresponding electrostatic contribution to overall binding affinity (K_a) could be derived from the dependence of the binding constant on salt concentration according to the eqn (1):

$$\log K_a = \log K_a^{nel} - Z\varphi \cdot \log [NaCl] \quad (1)$$

where Z is the number of DNA phosphates that interact with the protein, φ is the number of Na^+ cations per phosphate released upon protein binding. For B–DNA duplexes of 24 bp and shorter, the value for φ is approx. 0.64 [27]. The right-hand side of the equation divides overall binding affinity into the non-electrostatic part described by K_a^{nel} and a salt-dependent electrostatic part [28,29]. When the linear dependence of $\log K_a$ is extrapolated to the salt concentration of 1 M, the electrostatic term in eqn (1) can be removed: $\log K_a = \log K_a^{nel}$, i.e. the binding affinity is given only by non-electrostatic interactions. Similarly to the binding affinity, the overall binding energy defined as $\Delta G_a = -2.3RT \cdot \log K_a$ could be divided into electrostatic and non-electrostatic terms $\Delta G_a = \Delta G_a^{nel} + \Delta G_a^{el}$. The electrostatic term ΔG_a^{el} disappears when the salt concentration approaches 1 M and the overall energy of binding is given only by the non-electrostatic term, $\Delta G_a = \Delta G_a^{nel} = -2.3RT \cdot \log K_a^{nel}$.

Surface plasmon resonance

Sensorgrams were recorded on a Biacore 3000 instrument (GE Healthcare) using CM5 chips. More details are available in the Supplementary Online Data at http://www.BiochemJ.org/bj/419/bj4190221add.htm.

RESULTS

Stoichiometry of protein–DNA complexes

In order to estimate the binding ratio of AtTRBs and DNA, oligonucleotide substrates containing two or one putative binding sites were designed. The telomeric duplex R4 covers the length of four plant telomeric repeats and comprises two putative Myb-domain-binding sites. The shorter duplex, double-stranded DNA fragment R2, consists of two telomeric repeats and contains one Myb-domain-binding site. For comparative purposes, oligonucleotide duplex N, as a representative of non-telomeric DNA, was used in the present study (Figure 1).

Both AtTRB1 and AtTRB3 bind telomeric DNA with the stoichiometry of one protein monomer per one telomeric repeat

The binding stoichiometry was analysed by EMSA with samples containing a variable protein/DNA ratio. Figure 2 shows fluorescently visualized bands indicating the mobility of free and protein-bound DNA duplexes in non-denaturating acrylamide gels.

Increasing the concentration of protein shifted the free labelled DNA duplex to a new position corresponding to a protein–DNA complex. The band corresponding to the free duplex R4 disappeared when the AtTRB1/R4 ratio was 4:1 (Figure 2a). Similarly, the complete binding of AtTRB3 to substrate R4 was observed at the same protein/DNA ratio (Figure 2b). Both AtTRB1 and AtTRB3 bind telomeric DNA with the stoichiometry of one protein monomer per one telomeric repeat.

In order to characterize interaction stoichiometry of AtTRBs with telomeric DNA further, proteins were allowed to interact with the shorter substrate R2 bearing two telomeric repetitions (Figure 2c). The results of EMSA with R2 demonstrate that a 2-fold decrease in the length of DNA reduces the protein/DNA binding ratio proportionally. These results confirmed that the stoichiometry of binding is one monomer of AtTRB1 or AtTRB3 per one telomeric repeat. If we consider binding of protein in dimeric form, as was shown in our recent study [14], then two protein dimers bind one R4 substrate (four telomeric repeats) or, in other words, one dimer of AtTRB binds the fragment R2 (two telomeric repeats). On the basis of these data, we could rephrase our initial statement regarding stoichiometry to the following form: one dimer of AtTRB binds the region of two telomeric repeats.

AtTRB1 shows the same binding stoichiometry for telomeric and non-telomeric DNA sequences, whereas AtTRB3 exhibits different binding capacities for telomeric and non-telomeric DNA sequences

The effect of DNA sequence on binding ability of AtTRB1 and AtTRB3 was analysed by comparing the protein/DNA ratio needed for complete saturation of telomeric R4 and non-telomeric N substrate. In this respect, AtTRB1 behaves similarly in both cases; the binding stoichiometry of AtTRB1 remained the same, as demonstrated in Figure 2(a).

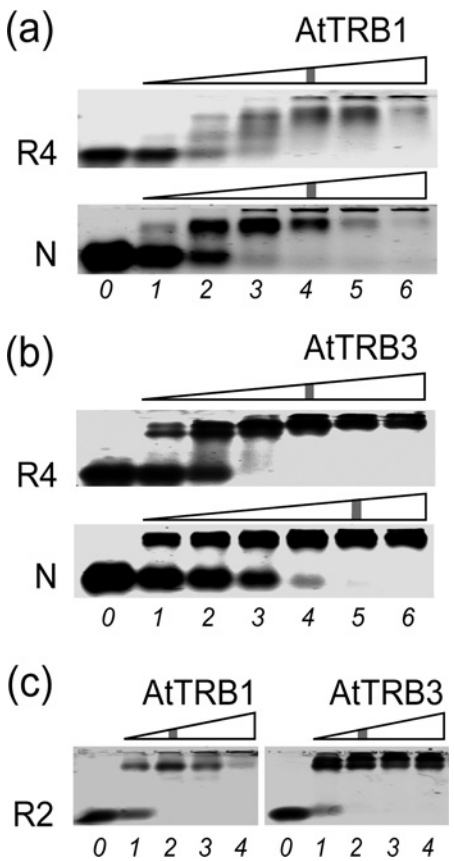


Figure 2 Non-denaturing EMSA

(**a**) AtTRB1 binding to fluorescently labelled oligonucleotide R4 with telomeric sequence and oligonucleotide N with non-telomeric sequence. (**b**) AtTRB3 binding to DNA oligonucleotides R4 and N. (**c**) AtTRB1 or AtTRB3 binding to oligonucleotide R2 with the sequence of two telomeric repetitions. The DNA oligonucleotides and AtTRBs were incubated with increasing amounts of protein. The numbers under electrophoretic lanes denote the stoichiometric protein/DNA ratio. The protein/DNA ratio corresponding to binding saturation is indicated with a grey line.

In contrast, AtTRB3 exhibits a markedly stronger dependence of the binding ability on DNA sequence that was manifested by a shift in ratio needed for saturation of the non-telomeric substrate N. The protein/DNA ratio was shifted to the higher values ($> 5:1$) in the case of duplex N than was the ratio for the telomeric duplex R4 (Figure 2b).

The difference in DNA-sequence-dependent saturation might be a result of different binding kinetics of AtTRB1 and AtTRB3. To assess this possibility, direct kinetic measurements were performed using FA.

Binding kinetics

The binding affinity of AtTRB variants to double-stranded DNA was analysed further by FA measurements. In these measurements, protein aliquots were added to the solution of labelled DNA duplex, and an increase of FA was observed. The equilibrium dissociation constants (K_d) obtained by analyses of anisotropy curves for binding are listed in Table 1.

AtTRB1 and AtTRB3 bind telomeric DNA with high affinity and specificity

The binding affinity of AtTRB1 to telomeric DNA is significantly higher in comparison with the binding to non-telomeric DNA. The titration curves obtained for AtTRB1 binding to DNA substrates

Table 1 Dissociation and association constants for binding of AtTRB1 and AtTRB3 to DNA

Values are means ± S.E.M. for three independent experiments in 50 mM sodium phosphate (pH 7.5) and 100 mM NaCl measured at 25 °C.

	R4		N		R2	
Protein	K_d (nM)	K_a (10^{-6} M^{-1})	K_d (nM)	K_a (10^{-6} M^{-1})	K_d (nM)	K_a (10^{-6} M^{-1})
AtTRB1	90 ± 20	11.0	1200 ± 300	0.83	210 ± 30	4.8
AtTRB3	400 ± 60	2.5	2900 ± 300	0.35	800 ± 100	1.3

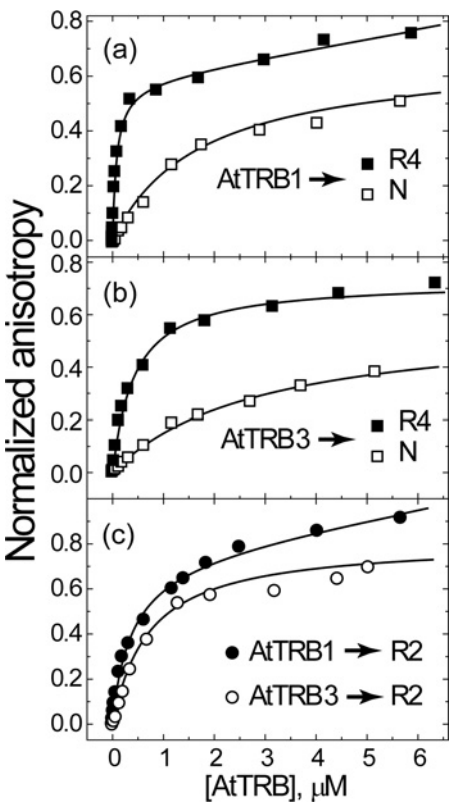


Figure 3 Binding of AtTRB1 and AtTRB3 to DNA duplexes

(**a**) FA measurements of binding of AtTRB1 to telomeric duplex R4 or non-telomeric duplex N. The binding at 20 nM DNA occurred in buffer containing 50 mM sodium phosphate (pH 7.5) and 100 mM NaCl. (**b**) FA measurements of binding of AtTRB3 to R4 or N duplex. Binding conditions were the same as in (**a**). (**c**) Binding isotherms of AtTRB1 and AtTRB3 with telomeric duplex R2 measured by FA. Binding conditions were the same as in (**a**).

R4 and N are shown in Figure 3(a). As expected, AtTRB1 shows significantly higher binding affinity to telomeric R4 than to the non-telomeric N DNA substrate. This can be clearly seen from the steeper rise of the curve corresponding to binding telomeric DNA. The evaluation of binding curves revealed K_d values of 90 and 1200 nM for R4 and N substrate respectively (Table 1). Comparison of dissociation constants thus demonstrates more than 13-fold higher affinity and binding specificity of AtTRB1 to DNA bearing telomeric sequences.

The binding affinity of AtTRB3 to telomeric sequence is higher in comparison with the binding to non-telomeric sequence, but the difference is less pronounced than in case of AtTRB1. AtTRB3 was allowed to bind either the telomeric substrate R4 or the non-telomeric duplex N (Figure 3b). The K_d values for the binding of AtTRB3 to R4 and N were 400 and 2900 nM respectively.

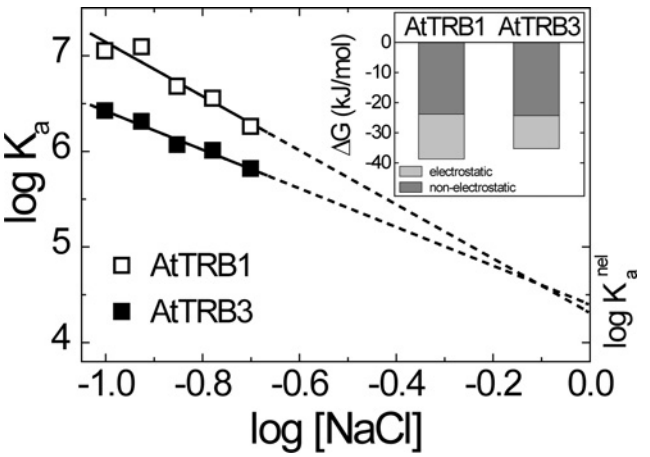


Figure 4 Dependence of the association constants for binding of AtTRB1 and AtTRB3 to substrate R4 on NaCl concentration

The inset shows the electrostatic and non-electrostatic components of the free energy of association of AtTRB1 or AtTRB3 with substrate R4.

AtTRB3 shows more than 7-fold higher affinity to telomeric DNA duplex than to non-telomeric DNA.

The absolute value of the dissociation constant was verified by SPR

In order to confirm the absolute values of binding constants obtained using FA, a reverse-order experiment was performed using SPR. In this experiment, AtTRB3 was immobilized on the chip surface, and duplex R4 was allowed to bind. The reverse arrangement of the SPR experiment changes interaction stoichiometry (one DNA duplex interacts with one immobilized protein, whereas four protein monomers bind one DNA duplex during FA measurements). This had been considered when the equilibrium binding constant was evaluated. The output of the non-linear fitting of SPR curves for different concentrations of DNA produces a K_d of 1700 nM, which agrees with the value determined previously with a factor of 2 at a similar salt concentration (see the Supplementary Online Data).

AtTRB1 and AtTRB3 show reduced binding affinity to R2 when compared with binding affinity to R4

When the length of DNA duplex is shortened from four to two telomeric repeats, the binding affinity decreases to the level of binding affinity recorded for the non-telomeric DNA. Even though there is one putative binding site present on the duplex R2, the binding affinity of AtTRB1 is quite low and is characterized by a K_d similar to that obtained for binding to duplex N. The shortening of telomeric DNA substrate has a similar effect on binding affinity of AtTRB3 (Table 1). The length reduction of telomeric DNA substrate thus results in a substantial fall in the binding affinity of both AtTRB1 and AtTRB3.

Electrostatic contribution to binding affinity

The binding of AtTRB1 or AtTRB3 to duplex R4 containing two putative binding sites induces the formation of four or three ion pairs respectively. Binding of both proteins to the substrate R4 was measured at different concentrations of NaCl. The change of binding parameters is set out in the double-log-plot of the association constants against salt concentration (Figure 4 and Table 2). From the slope, the parameter Z was calculated. Z denotes the number of newly formed ionic bonds between protein and DNA. This number is 4 (after rounding) for binding of AtTRB1, and 3 for AtTRB3. Thus approx. four ion pairs are formed upon binding of AtTRB1 and approx. three ion pairs upon binding of AtTRB3 to the telomeric DNA.

Table 2 Salt-concentration-dependence of association constants for binding of AtTRB1 and AtTRB3 to R4

Values are means ± S.E.M.

Protein	[NaCl] (mM)	log K_a	δlog K_a/δlog [NaCl]	log K_a^{nel}	Z
AtTRB1	100	7.04	2.8 ± 0.2	4.31 ± 0.2	4.4
	119	7.08			
	141	6.67			
	167	6.54			
	200	6.25			
AtTRB3	100	6.41	2.0 ± 0.1	4.4 ± 0.1	3.2
	119	6.31			
	141	6.06			
	167	6.00			
	200	5.81			

The binding energy is provided mainly by a non-electrostatic component in the case of both AtTRBs

Further evaluation of the salt-dependent binding constant was performed to obtain the non-electrostatic contribution to the binding affinity. The electrostatic and non-electrostatic components of the binding energy for AtTRB1 or AtTRB3 to R4 are shown in the inset of Figure 4. It is notable that the non-electrostatic components of binding energy ΔG_a^{nel} for the two proteins are identical within error range with magnitudes of 25 kJ · mol^{-1} for both AtTRB1 and AtTRB3. If this value is compared with the values of the overall binding energy 40 kJ · mol^{-1} for AtTRB1 and 37 kJ · mol^{-1} for AtTRB3, it can be concluded that the non-electrostatic interactions contribute to the total energy of binding by approx. 60 % for AtTRB1 and by approx. 70 % for AtTRB3. Hence, it is apparent that the major part of the binding energy originates from the non-electrostatic interactions.

The greater electrostatic component is responsible for a more favourable overall binding energy of AtTRB1 compared with AtTRB3

Further inspection of calculated energetic data allowed us to identify the main reason for different binding affinities between these similar proteins. It is demonstrated that the kinetics of protein–DNA interactions are different because of the electrostatic term of the binding energy (inset in Figure 4). In other words, the difference in the total binding energy for AtTRB1 and AtTRB3 is entirely given by the change in the electrostatic component of binding.

DISCUSSION

Kinetics and stoichiometry of binding

The present study shows that binding of AtTRB1 and AtTRB3 with the telomeric DNA proceeds with the stoichiometry of one protein monomer per one telomeric repeat. A higher protein/DNA ratio was observed only in case of AtTRB3 binding to non-telomeric DNA (Figure 2b). The shift in the ratio can be explained by the observed lower affinity of AtTRB3 for non-telomeric DNA. The decrease in binding affinity with the change from telomeric to non-telomeric sequence was confirmed also by our kinetic measurements (Table 1). All recently characterized AtTRBs form tightly bound homo- and hetero-dimers and multimers [14,15]. Relatively strong mutual interactions of AtTRBs were also

verified independently using SPR (results not shown) and their dimerization ability was demonstrated by gel chromatography (see the Supplementary Online Data). Therefore the feasibility of protein dimerization and stoichiometric data of the present study support the assumption that AtTRBs bind to DNA in dimeric form. In this respect, the AtTRBs behave similarly to human TRF1 and TRF2 [30–32], with the exception that TRFs do not form heterodimers.

Surprisingly, the affinity of AtTRB1 to telomeric substrate R4 is 4-fold higher than that of AtTRB3, although AtTRB1 and AtTRB3 are relatively similar in their primary sequences.

Interestingly, it has been found that K_d values observed in the present study for AtTRB1 and AtTRB3 correspond very well to K_d values obtained for the DNA-binding domain of human TRF1 and TRF2 when interacting with telomeric DNA [33]. Moreover, similarly to AtTRB1 and AtTRB3, human TRF1 binds telomeric DNA with a 4-fold higher affinity than that of TRF2.

In order to explain potential reasons for the different binding manner of AtTRB1 and AtTRB3, we compared our findings with available equilibrium kinetic data for the binding of the Myb domain. The K_d obtained for the binding of the Myb domain alone to telomeric DNA from NMR studies was in the range of 1 μM [18]. If we compare this value measured at physiological salt concentration with the values for the binding of full-length proteins measured in the present study at a corresponding NaCl concentration, the magnitude of K_d for AtTRB3 is slightly lower at 0.9 μM (see Table 1), and the K_d for AtTRB1 is significantly lower (0.2 μM). Both full-length proteins showed higher binding capacities than that reported for a Myb domain alone. Since the Myb domain sequence is highly conserved between AtTRB1 and AtTRB3, the higher binding affinity of AtTRB1 should originate from another part of the protein. The domain that may contribute to the tuning of binding affinity of AtTRBs to DNA is the H1/5 domain [13], as supported by our recent findings [14]. The conservation of the H1/5 domain between AtTRB1 and AtTRB3 is lower than that of the Myb domain and differs in a way that might allow the corresponding protein region to adopt a structure with a different net charge on the surface. The surface net charge is important for a long-range non-specific electrostatic attraction among proteins and DNA, whereas non-electrostatic interactions that are important for specific recognition of a DNA sequence comprise hydrogen bonds between outer groups of DNA and polar residues of the protein [18].

Electrostatic component of binding

Proteins controlling and regulating nucleic acid structure and function usually show both sequence-non-specific binding to DNA and a higher-affinity binding of their specific physiological DNA target. In general, protein–DNA binding takes place in two steps. In the first step, a non-specific, mainly electrostatic, binding to the phosphate backbone occurs; in the second step, the protein explores the DNA surface for specific non-electrostatic interactions such as hydrogen bonds [34].

Different contributions of electrostatic and non-electrostatic interactions to binding were observed for different classes of DNA-binding proteins. For example, telomere-binding protein α from *Oxytricha nova* induces the formation of two ion pairs upon binding to DNA, and the electrostatic contribution to the free energy of binding is approx. 15 % [25]. On the other hand, proteins containing a strongly positively charged scissor-grip motif for DNA recognition induce the formation of six ion pairs with the electrostatic contribution to the total free energy of binding being 45 % [29].

The different contribution of electrostatic attraction for binding of AtTRB1 and AtTRB3 was observed. We estimated the number of four and three ion pairs upon AtTRB1 or AtTRB3 binding to R4 and the corresponding electrostatic contribution to the total free energy of binding at 40 and 30 % respectively. This correlates well with data available for electrostatic interactions of other DNA-binding proteins. The DNA-binding event of AtTRBs is driven mainly by non-electrostatic interactions. On the whole, our results show that AtTRBs bind telomeric DNA primarily in a sequence-specific manner that is essential for the recognition of binding sites within telomeric DNA.

Kinetic data contribute to understanding of nucleoprotein complex arrangement

Analyses of our kinetic data together with available structural data may be also used to elucidate the arrangement of nucleoprotein complexes of AtTRBs with telomeric DNA.

In general, one might suppose that the same binding preferences to telomeric DNA are given primarily by the occurrence of the recognition sequence in DNA. For this reason, one would also expect the same binding kinetics for the telomeric DNA with one or two putative binding sites under the consideration of a non-co-operative independent binding. As follows from the previous assumptions, the duplex R2, containing one binding site, should have reached the saturation of binding sites faster (K_d would be lower) when compared with that for duplex R4, with two binding sites. However, our data show the opposite. The binding affinity of both examined proteins to duplex R2 is lower (K_d is shifted to higher values) than in the case of binding to R4. Our quantitative kinetics results confirmed a previously reported decrease in binding affinity of AtTRBs with the shortening of telomeric DNA substrate [13]. Moreover, the lower affinity to DNA containing only one putative binding site might be an indication of an insufficient space for the binding of an active protein. Importantly, it has been shown that the minimum length of DNA for Myb domain binding is approx. 13 bp [18]. If AtTRBs interacted with the DNA exclusively through the Myb domain and binding sites were positioned suitably within the sequence, the length of R2 duplex (14 bp) should have been sufficient for proper binding without a change in binding affinity. Since a significant fall in binding affinity was observed, the kinetics data suggest that there is also another domain taking part in the interaction. As a result, the binding affinity of AtTRBs to the 14 bp long and 28 bp long DNA duplex differs substantially. In our recent results, the H1/5 domain promotes interaction with DNA [14]. Presumably, the short length might prevent the H1/5 domain from properly interacting with the DNA. Hence, the constrained binding without H1/5 domain might be the main reason for the reduction of the overall binding affinity to substrate R2.

Although the picture of a molecular mechanism controlling telomerase activity is far from complete, it is important to consider how the protein-binding events measured in the present study relate to structural arrangements and subsequent interactions essential for the biology of telomeres. If we take into account the kinetic data and the dimerization ability of AtTRBs, a speculative protein arrangement on telomeric DNA could be considered (Figure 5).

The model of binding arrangement considers that the protein monomers form a dimer that binds two adjacent binding sites simultaneously. This type of interaction mode is quite common in the sequence-specific binding of proteins that take part in regulatory mechanisms [35]. This model, where two recognition sites on DNA are bound by one protein dimer, might explain well the fall of binding activity when the DNA substrate is

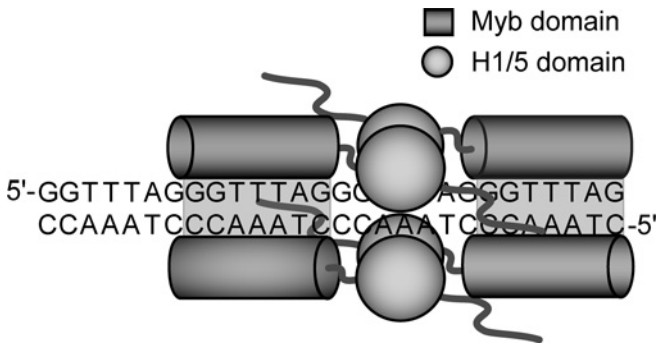


Figure 5 Speculative model of interaction of AtTRBs with telomeric DNA

Both homo- and hetero-dimers of AtTRB may participate in the interaction with telomeric DNA.

shortened from 28 to 14 bp as was observed for binding to the R4 and R2 duplex respectively. The introduced model is supported by stoichiometric and kinetic data presented here and it is also in accordance with our previous study demonstrating weaker binding to DNA containing fewer telomeric repeats [13]. The binding arrangement shown in Figure 5 also takes into consideration the multimerization ability of the H1/5 domain that could promote the arrangement of protein monomers in the DNA region between the binding sites. Moreover, the formation of homo- and hetero-multimers of SMH proteins and their ability to interact with other proteins (e.g. AtPOT1b [15,16]) contribute to a network of protein interactions that could be employed in the organization of telomere to form highly ordered chromatin structures, such as t-loops, in a similar way to human TRFs [31,32].

Thus, on the basis of results of the present study and the data available, we suggest that interactions of the two AtTRBs with telomeric DNA occur simultaneously with two binding sites. Therefore the minimal length of duplex DNA required for the proper binding of full-length AtTRB1 and AtTRB3 should harbour at least two putative binding sites that are bound by two dimers of AtTRBs. Consequently, SMH proteins are able to distinguish between short (< 10 bp) telomere-like sequences that are dispersed throughout the genome, e.g. in promoter regions [36], and longer tracts of telomere repeats occurring in telomeres.

There is still a considerable lack of general knowledge of intracellular arrangement, molecular crowding effects, association mechanisms and kinetics of protein–DNA-binding events in a living cell. Nevertheless, we can draw a speculative view of the *in vivo* consequences of our *in vitro* data, if we consider that the behaviour of a protein would not be markedly changed in the cellular environment. The access of AtTRBs to their telomeric target sites is restricted in both spatial and temporal ways by chromatin structure: the telomeric heterochromatin structure provides low accessibility upon its tight condensation, and thus the binding of specific proteins to DNA may occur preferentially in a short time slot between DNA replication and chromatin condensation [37].

AtTRBs might be first recruited by a weak non-specific binding to multiple chromosome regions. Then, once the specific target sites become accessible, highly specific binding occurs. On the other hand, the AtTRB molecules which are bound only by a highly dynamic non-specific interaction (in non-telomeric regions) can be easily displaced by other proteins binding with a higher affinity. Thus AtTRBs at non-telomeric sites do not impede other functional DNA–protein interactions.

In this way, non-specific binding could serve as a tool for increasing the local concentration of the proteins on DNA [34]. Accumulation of SMH proteins on DNA via non-specific electrostatic interactions may be important for their immediate availability for functional and specific binding to their telomere target sites.

Although further details of the binding interactions of proteins and their biological significance have yet to be determined, these results demonstrate the advantage of the approach employed in the present study by using a complete protein for *in vitro* studies rather than the commonly used Myb-domain-bearing fragment. Our data imply that AtTRB1 and AtTRB3 are telomere-specific proteins that bind telomeric DNA with distinct kinetics given by differences in their electrostatic interactions with DNA. To our knowledge, this is the first quantitative study of the plant-specific SMH family of proteins. The present paper demonstrates that the detailed quantification of protein–DNA interactions may provide new insights into the structural dynamics of telomeres.

ACKNOWLEDGEMENTS

We are grateful to M. Chester for the critical reading of the manuscript.

FUNDING

This work was supported by the Grant Agency of the Czech Republic [grant numbers 521/08/P452 and 204/08/H054], the Czech Ministry of Education [grant number LC06004], the Grant Agency of the Czech Academy of Sciences [grant number IAA500040801] and the institutional support [grant numbers MSM0021622415, MSM0021622413, AV0Z50040702 and AV0Z50040507].

REFERENCES

1 Zakian, V. A. (1995) Telomeres: beginning to understand the end. Science **270**, 1601–1607
2 Smogorzewska, A. and de Lange, T. (2004) Regulation of telomerase by telomeric proteins. Annu. Rev. Biochem. **73**, 177–208
3 Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H. and de Lange, T. (1999) Mammalian telomeres end in a large duplex loop. Cell **97**, 503–514
4 Cesare, A. J., Quinney, N., Willcox, S., Subramanian, D. and Griffith, J. D. (2003) Telomere looping in *P. sativum* (common garden pea). Plant J. **36**, 271–279
5 Smogorzewska, A., van Steensel, B., Bianchi, A., Oelmann, S., Schaefer, M. R., Schnapp, G. and de Lange, T. (2000) Control of human telomere length by TRF1 and TRF2. Mol. Cell. Biol. **20**, 1659–1668
6 Baumann, P. and Cech, T. R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. Science **292**, 1171–1175
7 Hockemeyer, D., Sfeir, A. J., Shay, J. W., Wright, W. E. and de Lange, T. (2005) POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. EMBO J. **24**, 2667–2678
8 Bilaud, T., Koering, C. E., Binet-Brasselet, E., Ancelin, K., Pollice, A., Gasser, S. M. and Gilson, E. (1996) The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. Nucleic Acids Res. **24**, 1294–1303
9 Kuchar, M. (2006) Plant telomere-binding proteins. Biol. Plant. **50**, 1–7
10 Yang, S. W., Kim, S. K. and Kim, W. T. (2004) Perturbation of NgTRF1 expression induces apoptosis-like cell death in tobacco BY-2 cells and implicates NgTRF1 in the control of telomere length and stability. Plant Cell **16**, 3370–3385
11 Hwang, M. G. and Cho, M. H. (2007) *Arabidopsis thaliana* telomeric DNA-binding protein 1 is required for telomere length homeostasis and its Myb-extension domain stabilizes plant telomeric DNA binding. Nucleic Acids Res. **35**, 1333–1342
12 Marian, C. O., Bordoli, S. J., Goltz, M., Santarella, R. A., Jackson, L. P., Danilevskaya, O., Beckstette, M., Meeley, R. and Bass, H. W. (2003) The maize single myb histone 1 gene, *Smh1*, belongs to a novel gene family and encodes a protein that binds telomere DNA repeats *in vitro*. Plant Physiol. **133**, 1336–1350
13 Schrumpfova, P., Kuchar, M., Mikova, G., Skrisovska, L., Kubicarova, T. and Fajkus, J. (2004) Characterization of two *Arabidopsis thaliana* myb-like proteins showing affinity to telomeric DNA sequence. Genome **47**, 316–324
14 Mozgova, I., Prochazkova Schrumpfova, P., Hofr, C. and Fajkus, J. (2008) Functional characterisation of domains in AtTRB1, a putative telomere-binding protein in *Arabidopsis thaliana*. Phytochemistry **69**, 1814–1819

15 Prochazkova Schrumpfova, P., Kuchar, M., Palecek, J. and Fajkus, J. (2008) Mapping of interaction domains of putative telomere-binding proteins AtTRB1 and AtPOT1b from *Arabidopsis thaliana*. FEBS Lett. **582**, 1400–1406
16 Kuchar, M. and Fajkus, J. (2004) Interactions of putative telomere-binding proteins in *Arabidopsis thaliana*: identification of functional TRF2 homolog in plants. FEBS Lett. **578**, 311–315
17 Shakirov, E. V., Surovtseva, Y. V., Osbun, N. and Shippen, D. E. (2005) The *Arabidopsis* Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. Mol. Cell. Biol. **25**, 7725–7733
18 Sue, S. C., Hsiao, H. H., Chung, B. C., Cheng, Y. H., Hsueh, K. L., Chen, C. M., Ho, C. H. and Huang, T. H. (2006) Solution structure of the *Arabidopsis thaliana* telomeric repeat-binding protein DNA binding domain: a new fold with an additional C-terminal helix. J. Mol. Biol. **356**, 72–85
19 Byun, M. Y., Hong, J. P. and Kim, W. T. (2008) Identification and characterization of three telomere repeat-binding factors in rice. Biochem. Biophys. Res. Commun. **372**, 85–90
20 Heyduk, T. and Lee, J. C. (1990) Application of fluorescence energy transfer and polarization to monitor *Escherichia coli* cAMP receptor protein and *lac* promoter interaction. Proc. Natl. Acad. Sci. U.S.A. **87**, 1744–1748
21 LeTilly, V. and Royer, C. A. (1993) Fluorescence anisotropy assays implicate protein-protein interactions in regulating trp repressor DNA binding. Biochemistry **32**, 7753–7758
22 Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. Protein Expression Purif. **41**, 207–234
23 Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–254
24 Murphy, J. H. and Trapane, T. L. (1996) Concentration and extinction coefficient determination for oligonucleotides and analogs using a general phosphate analysis. Anal. Biochem. **240**, 273–282
25 Buczek, P. and Horvath, M. P. (2006) Thermodynamic characterization of binding *Oxytricha nova* single strand telomere DNA with the α protein N-terminal domain. J. Mol. Biol. **359**, 1217–1234
26 Kuzmic, P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. Anal. Biochem. **237**, 260–273
27 Olmsted, M. C., Bond, J. P., Anderson, C. F. and Record, Jr, M. T. (1995) Grand canonical Monte Carlo molecular and thermodynamic predictions of ion effects on binding of an oligocation (L8+) to the center of DNA oligomers. Biophys. J. **68**, 634–647
28 Record, Jr, M. T., Zhang, W. and Anderson, C. F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. Adv. Protein Chem. **51**, 281–353
29 Dragan, A. I., Liu, Y., Makeyeva, E. N. and Privalov, P. L. (2004) DNA-binding domain of GCN4 induces bending of both the ATF/CREB and AP-1 binding sites of DNA. Nucleic Acids Res. **32**, 5192–5197
30 Bianchi, A., Smith, S., Chong, L., Elias, P. and de Lange, T. (1997) TRF1 is a dimer and bends telomeric DNA. EMBO J. **16**, 1785–1794
31 Fairall, L., Chapman, L., Moss, H., de Lange, T. and Rhodes, D. (2001) Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. Mol. Cell **8**, 351–361
32 Khan, S. J., Yanez, G., Seldeen, K., Wang, H., Lindsay, S. M. and Fletcher, T. M. (2007) Interactions of TRF2 with model telomeric ends. Biochem. Biophys. Res. Commun. **363**, 44–50
33 Hanaoka, S., Nagadoi, A. and Nishimura, Y. (2005) Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities. Protein Sci. **14**, 119–130
34 Revzin, A. (1990) The Biology of Nonspecific DNA–Protein Interactions, CRC Press, Boca Raton
35 von Hippel, P. H. (2007) From "simple" DNA–protein interactions to the macromolecular machines of gene expression. Annu. Rev. Biophys. Biomol. Struct. **36**, 79–105
36 Regad, F., Lebas, M. and Lescure, B. (1994) Interstitial telomeric repeats within the *Arabidopsis thaliana* genome. J. Mol. Biol. **239**, 163–169
37 Fajkus, J. and Trifonov, E. N. (2001) Columnar packing of telomeric nucleosomes. Biochem. Biophys. Res. Commun. **280**, 961–963

Received 7 November 2008/18 December 2008; accepted 22 December 2008
Published as BJ Immediate Publication 22 December 2008, doi:10.1042/BJ20082195

Biochem. J. (2009) **419**, 221–228 (Printed in Great Britain) doi:10.1042/BJ20082195

online data

SUPPLEMENTARY ONLINE DATA

Single-Myb-histone proteins from *Arabidopsis thaliana*: a quantitative study of telomere-binding specificity and kinetics

Ctirad HOFR*[1], Pavla ŠULTESOVÁ*, Michal ZIMMERMANN*, Iva MOZGOVÁ*, Petra PROCHÁZKOVÁ SCHRUMPFOVÁ*, Michaela WIMMEROVÁ† and Jiří FAJKUS*‡[1]

*Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, CZ-62500 Brno, Czech Republic, †National Centre for Biomolecular Research and Department of Biochemistry, Faculty of Science, Masaryk University, CZ-61137 Brno, Czech Republic, and ‡Laboratory of DNA–Molecular Complexes, Institute of Biophysics, Czech Academy of Sciences, CZ-61265 Brno, Czech Republic

EXPERIMENTAL

Gel-filtration chromatography

The molecular masses of the protein in monomeric and dimeric forms were estimated by size-exclusion gel-filtration chromatography through a Superdex 200 10/30 GL column (GE Healthcare), using a gel-filtration standard (Bio-Rad Laboratories) in a buffer containing 50 mM sodium phosphate (pH 7.5) and 300 mM NaCl. The molecular masses of proteins were estimated from a linear fit to the log M_r against elution volume plot generated with the protein standards. Supplementary Figure S1 shows the chromatograms.

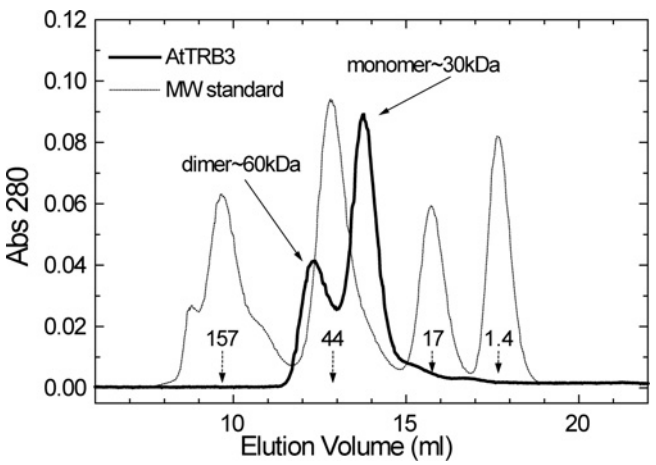


Figure S1 Size-exclusion chromatograms of protein AtTRB3 (continuous line) and molecular-mass standard (broken line)

Abs, absorbance; MW standard, molecular-mass standard. The numbers next to the arrows indicate determined molecular-mass values of monomeric and dimeric protein forms.

Surface plasmon resonance

All SPR experiments were performed with a Biacore 3000 instrument (GE Healthcare) at 25 °C using TBST (Tris-buffered saline with Tween 20: 10 mM Tris/HCl, pH 7.5, 150 mM NaCl, containing 0.005 % Tween 20) and a flow rate of 5 μl/min. AtTRB3 was immobilized on the research-grade CM5 sensor chip in a buffer containing 10 mM Hepes, 150 mM NaCl (pH 7.5) and 0.005 % Tween 20. Sensorgrams were run in the automatic subtraction mode using FC (flowcell) 1 as an unmodified reference. Data were collected for FC 2, FC 3 and FC 4, which contained various amounts of AtTRB3. Injections of DNA were made using the 'quickinject' injection mode, going from lowest to highest concentration samples, with a 5 min contact time and a 1200 s dissociation phase in all cases. Regeneration was achieved using several (two to five) 1 min pulses of 50 mM NaOH. All sensorgrams were obtained at 25 °C. Data were analysed by equilibrium analysis in addition to the kinetic analysis. The equilibrium response was plotted against the concentration of DNA and fitted to:

$$R = K_a[\mathrm{DNA}]R_{max}(K_a[\mathrm{DNA}] + 1)$$

where R is the equilibrium response at a specific concentration of DNA substrate, R_{max} is the response at saturation of the DNA substrate on the chip, K_a is the equilibrium association constant, which is the reciprocal of the dissociation constant K_d ($K_a = 1/K_d$). When assuming a non-co-operative binding model, the apparent K_d from SPR experiments should be divided by 4 to resemble different binding stoichiometry of FA and SPR experiments. The output of the non-linear fitting of SPR curves for different concentrations of DNA produces a K_d of 6.8 μM, which, divided by 4, gives 1.7 μM. This value agrees well with the value of K_d determined from FA measurements considering different buffer conditions. Supplementary Figure S2 shows the sensorgrams and the response curve.

Purification of AtTRBs

AtTRB1 and AtTRB3 were expressed in soluble forms in cytoplasm of *E. coli*. The purification strategy consisted of two affinity steps. A capture step by IMAC (immobilized metal-ion-affinity chromatography) was followed by a purification step using HAC (heparin-affinity chromatography). To confirm final purity, collected fractions were separated by SDS/PAGE (0.1 % SDS, 10 % acrylamide). Supplementary Figure S3 shows the gel-purified proteins.

[1] Correspondence may be addressed to either of these authors (email hofr@sci.muni.cz or fajkus@sci.muni.cz).

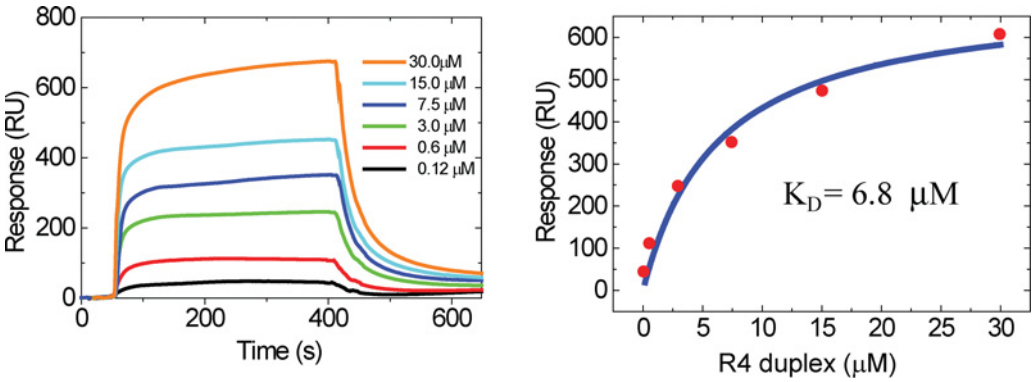


Figure S2 Binding of telomeric duplex R4 to immobilized AtTRB3

Response signals from the saturated region of the sensorgram have been used to calculate equilibrium dissociation constant K_d. RU, response units.

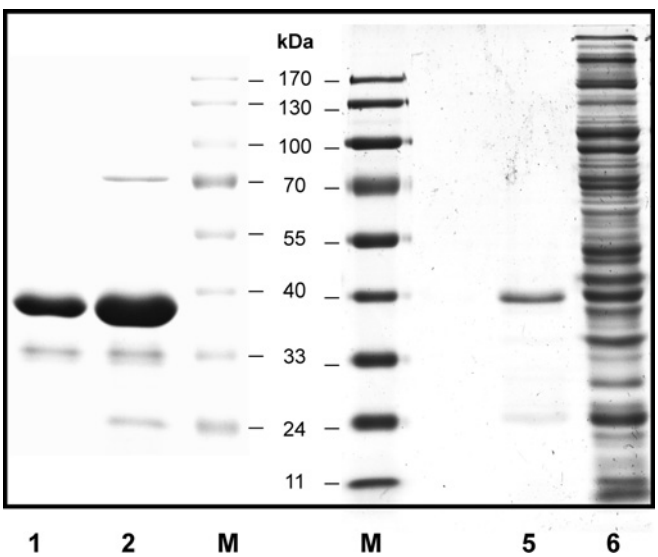


Figure S3 Analysis of purification steps using SDS/PAGE

Lane 1, collected fractions containing AtTRB3 after IMAC (immobilized metal-ion-affinity chromatography) and subsequent HAC (heparin-affinity chromatography) (10 μg); lane 2, collected fractions containing AtTRB3 after IMAC (15 μg); lanes M, molecular-mass markers (sizes are indicated in kDa); lane 5, collected fractions containing AtTRB1 after IMAC and subsequent HAC (3 μg); lane 6, clarified cytoplasmic extract with expressed AtTRB1 (35 μg). The proteins were stained with Coomassie Brilliant Blue.

Received 7 November 2008/18 December 2008; accepted 22 December 2008
Published as BJ Immediate Publication 22 December 2008, doi:10.1042/BJ20082195

Cell

The Solution Structure of the ADAR2 dsRBM-RNA Complex Reveals a Sequence-Specific Readout of the Minor Groove

Richard Stefl,[1,4,6] Florian C. Oberstrass,[1,6,7] Jennifer L. Hood,[3] Muriel Jourdan,[1,8] Michal Zimmermann,[5] Lenka Skrisovska,[1] Christophe Maris,[1] Li Peng,[2] Ctirad Hofr,[5] Ronald B. Emeson,[2] and Frédéric H.-T. Allain[1,*]

[1]Institute of Molecular Biology and Biophysics, ETH Zurich, CH-8093 Zürich, Switzerland
[2]Department of Pharmacology, Vanderbilt University, Nashville, TN 37232, USA
[3]Neuroscience Graduate Program, Vanderbilt University, Nashville, TN 37232, USA
[4]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, CZ-62500 Brno, Czechia
[5]Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, CZ-62500 Brno, Czechia
[6]These authors contributed equally to this work
[7]Present address: Department of Bioengineering, Stanford University, 318 Campus Drive, Stanford, CA 94305, USA
[8]Present address: Departement de Chimie Moleculaire, 38041 Grenoble Cedex09, France
*Correspondence: allain@mol.biol.ethz.ch
DOI 10.1016/j.cell.2010.09.026

SUMMARY

Sequence-dependent recognition of dsDNA-binding proteins is well understood, yet sequence-specific recognition of dsRNA by proteins remains largely unknown, despite their importance in RNA maturation pathways. Adenosine deaminases that act on RNA (ADARs) recode genomic information by the site-selective deamination of adenosine. Here, we report the solution structure of the ADAR2 double-stranded RNA-binding motifs (dsRBMs) bound to a stem-loop pre-mRNA encoding the R/G editing site of GluR-2. The structure provides a molecular basis for how dsRBMs recognize the shape, and also more surprisingly, the sequence of the dsRNA. The unexpected direct readout of the RNA primary sequence by dsRBMs is achieved via the minor groove of the dsRNA and this recognition is critical for both editing and binding affinity at the R/G site of GluR-2. More generally, our findings suggest a solution to the sequence-specific paradox faced by many dsRBM-containing proteins that are involved in post-transcriptional regulation of gene expression.

INTRODUCTION

ADARs convert adenosine-to-inosine (A-to-I) by hydrolytic deamination in numerous mRNA and pre-mRNA transcripts (Bass, 2002; Nishikura, 2006). Due to the similar base-pairing properties of both nucleosides, inosine is interpreted as guanosine by cellular machineries during the processes of translation and splicing. In this way, editing-mediated alterations in sequence can alter codon identity or base-pairing interactions within higher-order RNA structures (Bass, 2002; Nishikura, 2006). As a result, ADARs can create protein isoforms or regulate gene expression at the RNA level (Bass, 2002; Nishikura, 2006; Valente and Nishikura, 2005). ADARs are widely expressed in most cell types, yet their expression and activity in neuronal tissues has been shown to be important for proper nervous system function (Higuchi et al., 2000; Palladino et al., 2000). Recent high-throughput sequencing analysis of A-to-I editing identified over 55 editing sites within the coding regions of mRNAs, with 38 of these sites involving a codon change that specifies an alternative amino acid. Many of these changes involve RNA transcripts encoding proteins that are critical for nervous system function (Li et al., 2009).

ADARs from all characterized species have a modular domain organization consisting of one-to-three dsRBMs followed by a conserved C-terminal catalytic adenosine deaminase domain. The structures of the two dsRBMs and of the isolated catalytic domain of ADAR2 have been determined in their free states (Macbeth et al., 2005; Stefl et al., 2006). Among the best-studied ADAR substrates are pre-mRNAs encoding subunits of the α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA)-subtype of ionotropic glutamate receptor (GluR-2, GluR-3 and GluR-4; Higuchi et al., 2000, 1993; Melcher et al., 1996) that contain one or both of two highly edited and functionally relevant sites, namely the R/G and Q/R editing sites (Aruscavage and Bass, 2000; Lomeli et al., 1994; Melcher et al., 1996).

ADARs can edit RNA substrates either specifically or nonspecifically depending upon the structures of the RNA substrates (Bass, 2002). In vitro studies have shown editing of up to 50% of the adenosine residues in both strands using synthetic dsRNAs that are perfectly complementary (Cho et al., 2003; Lehmann and Bass, 2000). Such nonspecific editing can be explained by the presence of dsRBMs which are thought to bind dsRNA in a sequence-independent manner (Tian et al., 2004), yet it remains unclear how certain RNA substrates are edited in a site-specific fashion. Several studies have suggested that the

presence of noncanonical elements in these dsRNAs–such as mismatches, bulges, and loops–could be important for site-selective A-to-I conversion (Bass, 2002; Stefl et al., 2006; Tian et al., 2004).

The dsRBMs of ADARs are not only essential for editing (Stefl et al., 2006; Valente and Nishikura, 2007), but the dsRBM also represents the second most abundant family of RNA recognition motifs. In addition to RNA editing, dsRBMs are involved in numerous post-transcriptional regulatory processes and most prominently in micro RNA (miRNA) biogenesis and function and RNA export (Dreyfuss et al., 2002; Tian et al., 2004). The few solved structures of dsRBM-containing proteins bound to short, synthetic RNA duplexes have suggested that dsRBMs recognize the A-form helix of dsRNA in a sequence-independent manner, since the majority of dsRBM-RNA interactions involve direct contact with the 2′-hydroxyl groups of the ribose sugars and direct or water-mediated contacts with nonbridging oxygen residues of the phosphodiester backbone (Gan et al., 2006; Ramos et al., 2000; Ryter and Schultz, 1998; Wu et al., 2004), and that a subclass of dsRBMs prefer stem-loops over A-form helices (Ramos et al., 2000; Wu et al., 2004).

We previously determined that each of the two dsRBMs of ADAR2 bind to a distinct location on the GluR-2 RNA encompassing the R/G editing site and that the interdomain linker (amino acids 147-231) is unstructured both in the free protein and in the complex (Stefl et al., 2006). To better understand RNA substrate recognition by ADAR2, we have determined the solution structure of the RNA helix surrounding the editing site and the solution structure of the two dsRBMs of ADAR2 bound to the GluR-2 R/G site.

RESULTS

Structure of the GluR-2 R/G RNA Helix Surrounding the Editing Site

The GluR-2 R/G site (A8) is embedded within a 71 nt RNA stem-loop containing three base-pair mismatches and capped by a 5′-GCUAA-3′ pentaloop (Figure 1A). We previously determined the structure of the apical part of the stem-loop and showed that the pentaloop is structured and adopts a fold reminiscent of a UNCG-type family of tetraloops (Stefl and Allain, 2005). Here, we have investigated the structure of the RNA helix surrounding the editing site that contains two A-C mismatches, one at the editing site (A8) and a second one ten base-pairs downstream (A18, Figure 1B). Monitoring adenine C2 chemical shifts (a sensitive probe to monitor the protonation state of N1) during a pH titration, we observed that A8 and A18 are fully protonated below pH 6.5, partially protonated between pH 6.5–8.5, and unprotonated above pH 8.5 (Figures 1H and 1I). The pKa for the adenosines N1 can be estimated between 7 and 7.5 at 310 K, which is 3.3 units higher than the value determined for an isolated AMP (pKa of 4.0; Legault and Pardi, 1994). Using 863 nOe-derived distance restraints, we solved the structure of the free RNA in the protonated state (pH 6.2). The structure is well defined, even for the A-C mismatches (Figure 1E and Table 1) that are stacked inside the stem. Therefore, at pH 6.2, the R/G site has a regular A-form helix structure (Figure 1D) containing two A^+-C base-pairs adopting a wobble conformation, stabilized by two hydrogen bonds each (Figures 1F and 1G).

Structure of ADAR2 dsRBMs Bound to Their Respective RNA Targets

Considering the distinct RNA binding location found previously for each dsRBM (Stefl et al., 2006) and the high molecular weight (over 50 kDa) of the complex formed between the two dsRBMs of ADAR2 and the GluR-2 R/G substrate (Figure 1A), we adopted a modular approach to solve the structure of this complex in solution. To this end, we first solved the structure of dsRBM1 in complex with a modified GluR-2 upper stem–loop (USL, Figure 1C, and Figure S1 available online) and then the structure of dsRBM2 bound to the GluR-2 lower stem-loop that contains the editing site (LSL, Figure 1B, and Figure S2). The use of a GluR-2 R/G USL mutant to determine the structure of dsRBM1 in complex with RNA was dictated by the poor data quality that we obtained with the wild-type (WT) sequence. In changing the loop sequence to that found in the GluR-3 USL (Aruscavage and Bass, 2000), we obtained a smaller and more stable RNA which provided NMR data of higher quality.

A total of 1707 and 1929 nOe-derived distance restraints (including 36 intermolecular ones for each complex) for ADAR2 dsRBM1–GluR-2 R/G USL mutant and ADAR2 dsRBM2–GluR-2 R/G LSL complexes, respectively, were used to obtain well-defined structures (Figure 2 and Table 1). The two dsRBM-RNA complexes are stabilized by a combination of hydrophobic interactions, hydrogen bonding and electrostatic contacts. In both dsRBM–RNA complexes, the dsRBMs adopt the expected $\alpha\beta\beta\beta\alpha$ topology in which the two α helices are packed along the three-stranded antiparallel β sheet. The entire interaction surface spans 12-14 base-pairs covering two minor grooves and a major groove (Figure 2). In both complexes, three distinct regions of the dsRBMs are involved in interaction with RNA. The first region is the helix α1, which interacts with the first minor groove of the RNA. The second region is a well-conserved KKNAK-motif, located at the amino-terminal tip of helix α2 and the preceding loop, that contact the RNA with nonsequence specific contacts between lysine side-chains and the phosphate oxygens across the major groove of the RNA (Lys127, 128, and 131 for dsRBM1 and Lys281, 282, 285 for dsRBM2, Figure 2). In addition, the dipole moment of helices α2 creates a positive charge in the N-terminal tip of these helices that interacts with the negatively charged phosphate backbone. This second set of interactions is mediated by the main-chain amides of K127 and K281, which are hydrogen bonded with the phosphates oxygen of A24 and U11, respectively (Figure 2). The third region of contact is the β1-β2 loop which interacts with the second minor groove of the RNA. The overall architecture of these two complexes resembles other previously determined dsRBM–RNA structures (Blaszczyk et al., 2004; Gan et al., 2008; Gan et al., 2006; Ramos et al., 2000; Ryter and Schultz, 1998; Stefl et al., 2005a; Wu et al., 2004). However, a detailed inspection of the interaction regions revealed striking differences between the two complexes and other dsRBM-RNA complexes, particularly in the first and the third regions where both dsRBMs present unexpected sequence-specific contacts to the RNA minor grooves (Figure 2).

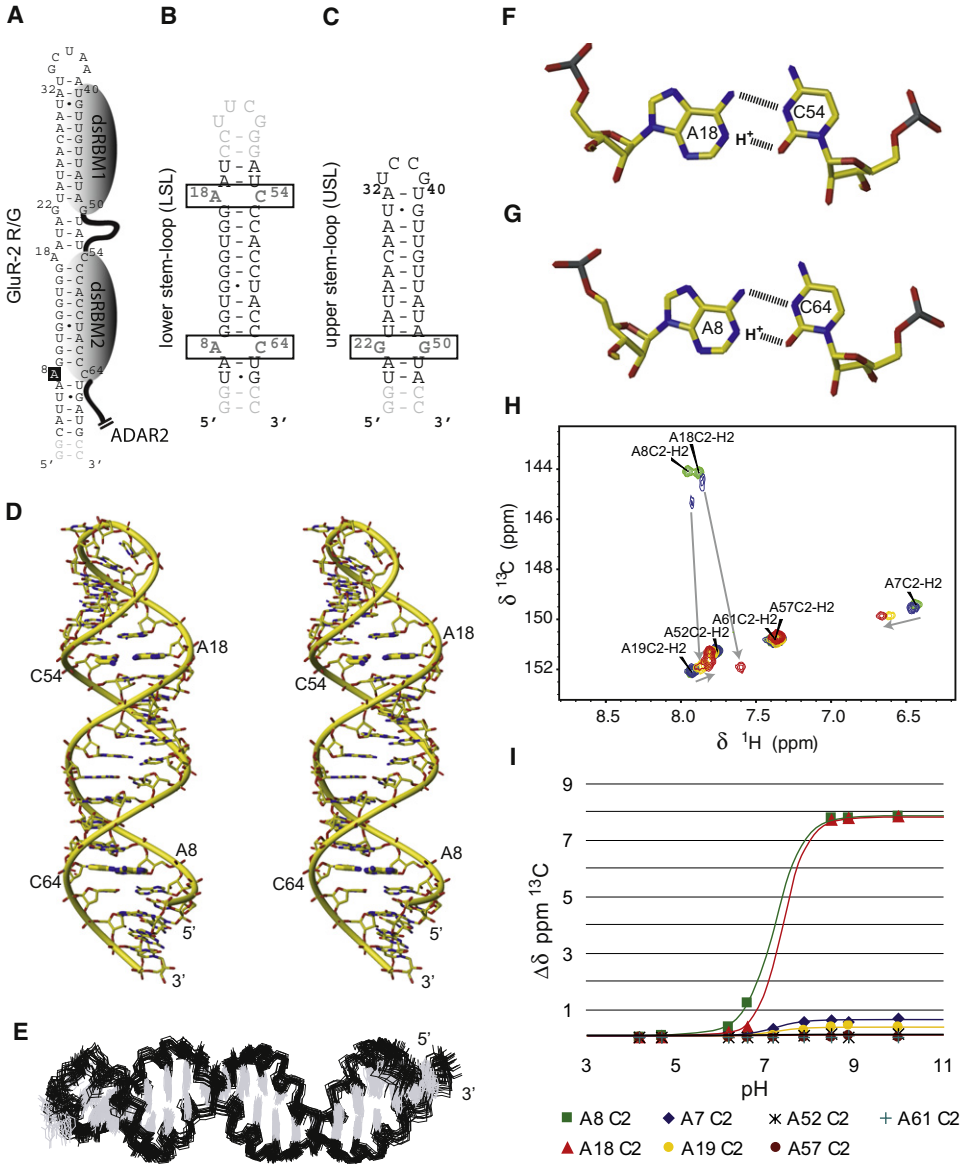


Figure 1. Secondary Structures of the RNAs and Solution Structure of GluR-2 R/G LSL RNA

(A) Secondary structure of GluR-2 R/G RNA. The indicated binding regions for the dsRBMs were proposed previously (Stefl et al., 2006).

(B) Secondary structure of the GluR-2 R/G lower stem-loop (LSL).

(C) Secondary structure of the GluR-2 R/G upper stem-loop (USL).

(D) Stereo view of the most representative structure of GluR-2 R/G LSL RNA. The A^+-C wobble base-pairs are highlighted in bold sticks.

(E–G) (E) Overlay of the 20 lowest energy structures of GluR-2 R/G LSL. The A^+-C wobble base-pairs A18-C54 (F) and A8-C64 (G) are shown.

(H) H2-C2 region of adenines in the ^{13}C-^{1}H-HSQC spectra of the GluR-B R/G LSL is shown at pH 4.7 (green peaks), 6.6 (blue peaks), 7.9 (orange peaks) and 8.9 (red peaks). The two adenines involved in the A^+-C wobble base-pair showed drastic perturbation.

(I) Diagram showing the pH-dependence of ^{13}C chemical shift changes of adenine C2′s.

Sequence-Specific Recognition by ADAR2 dsRBM1

In the ADAR2 dsRBM1–RNA complex, contacts from helix α1 are centered at the A32-U40 base-pair below the UCCG tetraloop (Figures 2A and 2C). Met84 makes a sequence-specific hydrophobic contact with H2 of A32 and Asn87 contacts the 2′-hydroxyl and O2 of U40. The Oε of Glu88 is hydrogen bonded to the amino group of the first cytosine of the tetraloop. In addition, Leu83 makes hydrophobic contacts with the sugar of G41. The entire helix α1 is tightly inserted in the minor groove created by the UCCG tetraloop and two adjacent base-pairs (Figure 2A). The β1-β2 loop of dsRBM1 binds the following minor groove of the RNA. This minor groove is widened as it has to accommodate base-pairing of two guanosines that make an N1 symmetrical G22-G50 mismatch (Figures 2A and 2D) that are the center of this interaction. Val104 side-chain contacts the H8 of G50 (that adopts a *syn* conformation) and a sequence-specific hydrogen

Table 1. NMR and Refinement Statistics for the GluR-2 R/G Upper Stem-Loop RNA Bound to ADAR2 dsRBM1, the Free GluR-2 R/G Lower Stem-Loop RNA, and Its Complex with ADAR2 dsRBM2, and the RDC-Reconstructed Complex of the Full-Length GluR-2 R/G Stem-Loop RNA Bound to ADAR2 dsRBM12

	USL RNA – dsRBM1 Complex		LSL RNA	LSL – dsRBM2 Complex		SL RNA – dsRBM12 complex	
	USL RNA	dsRBM1		LSL RNA	dsRBM2	SL RNA	dsRBM12
NMR Distance and Dihedral Constraints and RDCs							
Distance restraints							
Total NOE	645	927	781	702	1054	1252	1981
Intraresidue	309	201	389	365	216	620	417
Interresidue	336	726	392	337	838	631	1564
Sequential (|i-j| = 1)	270	252	352	306	241	555	493
Nonsequential (|i-j| > 1)	66	474	40	31	597	76	1071
Hydrogen bonds	35	64	81[a]	75	62	132	126
Protein–RNA intermolecular	36			36		72	
Total dihedral angle restraints	180		252	267			
RNA							
Sugar pucker	34		84	84			
Backbone[b]	146		168	183			
RDC restraints							45[d]
Structure Statistics[c]							
Violations (mean and SD)							
Number of distance restraint violations > 0.2 Å	8.45 ± 2.50		0	1.10 ± 1.25		14.31 ± 3.86	
Number of dihedral angle restraint violations > 5°	0.7 ± 0.47		0	0		5.30 ± 3.32	
Max. dihedral angle restraint violation (°)	5.82 ± 1.22		3.28 ± 0.77	2.69 ± 1.12		15.51 ± 2.36	
Max. distance constraint violation (Å)	0.29 ± 0.03		0.16 ± 0.01	0.23 ± 0.06		0.32 ± 0.05	
Deviations from idealized geometry[d]							
Bond lengths (Å)	0.0042 ± 0.00007		0.0046 ± 0.00005	0.0041 ± 0.00005		0.0048 ± 0.00005	
Bond angles (°)	1.989 ± 0.011		2.137 ± 0.017	1.903 ± 0.011		1.995 ± 0.008	
RDCs violations							
Absolute RDC violations (Hz)						1.12 ± 0.82	
Average pairwise r.m.s.d (Å)[c]							
Protein (79-142) for dsRBM1; (221-282) for dsRBM2							
Heavy atoms	1.11 ± 0.17			1.01 ± 0.12		1.60 ± 0.36	
Backbone atoms	0.59 ± 0.14			0.37 ± 0.08		1.22 ± 0.42	
RNA							
All RNA heavy atoms	0.60 ± 0.16		1.15 ± 0.35	1.48 ± 0.51		1.30 ± 0.40	
Complex							
All complex heavy atoms	1.01 ± 0.15			1.49 ± 0.39		1.75 ± 0.31	

[a] In the final structure calculations of the free RNA, H-bond restraints were applied in the two A-C mismatches. This is based on initial structures and on the protonation state of A8/A18. For the structures of the RNA in complex no H-bond restraints for the two A-C mismatches have been applied.
[b] Based on A-form geometry derived from high-resolution crystal structures: α(270°–330°), β(150°–210°), γ(30°–90°), δ(50°–110°), ε(180°–240°), and ζ (260°–320°). These restraints were used only for the double-helical region. No angle restraints were imposed on the two A-C mismatches and the loops.
[c] Calculated for an ensemble of the 20 lowest energy structures.
[d] 16 RDCs of dsRBM1 and 29 RDCs of dsRBM2.

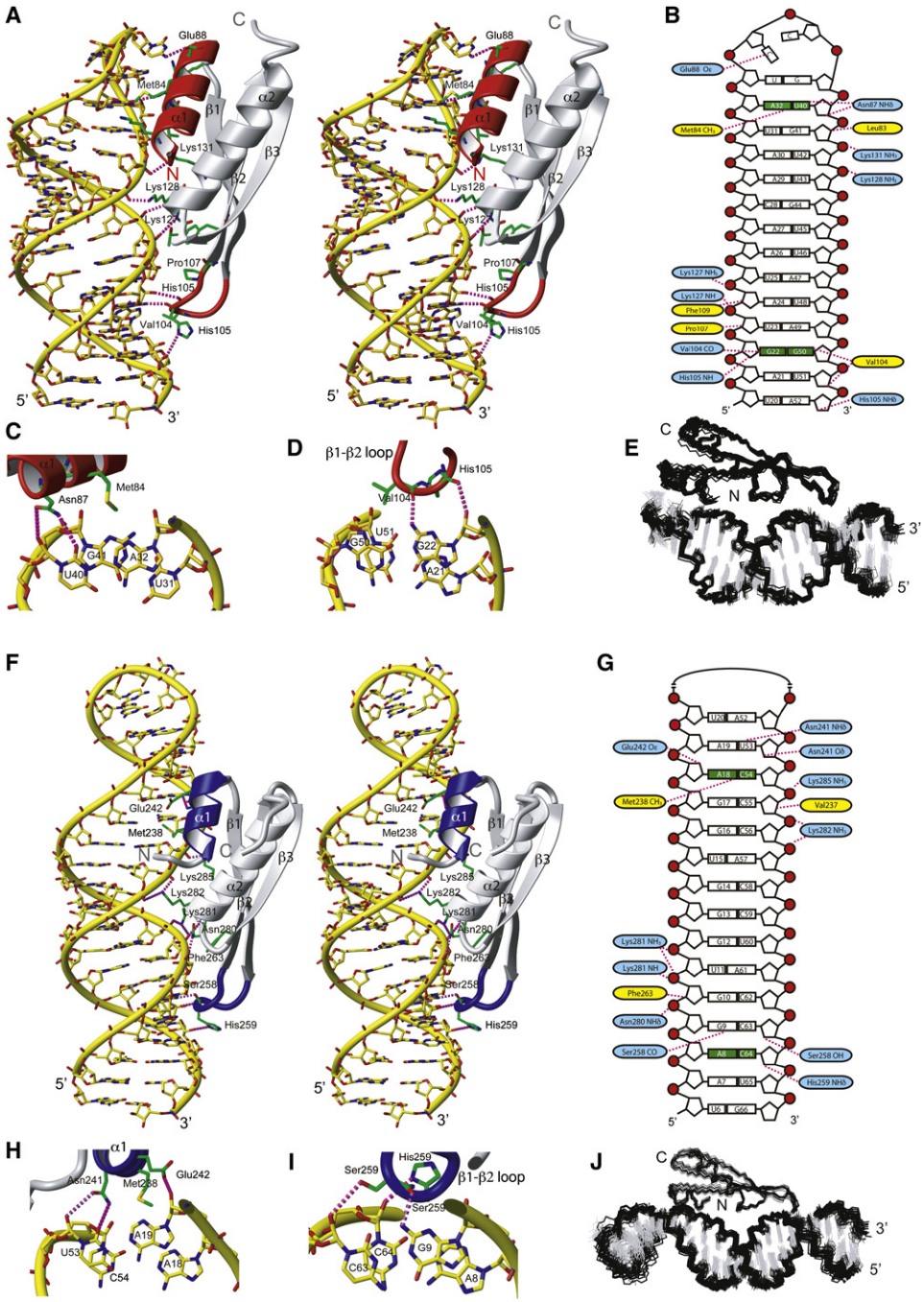


Figure 2. RNA Recognition by ADAR2 dsRBM1 and dsRBM2

(A) Stereo view of the most representative structure of dsRBM1 bound to USL RNA. The RNA is represented as a yellow stick model and the protein is shown as a ribbon model with residues that contact the RNA shown in green. Helix α1 and the β1-β2 loop that mediate the sequence-specific contacts are colored in red. Hydrogen bonds are indicated by magenta dotted lines. (B) Scheme showing contacts between dsRBM1 and the USL RNA. Protein residues that form hydrogen bonds to the RNA are shown in blue and the one having hydrophobic interactions are in yellow. Close-up view of minor groove sequence-specific recognitions mediated by helix α1 (C) and the β1-β2 loop (D) of dsRBM1. (E) Overlay of the 20 lowest energy structures of the dsRBM1-USL complex. (F) Stereoview of the most representative structure of the dsRBM2 bound to LSL RNA. Helix α1 and the β1-β2 loop that mediate the sequence-specific contacts are colored in blue. (G) Scheme showing contacts between dsRBM2 and the LSL RNA. Close-up view of the minor groove sequence-specific recognitions mediated by helix α1 (H) and the β1-β2 loop (I). (J) Overlay of the 20 lowest energy structures of the dsRBM2-LSL complex. For NMR data of these two complexes, see also Figure S1 and Figure S2.

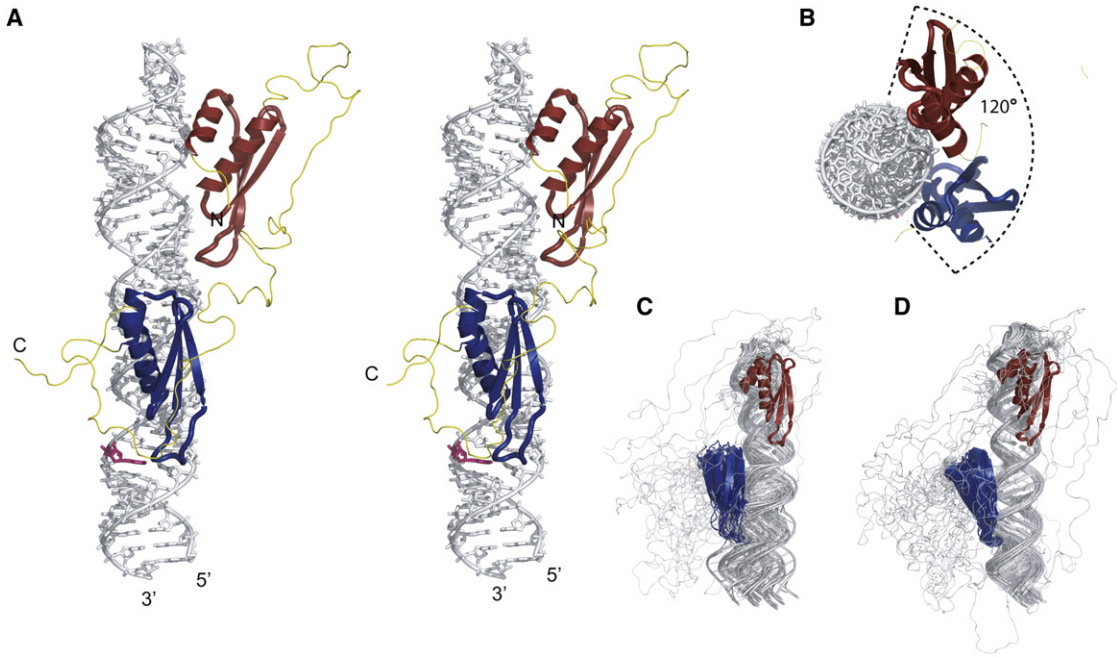


Figure 3. Structure of ADAR2 dsRBM12 Bound to GluR-2 R/G

(A) Stereo view of the most representative RDC-reconstructed structure of the ADAR2 dsRBM12 bound to GluR-2 R/G. The RNA is represented as a stick model (in gray; the edited adenosine is highlighted in pink) and the protein is shown as a ribbon model (dsRBM1 in red; dsRBM2 in blue; linker in yellow). (B) Top view of the complex. Overlay of the 20 lowest energy structures calculated without (C) and with RDCs (D), superimposed on dsRBM1.

bond is formed between the main-chain carbonyl of V104 and the amino group of G22. The widened minor groove accommodates additional interactions between three side-chains (Phe109, Pro107, His105) and the sugars of the base-pairs above and below. Altogether, dsRBM1 binds the RNA stem-loop at a single register via two sequence-specific contacts at two consecutive RNA minor grooves: a hydrogen bond to the amino group of the G22 in the GG mismatch via the β1-β2 loop and an hydrophobic contact to the adenine H2 of A32 via Met84 in helix α1.

Sequence-Specific Recognition by ADAR2 dsRBM2

The dsRBM2 of ADAR2 is adjacent to the deaminase domain and is essential for A-to-I editing at the R/G site (Stefl et al., 2006; Xu et al., 2006). In the ADAR2 dsRBM2–GluR-2 R/G LSL complex, Asn241, Glu242, Met238, Val 237 of helix α1 contact the minor groove region centered at the A18-C54 mismatch (Figures 2F and 2H). At pH 7.6, where the protein-RNA complex has been determined, this mismatch is unprotonated and Met238 makes a sequence-specific hydrophobic contact with A18 H2. Contacts to the base-pair above and below by Asn241 and Glu242, and by Val 237, respectively, further stabilize the interaction of helix α1 in this region (Figure 2H). The β1−β2 loop of dsRBM2 interacts with the second minor groove. The contacts are centered at the G9-C63 Watson-Crick base-pair located above the A8-C64 mismatch containing the editing site. A sequence-specific hydrogen bond is formed between the main-chain carbonyl of Ser258 and the amino of G9 (Figures 2F and 2I). Additionally, nonsequence specific contacts between the side-chains of Ser 258, His 259 and Phe 263 and the G9-C63 base-pair and the base-pairs above and below increase the stability of the interaction with the RNA minor groove (Figure 2G). In the vicinity of the editing site, dsRBM2 contacts C63, while A8 is not contacted by any residue from the β1-β2 loop therefore making A8 accessible to the deaminase domain. Altogether, dsRBM2 similar to dsRBM1, recognizes the RNA helix via two sequence-specific contacts at two consecutive RNA minor grooves: a hydrogen bond to the amino group of the G9 at the GC 3′ to the editing site via the β1-β2 loop and a hydrophobic contact to the adenine H2 of A18 via Met238 in helix α1. In the NMR spectra (*data not shown*), we could observe intermolecular nOes corresponding to dsRBM2 being positioned at a second binding register one base-pair above (although with only 20% occupancy). In this case the β1-β2 loop contact G10 and Met 238 contact A19. Although two consecutive binding sites for dsRBM2 are observed here, they both confirm the sequence-specific nature of the dsRBM2-RNA interaction.

Structure of ADAR2 dsRBM12 in Complex with GluR-2 R/G RNA

Next, we determined the structure of ADAR2 dsRBM12 in complex with GluR-2 R/G RNA (Figures 3A and 3B). To calculate an atomic model of this complex, we used the distance constraints measured in the two sub-complexes described above (Figure 3C). This strategy could be used considering (1) the distinct RNA binding location for each dsRBMs, with no mutual interactions (Stefl et al., 2006), (2) the flexible unstructured linker connecting dsRBM1 and dsRBM2 in the complex

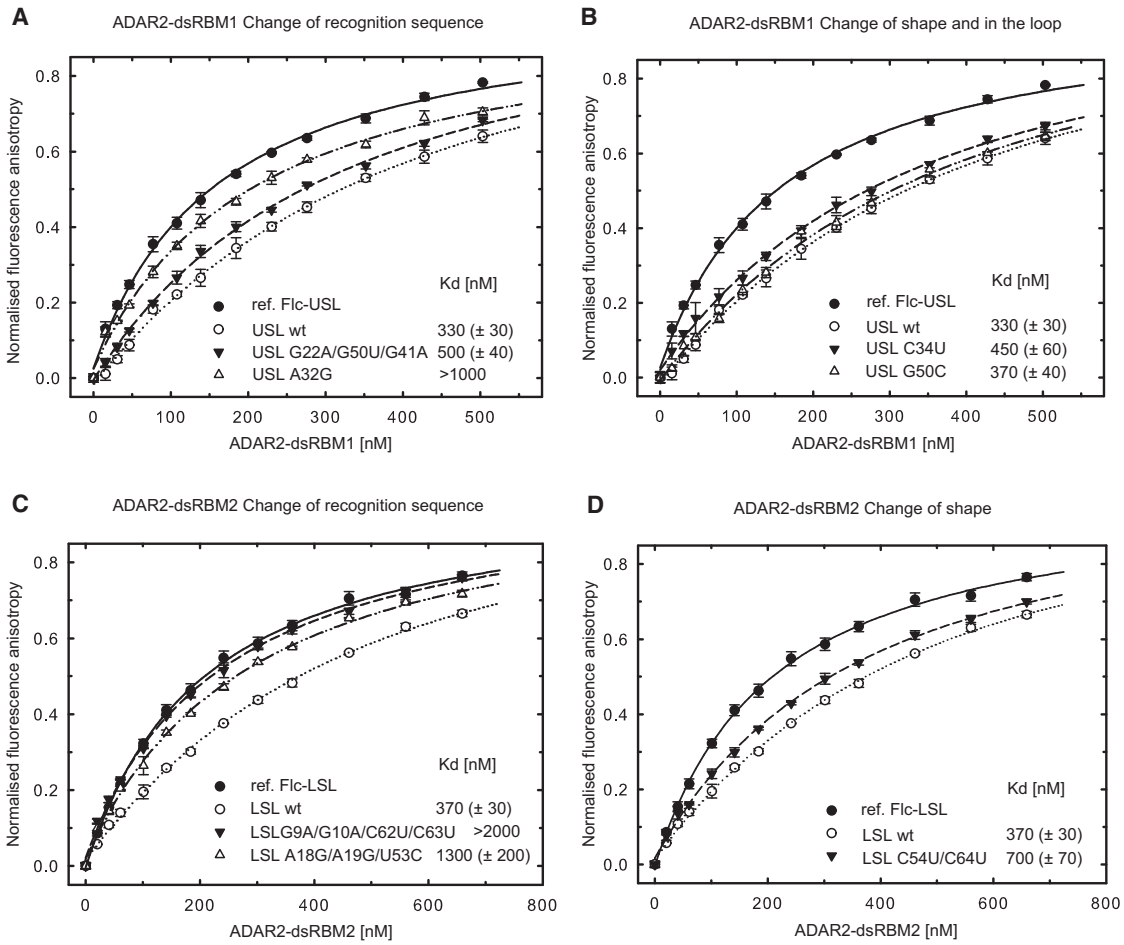


Figure 4. ADAR2 dsRBMs Bind Preferentially to RNAs that Contains Their Sequence-Specific Recognition Motifs

(A) ADAR2 dsRBM1 was titrated with fluorescently labeled USL and binding was measured by fluorescence anisotropy (black circles; fluorescein labeled reference, Flc-USL). The same experiment was then carried out in the presence of competing unlabeled USL wt (○), USL G22A/G50U/G41A mutant (▼), and USL A32G mutant (△). Equilibrium dissociation constants (K_d) were calculated from the best fit to the data as described in Experimental Procedures.

(B) The same assay as shown in (A) but for USL C34U mutant (▼) and USL G50C mutant (△).

(C) ADAR2 dsRBM2 was titrated with fluorescently labeled LSL and binding was measured by fluorescence anisotropy (●; fluorescein labeled reference, Flc-LSL). The same experiment was then carried out in the presence of competing unlabeled LSL wt (○), LSL G9A/G10A/C62U/C63U mutant (▼), and LSL A18G/A19G/U53C mutant (△).

(D) The same assay as shown in (C) but for LSL C54U/C64U mutant (▼).Wild-type and mutant sequences are shown in Figure S3.

(Stefl et al., 2006) and (3) an overlap in the RNA sequence of the joint region of the subcomplexes (Figure 1). Long-range structural constraints for this elongated complex were derived from residual dipolar couplings (RDCs) measured with a deuterated protein on the full-length complex (dsRBM12 bound to GluR-2 R/G RNA, Figure 1A). The pentaloop which is not contacted by dsRBM1 was modeled using the structure that was determined previously (Stefl and Allain, 2005). With this strategy, we could then determine a precise solution structure of this 50 kDa complex using 45 ^{15}N-^{1}H RDCs (Figure 3D, Table 1). In the structure, the two dsRBMs bind one face of the RNA covering approximately 120 degrees of the space around the RNA helix (Figure 3B). This suggests that the binding of an additional molecule of ADAR2 would be sterically possible, consistent with studies indicating that ADAR2 dimerization is necessary for RNA editing (Chilibeck et al., 2006; Cho et al., 2003; Gallo et al., 2003; Valente and Nishikura, 2007).

Sequence-Specific Contacts of ADAR2 dsRBMs Are Important for Binding Affinity

To confirm the ADAR2 dsRBMs sequence-specific preference in a quantitative solution binding assay, we performed fluorescence anisotropy (FA) experiments by titrating dsRBM1 and dsRBM2 against labeled USL and LSL RNAs, respectively. Unlabeled wild-type and mutant RNAs (Figure S3) were used for competition experiments as described in Experimental procedures. The equilibrium dissociation constants were calculated from the displacement of the binding curves (Figure 4). We designed two sets of mutations, one set was designed to change the recognition sequence of USL and LSL RNAs (Figures 4A and

4C and Figure S3) and a second set was designed to maintain the recognition sequence, but change the RNA shape via mismatches of USL and LSL RNAs into Watson-Crick base-pairs (Figures 4B and 4D and Figure S3) to measure their effect on overall binding affinity. In mutating any of the bases that are recognized in a sequence-specific manner by dsRBM1 in USL (G22, A32 or C34), the apparent affinity is reduced compared to the wild-type (Figures 4A and 4B). However when the G22-G50 mismatch is replaced by a Watson-Crick G22-C50 pair, the affinity is almost identical to wild-type RNA, confirming that dsRBM1 recognizes the sequence rather than the shape of the RNA helix (note that G41 was mutated in the first RNA mutant to prevent the sequence-specific recognition of G41 by dsRBM1). Similarly for the LSL, when G9 or A18 are mutated, dsRBM2 binding is reduced more than five-fold (Figure 4C), yet when the two AC mismatches are replaced by Watson-Crick AU pairs, the affinity is only reduced by two-fold (Figure 4D). In this latter context, the sequence-specific contacts are the same for the WT and mutant RNAs, but the presence of a more deformable A18-C54 base-pair in the WT structure could explain the higher affinity of dsRBM2 to the WT RNA (note that additional mutations were introduced in the first two RNA mutants of LSL to abolish the two binding registers found in the wild-type LSL). Altogether, the FA data strongly support the idea that the sequence-specific interactions observed in the structures of ADAR2 dsRBMs-dsRNA are important for the affinity of both dsRBMs and that they finely tune the preferential binding to these recognition motifs.

Sequence-Specific Contacts of ADAR2 dsRBMs Are Important for Editing

To test the functional importance of the four sequence-specific contacts identified in the ADAR2 dsRBM12-GluR-2 R/G RNA complex, single amino acid mutants in helix α1 (M84 or M238) were mutated to alanine or double mutants in the β1−β2 loop in either dsRBM1 or dsRBM2 were evaluated for their ability to edit the wild-type GluR-2 R/G site (Figure 5A). It was necessary to generate double mutants around the carbonyls of V104 in dsRBM1 and S258 in dsRBM2 to change the structure of the main-chain of this loop. All four mutants showed a significant decrease in RNA editing ranging from a near ablation of editing (S258A,H259A in the β1−β2 loop of dsRBM2), to 20% editing (V104A,H105A in the β1−β2 loop of dsRBM1 and M84A in helix α1 of dsRBM1), to 30% editing (M238A in helix α1 of dsRBM2) of that demonstrated by the wild-type protein. These data clearly show that the loss of the sequence-specific contacts of any of the two dsRBMs strongly decreases editing at the R/G site with the contact mediated by the β1−β2 loop of dsRBM2 more strongly affecting editing than the other contacts. In agreement with deletion studies of ADAR2 (Macbeth et al., 2004; Stefl et al., 2006), the S258A,H259A mutations have a stronger effect, likely due to the binding of the β1−β2 loop of dsRBM2 near the editing site.

Converse experiments in which mutations in the sequence-specific recognition motifs of dsRBM2 (mut1 and mut2), dsRBM1 (mut4) or both (mut3) within the GluR-2 RNA (Figure S4) were assessed for their ability to affect R/G editing by wild-type ADAR2 revealed a significant decrease in maximal editing rates (V_{max}) for all RNA mutants tested (Figure 5B) providing further support for the functional significance of these contacts. Best-fit kinetic curves for wild-type and mutant RNAs corresponded to a model of substrate inhibition, consistent with previously observed kinetic models for ADARs in which the formation of a ternary complex containing an ADAR dimer and RNA substrate is required for efficient adenosine deamination.

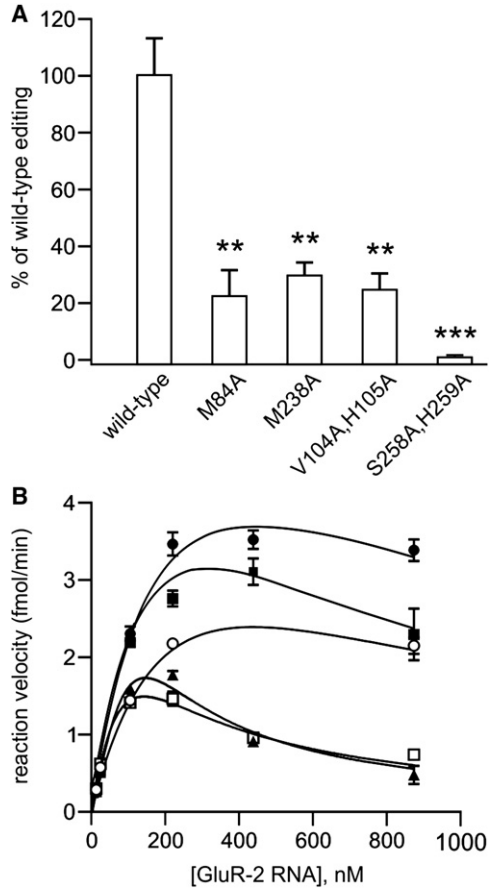


Figure 5. Sequence-Specific Contacts of ADAR dsRBMs Are Important for Editing Activity

(A) Quantitative analysis of in vitro editing efficiency for ADAR2 dsRBM double mutants; all mutants were assayed in duplicate for in vitro editing activity at the GluR-2 R/G site using three independent nuclear extracts (mean ± SEM; *$p < 0.05$, **$p < 0.005$; ***$p < 0.001$).

(B) Kinetic analysis of wild-type ADAR2 editing with GluR-2 R/G mutants. Increasing concentrations of GluR-2 RNAs (see Figure S4; wild-type ●; mut 1 ■, mut 2 ○, mut 3 ▲, mut 4 □) were incubated with wild-type rat ADAR2 protein as described above; all mutant RNAs were assayed in triplicate for determination of in vitro reaction velocity (mean ± SEM). Nonlinear fitting of kinetic curves corresponded to a model of substrate inhibition (R^2 = 0.91-0.98 for all RNAs) with V_{max} values corresponding to 3.92, 3.84, 2.08, 1.20, and 1.29 fmol/min for wild-type, mut1, mut2, mut3, and mut4, respectively. Wild-type and mutant sequences are shown in Figure S4.

DISCUSSION

In solving the structure of ADAR2 dsRBMs bound to the GluR-2 R/G site, we demonstrated that despite forty-four possible

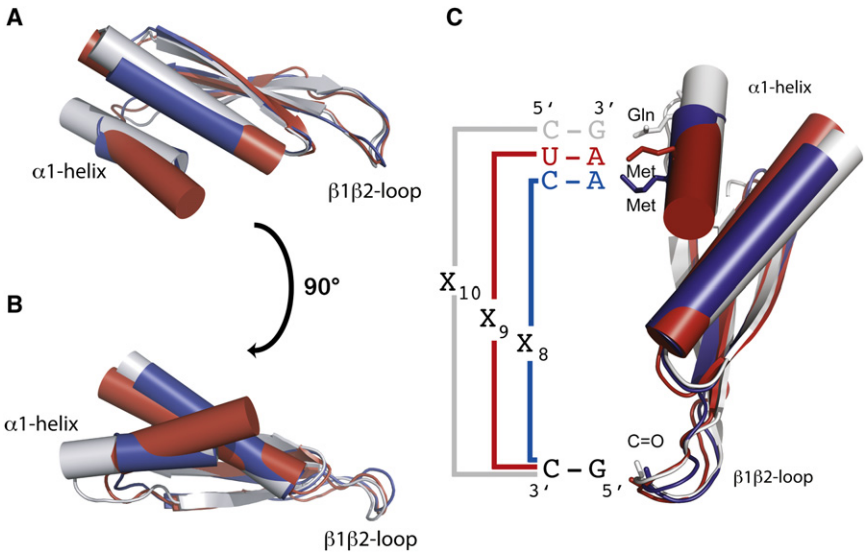


Figure 6. RNA Recognition Code of Various dsRBMs

(A) and (B) Overlay of the ADAR2 dsRBM1 (in blue), ADAR2 dsRBM2 (in red), and Aquifex aeolicus RNaseIII dsRBM (in gray) structures highlights the variability of helix α1 within the dsRBM fold and its importance for the determination of the register length between the two specific contacts on the RNA helix (C). For Aquifex aeolicus RNaseIII dsRBM–dsRNA interactions, see also Figure S5 and for sequence alignments of different dsRBMs, see also Figure S6.

binding sites on the GluR-2 R/G RNA stem-loop (considering a 32 base-pair stem, a 10 base-pair register between the two sequence-specific contacts and two possible orientations for the dsRBM), each dsRBM binds at a very specific register on this large RNA molecule. This binding is achieved by a direct readout of the RNA sequence in the minor groove of the A-form helix. The two dsRBMs of ADAR2 use helix α1 and the β1-β2 loop as molecular rulers to find their binding register in the RNA minor groove of the GluR-2 R/G RNA. Through the β1−β2 loop, the carbonyl oxygens of Val104 in dsRBM1 and Ser258 in dsRBM2 contact the amino groups of base-paired guanines, G22 and G9 respectively. The same type of sequence-specific RNA recognition of GC or GU base-pairs in the minor groove of RNA helices have been observed in several ribosomal proteins of the large subunit (Klein et al., 2004) and in some tRNA synthetases bound to RNA (Rould et al., 1989) although the fold of these proteins and the overall binding mode are different from a dsRBM. Through helix α1, the side-chain methyl groups of Met84 in dsRBM1 and of Met238 in dsRBM2 are in contact with the H2s of A32 and A18, respectively. Recognition of these two anchoring points in the minor-groove, separated by 9 and 8 base-pairs for dsRBM1 and dsRBM2, respectively, illustrates how the two dsRBMs find their sequence-specific binding registers, demonstrating that these dsRBMs have more sequence-specificity than previously thought. Interestingly, in each complex, one of the two anchoring points involves a mismatched base-pair (the G22-G50 base-pair for dsRBM1 and the A18-C54 base-pair for dsRBM2). It is therefore possible that the highly exposed amino or C2H2 groups of these mismatches in the minor groove further assist the dsRBMs of ADAR2 to find their binding register, supporting earlier findings that these two mismatches are important for positioning ADAR2 at the R/G site (Ohman et al., 2000). In addition to sequence-specific interactions between ADAR2 dsRBMs and its GluR-2 target, K127 (dsRBM1) and K281 (dsRBM2) make contacts with phosphate oxygens across the major groove of the RNA (Figure 2). These basic amino acid moieties are conserved in the loop between the β3 and α2 regions for all dsRBMs (Tian et al., 2004) and mutation of these residues in PKR and Staufen have been shown to ablate dsRNA-binding activity (McMillan et al., 1995; Ramos et al., 2000), indicating the importance of both sequence-specific and sequence–independent recognition of the RNA substrate for site-specific adenosine deamination.

Prior to this work, the structures of only four dsRBM-containing proteins in complex with RNA had been determined by X-ray crystallography (XlrbpA and *Aquifex aeolicus* (Aa) RNaseIII) or NMR spectroscopy (Staufen and Rnt1p; Gan et al., 2006; Ramos et al., 2000; Ryter and Schultz, 1998; Wu et al., 2004). In the two solution structures, the dsRBMs appear to recognize primarily the loop of the RNA while in the two crystal structures the dsRBMs are found bound across the junction between coaxially stacked helices. Lack of clear sequence-specific contacts led to the general opinion that dsRBMs are shape-specific rather than sequence-specific RNA binding domains (Stefl et al., 2005a). The two dsRBM-RNA complexes of ADAR2 reported here have revealed that dsRBMs recognize not only the shape of the RNA (a stem-loop for dsRBM1 and an A-form helix for dsRBM2), but also more surprisingly the sequence of the RNA. Interestingly, in a recent crystal structure of an Aa RNaseIII dsRBM bound to a stem-loop, sequence-specific contacts in the minor groove via helix α1 and the β1−β2 loop have been observed (Gan et al., 2008). The helix α1 in Aa RNaseIII is elongated by one turn compared to the helix α1 of the dsRBMs of ADAR2 and a Gln side-chain recognizes a guanine by two sequence-specific hydrogen bonds (Figure S5). The contact mediated by the β1−β2 loop in Aa RNaseIII are similar to the dsRBMs in ADAR2. The β1−β2 loop has the same length (six amino acids) and the main-chain carbonyl of the third residue of the loop is hydrogen bonded to a guanine amino of a GU base-pair. Despite similarities in the mode of binding, the three dsRBMs recognize different sequences and different register lengths. The dsRBM of Aa RNaseIII preferentially recognizes an RNA helix containing a G-X_{10}-G sequence while the dsRBM1 and dsRBM2 of ADAR2 preferentially recognize G-X_9-A and G-X_8-A sequences, respectively (Figure 6). The length and the positioning of helix α1 relative to the dsRBM fold appear to be the key structural elements that determine the register length of the different dsRBMs (Figure 6C).

Our findings regarding the RNA binding specificity of dsRBMs have important implications for the sequence-specificity paradox of ADAR2, but also of many other dsRBM-containing proteins that continue to puzzle investigators (Tian et al., 2004). Apparent differences in the sequences of dsRBMs between mammalian ADAR2 and ADAR1 (Figure S6), where ADAR1 dsRBMs appears to have a longer helix α1 and lack the ADAR2 equivalent of Met 84 and Met 238, could explain why ADAR1 and ADAR2 have different substrate specificities (Bass, 2002; Lehmann and Bass, 2000). Furthermore, our structure shows how dsRBM2 of ADAR2 binds the GluR-2 R/G site near the editing site in recognizing the amino group of the guanosine 3′ to the edited A. This would explain the strong preference for a guanosine moiety 3′ to the edited adenosine that is found in a great majority of substrates selectively edited by ADAR2 (Bass, 2002; Lehmann and Bass, 2000; Li et al., 2009; Riedmann et al., 2008) and more recently in long double-stranded RNA (Eggington and Bass, personal communication). This sequence preference disappears when the dsRBMs are deleted from ADAR2 (Eggington and Bass, personal communication) further supporting that this sequence requirement is due to dsRBM binding. Finally, in interacting with the guanosine 3′ to the edited adenosine and to the nucleotide that base-pairs with the editing site, dsRBM2 not only brings the deaminase domain in close proximity to the editing site, but also does not prevent access of the adenosine to the deaminase domain. When this precise positioning is impaired, specific editing is nearly abolished (see the effect of the S258A, H259A mutant) which emphasizes the functional importance of sequence-specific recognition of RNA by dsRBMs for A-to-I editing.

The sequence-specific contacts that we observed with the dsRBMs ADAR2 are interesting when comparing sequence alignments of several dsRBM structures that have been determined (Figure S6). This alignment reveals a surprisingly high variability in the length and amino acid sequence composition of the two regions of the dsRBMs mediating the sequence-specific interactions with the RNA, namely the helix α1 and the β1−β2 loop. This strongly suggests that dsRBMs are likely to have different binding specificity in agreement with reports indicating that dsRBMs from different proteins are not functionally interchangeable (Liu et al., 2000; Parker et al., 2008). Similar to ADAR2, many dsRBM-containing proteins involved in miRNA and siRNA processing and function are likely to bind RNA in a sequence-specific manner, that would modulate their target selection and mechanism of action. For example, DICER was shown to compete with ADARs for the same RNA substrates (Kawahara et al., 2007; Yang et al., 2006). Interestingly, ADARs modulate the processing of miRNA precursors not only by A-to-I modifications that alter the secondary structure of pri-miRNA (Kawahara et al., 2007; Tonkin and Bass, 2003; Yang et al., 2006), but also simply by RNA-binding alone to pri-miRNAs, as recently shown with catalytically inactive ADARs (Heale et al., 2009). This latter function for ADARs, as regulators of pri-miRNA processing, closely resemble that found for single-stranded sequence-specific RNA-binding proteins such as Lin28, hnRNP A1 or KSRP (Guil and Caceres, 2007; Heo et al., 2008, 2009; Michlewski et al., 2008; Newman et al., 2008; Trabucchi et al., 2009). Furthermore, RNAi activity has been shown to coincide with siRNA sequence motifs (Katoh and Suzuki, 2007). Altogether it is becoming clear that sequence-specific recognition mediated by dsRBMs is functionally important for dsRBM containing proteins. We have demonstrated here with ADAR2 how such sequence-specific recognition is mediated in dsRBMs and how this is relevant for RNA editing. Future work will be required to elucidate the variations in dsRNA-binding specificity and their functional relevance for numerous other members of the dsRBM-containing protein family.

EXPERIMENTAL PROCEDURES

Preparation of Proteins

Details on cloning, expression and purification of the ADAR2 dsRBM1, ADAR2 dsRBM2, and ADAR2 dsRBM12 constructs have been described previously (Stefl et al., 2005b, 2006).

NMR Spectroscopy

All NMR spectra were acquired at 310 K. Spectra were recorded at 500, 600, and 900 MHz Bruker spectrometers. All spectra were processed with XWINNMR or Topspin1.3/2.0 (Bruker BioSpin) and analyzed with Sparky 3.0 (Goddard T.G. and Kellner D.G., University of California, San Francisco). The ^{1}H, ^{13}C and ^{15}N chemical shifts of the protein in complex, were assigned by standard methods (Sattler et al., 1999). The ^{1}H-^{15}N HSQC and ^{1}H-^{13}C HSQC spectra of dsRBM1 and dsRBM2 in free and bound forms are shown in Figure S1 and Figure S2. All distance restraints were derived from 3D ^{15}N,^{13}C-edited NOESYs and 2D ^{1}H-^{1}H NOESY (t_m = 150 ms) collected at 900 MHz. RNA exchangeable proton resonances were assigned using ^{1}H-^{1}H NOESY spectrum (t_m = 200 ms) at 278 K. Nonexchangeable proton resonances were assigned using ^{1}H-^{1}H, NOESY, ^{1}H-^{1}H TOCSY, ^{1}H-^{13}C HSQC, 3D ^{13}C-edited NOESY, 2D ^{1}H-^{1}H double-half-filtered NOESY (t_m = 150 ms) (Peterson et al., 2004) and 3D ^{13}C F_1-edited, F_3-filtered NOESY-HSQC spectrum (t_m = 150 ms) (Zwahlen et al., 1997) in 99.99% $^{2}H_2O$ (v/v). The NOEs were semiquantitatively classified based on their intensities in the 2D and 3D NOESY spectra. Hydrogen bond distance restraints were used for base-pairs, when the imino-protons were observed experimentally. The assignments of intermolecular NOEs were based on 3D ^{13}C F_1-edited, F_3-filtered NOESY-HSQC spectrum (t_m = 150 ms), 2D ^{1}H-^{1}H F_1-^{13}C-filtered F_2-^{13}C-edited NOESY (t_m = 150 ms) on the protein-RNA complexes with either the protein or the RNA ^{13}C-^{15}N labeled. In case of dsRBM2–GluR-2 R/G LSL RNA complex, we observed an extra set of five weaker intermolecular nOes, which were discarded from structure calculation. These intermolecular restraints cannot be explained with the presented structure of dsRBM2-GluR-2 R/G LSL RNA complex. They originate from a minor conformation in which the protein is shifted up by one base pair toward the UUCG tetraloop.

Structure Calculation and Refinement

Distance constraints for the proteins bound to RNA where generated by the ATNOS/CANDID package (Herrmann et al., 2002). The accuracy of the list of automatically generated distance constraints was manually checked. Distance constraints for the free and bound RNAs as well as for the intermolecular NOEs were assigned manually. Preliminary structures of the free RNA and the protein-RNA complexes were obtained by a simulated annealing protocol in CYANA (Guntert et al., 1997; Herrmann et al., 2002). To impose better convergence of the ensemble, an artificial torsion angles for the canonical dsRNA regions were used as described previously (Oberstrass et al., 2006). Additional angle restraints to maintain proper local geometries were used (Tsui et al., 2000). The final refinement of all structures was performed using a 20 ps simulated annealing protocol in AMBER (Case et al., 2002) as described in the Supplemental Information. From 40 refined structures, the twenty conformers with the lowest AMBER energy were selected to form the final ensemble of structures. Structural quality was assessed using PROCHECK (Laskowski et al., 1996). Figures were prepared with MOLMOL (Koradi et al., 1996) and Pymol (DeLano, 2002).

Fluorescence Anisotropy

Fluorescence anisotropy was measured on a FluoroMax-4 spectrofluorometer (Horiba Jobin-Yvon, USA) equipped with a thermostated cell holder and an automatic titrator. All measurements were conducted in 50mM sodium phosphate buffer (pH 7.0) and 100mM NaCl at 10°C. To avoid any effects caused by 5′-end labeling of RNAs, the experiments were designed as a competition assay. At first, a reference measurement was carried out in which 1400 μl of 10nM fluorescein labeled wild-type RNA was titrated by the protein. Then, the same titration experiment was repeated in the presence of 500nM unlabeled RNA (either wild-type or mutants; Vasiljeva et al., 2008). Total volume of protein added to each reaction was 33 μl. The fitting was performed using DynaFit software (Kuzmic, 1996, 2006). Initially, the K_d for the reference protein–labeled RNA complex was determined. The obtained K_d value was then used as a fixed parameter when fitting the competition data. A 1:1 binding stoichiometry was assumed in all cases. The data were normalized for visualization purposes.

Quantitative Analysis of In Vitro RNA Editing

For in vitro editing reactions, a 116 nt RNA encoding a portion of the mouse GluR-2 pre-mRNA with the complete R/G duplex was transcribed in vitro (Stefl et al., 2006) and incubated with wild-type or mutant ADAR2 proteins derived from nuclear extracts obtained from transiently transfected HEK293 cells (Sansam et al., 2003). Equivalent amounts of wild-type and mutant ADAR2 protein, as determined by Western blotting, were incubated with 40 ng of the R/G transcript at 30°C for 20 min. These incubation conditions were determined empirically by performing time-course analyses with wild-type ADAR2 protein to ensure that the assay was in the linear range (*data not shown*). The reaction was stopped and the R/G transcript isolated by direct addition of TRI Reagent (Molecular Research Center) at the end of the incubation period. For quantification of RNA editing, the in vitro reaction product was reverse transcribed using AMV Reverse Transcriptase (Promega) and an antisense primer (5′-CGGCCAATCGTACGTACCTCCGGCCGAATTCTACAAACC GTTAAGAGTCTTA-3′) with a unique 5′-extension (underlined). The resulting amplicon was diluted 1:1000 and 1 μl was subsequently amplified by PCR using sense (5′-CCGGGAGCTCATCGCCACACCTAAAGGATCC-3′) and antisense (5′-CGGCCAATCGTACGTACCTCC-3′) primers corresponding to GluR-2 and the unique 5′-extension sequences, respectively. PCR amplicons were purified using the Wizard SV PCR and Gel Cleanup System (Promega) and digested with Mse I (New England Biolabs) to generate 100 and 70 bp products representing edited and nonedited transcripts, respectively. The resulting digestion products were resolved on a 4% Agarose gel and editing efficiency was quantified by phosphorimager analysis (GE Healthcare).

In vitro editing reactions using GluR-2 R/G mutant RNAs were performed as described above with equivalent amounts of wild-type ADAR2 protein derived from nuclear extracts obtained from transiently transfected HEK293 cells (Sansam et al., 2003). Wild-type and mutant transcripts were trace labeled with [α-^{32}P]-UTP and concentrations of in vitro transcribed RNAs were determined using a Perkin-Elmer Tri-Carb 2800TR scintillation spectrometer based upon the calculated specific activity for each transcript.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and six figures and can be found with this article online at doi:10.1016/j.cell.2010.09.026.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation (Nr. 3100A0-118118) and the SNF-NCCR structural biology to F.H.T.A and the National Institutes of Health (R01 NS33323) to R.B.E. R.S. is supported by the Ministry of Education of the Czech Republic (MSM0021622413, Ingo LA08008), GACR (204/08/1212, 305/10/1490), GAAV (IAA401630903), HHMI/EMBO start-up grant, and HFSP Career Development Award. M.Z. and C.H. are supported by GACR (204/08/H054) and by the Ministry of Education of the Czech Republic (MSM0021622415). M.Z. is in receipt of a Brno City Scholarship for Talented Ph.D. Students. The coordinates of the structures of GluR-2 LSL RNA, ADAR2 dsRBM1 bound to GluR-2 USL RNA, ADAR2 dsRBM2 bound to GluR-2 LSL RNA and ADAR2 dsRBM12 bound to GluR-2 have been deposited in the Protein Data Bank with accession codes 2l2j, 2l3c, 2l2k, and 2l3j, respectively.

Received: September 21, 2009
Revised: May 26, 2010
Accepted: August 30, 2010
Published: October 14, 2010

REFERENCES

Aruscavage, P.J., and Bass, B.L. (2000). A phylogenetic analysis reveals an unusual sequence conservation within introns involved in RNA editing. RNA *6*, 257–269.

Bass, B.L. (2002). RNA editing by adenosine deaminases that act on RNA. Annu. Rev. Biochem. *71*, 817–846.

Blaszczyk, J., Gan, J., Tropea, J.E., Court, D.L., Waugh, D.S., and Ji, X. (2004). Noncatalytic assembly of ribonuclease III with double-stranded RNA. Structure *12*, 457–466.

Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham, T.E., III, Wang, J., Ross, W.S., Simmerling, C.L., Darden, T.A., Merz, K.M., Stanton, R.V., et al. (2002). AMBER 7 (San Francisco: University of California).

Chilibeck, K.A., Wu, T., Liang, C., Schellenberg, M.J., Gesner, E.M., Lynch, J.M., and MacMillan, A.M. (2006). FRET analysis of in vivo dimerization by RNA-editing enzymes. J. Biol. Chem. *281*, 16530–16535.

Cho, D.S., Yang, W., Lee, J.T., Shiekhattar, R., Murray, J.M., and Nishikura, K. (2003). Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. J. Biol. Chem. *278*, 17093–17102.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (Palo Alto, CA, USA: DeLano Scientific).

Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nat. Rev. Mol. Cell Biol. *3*, 195–205.

Gallo, A., Keegan, L.P., Ring, G.M., and O'Connell, M.A. (2003). An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. EMBO J. *22*, 3421–3430.

Gan, J., Shaw, G., Tropea, J.E., Waugh, D.S., Court, D.L., and Ji, X. (2008). A stepwise model for double-stranded RNA processing by ribonuclease III. Mol. Microbiol. *67*, 143–154.

Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. Cell *124*, 355–366.

Guil, S., and Caceres, J.F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nat. Struct. Mol. Biol. *14*, 591–596.

Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. J. Mol. Biol. *273*, 283–298.

Heale, B.S., Keegan, L.P., McGurk, L., Michlewski, G., Brindle, J., Stanton, C.M., Caceres, J.F., and O'Connell, M.A. (2009). Editing independent effects of ADARs on the miRNA/siRNA pathways (EMBO J.).

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol. Cell *32*, 276–284.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696–708.

Herrmann, T., Guntert, P., and Wuthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J. Mol. Biol. *319*, 209–227.

Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P.H. (2000). Point mutation in an AMPA

receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature *406*, 78–81.

Higuchi, M., Single, F.N., Kohler, M., Sommer, B., Sprengel, R., and Seeburg, P.H. (1993). RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. Cell *75*, 1361–1370.

Katoh, T., and Suzuki, T. (2007). Specific residues at every third position of siRNA shape its efficient RNAi activity. Nucleic Acids Res. *35*, e27.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., and Nishikura, K. (2007). Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. Science *315*, 1137–1140.

Klein, D.J., Moore, P.B., and Steitz, T.A. (2004). The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. J. Mol. Biol. *340*, 141–177.

Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph *14*, 51–55, 29-32.

Kuzmic, P. (1996). Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. Anal. Biochem. *237*, 260–273.

Kuzmic, P. (2006). A generalized numerical approach to rapid-equilibrium enzyme kinetics: application to 17beta-HSD. Mol. Cell. Endocrinol. *248*, 172–181.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Legault, P., and Pardi, A. (1994). In-Situ Probing of Adenine Protonation in Rna by C-13 Nmr. J. Am. Chem. Soc. *116*, 8390–8391.

Lehmann, K.A., and Bass, B.L. (2000). Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. Biochemistry *39*, 12875–12884.

Li, J.B., Levanon, E.Y., Yoon, J.K., Aach, J., Xie, B., Leproust, E., Zhang, K., Gao, Y., and Church, G.M. (2009). Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. Science *324*, 1210–1213.

Liu, Y., Lei, M., and Samuel, C.E. (2000). Chimeric double-stranded RNA-specific adenosine deaminase ADAR1 proteins reveal functional selectivity of double-stranded RNA-binding domains from ADAR1 and protein kinase PKR. Proc. Natl. Acad. Sci. USA *97*, 12541–12546.

Lomeli, H., Mosbacher, J., Melcher, T., Hoger, T., Geiger, J.R., Kuner, T., Monyer, H., Higuchi, M., Bach, A., and Seeburg, P.H. (1994). Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. Science *266*, 1709–1713.

Macbeth, M.R., Lingam, A.T., and Bass, B.L. (2004). Evidence for auto-inhibition by the N terminus of hADAR2 and activation by dsRNA binding. RNA *10*, 1563–1571.

Macbeth, M.R., Schubert, H.L., Vandemark, A.P., Lingam, A.T., Hill, C.P., and Bass, B.L. (2005). Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. Science *309*, 1534–1539.

McMillan, N.A., Carpick, B.W., Hollis, B., Toone, W.M., Zamanian-Daryoush, M., and Williams, B.R. (1995). Mutational analysis of the double-stranded RNA (dsRNA) binding domain of the dsRNA-activated protein kinase, PKR. J. Biol. Chem. *270*, 2601–2606.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P.H., and Higuchi, M. (1996). A mammalian RNA editing enzyme. Nature *379*, 460–464.

Michlewski, G., Guil, S., Semple, C.A., and Caceres, J.F. (2008). Posttranscriptional regulation of miRNAs harboring conserved terminal loops. Mol. Cell *32*, 383–393.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA *14*, 1539–1549.

Nishikura, K. (2006). Editor meets silencer: crosstalk between RNA editing and RNA interference. Nat. Rev. Mol. Cell Biol. *7*, 919–931.

Oberstrass, F.C., Lee, A., Stefl, R., Janis, M., Chanfreau, G., and Allain, F.H. (2006). Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. Nat. Struct. Mol. Biol. *13*, 160–167.

Ohman, M., Kallman, A.M., and Bass, B.L. (2000). In vitro analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. RNA *6*, 687–697.

Palladino, M.J., Keegan, L.P., O'Connell, M.A., and Reenan, R.A. (2000). A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. Cell *102*, 437–449.

Parker, G.S., Maity, T.S., and Bass, B.L. (2008). dsRNA binding properties of RDE-4 and TRBP reflect their distinct roles in RNAi. J. Mol. Biol. *384*, 967–979.

Peterson, R.D., Theimer, C.A., Wu, H., and Feigon, J. (2004). New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes. J. Biomol. NMR *28*, 59–67.

Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D., and Varani, G. (2000). RNA recognition by a Staufen double-stranded RNA-binding domain. EMBO J. *19*, 997–1009.

Riedmann, E.M., Schopoff, S., Hartner, J.C., and Jantsch, M.F. (2008). Specificity of ADAR-mediated RNA editing in newly identified targets. RNA *14*, 1110–1118.

Rould, M.A., Perona, J.J., Soll, D., and Steitz, T.A. (1989). Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. Science *246*, 1135–1142.

Ryter, J.M., and Schultz, S.C. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. EMBO J. *17*, 7505–7513.

Sansam, C.L., Wells, K.S., and Emeson, R.B. (2003). Modulation of RNA editing by functional nucleolar sequestration of ADAR2. Proc. Natl. Acad. Sci. USA *100*, 14018–14023.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucl. Magn. Reson. Spectrosc. *34*, 93–158.

Stefl, R., and Allain, F.H. (2005). A novel RNA pentaloop fold involved in targeting ADAR2. RNA *11*, 592–597.

Stefl, R., Skrisovska, L., and Allain, F.H. (2005a). RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO Rep. *6*, 33–38.

Stefl, R., Skrisovska, L., Xu, M., Emeson, R.B., and Allain, F.H. (2005b). Resonance assignments of the double-stranded RNA-binding domains of adenosine deaminase acting on RNA 2 (ADAR2). J. Biomol. NMR *31*, 71–72.

Stefl, R., Xu, M., Skrisovska, L., Emeson, R.B., and Allain, F.H. (2006). Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. Structure *14*, 345–355.

Tian, B., Bevilacqua, P.C., Diegelman-Parente, A., and Mathews, M.B. (2004). The double-stranded-RNA-binding motif: interference and much more. Nat. Rev. Mol. Cell Biol. *5*, 1013–1023.

Tonkin, L.A., and Bass, B.L. (2003). Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. Science *302*, 1725.

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A.D., Filipowicz, W., Ramos, A., Gherzi, R., and Rosenfeld, M.G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature *459*, 1010–1014.

Tsui, V., Zhu, L., Huang, T.H., Wright, P.E., and Case, D.A. (2000). Assessment of zinc finger orientations by residual dipolar coupling constants. J. Biomol. NMR *16*, 9–21.

Valente, L., and Nishikura, K. (2005). ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. Prog. Nucleic Acid Res. Mol. Biol. *79*, 299–338.

Valente, L., and Nishikura, K. (2007). RNA binding-independent dimerization of adenosine deaminases acting on RNA and dominant negative effects of nonfunctional subunits on dimer functions. J. Biol. Chem. *282*, 16054–16061.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat. Struct. Mol. Biol. *15*, 795–804.

Wu, H., Henras, A., Chanfreau, G., and Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. Proc. Natl. Acad. Sci. USA *101*, 8307–8312.

Xu, M., Wells, K.S., and Emeson, R.B. (2006). Substrate-dependent contribution of double-stranded RNA-binding motifs to ADAR2 function. Mol. Biol. Cell *17*, 3211–3220.

Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R., and Nishikura, K. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nat. Struct. Mol. Biol. *13*, 13–21.

Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R., and Kay, L.E. (1997). Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage lambda N-peptide/boxB RNA complex. J. Am. Chem. Soc. *119*, 6711–6721.

⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 5, pp. 3645–3657, February 4, 2011

Recognition of Transcription Termination Signal by the Nuclear Polyadenylated RNA-binding (NAB) 3 Protein*

Received for publication, June 25, 2010, and in revised form, November 9, 2010 Published, JBC Papers in Press, November 17, 2010, DOI 10.1074/jbc.M110.158774

Fruzsina Hobor[‡1,2], Roberto Pergoli[‡1,3], Karel Kubicek[‡], Dominika Hrossova[‡], Veronika Bacikova[‡], Michal Zimmermann[§2,4], Josef Pasulka[‡], Ctirad Hofr[§], Stepanka Vanacova[‡], and Richard Stefl[‡5]

From the [‡]National Centre for Biomolecular Research and [§]Department of Functional Genomics and Proteomics, Central European Institute of Technology, Faculty of Science, Masaryk University, Brno CZ-62500, Czechia

Non-coding RNA polymerase II transcripts are processed by the poly(A)-independent termination pathway that requires the Nrd1 complex. The Nrd1 complex includes two RNA-binding proteins, the nuclear polyadenylated RNA-binding (Nab) 3 and the nuclear pre-mRNA down-regulation (Nrd) 1 that bind their specific termination elements. Here we report the solution structure of the RNA-recognition motif (RRM) of Nab3 in complex with a UCUU oligonucleotide, representing the Nab3 termination element. The structure shows that the first three nucleotides of UCUU are accommodated on the *β*-sheet surface of Nab3 RRM, but reveals a sequence-specific recognition only for the central cytidine and uridine. The specific contacts we identified are important for binding affinity *in vitro* as well as for yeast viability. Furthermore, we show that both RNA-binding motifs of Nab3 and Nrd1 alone bind their termination elements with a weak affinity. Interestingly, when Nab3 and Nrd1 form a heterodimer, the affinity to RNA is significantly increased due to the cooperative binding. These findings are in accordance with the model of their function in the poly(A) independent termination, in which binding to the combined and/or repetitive termination elements elicits efficient termination.

RNA Polymerase II (RNA Pol II)[6] transcribes messenger RNA (mRNA), but also a subset of small nuclear and small nucleolar RNAs (snRNAs/snoRNAs), micro-RNA precursors, and a class of intergenic and antisense RNAs (1). RNA Pol II uses two different mechanisms for transcription termination of these "coding" and "non-coding" RNAs. Although the RNA Pol II termination of mRNA requires a large multiprotein complex that recognizes the poly(A) signal in the nascent transcript (2), the termination of the non-coding RNAs requires no poly(A) signal (2–4).

In the poly(A)-independent mechanism, transcription termination requires a specific factor, the Nrd1 complex. This complex consists of three proteins: the nuclear pre-mRNA down-regulation (Nrd) 1 protein, the nuclear polyadenylated RNA-binding (Nab) 3 protein, and the putative RNA helicase Sen1 (5–7). The Nrd1 complex interacts with the exosome, a complex of 10–12 exoribonucleolytic and RNA-binding proteins (8) and the Trf4-Air2-Mtr4 polyadenylation (TRAMP) complex (9–11), which are involved in the 3′ end processing of non-coding RNA transcripts (3, 4, 7).

In yeast, transcription termination mediated by the Nrd1 complex requires binding to both the nascent RNA and the carboxyl-terminal domain of RNA Pol II, which consists of 26 repeats of the sequence Tyr^1-Ser^2-Pro^3-Thr^4-Ser^5-Pro^6-Ser^7 (1, 12). Interestingly, the Nrd1 complex binds the carboxyl-terminal domain when it is phosphorylated at Ser^5, a typical feature of the early elongation phase of the transcription cycle. The Ser^5-phosphorylated carboxyl-terminal domain is recognized by the carboxyl-terminal domain-interacting domain of Nrd1 (13, 14). The RNA-binding subunits of the Nrd1 complex, Nrd1 and Nab3, recognize their specific RNA sequences (called terminator elements) in the nascent transcripts of RNA Pol II. It is believed that this specific binding of Nrd1 complex to the terminator elements is the initial step in the assembly of termination machinery.

A number of studies narrowed the sequence regions with terminator elements (5, 6, 15–17) that were subsequently identified as GUAR (where R stands for purine) and UCUU sequences (18). GUAR and UCUU terminator elements are recognized by Nrd1 and Nab3, respectively, via their fragments encompassing RNA recognition motifs (RRMs) (18). These terminator sequences are located downstream of snRNA and snoRNA genes (18) although their relative orientation and spacing are not highly conserved. In addition, it was demonstrated that Nrd1 and Nab3 form a stable het-

* This work was supported in part by a Howard Hughes Medical Institute/European Molecular Biology Organization start-up grant, a Human Frontier Science Program Career Development Award, Ministry of Education of the Czech Republic Grants MSM0021622413, MSM0021622415, and Ingo LA08008, Czech Science Foundation Grants 204/08/1212 and 305/10/1490, Grant Agency of the Academy of Sciences of the Czech Republic IAA401630903, Wellcome Trust Grant 084316/Z/07/Z, and EMBO Installation Grant 1642.

The atomic coordinates and structure factors (codes 2KVI and 2L41) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

⌘ *Author's Choice*—Final version full access.

[1] Both authors contributed equally to this work.

[2] Supported by Brno City Municipality Scholarships for Talented Ph.D. Students.

[3] Supported by the European Community FP-7 Grant 205872.

[4] Supported by the GACR Grant 204/08/H054.

[5] To whom correspondence should be addressed: University Campus Bohunice, Kamenice 5/A4, Brno, CZ-62500 Brno, Czechia. Tel.: 420549492436; Fax: 420549492556; E-mail: stefl@chemi.muni.cz.

[6] The abbreviations used are: RNA Pol II, RNA polymerase II; FA, fluorescence anisotropy; Nab3, nuclear polyadenylated RNA-binding 3; Nrd1, nuclear pre-mRNA down-regulation 1; RRM, RNA-recognition motif; snoRNAs, small nucleoar RNAs; HSQC, heteronuclear single quantum coherence; RNP, ribonucleoprotein; TAMRA, *N*,*N*,*N*′,*N*′-tetramethyl-6-carboxyrhodamine; PTB, polypyrimidine tract-binding protein.

erodimer and bind to snoRNA terminators that contain multiple Nrd1- and Nab3-binding sequences (19).

Both Nrd1 and Nab3 contain RRM that likely mediates the binding to their specific RNA sequences. The RRM is the most abundant RNA-binding domain in higher vertebrates; *e.g.* the RRM is present in about 2% of human genes (20). It is a small protein domain of ~90 amino acids with a typical $\beta\alpha\beta\beta\alpha\beta$ topology that forms a four-stranded β-sheet packed against two α-helices (21–23). The structure of this domain is relatively well defined despite a little sequence conservation among various RRMs. The solved structures of RRM bound to RNA show the complexity of protein-RNA recognition mediated by the RRM, which often involves not only RRM-RNA interactions but also RRM-RRM and other RRM-protein interactions. The main protein surface of the RRM involved in the interaction with the RNA is the four-to-five-stranded β-sheet, which typically contacts two or three nucleotides. Frequently, RRM-containing proteins bind more than three nucleotides and recognize longer single-stranded RNA or even internal RNA loops by employing of β-strand loops and N- or C-terminal flanking regions of RRMs (21–23).

To better understand the structural basis behind the poly(A) independent transcription termination pathway, we initiated an NMR study of *Saccharomyces cerevisiae* Nab3. Here, we present the three-dimensional solution structure of the Nab3 RRM in free form and in complex with the 5′-UCUU-3′ RNA substrate. The structure of the complex reveals recognition of the YCU sequence (where Y stands for pyrimidine) by the Nab3 RRM. We confirmed the sequence-specific intermolecular contacts by site-directed mutagenesis and fluorescence anisotropy (FA) measurements, and their physiological role was also confirmed by yeast phenotypic analyses. Finally, we demonstrate that the weak RNA binding of the isolated RRMs of Nab3 and Nrd1 is greatly enhanced when Nab3 and Nrd1 form a heterodimer and bind the RNA cooperatively.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification of Proteins—The coding sequence corresponding to the RRM of the *Nab3* gene from *S. cerevisiae* (961–1245) was amplified by polymerase chain reaction (PCR), and cloned into a pET22b expression vector (Novagen) via NdeI and XhoI restriction sites. The resulting C-terminal His_6-tagged construct was verified by DNA sequencing. The protein was overexpressed in *Escherichia coli* BL21-Codon Plus (DE3)-RIPL (Stratagene), transformed with the pET22b-RRM Nab3 construct at 37 °C in M9 minimal medium, supplemented with 50 mg/liter of ampicillin. For isotope labeling, the medium was supplemented with $^{15}NH_4Cl$ and [U-$^{13}C_6$]glucose. Cells were grown at 37 °C to A_{600} ~1 and induced with 1 mM isopropyl β-D-thiogalactoside. Cells were harvested by centrifugation (6000 × *g* for 10 min), resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM β-mercaptoethanol, pH 8), and disrupted by sonication. The cell debris was cleared by centrifugation (14,000 × *g* for 60 min). Soluble lysate was loaded on a nickel-nitrilotriacetic acid column (Qiagen), equilibrated with lysis buffer, washed with a high salt buffer (50 mM sodium phosphate, 500 mM NaCl, 10 mM β-mercaptoethanol, 5 mM imidazole, pH 8), and eluted with imidazole gradient (50–500 mM) of elution buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM β-mercaptoethanol, pH 8). The protein was subsequently loaded on a Superdex 75 gel filtration column (GE Healthcare), equilibrated with lysis buffer. The protein fractions from gel filtration were dialyzed against lysis buffer. The purified protein was 99% pure, as judged by Coomassie-stained SDS-PAGE. For NMR measurements the pure protein was concentrated to 2.5 mM in 550 μl of 50 mM sodium phosphate (pH 8.0), containing 300 mM NaCl, and 10 mM β-mercaptoethanol. The cloning, expression, and purification of Nrd1 RRM-(340–410) were carried out in the same way as for Nab3 RRM.

The expression and purification of the Nrd1-Nab3 heterodimer have been done in a similar manner as reported previously (19). To improve the yield of expression, we used *E. coli* BL21-Codon Plus (DE3)-RIPL (Stratagene). We used the following final buffer (50 mM Tris (pH 8.0), containing 150 mM NaCl, and 10 mM β-mercaptoethanol) to have the same conditions for all fluorescence anisotropy measurements. Prior to RNA titration, all proteins were tested for the residual RNase activity using RNaseAlert Lab Test (Ambion). RNA oligonucleotides were purchased from Thermo Fisher Scientific/Dharmacon and Sigma.

Generation of Nab3 RRM Mutants—Site-specific mutagenesis was performed using the QuikChange site-directed mutagenesis kit (Stratagene) with complementary sense and antisense (AS) oligonucleotide primers as follows: R331A (S), 5′-gcacaatattcctccgaagtcagcattattcattggtaatttgccg-3′ and (AS), 5′-cggcaaattaccaatgaataatgctgacttcggaggaatattgtgc-3′; N361A (S), 5′-tccatacggtcatatcatgcaaatcgctatcaaaaatgcctttggattcatt-3′ and (AS), 5′-aatgaatccaaaggcatttttgatagcgatttgcatgatatgaccgtatgga-3′; E397A (S), 5′-gcaaaaagttgatcctggcagtttctagctcgaatgc-3′ and (AS), 5′-gcattcgagctagaaactgccaggatcaacttttgc-3′; E397K (S), 5′-tggcaaaaagttgatcctgaaagtttctagctcgaatgc-3′ and (AS) 5′-gcattcgagctagaaactttcaggatcaacttttgcca-3′; S399A (S), 5′-agttgatcctggaagttgctagctcgaatgctcgt-3′ and (AS), 5′-acgagcattcgagctagcaacttccaggatcaact-3′; S399K (S), 5′-actttggcaaaaagttgatcctggaagttaaaagctcgaatgctcgtcc-3′ and (AS), 5′-ggacgagcattcgagcttttaacttccaggatcaacttttgccaaagt-3′. All mutations were verified by DNA sequence analysis.

NMR Spectroscopy—All NMR spectra of 2.5 mM uniformly ^{15}N,^{13}C-labeled Nab3 RRM in 50 mM sodium phosphate buffer (pH 8.0), 300 mM NaCl, 10 mM β-mercaptoethanol (90% H_2O/10% D_2O) were recorded on Bruker AVANCE 600 and 900 MHz spectrometers equipped with a cryoprobe at a sample temperature of 30 °C. All spectra were processed with Topspin 2.1 (Bruker BioSpin) and analyzed with Sparky 3.0 (T. G. Goddard and D. G. Kneller, University of California, San Francisco). The ^1H, ^{13}C, and ^{15}N chemical shifts of Nab3 RRM were assigned as described previously (24). All distance restraints were derived from the three-dimensional ^{15}N- and ^{13}C-edited NOESYs and two-dimensional ^1H-^1H-labeled NOESY (with mixing time of 150 ms) collected at 900 MHz spectrometer.

Akin to the free Nab3 RRM, the backbone resonance assignments of Nab3 RRM in the bound form were achieved

using three-dimensional triple resonance experiments: HNCA, HNCACB, and CBCA(CO)NH (24–26). Resonances of the aliphatic side chains were assigned by a combination of three-dimensional HCCH TOCSY, three-dimensional HNHA, and three-dimensional ^{13}C-edited NOESY spectra. Resonances of the aromatic side chains were assigned using two-dimensional homonuclear NOESY, three-dimensional $^{13}C_{arom}$-edited NOESY, and two-dimensional (HB)CB(CGCD)HD spectra. All distance restraints were derived from the three-dimensional ^{15}N- and ^{13}C-edited NOESYs (with mixing time of 150 ms) collected at 900 MHz spectrometer. The RNA resonances in complex were assigned using a combination of standard through-space and through-bond experiments (27, 28). We could not assign all sugar resonances unambiguously due to a high resonance overlap in the spectra (only unlabeled RNA was used). The sugar pucker conformation for all nucleotides is *C2′-endo*, as identified in the two-dimensional homonuclear TOCSY spectrum (strong cross-peaks between the H1′ and H2′ resonances). Intermolecular distance constraints were obtained from the three-dimensional ^{13}C F_1-edited, F_3-filtered NOESY-HSQC experiment (29), which was recorded in H_2O (with WATERGATE water suppression) as a two-dimensional filter NOESY omitting ^{13}C chemical shift evolution.

Structure Calculations—The preliminary structure determinations of the free and bound Nab3 RRM were performed with the automated NOE assignment module implemented in the CYANA program (30). This automated NOE assignment procedure is a re-implementation of the former CANDID algorithm (31) on the basis of a probabilistic treatment of the NOE assignment. CYANA carries out automated assignment and distance calibration of NOE intensities, removal of meaningless restraints, structure calculation with torsion angle dynamics, and automatic upper distance limit violation analysis. The resultant NOE cross-peak assignments were subsequently confirmed by visual inspection of the spectra. The predicted protein backbone ϕ and ψ torsion angle from the chemical shifts (32) for the secondary structure elements were also included in the calculations. In the next step, CYANA-generated restraints along with manually assigned protein-RNA intermolecular restraints were used for further refinement of the preliminary structures with AMBER 10.0 software (33). This calculations employed a modified version (AMBER ff99SB) of the force field described by Cornell *et al.* (34) along with a refinement protocol described in Padrta *et al.* (35), and an explicit solvent. From 40 refined structures, the 20 conformers with the lowest AMBER energy were selected to form the final ensemble of structures. Structural quality was assessed using PROCHECK (36) and WHAT IF (37). Molecular graphics were generated using MOLMOL (38) and PyMOL (57).

Fluorescence Anisotropy Measurements—The equilibrium binding of Nab3 RRM to different oligonucleotides was analyzed by fluorescence anisotropy. The RNA oligonucleotides were either 5′-labeled with TAMRA or fluorescein attached via a hexyl linker. The measurements were conducted on a FluoroMax-4 spectrofluorometer (Horiba Jobin-Yvon Edison, NJ). The instrument was equipped with a thermostatted cell holder with a Neslab RTE7 water bath (Thermo Scientific). The whole system was operated using FluorEssence software (version 2.5.3.0, Horiba Jobin-Yvon). The TAMRA fluorophore was excited at 561 nm and its emission was collected at 581 nm. The widths of both excitation and emission monochromatic slits were 8 nm and integration time was set to 3 s. The fluorescein fluorophore was excited at 488 nm and its emission was collected at 520 nm. For measurement with the individual domains (Nrd1-(340–410) and Nab3-(331–415)) the width of both excitation and emission monochromatic slits were 7 nm and the integration time was set to 3 s. 10 nM labeled oligonucleotide (volume 1.4 ml) was titrated with increasing amounts of the protein in 50 mM sodium phosphate buffer (pH 7.5), supplemented with 150 mM NaCl and 10 mM β-mercaptoethanol. For measurement with the heterodimer (Nab3-(191–565)–Nrd1-(1–548)), 1 nM fluorescein-labeled RNA was used. Both excitation and emission monochromatic slits were 14 nm, the integration time was set to 3 s.

In all measurements, an identical concentration of the oligonucleotide was included in the protein stock solution to prevent dilution of the RNA during titration. All experiments were carried out at 25 °C in a stirred 1.5-ml quartz cuvette. Protein aliquots were added stepwise until the cuvette was filled. After this point, a certain volume of the sample was always removed from the cuvette before addition of a protein aliquot of the same volume. A fixed delay of 30 s was set between each aliquot addition and start of the measurement to allow the reaction to reach equilibrium. This delay was sufficient, as no further change in anisotropy was observed. Each data point is an average of five measurements. Neither TAMRA nor fluorescein labels showed binding with any of the protein constructs.

The data were analyzed in SigmaPlot 11 software (Systat Software). The experimental isotherms were fit to a single-site binding model according to Heyduk and Lee (39) using nonlinear least squares regression. The data were normalized for visualization purposes.

Yeast Strains and Plasmids—The plasmid for the expression of wild-type Nab3 in yeast contain 550 bp of NAB3 promoter upstream of the AUG and a 286-bp sequence downstream of the stop codon. SV320 contains the wild-type NAB3 inserted in pRS415 (a LEU2 CEN plasmid) (40). Plasmids SV321-SV326 contain point mutants R331A, N361A, E397A, E397K, S399A, and S399K, respectively (see above for the primers). Strain DLY889 containing the endogenous NAB3 under control of the GAL1 promoter (3) was transformed with plasmids containing either wild-type Nab3 or Nab3 RRM point mutants with a LEU2 selectable marker (SV320, SV321, SV322, SV323, SV324, SV325, and SV326, respectively). The resulting strains were used for growth tests and Western blot analyses.

Growth Test Analyses—To test whether the mutated residues were essential for growth, the resulting transformants were grown in SC-LEU-HIS + 2% galactose at 30 °C to an A_{600} 1.0. The cultures were then serially diluted in 96-well plates by a factor of 10, and spotted onto SC-LEU-HIS medium containing 2% glucose to repress the expression of the endogenous NAB3 or control medium (SC-HIS + 2% galac-

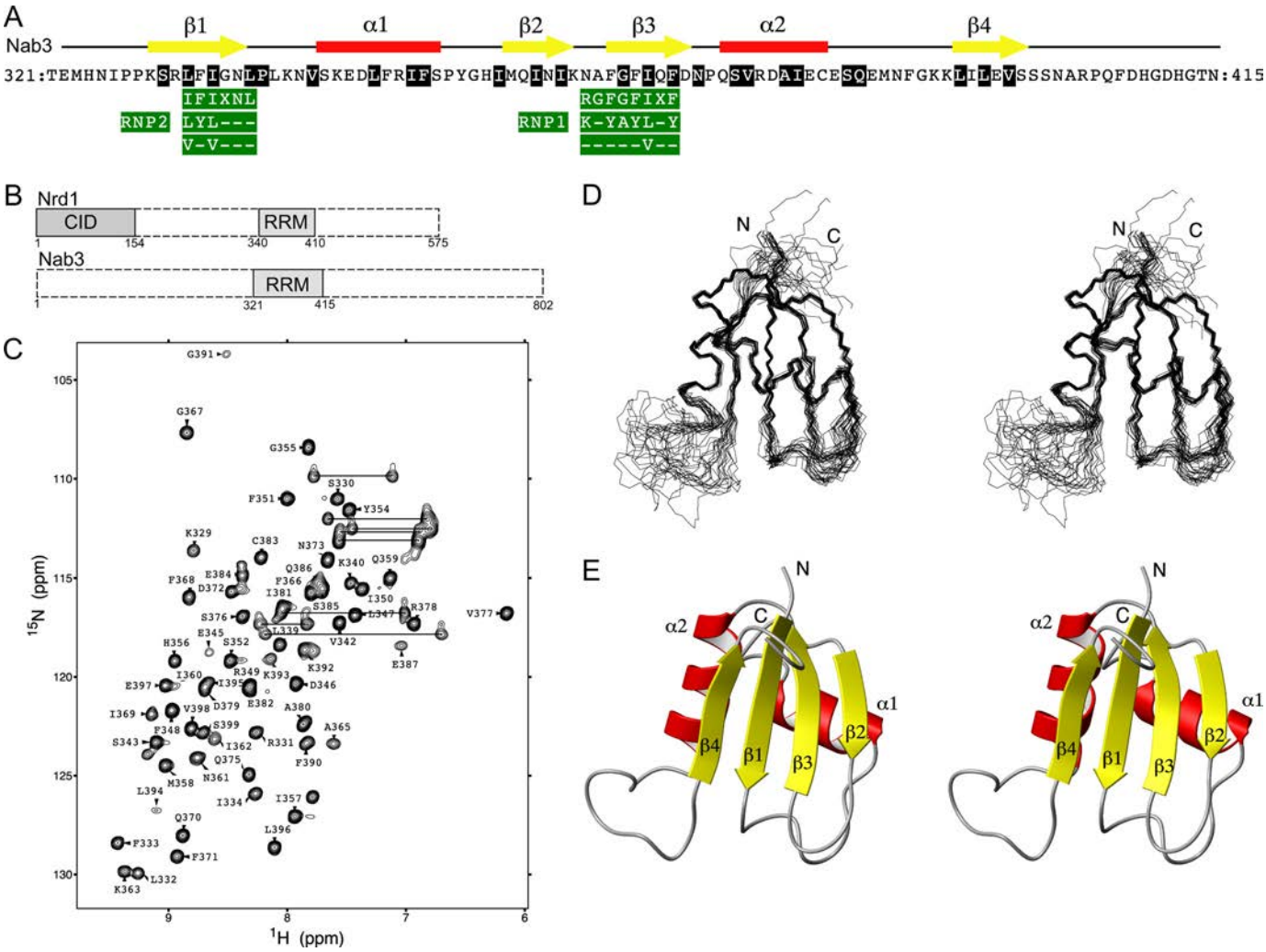


FIGURE 1. **Overview of the RRM of Nab3 sequence, topology, NMR spectra, solution structure, and domain structure in Nab3 and Nrd1.** *A*, amino acid sequence of the *S. cerevisiae* Nab3 RRM along with its secondary structure elements and general consensus of RNP1 and RNP2 motifs. *B*, a schematic drawing of the domain structure of Nab3 and Nrd1. *C*, two-dimensional ^{1}H-^{15}N HSQC spectrum of 2.5 mM uniformly ^{15}N,^{13}C-labeled Nab3 RRM in 50 mM sodium phosphate buffer (pH 8.0), 300 mM NaCl, and 10 mM β-mercaptoethanol (90% H_2O, 10% D_2O). The spectrum was acquired at 303 K on a Bruker Avance 600 MHz spectrometer. The assignments are labeled by the one-letter code of amino acids accompanied by a sequence number. The side chain resonances of asparagine and glutamine are connected by *horizontal lines*. *D*, stereo view of the 20 lowest energy structures of Nab3 RRM. The protein backbone is shown as a wire model. *E*, stereo view of the representative (the lowest energy) structure of Nab3 RRM shown as a *ribbon diagram*. The figure was generated with MOLMOL (38).

tose). These plates were incubated at 24, 30, and 37 °C for 3 days.

Western Blot Analysis—Protein extracts were prepared from cultures grown either on galactose containing medium and cultures shifted to glucose containing medium (as described above). Proteins were resolved on a 12% SDS-PAGE gel, transferred to nitrocellulose membrane by a semi-dry electroblotter (Bio-Rad), and probed for the presence of Nab3p with the mAb 2F12 (41), or antibodies directed against the HA epitope (sc805, Santa Cruz Biotechnology) present on the endogenous Nab3. For loading control we used the antibodies against Air2 protein (42).

RESULTS

Structure of Nab3 RRM—The RRM of *S. cerevisiae* Nab3 was examined by NMR spectroscopy (Fig. 1, *A* and *B*). The ^{1}H, ^{13}C, and ^{15}N chemical shift assignments were obtained as described previously (24). All NMR experiments were measured at a high salt concentration (300 mM NaCl, 50 mM sodium phosphate (pH 8.0), and 10 mM β-mercaptoethanol) to prevent protein precipitation. The solution structure determination of Nab3 RRM employed homonuclear and heteronuclear NMR techniques. The ^{15}N-^{1}H HSQC experiment shows a well dispersed spectrum (Fig. 1*C*), indicating a folded domain. There are a number of missing peaks in this spectrum that mainly correspond to the N- and C-terminal regions of the studied protein construct (outside of the RRM domain). These regions were included in the study as they often form additional structural elements (β-strand or α-helix) in RRMs or contribute to the RNA binding. In addition, several residues in the loops showed no NMR signals. These missing signals are likely a result from the relatively high pH used in the NMR study that was necessary to prevent the precipitation of Nab3 RRM. The three-dimensional structure of Nab3 RRM was determined by combined automated NOESY cross-peak assignment (30) and structure calculations with

TABLE 1
NMR and refinement statistics for Nab3 RRM and Nab3 RRM–UCUU complex

NMR distance and dihedral angle restraints	Nab3 RRM	Nab3 RRM–UCUU complex
Distance restraints		
Total NOEs	857	852
Intra-residue	201	228
Inter-residue		
Sequential ($\|i-j\| = 1$)	218	187
Medium range ($1 \langle \|i-j\| \rangle 5$)	164	116
Long range ($\|i-j$[bar ≥ 5)	274	310
Hydrogen bond restraints	25	25
Intermolecular		11
Dihedral angle restraints		
ϕ and ψ	76	88[a]
Structure statistics[b]		
Residual NOE violations (mean ± S.D.)		
Number > 0.20 Å	1.5 (± 0.83)	4 (± 2)
Maximum (Å)	0.24 (± 0.03)	0.46 (± 0.08)
Residual dihedral angle violations		
Number > 10.0°	0	0
Maximum (°)	0	0
Ramachandran plot statistics[b,c,d]		
Residues in most favored regions (%)	90.8	88.6
Residues in additionally allowed regions (%)	9.1	9.6
Residues in generously allowed regions (%)	0.1	1.7
Residues in disallowed regions (%)	0.0	0.1
Deviations from idealized geometry		
Bond length (Å)	0.0011 ± 0.0001	0.0010 ± 0.0001
Bond angles (Å)	1.46 ± 0.02	1.48 ± 0.02
Average root mean square deviation to mean structure (Å)[b]		
Protein		
Backbone atoms[b]	0.57 ± 0.12	0.49 ± 0.10
Heavy atoms[b]	1.46 ± 0.16	1.24 ± 0.12
RNA		
All RNA heavy atoms[e]		0.98 ± 0.21
Complex		
Protein[b] and RNA heavy atoms[e]		1.27 ± 0.12
WHAT IF[f] structure *Z*-scores[d,g]		
Packing quality	−2.3	−1.9
Ramachandran plot appearance	−3.5	−3.1

[a] Includes *C2′-endo* sugar pucker and *anti* conformation of the glycosidic bond was used for all nucleotides (56).
[b] Calculated for an ensemble of the 20 lowest energy structures.
[c] Based on PROCHECK analysis (32).
[d] Calculated for the structured part of the protein construct.
[e] Calculated for $U_1C_2U_3$.
[f] Based on WHAT IF analysis (37).
[g] *Z*-score (54, 55) is defined as the deviation from the average value for this indicator observed in a database of high-resolution crystal structures.

torsion angle dynamics implemented in the program CYANA 2.1 (43), followed by refinement in explicit solvent using AMBER 10 (33). An ensemble of the 20 lowest energy structures along with the best energy structure are shown in Fig. 1, *D* and *E*, respectively. These structures have an average backbone root mean square deviation of 0.57 ± 0.12 Å for the secondary structure elements. A full summary of structural statistics including the backbone ϕ-ψ angle distribution is given in Table 1.

The three-dimensional structure of Nab3 RRM adopts a compact fold with an $\beta1\alpha1\beta2\beta3\alpha2\beta4$ topology that is similar to the canonical fold of RRM family (21, 22). The fold is composed of two α-helices and a 3_{10} helix that are packed along a face of a four-stranded antiparallel β-sheet. A central hydrophobic core composed of the residues shown in Fig. 1*A* stabilizes the fold of the domain. Nab3 RRM contains a well conserved signature of the RRM family, RNP1 and RNP2 sequences (44–46). These two conserved amino acid sequences found between Leu^{332}–Leu^{337} and Asn^{364}–Phe^{371} are located on the $\beta3$- and $\beta1$-strands, respectively. Their sequence compositions correspond to the general RNP2 and RNP1 consensus (ILV)-(FY)-(ILV)-*X*-N-L and (RK)-G-(FY)-(GA)-(FY)-(ILV)-*X*-(FY), respectively, except for the first two amino acids of the RNP1 (Fig. 1*A*). Nab3 RRM has asparagine and alanine in these positions (Fig. 1*A*). The presence of aromatic residues in RNP1 and RNP2 sequences, which usually mediates the stacking interaction with RNA bases, along with a number of basic and polar residues on the β-sheet surface, indicate a potential role of Nab3 RRM in RNA binding.

Characterization of the Nab3-UCUU Interactions by NMR—To investigate the interaction and binding mode between Nab3 RRM and RNA, we carried out an NMR chemical shift perturbation study with a UCUU element, which has been shown to elicit transcription termination via the Nrd1 pathway. In the RNA titration experiment, we observed that the protein amide resonances moved upon RNA binding from their initial positions, corresponding to the free form, in a stepwise directional manner until they reach their final positions that correspond to the fully bound state, with stoichiometry of 1:1 (Fig. 2). Additional RNA aliquots resulting in excess RNA resulted in no further change of chemical shifts, confirming the 1:1 stoichiometry of the complex. These titration data suggest that protein amide resonances are in a fast exchange regime between

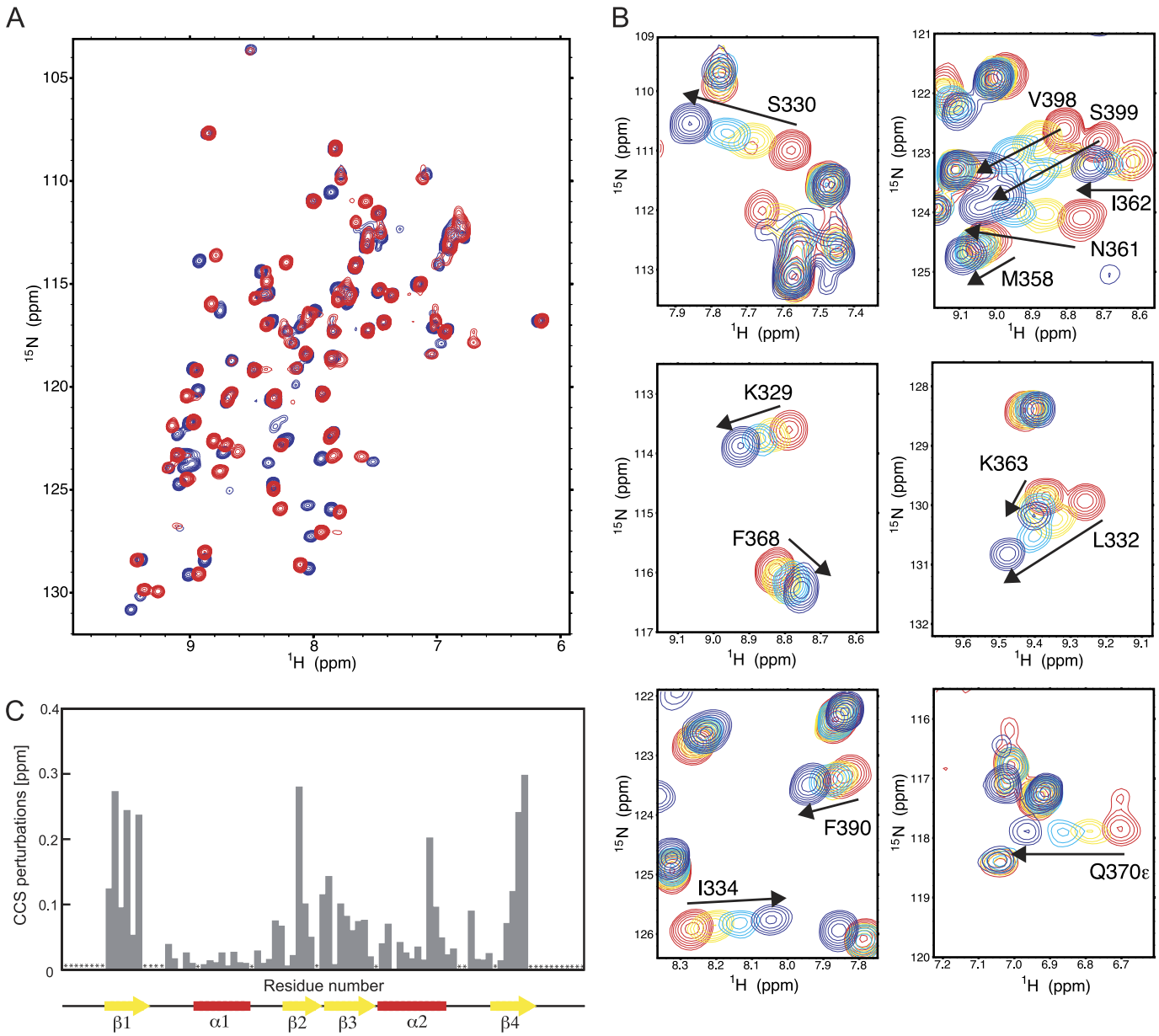


FIGURE 2. **NMR titration experiments of Nab3 RRM with UCUU RNA.** *A*, ^{1}H-^{15}N HSQC spectra of Nab3 RRM alone (in *red*) and in the presence of 1 eq of 5′-UCUU-3′ (in *blue*) at 303 K. *B*, close-up views of the ^{1}H-^{15}N HSQC spectra, showing selected chemical shift changes during the titration. *C*, quantification of chemical shift perturbations of Nab3 RRM upon binding to UCUU RNA. The combined chemical shift perturbations $([\omega_{HN}\Delta\delta_{HN}]^2 + [\omega_N\Delta\delta_N]^2)^{1/2}$, where $\omega_{HN} = 1$ and $\omega_N = 0.154$ are weight factors of the nucleus (52), are plotted *versus* the amino acid residue number. Large changes occur on the β-sheet surface. The assignments of residues indicated by *asterisks* could not be obtained for neither the free nor bound protein, or indicates proline residues.

their free and bound forms relative to NMR time scale. The binding of UCUU to the RRM of Nab3 induces chemical shift perturbation of the residues shown in Fig. 2.

These chemical shift changes indicate that the above mentioned residues are involved in binding to the RNA, or alternatively, could undergo a conformational change upon RNA binding. Mapping the perturbed residues on the sequence of Nab3 RRM delineates that the Nab3 RRM binds the RNA through its β-sheet surface and also through the $\beta2\beta3$ loop (Fig. 2*C*).

Structure of Nab3 RRM in Complex with UCUU—When solving the structure of Nab3 RRM bound to RNA, we extensively tested different lengths of RNA, buffer conditions, and temperatures with the aim to optimize the NMR spectral quality of the complexes. Longer RNA substrates, a UCUU core motif with flanking sequences, resulted in the significant broadening of NMR signals of the complexes. Interestingly, we obtained the NMR spectra of better quality (for both protein and RNA in complex) with a four-nucleotide UCUU despite the fact that this RNA has lower affinity to Nab3 RRM compared with the longer substrates (see below). Similar improvement of the NMR spectral quality by using of a minimal specific RNA sequence has also been observed for other protein-RNA complexes investigated by NMR (47–49). Therefore, we pursued the structure determination of the Nab3 RRM-UCUU complex.

The RRM of Nab3 in complex with RNA display the canonical RRM-fold with an $\beta1\alpha1\beta2\beta3\alpha2\beta4$ topology and is similar to that of the unbound form (Figs. 1 and 3). Akin to the free

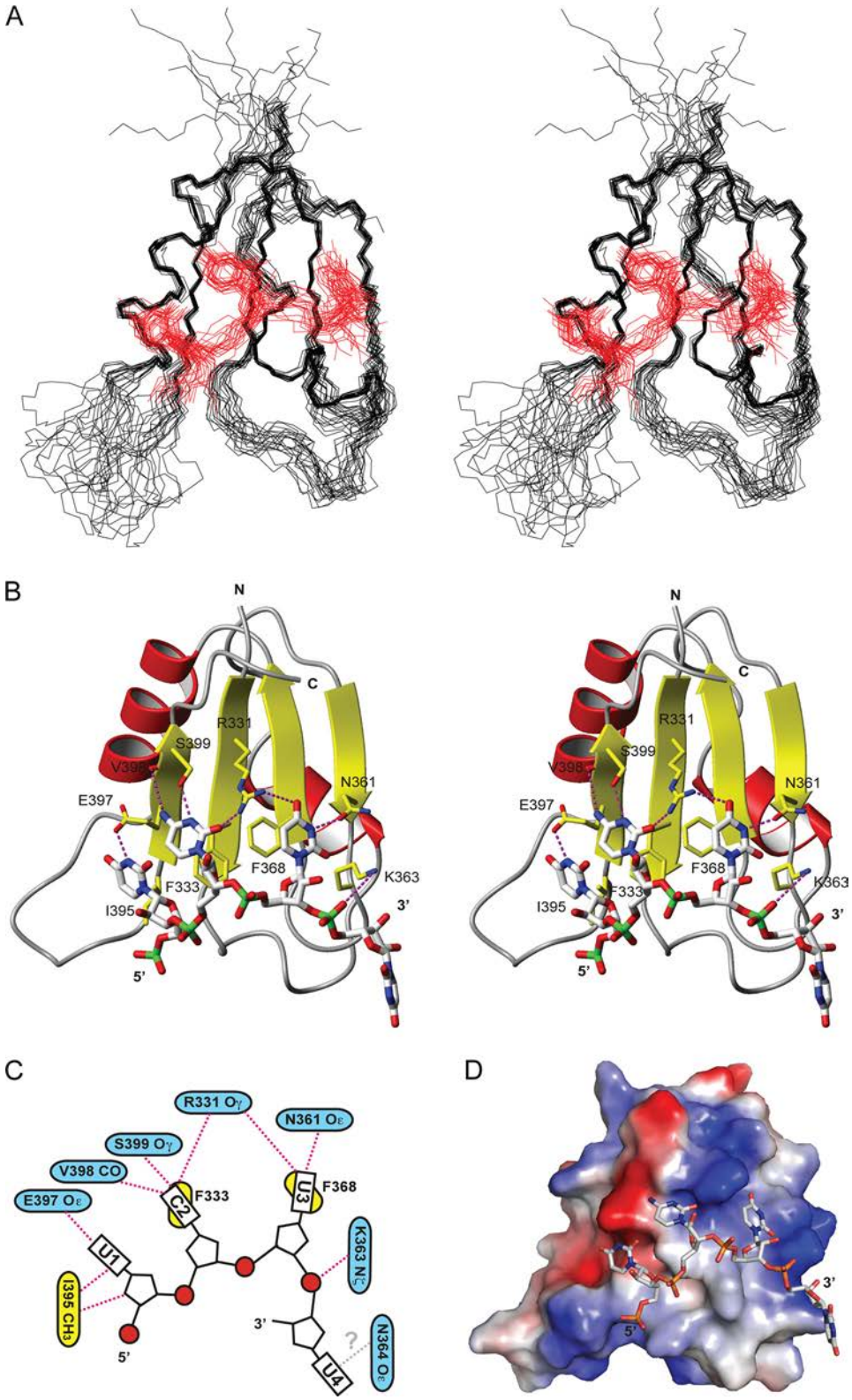


FIGURE 3. **Overview of the solution structure of the Nab3 RRM in complex with UCUU.** *A*, stereo view of the 20 lowest energy structures of the Nab3 RRM-UCUU complex. The protein backbone is shown as a wire model in *black*. The RNA heavy atoms are shown as a wire model in *red*. *B*, stereo view of the representative (the lowest energy) structure of the Nab3 RRM-UCUU complex. The RNA is represented as a *white stick model* and the protein is shown as a *ribbon* model with residues that contact the RNA shown in *yellow*. Putative hydrogen bonds are shown by *dotted magenta lines*. *C*, scheme showing contacts between Nab3 RRM and the UCUU RNA. Protein residues that form putative hydrogen bonds to the RNA are shown in *blue* and the one having hydrophobic interactions are in *yellow*. A hypothetic recognition of U4 is labeled by a *gray question mark*. *D*, solvent-accessible surface representation of Nab3 RRM colored by electrostatic potential (*blue*, positive; *red*, negative) and stick representation for the RNA of the representative structure of the complex. Figures were generated with MOLMOL (38).

form of Nab3 RRM, the N- and C-terminal regions as well as the long $\alpha 2\beta 4$ loop are structurally undefined due to the lack of experimental data. The UCUU RNA adopts a single-stranded conformation and the first three nucleotides are positioned over the whole β-sheet surface in a canonical arrangement in which the 5′ end is located on the first half of the β-sheet ($\beta 4\beta 1$) and the 3′ end on the second half ($\beta 3\beta 2$) (21, 22) (Fig. 2). The overall position of RNA on the β-sheet coincides with the perturbed residues from the titration experiment (Figs. 2 and 3). All bases have an *anti* conformation of the glycosidic bond and *C2′-endo* conformation of the sugar pucker.

The NMR spectra provided a limited number of intermolecular NOEs (11 unambiguous intermolecular NOEs) that loosely define the position of $U_1C_2U_3$ on the β-sheet surface of Nab3 RRM (Fig. 3*A*), but are sufficient to reveal the molecular basis of $U_1C_2U_3$ recognition by Nab3 RRM (Fig. 3, *B* and *C*; the protein-RNA hydrogen bonds described below are inferred from the final ensemble of structures and thus they should be considered as putative hydrogen bonds). We could not define the position of U_4 due to the lack of intermolecular NOEs. Based on NMR titration data, we speculate that U_4 could be recognized by the asparagine side chain or a main chain of the $\beta 2$–$\beta 3$ loop that are in proximity to the base of U_4, as displayed in the representative structure (Fig. 3*B*). In our NMR structure, C_2 and U_3 are involved in base stacking with the aromatic rings of Phe^{333} and Phe^{368}, respectively. The Watson-Crick edge of C_2 is recognized by the main chain carbonyl group of Val^{398} and the hydroxyl group of Ser^{399} that form hydrogen bonds with the amino and imino groups of C_2, respectively. One-half of the 20 structures in the final ensemble has Arg^{331} in a position in which it contacts the O_2 oxygen of C_2 (for the importance of Arg^{331} see below). In addition, it is likely also that Ser^{400} could be involved in the recognition of C_2 as the resonances of this residue broadened beyond detection upon RNA binding. Our structure also rules out the possibility that a purine could be accommodated in the C_2 position due to a steric restriction imposed by the Glu^{397} side chain.

The recognition of U_3 is mediated by the Arg^{331} and Asn^{361} side chains. The side chain NH_2 group of Arg^{331} contacts the O4 carbonyl functional group of the base and the side chain carbonyl group of Asn^{361} forms a hydrogen bond with the imino proton of U_3. Akin to C_2, the position of U_3 cannot be exchanged by a purine due to a steric hindrance of Arg^{331}. The sugar of the U_3 residue is further contacted by the aliphatic region of the Lys^{363} side chain.

In contrast to C_2 and U_3, the recognition of U_1 is less evident from the structure. A single hydrogen bond is formed between the imino proton of U_1 and the Oϵ of Glu^{397}. There are also hydrophobic contacts between the sugar and the base of U_1 and the side chain of Ile^{395}. However, these contacts do not explain fully the sequence specificity of a uridine nucleotide. A cytidine nucleotide in this position could also form a similar interaction with the glutamate. Altogether, our NMR structure indicates that Nab3 RRM recognizes the YCU sequence (where Y stands for pyrimidine).

Nab3 RRM Binds the Nab3 Termination Element with Low Affinity—The FA measurements were carried out to further characterize the binding of Nab3 RRM to various RNA substrates. In FA measurements, formation of the protein-RNA complex is monitored directly from an increase of the FA value that occurs when the protein binds fluorescently labeled RNA. Binding curves were recorded in the course of titration experiments, where protein aliquots were added to 10 nM fluorescently labeled RNAs (Fig. 4).

First, we assayed the binding affinity of Nab3 RRM to UCUU that has been reported as the minimal Nab3 terminator element. A tetranucleotide GUAA was used as a nonspecific control substrate. The comparison of the anisotropy data for specific and nonspecific four-nucleotide substrates along with determined equilibrium dissociation constants (K_d) are shown in Fig. 4*A*. Corresponding logarithmic values of K_a ($K_a = 1/K_d$) are shown in graph in Fig. 4*E*. Nab3 RRM binds the specific recognition sequence UCUU with more than 6-fold higher affinity compared with the nonspecific substrate GUAA.

As terminator elements often occur in multiple repeats, we tested a longer substrate with three UCUU repeats. The substrate sequence was derived from the snR47 that is terminated by the Nrd1 pathway (Fig. 4*F*). Nab3 RRM binds the three UCUU-containing substrate with a K_d of 48 ± 2 μM, 1 order of magnitude stronger than we observed for a single UCUU motif (Fig. 4*B*). Furthermore, we assayed snR13, which is another naturally occurring Nrd1-dependent terminator that contains two UCUU and one CCU motifs (Fig. 4*F*). As expected, Nab3 RRM binds this substrate with a K_d of 46 ± 1 μM, a similar binding affinity to that of snR47 (Fig. 4*B*). As a control, we used SL RNA, which has a similar size and its sequence lacks UCUU or even CU recognition motifs (Fig. 4*F*). For this nonspecific substrate, Nab3 RRM binding is reduced more than 30-fold compared with the snR13 and snR47 terminators (Fig. 4, *B* and *E*).

Nrd1 RRM Binds the Nrd1 Termination Element with Low Affinity—As Nab3 RRM binds its termination motif with a low affinity, we therefore decided to investigate the RNA-binding properties of Nrd1 RRM, which is the second RNA-binding domain occurring in the Nrd1-dependent termination complex (Fig. 1*B*). Akin to Nab3 RRM, we assayed the binding affinity of Nrd1 RRM to RNA using FA. We found that Nrd1 RRM-(340–410) binds to the minimal termination RNA sequence GUAA (GUAR is known as the Nrd1-termination element (where R stands for purine) (18) with a low affinity in the mid-micromolar range (Fig. 4*C*; K_d of 66 ± 1 μM). Next, we assayed a longer RNA substrate, snR13, which contains two GUAR motifs (Fig. 4*F*). The titration curve for snR13 yielded a K_d of 11 ± 1 μM (Fig. 4*C*).

A Complex of Nrd1-Nab3 Binds RNA with High Affinity—As the Nrd1 and Nab3 proteins form a heterodimer *in vivo* and *in vitro* (19), we assayed RNA binding of the co-expressed Nrd1-Nab3 heterodimer (19) using FA. As a substrate, we used RNA derived from the snR13 terminator that contains two copies of each recognition element (Fig. 4*F*). This RNA substrate binds the Nrd1-Nab3 heterodimer with a K_d of 2.4 ± 0.3 nM (Fig. 4*D*), about 4–5 orders of magnitude stron-

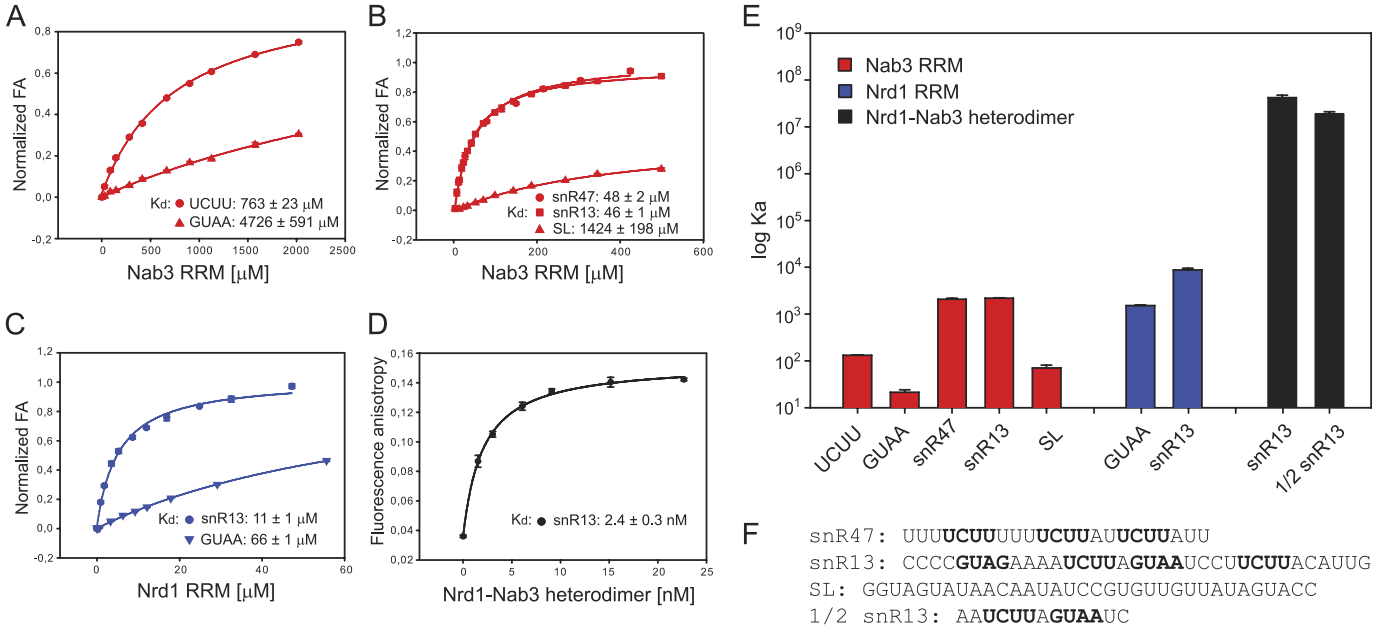


FIGURE 4. **Equilibrium binding of Nab3 RRM, Nrd1 RRM, and Nrd1-Nab3 heterodimer with fluorescently labeled RNA monitored by fluorescence anisotropy.** *A*, Nab3 RRM was titrated with UCUU and GUAA (each 10 nM), and their binding isotherms are shown as *red circles* and *triangles*, respectively. *B*, Nab3 RRM was titrated with snR47, snR13, and SL substrates (each 10 nM), and their binding isotherms are shown as *red circles*, *squares*, and *triangles*, respectively. *C*, Nrd1 RRM was titrated with GUAA and snR13 (each 10 nM) and their binding isotherms are shown as *blue inverted triangles* and *circles*, respectively. *D*, Nrd1-Nab3 heterodimer was titrated with snR13 (100 pM). *E*, summary of the association constants (K_a) for the RRM of Nrd1 (in *blue*) and Nab3 (in *red*) in their free forms as well as for the Nrd1-Nab3 heterodimer (in *black*). Logarithmic scale of K_a is shown to cover a wide range of affinities. *F*, RNA sequences used in the affinity measurements. The buffers contained the same ion strength and pH values for all proteins. Equilibrium dissociation constant (K_d) was calculated from the best fit to the data using a single-site binding isotherm. *Error* is denoted as S.E. The data were normalized for visualization purposes (*A–C*).

ger compared with the individual RRMs of Nrd1 and Nab3 (Fig. 4*E*). Another snR13-derived substrate, ½ snR13, containing only one copy of each recognition element (Fig. 4*F*), binds the Nrd1-Nab3 heterodimer with only a slightly lower affinity (K_d of 5.3 ± 0.6 nM). However, when one of the recognition elements (either Nrd1 or Nab3) is removed from the ½ snR13 substrate, the affinity is reduced such that the dissociation constant cannot be determined (saturation of the binding curve cannot not be reached due to the low solubility limit of the Nrd1-Nab3 heterodimer; data not shown).

Impacts of Nab3 RRM Point Mutations on RNA Binding Affinity—It has been shown previously that mutations of C_2 and U_3 in the UCUU motif reduce the binding affinity to Nab3 (18) or to the Nrd1-Nab3 heterodimer (19). We performed the converse experiments in which we assessed Nab3 RRM mutants for their ability to bind snR47 RNA in a quantitative solution binding assay by fluorescence anisotropy titration experiments. We mutated the non-canonical amino acid residues on the β-sheet surface (R331A, N361A, S399A, and E397A) that specifically recognize the bases of the $U_1C_2U_3$ sequence (Fig. 3, *B* and *C*). These residues surround the conserved residues of RNP1 and RNP2 consensuses. We found that mutants R331A and S399A showed a 3–4-fold decrease in binding affinity of that demonstrated by the wild-type protein (Fig. 5*A*). In contrast, mutants N361A and E397A showed binding affinity similar to the wild-type Nab3 RRM (Fig. 5*A*).

Functional Significance of the Nab3 RRM Residues That Contact RNA—To address the importance of the specific contacts identified in the Nab3 RRM-UCUU complex for Nab3 function *in vivo*, single amino acid mutants (R331A, S399A, S399K, E397A, and E397K) were prepared in a yeast expression vector and introduced into a yeast strain in which the endogenous NAB3 promoter was replaced with the GAL1 promoter (3). To test whether the mutated residues were essential for growth, the resulting transformants were spotted onto glucose containing plates. The shift to glucose represses the expression of the GAL1-driven endogenous NAB3, which completely impairs cell viability (Fig. 5*B*). This lethality was rescued by wild-type Nab3 (Fig. 5*B*). Mutating the three residues involved in the recognition of C_2 and U_3 (Fig. 3), R331A, N361A and S399A, or S399K, caused lethality (Fig. 5*B*), providing further support for the functional significance of these contacts. In contrary, the Glu^{397} mutant displayed only slow growth at 24 and 37 °C.

To further confirm that mutant lethality was not due to impaired protein expression, we performed Western blot analysis with antibodies directed against Nab3 and antibodies against the HA tag present on the endogenous Nab3 only. We detected similar levels of Nab3 in all proteins expressed episomally, whereas no HA-tagged endogenous Nab3 was detected in yeast grown on glucose medium (Fig. 5*C*).

DISCUSSION

RNA Recognition by Nab3 RRM and Its Comparison with Other RRMs—We have solved the structure of Nab3 RRM in free form and in complex with UCUU. In both forms, the fold of Nab3 RRM is very similar, with minor adjustments upon RNA binding, and resembles canonical RRM (21–23). However, it has shorter α-helices than observed in the structures

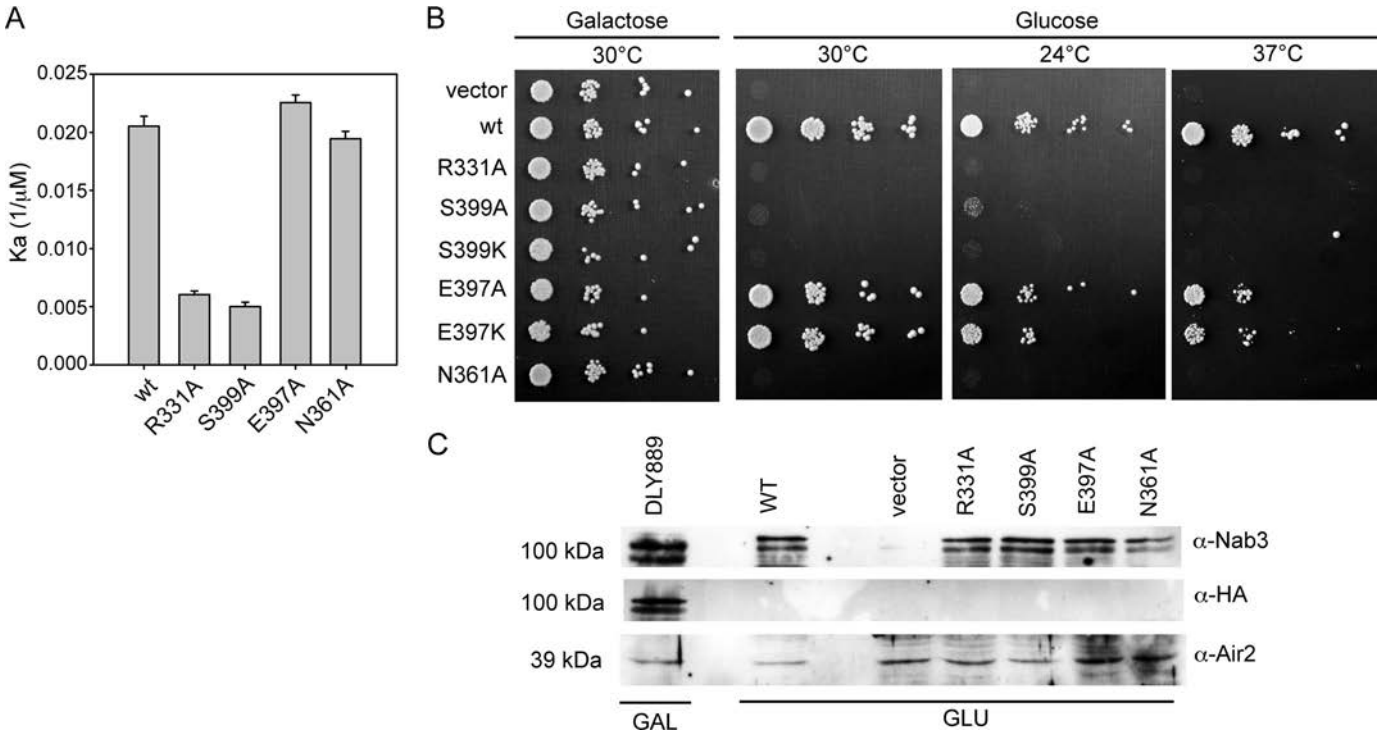


FIGURE 5. **The important residues of Nab3 RRM that are required for RNA binding and cell viability.** *A*, equilibrium binding of Nab3 RRM mutants with fluorescently labeled RNA monitored by fluorescence anisotropy. The Nab3 RRM R331A, N361A, S399A, and E397A mutants along with the wild-type of Nab3 RRM were titrated with fluorescently labeled snR47 substrate. Equilibrium association constants (K_a) are shown for individual mutants with S.E. *B*, residues Arg^{331}, Ser^{399}, and Asn^{361} are required for yeast viability. The indicated Nab3 RRM mutants were expressed episomally from pRS415 plasmids in the yeast strain with the endogenous NAB3 driven by the GAL1 promoter. Mutant strains were spotted on plates containing 2% glucose and a control galactose plate and incubated for 3 days at the indicated temperatures. Growth on glucose-containing plates leads to the repression of GAL1-driven wild-type Nab3, and thus shows the functionality of the different Nab3 mutants. *Vector* is a control where the GAL1::NAB3 strain contains an empty pRS415 plasmid, *wt* is the wild-type NAB3. *C*, expression of Nab3 proteins from pRS415 in glucose-containing medium. Western blot analysis was perfomed with protein extracts from the original GAL1::NAB3 strain (DLY889) grown in galactose-containing medium and extracts from DLY889 transformed with plasmids carrying wild-type and mutant NAB3 grown for 20 h in glucose-containing medium. *Air2* was used as a loading control.

of canonical RRMs, and an extra 3_{10} helix between the α1 helix and β2 strand (22).

Nab3 RRM binds UCUU in a canonical manner as observed in other RRM-RNA complexes (21–23). Specifically, Nab3 RRM contains a well conserved signature of the RRM family, RNP1 and RNP2 sequences (44–46). Two conserved phenylalanine residues are used in RNP1 (Phe^{333}) and RNP2 (Phe^{368}) to mediate the stacking interaction with RNA bases C_2 and U_3, respectively. These aromatic residues are surrounded by basic and polar amino acid residues that mediate the sequence-specific recognition of C_2U_3 and in part of U_1. The specifically recognized nucleotides are accommodated on the β-sheet only and neither loops nor N-/C-terminal regions to the RRM are involved in the recognition process.

Interestingly, the binding preference for CU has also been reported for the polypyrimidine tract-binding protein (PTB) (49). Unlike Nab3 RRM, PTB RRMs display additional topology elements to the canonical RRM involved in RNA recognition and significantly differ from RNP1 and RNP2 consensus (Fig. 6*A*). In particular, PTB RRMs lack aromatic residues in RNP1 and RNP2 that usually make extensive stacking interactions with the RNA bases and sugars as in the case of Nab3 RRM (Fig. 6). In addition, the structures of Nab3 and PTB RRMs revealed that these domains are sequentially unrelated on the entire RNA interaction surface, except for the last serine residue of β-strand 4 that is present in all PTB and Nab3 RRMs (Fig. 6*A*). The serine residue is involved in recognition of a cytosine in the structures of PTB (49) and Nab3 bound to RNA (Fig. 6). Furthermore, in the structures of PTB RRM1-RNA and Nab3 RRM-RNA, recognition of the 5′-end uridine is mediated in a similar way in which its imino proton is contacted by glutamine and glutamate, respectively (Fig. 6). In both structures, a cytidine could be tolerated instead of a uridine in this position. In contrast, recognition of the 3′-end uridine is mediated differently in these two structures. Whereas Nab3 RRM utilizes the side chains of arginine (in β1) and asparagine (in β2) to recognize the uridine, PTB RRM1 uses the main chain of leucine and lysine located in the C-terminal extension to the RRM to facilitate the uridine binding.

The importance of the serine residue, which is discussed above, is demonstrated by our affinity measurements with the Nab3 RRM S399A mutant that decreases the binding affinity to its UCUU-containing substrate (Fig. 5*A*). Correspondingly, *in vivo* analyses of Nab3 S399A or S399K mutants show that the serine residue is essential for yeast viability (Fig. 5*B*). Similarly, the decreased affinity of the Nab3 RRM R331A mutant to the RNA is manifested by the lethal effect in yeast (Fig. 5). Interestingly, Asn^{361}, which specifically recognizes U_3 along with Arg^{331}, shows lethality *in vivo* for the N361A mutant, whereas the affinity of this mutant is only slightly reduced compared with the wild-type (Fig. 5). Furthermore, the

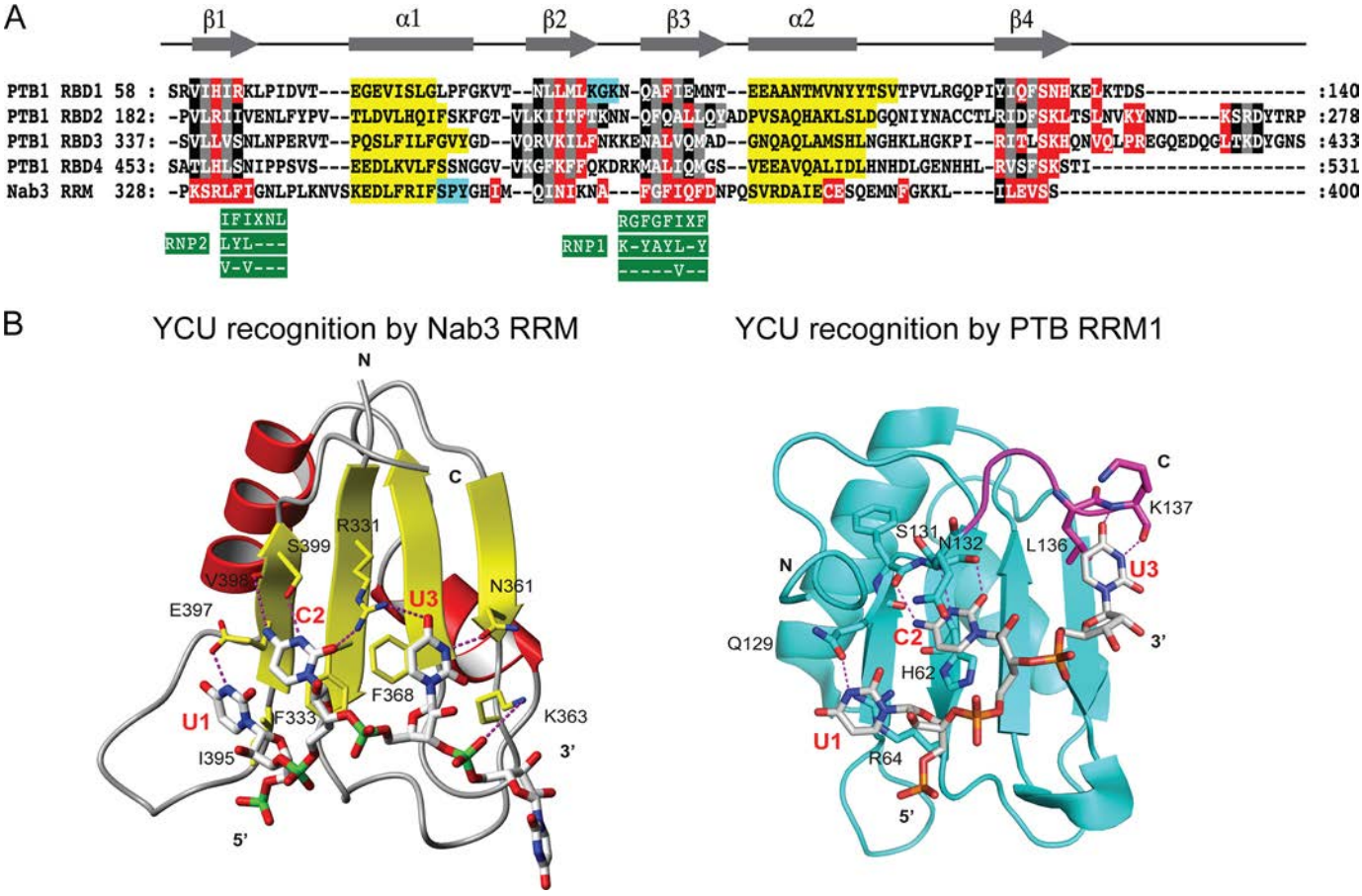


FIGURE 6. **Recognition of YCU by Nab3 and PTB RRMs.** *A*, sequence alignment of PTB RRM1, -2, -3, and -4 and Nab3 RRM whose structures have been solved. The alignment was performed using ClustalW (53) and manually optimized using the three-dimensional structural information (49). For the RRMs of PTB, amino acids interacting with the RNA are shown in *red boxes*, residues in *gray* and *black boxes* are located in the β-sheet and residues in *yellow* and *cyan boxes* are in the α- or 3_{10} helices, respectively (49). Residues in *gray boxes* form the hydrophobic core of the domains. For the RRM of Nab3, residues in *red boxes* are significantly perturbed upon RNA binding. *B*, comparison of Nab3 RRM (*left*: in *yellow* and *red schematics*) and PTB RRM1 (*right*: in *cyan schematics*) binding to UCU nucleotides (represented as a *stick model*). The protein residues that mediate the specific recognition are highlighted as a stick model. The C-terminal region of PTB RRM1 that mediates the recognition of U_3 is shown in *magenta*.

N361A mutant yeast strain shows accumulation of cryptic unstable transcripts when shifted to glucose,[7] supporting the functional significance of this contact in the termination via the Nrd1 pathway. The Glu[397] mutant does not have an impact on RNA binding *in vitro* and causes only slow growth at higher temperatures, indicating a minor role of this contact for the function of Nab3. Altogether, these functional data corroborate with our structural findings that Nab3 RRM specifically recognizes the YCU sequence.

Association between Nab3 and Nrd1 Increases the Affinity to the Termination Sequences—Our FA experiments showed that Nab3 RRM binds UCUU with the high micromolar range of equilibrium dissociation constant (K_d), yet the Nab3 RRM binds nonspecific four-nucleotide RNA with a K_d in the low millimolar range (Fig. 4). The observed affinity for the specific RNA substrate is weaker than the affinity usually observed for single canonical RRMs (21, 47). The affinity of Nab3 RRM binding to UCUU is lower by 2-fold compared with the affinity of PTB RRM1 to CUCU that also recognizes the YCU motif (Fig. 6) (47). Interestingly, the apparent K_d of Nab3 RRM binding to longer RNAs with multiple UCUU motifs (snR47 and snR13) are ~50 μM, more than 1 order of magnitude stronger than we observed for a single UCUU motif (Fig. 4). The increased binding affinity likely originates from the presence of multiple binding sites and due to the presence of flanking sequences to the UCUU motif in snR47 and snR13 RNA substrates. The effect of flanking sequences is likely nonspecific, mediated by electrostatic interactions of additional phosphate groups. A similar increase in the binding affinity, when multiple binding motifs are present, has also been observed for the RRMs of PTB (47). The formation of multiple complexes between the snR47-derived RNA substrate and an RRM-containing Nab3 construct (Nab3-(277–565)) has also been observed previously using electrophoretic mobility shift assay (18). It has been also demonstrated that mutations in the 2nd, 3rd, and 4th position of the UCUU motif decrease the binding affinity to the Nab3-(277–565) (18). Furthermore, our data showed that Nrd1, the second RNA-binding subunit of the Nrd1 complex containing a single RRM, binds its termination element, GUAA, with a low affinity (K_d of 66 ± 1 μM; Fig. 4).

Therefore, it is very likely that Nrd1 and Nab3 bind RNA in a cooperative manner to achieve the nanomolar range of af-

[7] F. Hobor, D. Hrossova, S. Vanacova, and R. Stefl, unpublished data.

finity to snR13 RNA, as previously estimated using electrophoretic mobility shift assay (EMSA) (19). Indeed, our FA measurements showed that the Nrd1-Nab3 heterodimer binds snR13 RNA with a K_d in the low nanomolar range, 4–5 orders of magnitude stronger compared with the individual RRMs of Nrd1 and Nab3 (Fig. 4*E*). The RRMs of Nrd1 and Nab3 do not bind each other (data not shown) and the regions that mediate the formation of the heterodimer are located near the N-terminal of each RRM (13).[7] It is difficult to assay the binding affinity of full-length Nrd1 and Nab3 individually due to their instability (19); the Nrd1 alone rapidly aggregates.[7] However, other regions outside of the RRMs of Nrd1 and Nab3 are not expected to contribute significantly to RNA binding as they do not contain an identifiable RNA-binding domain. In a similar way, the cooperative RNA binding of two RRM-containing proteins $U2AF^{65}$ and $U2AF^{35}$ is utilized to enhance the affinity and selectivity in the process of defining the site of spliceosomal assembly (50, 51).

Implication for Poly(A) Independent Transcription Termination—We have shown that the RRM of Nab3 binds specifically the YCU sequence (where Y stands for pyrimidine). This is in good agreement with previous functional data that led to the proposal of the UCUU sequence as the Nab3 termination element (6, 18). However, the first position of the UCUU motif is not fully conserved in some snRNA downstream sequences (18); it can be either U or C that perfectly matches our structural findings. Furthermore, the last position of UCUU is also not fully conserved but we cannot explain the recognition of another nucleotide due to the lack of experimental data for the recognition of this nucleotide in our structure. Considering the specificity only for the CU dinucleotide and relatively weak affinity of Nab3 RRM to YCU-like sequences, it is evident that Nab3 alone cannot recruit the Nrd1 complex to the correct termination sites. Indeed, we show the first quantitative evidence that the association of Nrd1 and Nab3 (each protein contains a single RRM) facilitates high affinity binding and sequence selectivity. It remains to be seen whether additional sequence elements are recognized, in addition to YCU and GUAR, upon the association of Nrd1-Nab3.

Acknowledgments—We thank to Dr. D. Libri for the GAL1::Nab3 strain and Dr. J. L. Corden for the generous gift of the Nab3 2F12 antibodies and the pST39 plasmid, containing Nab3-(191–565) and Nrd1-(1–548). We are grateful to Massimo Lucci for assistance with NMR measurements at CERM Florence, Italy. The NOESY spectra including the intermolecular filter experiments were obtained at the CERM NMR facility supported by European Union-NMR program Grant RII3–026145.

REFERENCES

1. Richard, P., and Manley, J. L. (2009) *Genes Dev.* **23,** 1247–1269
2. Bentley, D. (2002) *Curr. Opin. Cell Biol.* **14,** 336–342
3. Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. (2006) *Mol. Cell* **23,** 853–864
4. Arigo, J. T., Eyler, D. E., Carroll, K. L., and Corden, J. L. (2006) *Mol. Cell* **23,** 841–851
5. Steinmetz, E. J., and Brow, D. A. (1996) *Mol. Cell. Biol.* **16,** 6993–7003
6. Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001) *Nature* **413,** 327–331
7. Vasiljeva, L., and Buratowski, S. (2006) *Mol. Cell* **21,** 239–248
8. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997) *Cell* **91,** 457–466
9. Vanácová, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005) *PLoS Biol.* **3,** e189
10. LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005) *Cell* **121,** 713–724
11. Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., Libri, D., and Jacquier, A. (2005) *Cell* **121,** 725–737
12. Proudfoot, N. J., Furger, A., and Dye, M. J. (2002) *Cell* **108,** 501–512
13. Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008) *Nat. Struct. Mol. Biol.* **15,** 795–804
14. Gudipati, R. K., Villa, T., Boulay, J., and Libri, D. (2008) *Nat. Struct. Mol. Biol.* **15,** 786–794
15. Steinmetz, E. J., and Brow, D. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95,** 6699–6704
16. Conrad, N. K., Wilson, S. M., Steinmetz, E. J., Patturajan, M., Brow, D. A., Swanson, M. S., and Corden, J. L. (2000) *Genetics* **154,** 557–571
17. Morlando, M., Greco, P., Dichtl, B., Fatica, A., Keller, W., and Bozzoni, I. (2002) *Mol. Cell Biol.* **22,** 1379–1389
18. Carroll, K. L., Pradhan, D. A., Granek, J. A., Clarke, N. D., and Corden, J. L. (2004) *Mol. Cell Biol.* **24,** 6241–6252
19. Carroll, K. L., Ghirlando, R., Ames, J. M., and Corden, J. L. (2007) *RNA* **13,** 361–373
20. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigó, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D.,

Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001) *Science* **291,** 1304–1351
21. Maris, C., Dominguez, C., and Allain, F. H. (2005) *FEBS J.* **272,** 2118–2131
22. Stefl, R., Skrisovska, L., and Allain, F. H. (2005) *EMBO Rep.* **6,** 33–38
23. Cléry, A., Blatter, M., and Allain, F. H. (2008) *Curr. Opin. Struct. Biol.* **18,** 290–298
24. Pergoli, R., Kubicek, K., Hobor, F., Pasulka, J., and Stefl, R. (2010) *Biomol. NMR Assign.* **4,** 119–121
25. Bax, A., and Grzesiek, S. (1993) *Acc. Chem. Res.* **26,** 131–138
26. Sattler, M., Schleucher, J., and Griesinger, C. (1999) *Prog. Nucl. Magn. Reson. Spectrosc.* **34,** 93–158
27. Varani, G., Aboulela, F., and Allain, F. H. (1996) *Prog. Nucl. Magn. Reson. Spectrosc.* **29,** 51–127
28. Peterson, R. D., Theimer, C. A., Wu, H., and Feigon, J. (2004) *J. Biomol. NMR* **28,** 59–67
29. Zwahlen, C., Legault, P., Vincent, S. J., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) *J. Am. Chem. Soc.* **119,** 6711–6721
30. Güntert, P. (2004) *Methods Mol. Biol.* **278,** 353–378
31. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) *J. Mol. Biol.* **319,** 209–227
32. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) *J. Biomol. NMR* **44,** 213–223
33. Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., and Kollman, P. A. (2008) *Amber 10*, University of California, San Francisco, CA
34. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc.* **117,** 5179–5197
35. Padrta, P., Stefl, R., Králík, L., Zídek, L., and Sklenár, V. (2002) *J. Biomol. NMR* **24,** 1–14
36. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
37. Vriend, G. (1990) *J. Mol. Graph.* **8,** 52–56, 29
38. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph.* **14,** 51–55, 29–32
39. Heyduk, T., and Lee, J. C. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87,** 1744–1748
40. Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122,** 19–27
41. Wilson, S. M., Datar, K. V., Paddy, M. R., Swedlow, J. R., and Swanson, M. S. (1994) *J. Cell Biol.* **127,** 1173–1184
42. San Paolo, S., Vanacova, S., Schenk, L., Scherrer, T., Blank, D., Keller, W., and Gerber, A. P. (2009) *PLoS Genet.* **5,** e1000555
43. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298
44. Swanson, M. S., Nakagawa, T. Y., LeVan, K., and Dreyfuss, G. (1987) *Mol. Cell. Biol.* **7,** 1731–1739
45. Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K., and Dreyfuss, G. (1986) *Mol. Cell. Biol.* **6,** 2932–2943
46. Dreyfuss, G., Swanson, M. S., and Piñol-Roma, S. (1988) *Trends Biochem. Sci.* **13,** 86–91
47. Auweter, S. D., Oberstrass, F. C., and Allain, F. H. (2007) *J. Mol. Biol.* **367,** 174–186
48. Hargous, Y., Hautbergue, G. M., Tintaru, A. M., Skrisovska, L., Golovanov, A. P., Stevenin, J., Lian, L. Y., Wilson, S. A., and Allain, F. H. (2006) *EMBO J.* **25,** 5126–5137
49. Oberstrass, F. C., Auweter, S. D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D. L., and Allain, F. H. (2005) *Science* **309,** 2054–2057
50. Kielkopf, C. L., Rodionova, N. A., Green, M. R., and Burley, S. K. (2001) *Cell* **106,** 595–605
51. Singh, R., Valcárcel, J., and Green, M. R. (1995) *Science* **268,** 1173–1176
52. Mulder, F. A., Schipper, D., Bott, R., and Boelens, R. (1999) *J. Mol. Biol.* **292,** 111–123
53. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) *Nucleic Acids Res.* **31,** 3497–3500
54. Spronk, C. A. E. M., Nabuurs, S. B., Krieger, E., Vriend, G., and Vuister, G. W. (2004) *Prog. Nucleic Magn. Res. Spectro.* **45,** 315–337
55. Hooft, R. W., Sander, C., and Vriend, G. (1997) *Comput. Appl. Biosci.* **13,** 425–430
56. Trantírek, L., Stefl, R., Masse, J. E., Feigon, J., and Sklenár, V. (2002) *J. Biomol. NMR* **23,** 1–12
57. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System,* DeLano Scientific, Palo Alto, CA

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 38, pp. 32206–32215, September 14, 2012

Synergism of the Two Myb Domains of Tay1 Protein Results in High Affinity Binding to Telomeres*[S]

Received for publication, May 28, 2012, and in revised form, July 3, 2012 Published, JBC Papers in Press, July 18, 2012, DOI 10.1074/jbc.M112.385591

Katarina Visacka[‡], Ctirad Hofr[§], Smaranda Willcox[¶], Ivona Necasova[§], Jana Pavlouskova[§], Regina Sepsiova[‡], Michaela Wimmerova[§], Lucia Simonicova[‡], Jozef Nosek[‡], Jiri Fajkus[§], Jack D. Griffith[¶], and Lubomir Tomaska[‡1]

From the [‡]Departments of Genetics and Biochemistry, Comenius University, Faculty of Natural Sciences, Mlynska dolina, 842 15 Bratislava, Slovakia, [§]Faculty of Science and Central European Institute of Technology, Masaryk University, Kamenice 5, Brno, CZ-62500, Czech Republic, and [¶]Lineberger Comprehensive Cancer Center and Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina 27599

Background: In contrast to mammalian TRF1 and TRF2, yeast telomeric protein *Yl*Tay1 possesses two Myb domains.
Results: Kinetic and thermodynamic analyses revealed binding properties of individual Myb domains of *Yl*Tay1p.
Conclusion: The combined presence of the two Myb domains synergistically increases the affinity of *Yl*Tay1p to telomeric DNA.
Significance: The study demonstrates evolutionary tinkering with telomere-associated proteins.

Double-stranded regions of the telomeres are recognized by proteins containing Myb-like domains conferring specificity toward telomeric repeats. Although biochemical and structural studies revealed basic molecular principles involved in DNA binding, relatively little is known about evolutionary pathways leading to various types of Myb domain-containing proteins in divergent species of eukaryotes. Recently we identified a novel type of telomere-binding protein *Yl*Tay1p from the yeast *Yarrowia lipolytica* containing two Myb domains (Myb1, Myb2) very similar to the Myb domain of mammalian TRF1 and TRF2. In this study we prepared mutant versions of *Yl*Tay1p lacking Myb1, Myb2, or both Myb domains and found that *Yl*Tay1p carrying either Myb domain exhibits preferential affinity to both *Y. lipolytica* $(GGGTTAGTCA)_n$ and human $(TTAGGG)_n$ telomeric sequences. Quantitative measurements of the protein binding to telomeric DNA revealed that the presence of both Myb domains is required for a high affinity of *Yl*Tay1p to either telomeric repeat. Additionally, we performed detailed thermodynamic analysis of the *Yl*Tay1p interaction with its cognate telomeric DNA, which is to our knowledge the first energetic description of a full-length telomeric-protein binding to DNA. Interestingly, when compared with human TRF1 and TRF2 proteins, *Yl*Tay1p exhibited higher affinity not only for *Y. lipolytica* telomeres but also for human telomeric sequences. The duplication of the Myb domain region in *Yl*Tay1p thus produces a synergistic effect on its affinity toward the cognate telomeric sequence, alleviating the need for homodimerization observed in TRF-like proteins possessing a single Myb domain.

The ends of linear chromosomes are capped by nucleoprotein complexes (telomeres) providing a solution to the end-replication problem and shielding the chromosomal termini from recognition by exonucleases and DNA repair machinery (1, 2). The major players in mediating telomeric functions are proteins directly associated with chromosomal ends. These include (i) telomerase, a unique reverse transcriptase carrying its own RNA template representing the major molecular tool for maintaining telomere length (3), (ii) single-stranded (ss) DNA-binding proteins (*e.g.* human and fission yeast Pot1 or budding yeast Cdc13) tightly attached to 3′ single-stranded telomeric overhang through 1–3 OB-folds (4), and (iii) Myb/homeodomain-containing proteins (*e.g.* mammalian TRF1, TRF2, fission yeast Taz1, and budding yeast Rap1) binding to the double-stranded (ds) repetitive part of telomeres (1, 5). Other proteins (*e.g.* Rap1, Tin2, and Tpp1 in Metazoa, Rap1, Poz1, Tpz1, and Ccq1 in fission yeasts) associate with telomeres indirectly via protein-protein interactions, and together with DNA-binding proteins they form a protective complex called shelterin (6). Finally, some proteins (*e.g.* tankyrase, various nucleases, and DNA repair proteins) form transient contacts with telomeres, whose frequency and duration depend on the state of a particular telomere (1).

The sequences of telomere-associated proteins undergo relatively fast evolutionary diversification (7), thus hampering the identification of their counterparts in distantly related organisms needed to uncover general principles of telomere maintenance. These efforts are made easier by the fact that despite their dissimilarities at the level of amino acid sequences, telomere-associated proteins share common structural elements (8). All known ds telomeric DNA-binding proteins contain at least one conserved Myb/homeodomain, although the rest of the proteins have very different sequence and domain topology (1, 8). Thus, Myb domains of known telomeric proteins can be used effectively as queries for searches within whole genomic sequences of organisms distantly related to established models for telomere biology (9–11).

* This work was supported, in whole or in part, by grants from Slovak Research and Development Agency 0035-11 (to L. T.) and 0123-10 (to J. N.), Scientific Grant Agency 1/0311/12 (to L. T.) and 1/0405/11 (to J. N.), and by National Institutes of Health Grants ES 03773 and GM 31819 (to J. D. G.).
[S] This article contains supplemental Table 1 and Figs. 1–3.
[1] To whom correspondence should be addressed: Dept. of Genetics, Comenius University, Faculty of Natural Sciences, Mlynska dolina B-1, 842 15 Bratislava, Slovak Republic. Tel.: 421-2-60296-433; Fax: 421 2 60296 434; E-mail: tomaska@fns.uniba.sk.

Yeasts, especially *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, and *Candida albicans* have proved to be invaluable models for the studies of telomeres (12). The ascomycetous fungi are a highly heterogeneous group of microorganisms comprising more than 1000 known species resulting from diverse evolutionary trajectories (13), making them ideal for comparative analysis of telomeres. With this concept we initiated studies of nuclear telomeres in nonconventional yeast species including *Yarrowia lipolytica*. We characterized *Y. lipolytica* mutants lacking telomerase and found that they rapidly lose telomeric repeats and survive due to structural changes at the chromosomal ends (14). Characterization of proteins associated with the telomeres of this yeast species led to identification of a protein Tay1 (telomere associated in *Yarrowia lipolytica* 1), which (with the exception of its two Myb domains) does not resemble any known dsDNA-binding telomeric protein (15). *Yl*Tay1 protein and its homologs found in *S. pombe* (Mug152) and filamentous fungi thus represent a novel group of proteins protecting the ds portion of telomeres. It was shown that *Yl*Tay1p exhibits a preference for telomeric DNA *in vitro* and seems to bind to the substrate DNA as a dimer (15). Interestingly, the sequences of both Myb domains (Myb1, Myb2) are more similar to the Myb domains of mammalian TRF1/TRF2 proteins than to Myb domains of yeast telomeric proteins Rap1 (*S. cerevisiae*, *C. albicans*, *K. lactis*) or Taz1 (*S. pombe*). However, the contribution of individual Myb domains to a high affinity binding to telomeric repeats was not examined, and the binding properties of *Yl*Tay1 were not compared with those of its mammalian counterparts.

Here, using various biochemical and biophysical methods, we investigated the contribution of each Myb domain to the specificity of *Yl*Tay1 toward telomeric DNA. We found that although the individual Myb domains exhibit preferential binding to telomeric sequences, their combined presence dramatically increases the affinity of the protein for telomeric repeats. Interestingly, *Yl*Tay1 protein exhibits a substantially higher affinity for human telomeric repeats than either TRF1 or TRF2, raising both evolutionary and technical implications.

EXPERIMENTAL PROCEDURES

Microbial Strains—*S. cerevisiae* W303–1A (*MATa*, *ade2-1*, *his3-11*, *his3-15*, *leu2-3*, *leu2-112*, *trp1-1*, *ura3-1*, *can1-100*). Yeast cultures were grown in YPD medium (1% (w/v) yeast extract (Difco), 2% (w/v) Bacto-peptone (Difco), 2% (w/v) glucose) at 28 °C. *Escherichia coli* strain DH5α (Invitrogen; F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK^-, mK^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1) was used for cloning of recombinant DNA, and strain BL21- (DE3) (Invitrogen; F^- *omp*T *hsd*SB(rB^-, mB^-), *gal*, *dcm* (DE3)) was used for production of recombinant *Yl*Tay1–6HN proteins.

Construction of Yeast and Bacterial Expression Vectors Carrying Versions of the YlTAY1 Gene—Recombinant DNA techniques were carried out by standard procedures (16). The oligonucleotides (supplemental Table 1) were synthesized by MWG Operon, Metabion, or Sigma. For expression in *S. cerevisiae*, *YlTAY1* was amplified with primers YlTay1_pYES_UP and YlTay1_pYES_DOWN using pTay1–6HN plasmid DNA (15) as a template. The primers carried HindIII and XhoI restriction sites in regions flanking the start and the stop codons of the *YlTAY1* open reading frame, respectively. To facilitate efficient translation initiation, the primer YlTay1_pYES_UP contained the Kozak sequence (5′-AAAAAA-3′) immediately upstream of the initiation ATG codon. The resulting PCR product was gel-purified using the Zymoclean Gel Recovery kit (Zymo Research), digested with HindIII and XhoI, ligated into the pYES2/CT vector linearized with the same restriction enzymes, and transformed into *E. coli* DH5α grown on LB solid media with 100 μg/ml ampicillin. The resulting plasmid pYES-Tay1 carries *YlTAY1* under the control of the *GAL1* promoter.

The mutant versions of pYES-Tay1 lacking Myb1 (Δ1) and Myb2 (Δ2) were prepared by inverse PCR with primers ΔMyb1_UP and ΔMyb1_DOWN (for Δ1) or ΔMyb2_UP and ΔMyb2_DOWN (for Δ2) using pYES-Tay1 DNA as a template. The resulting PCR fragments were gel-purified, ligated using T4 DNA ligase, and transformed into *E. coli* DH5α grown on LB solid media with 100 μg/ml ampicillin. The mutant lacking both Myb domains (ΔΔ) was amplified by inverse PCR using ΔMyb2_UP and ΔMyb2_DOWN using pYES-Tay1-Δ1 as a template and cloned analogously as the single mutants. The mutant versions of pYES-Tay1 carrying substitutions at selected sites within either Myb1 or Myb2 domains were prepared by inversion PCR with the corresponding couple of primers (their names indicate the mutated amino acids; supplemental Table 1) using pYES-Tay1 DNA as a template. The bacterial expression vectors carrying mutant versions of *Yl*Tay1 lacking Myb domains (pTay1–6HN-Δ1, -Δ2, and -ΔΔ) were prepared analogously as described for pYES-Tay1 plasmids with the exception of using pTay1–6HN DNA as a template.

All PCR reactions were performed using 1 unit of Phusion Hot Start High-Fidelity DNA polymerase (Finnzymes) and contained each dNTP at 200 μM, corresponding primers at 1 μM and 10 ng of template DNA. The conditions of PCR reactions were adjusted according to the sequences of the primers and lengths of the final PCR product.

Yeast Growth Conditions and Transformation—Plasmids carrying WT (pYES-Tay1) or mutant versions of *YlTAY1* gene (pYES-Tay1-Δ1, -Δ2, and -ΔΔ) were transformed into *S. cerevisiae* strain W303–1A using the lithium acetate method (17), and transformants were selected for uracil prototrophy on solid SD media (0.17% (w/v) yeast nitrogen base, 0.5% (w/v) $(NH_4)_2SO_4$, 2% (w/v) glucose) containing corresponding amino acids and bases with the exception of uracil. Cells from single colonies were grown on SD media, then inoculated into liquid SGal (the same composition as SD except that instead of glucose, 2% (w/v) galactose was used as a sole carbon source) media and cultivated for 6 h at 28 °C. Three microliters were then spotted from 10-fold serial dilutions on either SD (repressible conditions) or SGal (inducible conditions) solid media, and the growth was inspected after a 3-day cultivation at 28 °C.

Purification of Recombinant YlTay1 Proteins—The expression and purification of various versions of *Yl*Tay1 protein was performed essentially as described in Kramara *et al.* (15) with several modifications. Bacterial expression plasmids carrying wild-type (pTay1–6HN-WT) as well as mutant versions (pTay1–6HN-Δ1, -Δ2, and -ΔΔ) of *YlTAY1* were transformed into One Shot BL21(DE3) cells, and the transformants were

grown on LB plates containing 100 μg/ml ampicillin. The cells were then inoculated into 30 ml of 2×YT media (1.6% (w/v) Bacto-Tryptone, 1% (w/v) Bacto-peptone, 1% (w/v) NaCl (pH 7.1)) containing 2% (w/v) glucose, 100 μg/ml ampicillin and cultivated overnight (15 h) at 37 °C and 225 rpm. The cells were centrifuged for 5 min at 3000 rpm (Sorvall RT 7 Plus) at 25 °C, washed once with 2×YT, inoculated into 1 liter of 2×YT containing 100 μg/ml ampicillin, and cultivated at 37 °C and 275 rpm until the A_{600} reached a value of 0.7–0.8. The culture was cooled to 28 °C followed by the addition of isopropyl β-D-1-thiogalactopyranoside (final concentration 1 mM) and cultivation for additional 3 h at 28 °C. The culture was then centrifuged for 15 min at 5000 rpm at 4 °C (F10–6 × 500y rotor in Sorvall RC 6+), the cells were washed once with 200 ml of ice-cold phosphate-buffered saline, and the pellet was frozen at −20 °C. The pellet was thawed on ice (30–45 min) and resuspended in a final volume of 30 ml of buffer A (20 mM HEPES-NaOH (pH 7.3), 300 mM NaCl, 1 mM DTT) containing 1× Complete (EDTA-free) protease inhibitors (Roche Applied Science), 10 mM $MgCl_2$, 50 units of DNase I (Invitrogen), and 2 μg of PureLink RNase A (Invitrogen). Lysozyme (Sigma) was added to a final concentration of 1 mg/ml, and the suspension was incubated for 15 min on ice with occasional shaking. The cells were broken by sonication (5 × 20 s at a setting of 7 (Branson Sonifer 450)). Each cycle of sonication was followed by 40 s of incubation on ice. Triton X-100 was added to a final concentration of 0.1% (v/v), and the suspension was incubated for an additional 15 min on ice. The insoluble material was pelleted by 30 min of centrifugation at 15,000 rpm at 4 °C (F21–8 × 50y in Sorvall RC 6+). The supernatant was mixed with a 0.5-ml bed volume of the His-Select(R) Cobalt Affinity Gel (Sigma) equilibrated with 3 × 10 volumes of buffer A. The whole suspension was transferred to a 50-ml Falcon tube and incubated for 60–90 min rocking at 7 °C. The beads were then washed 3 times with 20 volumes of buffer A containing 0.1% (v/v) Triton X-100 followed by 5 × 10 volumes of the wash buffer I (20 mM HEPES-NaOH (pH 7.3), 300 mM NaCl, 10 mM imidazole (pH 7.7)), and 5 × 10 volumes wash buffer II (50 mM sodium phosphate buffer (pH 7.0), 50 mM NaCl, 10 mM imidazole (pH 7.7)). The beads were then transferred to a chromatographic column, and the bound proteins were eluted with 6 × 1 ml of elution buffer (50 mM sodium phosphate buffer (pH 7.0), 50 mM NaCl, 500 mM imidazole (pH 7.7)). The fractions containing *Yl*Tay1 protein (or its mutant versions) were loaded onto 5-ml PD MidiTrap G-25 columns (GE Healthcare) prewashed three times with elution buffer without imidazole. The presence and purity of proteins were verified by 10% SDS-PAGE stained with Coomassie Brilliant Blue R-250. Concentrations of proteins were determined by the Bradford assay (Bio-Rad), and proteins were stored in 100-μl aliquots at −80 °C.

Electrophoretic Mobility Shift Assay (EMSA)—The double-stranded YlTEL50 probe (carrying 5 *Y. lipolytica* telomeric repeats)) used for EMSA was prepared by digestion of 15 μg of pMH25 plasmid (15) in a final volume of 50 μl using 50 units of EcoRI (New England Biolabs) followed by dephosphorylation using 50 units of calf intestinal phosphatase. The resulting 50-bp fragment was gel-purified using the ZymoClean gel extraction kit (ZymoResearch). The gel-isolated YlTEL50 DNA was labeled with 10 units of T4 polynucleotide kinase (Fermentas) and 50 μCi of [γ-^{32}P]ATP (final concentration 0.5 μM) for 60 min at 37 °C in a final volume of 20 μl. Reactions were diluted to 50 μl with water, and the labeled oligonucleotide was purified using G-50 Sephadex (GE Healthcare). The binding reaction and gel electrophoresis were performed as described (15).

Electron Microscopy—The typical DNA binding reaction for electron microscopy was performed in 10 μl of 1× HN buffer containing 5 ng/μl substrate DNA and 7–10 ng/μl purified *Yl*Tay1p. The reactions were carried out at room temperature for 15 min followed by the addition of 10 μl of 1.2% glutaraldehyde and incubation at room temperature for additional 6 min. To remove the unbound proteins and fixative, the samples were diluted to 50 μl in HN buffer and passed over 2-ml columns of 6% agarose beads (ABT Inc., Burgos, Spain) equilibrated with TE buffer (10 mM Tris-HCl, (pH 7.4), 0.1 mM EDTA-NaOH). Aliquots of the fractions containing the complexes were mixed with a buffer containing spermidine and adsorbed onto copper grids coated with a thin carbon film glow-charged shortly before sample application. After adsorption of the samples for 2–3 min, the grids were dehydrated through a graded ethanol series and rotary shadowcast with tungsten at 10^{-7} torr (18). Samples were examined in an FEI T12 TEM equipped with a Gatan 2kx2k SC200 CCD camera. Dimensions of particles in the images saved from the CCD cameras were analyzed using Digital Micrograph software (Gatan, Inc.). Adobe Photoshop software was used to arrange images into panels for publication.

Isothermal Titration Calorimetry (ITC)[2]—ITC measurements were performed on a VP-ITC instrument (Microcal; GE Healthcare) at 25 °C. Solutions of protein and DNA were prepared with the same batch of buffer (50 mM sodium phosphate, 50 mM NaCl (pH 7.0)) to minimize artifacts due to minor differences in buffer composition. The duplex portion of the telomeric DNA was prepared by hybridizing the complementary strands at a 1:1 molar ratio. The proper formation of the duplex was checked by ion-exchange chromatography using Mono Q HR 5/5 column (GE Healthcare) with a NaCl gradient from 0.1 to 2.0 M. The protein solution (11 μM) was degassed and placed in the cell (1.423 ml). The DNA solution (50 μM) in a syringe was gradually added to the protein solution in 25 injections of 10 μl at intervals of 5 min while stirring at 242 rpm. The experimental data were fitted in Origin 7.0 software supplied with the instrument using a one-site binding model. From the fit, the binding enthalpy change (ΔH), association constant (K_a) and binding stoichiometry (n) were obtained. Binding free energy (ΔG) and entropy (ΔS) contributions were determined from the standard equation $\Delta G = -\mathrm{RT}\ln K_a = \Delta H \times T\Delta S$. The uncertainties in the parameters obtained are given by standard deviation of three independent measurements and the minor inaccuracy in determination of protein and DNA concentrations (relative error contribution 10%).

Fluorescence Anisotropy—The equilibrium binding of *Yl*Tay1 protein variants to DNA oligonucleotide duplexes containing 1.5 telomeric repeat was analyzed by fluorescence anisotropy. The DNA oligonucleotide labeled with the Alexa Fluor 488 (Invit-

[2] The abbreviations used are: ITC, isothermal titration calorimetry; FA, fluorescent anisotropy.

rogen) was allowed to hybridize with the complementary oligonucleotide at an equimolar ratio. The complete formation of the duplex was verified by PAGE. The measurements of fluorescence anisotropy were conducted on a FluoroMax-4 spectrofluorometer (Horiba Jobin-Yvon, Edison, NJ) equipped with a thermostable cell holder and magnetic stirrer. Samples were excited with vertically polarized light at 490 nm, and both vertical and horizontal emissions were recorded at 520 nm. The integration time was 3 s. All measurements were conducted at 25 °C in 50 mM sodium phosphate buffer (pH 7.0) containing 50 mM NaCl. A fixed delay of 120 s was set between each protein aliquot addition and start of the measurement to allow the binding reaction to reach equilibrium. This delay was sufficient, as no further change in anisotropy was observed. Each data point is an average of three measurements. The experimental binding isotherms were analyzed by nonlinear least squares regression in SigmaPlot 11 software (Systat Software) using a single-site binding model according to Heyduk and Lee (19) and confirmed by numerical approach using DynaFit software (20).

RESULTS

Heterologous Expression in S. cerevisiae Is a Convenient System for Assessing the Role of Myb Domains on the Activity of YlTay1p—In silico analysis of the amino acid sequence of *Yl*Tay1 protein revealed that it contains two putative Myb domains (Ref. 15, Fig. 1*A*). To investigate their contribution to the biological effects of the protein, we took advantage of a serendipitous finding that the expression of the *YlTAY1* gene from a strong (*ADH1*) promoter inhibits growth of *S. cerevisiae*.[3] To test the possibility that the inhibitory effect of *YlTAY1* expression is due to its DNA binding activity, we prepared an expression vector with the *YlTAY1* gene and its mutant versions under the inducible *GAL1* promoter. As expected, whereas the growth of the transformants carrying the wild-type *YlTAY1* (WT) on glucose was similar to the control strain, their growth was dramatically reduced on media containing galactose as a sole carbon source (Fig. 1*B*). Next, we tested the mutant version of *Yl*Tay1p lacking both putative Myb domains (ΔΔ) and found that their growth on galactose was restored, indicating that the DNA binding activity of *Yl*Tay1p is responsible for the growth inhibition (Fig. 1*B*). Importantly, expression of the mutant versions of *Yl*Tay1p lacking either Myb1 (Δ1) or Myb2 domain (Δ2) did not interfere with the growth of *S. cerevisiae*, demonstrating that both Myb domains are essential for inhibition of growth mediated by the expression of *Yl*Tay1p on *S. cerevisiae*.

Although the molecular mechanism responsible for the inhibitory effect of *Yl*Tay1p on *S. cerevisiae* is not clear (see "Discussion"), the fact that it is dependent on the two Myb domains enables studies assessing the importance of specific amino acid residues for DNA binding of *Yl*Tay1p. With this aim we constructed a series of mutant versions of *YlTAY1* carrying mutations in the conserved positions of the Myb domains. The mutations were designed based on the phenotypes of the versions of hTRF2 mutated in different positions of its Myb domain (21, 22) (Table 1). The mutated genes were then expressed under the *GAL1* promoter to assess their effect on growth of *S. cerevisiae*. We found that although two mutant versions of *Yl*Tay1p (E159K and E246K) inhibited growth of *S. cerevisiae* cells on galactose, four versions (I178A, R189E, I264A, and R283E) lost this ability (Fig. 1*C*, Table 1). The same phenotype observed for deletion mutants and mutants carrying point mutations indicates that the deletion mutants (Δ1, Δ2, ΔΔ) can be used for a detailed biochemical analysis of the DNA binding properties exhibited by individual Myb domains. Furthermore, these results demonstrate that this heterologous experimental system has a potential for fine mapping of the

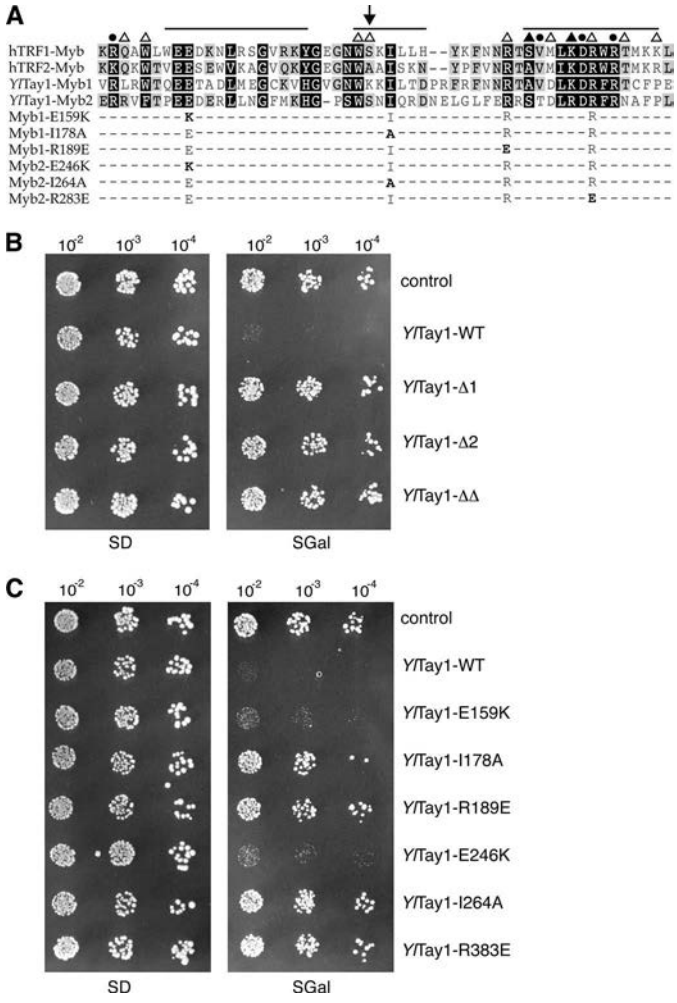


FIGURE 1. **YlTay1p contains two TRF-like Myb domains, both necessary to confer growth inhibition in *S. cerevisiae*.** *A*, shown is conserved features of Myb domains of human TRF1 and TRF2 proteins and two Myb domains of *Yl*Tay1p. Amino acid alignment was produced by ClustalW. Indicated are the positions mutated in the corresponding mutant variants of *Yl*Tay1p, chosen according to the results of Hanaoka *et al.* (21) and Konishi and de Lange (22). *Closed circles*, amino acids that interact with DNA bases; *open triangles*, amino acids that interact with the sugar-phosphate; *closed triangles*, amino acids that interact with both DNA bases and sugar-phosphate; *solid lines*, helical regions (positions of the residues in human TRF1 are adopted from Hanaoka *et al.* (21)). *Arrow*, see "Discussion." *B*, the inhibitory effect of galactose-induced expression of *YlTAY1* gene from *GAL1* promoter on growth of *S. cerevisiae* can be alleviated by deletion of Myb1, Myb2, or both Myb domains. The *scheme on left* illustrates the nomenclature of the deletion mutants lacking Myb1 and/or Myb2 used in this study. *C*, the inhibitory effect of expression of *YlTAY1* gene on growth of *S. cerevisiae* is alleviated by point mutations in some of the conserved amino acids.

[3] A. Mihalikova, S. Kinsky, and L. Tomaska, unpublished results.

TABLE 1

The positions mutated in *Yl*Tay1p and corresponding mutations in hTRF1/hTRF2

The positions were selected based on the studies of Hanaoka *et al.* (21) and Konishi and de Lange (22).

Mutation in *Yl*Tay1p	Corresponding mutation in hTRF2/mTRF2	Inhibition of growth in *S. cerevisiae*
WT		Yes
Δ1		No
Δ2		No
ΔΔ		No
Myb1_E159K	E454K/E449K	Yes
Myb1_I178A	I473A/I468A	No
Myb1_R189E	R482E/R477E	No
Myb2_E246K	E454K/E449K	Yes
Myb2_I264A	I473A/I468A	No
Myb2_R283E	R490E/R485E	No

*Yl*Tay1p protein (for example by mutagenic PCR) because it provides a convenient read-out for mutations disabling DNA binding.

Each Myb Domain of YlTay1p Mediates Preferential Binding to Telomeric Repeats—The experiments in *S. cerevisiae* indicated that both Myb domains of *Yl*Tay1p are important for mediating the effect of the protein, but they did not reveal whether loss of a single Myb domain results in the inability of *Yl*Tay1p to bind DNA, or alternatively, if it is due to changes in its specificity toward telomeric sequences. To address this issue, we prepared a set of bacterial expression vectors carrying the three-deletion versions of *Yl*Tay1p (Δ1, Δ2, ΔΔ). The WT as well as mutant proteins were produced in *E. coli*, purified by affinity chromatography (Fig. 2*A*), and assayed for their DNA binding activities *in vitro* using EMSAs. These experiments demonstrated that, similar to WT protein, both single mutants (Δ1, Δ2) exhibit DNA binding (Fig. 2*B*), indicating that a single Myb domain is fully competent for DNA binding. The double mutant (ΔΔ) did not exhibit any observable binding, demonstrating that at least one Myb domain is necessary for DNA binding *in vitro* (Fig. 2*B*).

To assess the ability of various versions of *Yl*Tay1p to bind to telomeric repeats, the purified recombinant proteins were incubated with the plasmid pYLTEL81 digested with *Bfu*AI, thus producing linear DNA molecules carrying 81 *Y. lipolytica* telomeric repeats at one end (15). Visualization of the DNA-protein complexes by electron microscopy clearly showed that in contrast to the double-mutant lacking both Myb domains (ΔΔ), all three versions of the *Yl*Tay1 protein carrying at least one Myb domain (WT, Δ1, Δ2) bound almost exclusively to the telomeric tract, with only a low binding activity observed toward the internal, non-telomeric region of the plasmid (Fig. 2*C*). Thus, single Myb domains (Myb1, Myb2) of *Yl*Tay1p not only have the ability to bind DNA but also exhibit preferential binding to telomeric sequences.

Measurements of Thermodynamic Parameters of YlTay1p by Isothermal Titration Calorimetry Reveal Its High Affinity Binding to Telomeric Repeats in Dimeric Form—To gain insight into the affinity and thermodynamics of wild-type *Yl*Tay1p binding to the DNA substrate, we performed measurements of *Yl*Tay1p binding to the 1.5 of the *Y. lipolytica* telomeric repeat (YlTR1.5; 5′-TTAGTCAGGGTTAGT-3′) by ITC. We used 1.5 telomeric repeats because EMSA analysis indicated that it represents the minimal binding site for *Yl*Tay1p (15). A typical ITC profile

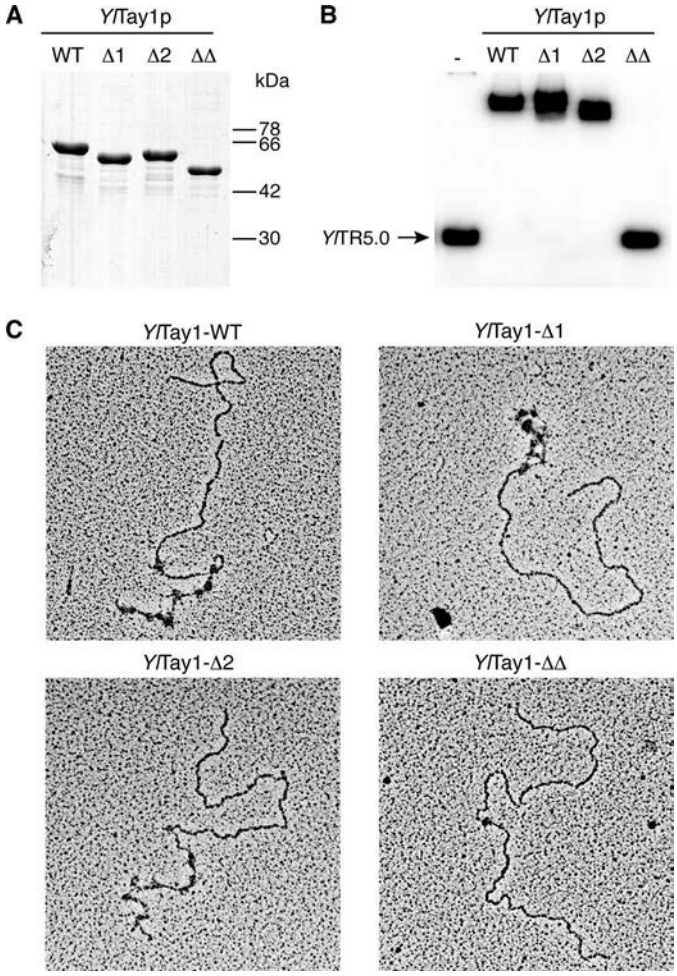


FIGURE 2. **Both Myb domains of *Yl*Tay1p are able to bind DNA and exhibit preference for telomeric repeats.** *A*, shown is SDS-PAGE analysis of purified 6xHN-*Yl*Tay1 protein and its mutated versions lacking Myb1 (Δ1), Myb2 (Δ2), and both (ΔΔ) Myb domains. *B*, electrophoretic-mobility shift assay demonstrates that both Myb domains of *Yl*Tay1p are able to bind telomeric DNA. *YlTR5.0*, double-stranded DNA probe carrying 5 *Y. lipolytica* telomeric repeats. *C*, shown is electron microscopic analysis of the binding of WT and mutant versions of *Yl*Tay1p to a model telomere containing 81 *Y. lipolytica* telomeric repeats. The total length of the linearized plasmid is 3500 bp; the length of the telomeric tract is 810 bp. The bound protein is represented by *dark particles at the terminal parts of DNA molecules*.

obtained at 25 °C is shown in Fig. 3, *top panel*. Exothermic heat pulses were observed after injections of DNA into the protein solution. Each area of the exothermic peak was integrated and corresponded to actual concentrations of the reacting molecules. The values of molar heat of binding for each injection were plotted as a function of DNA:protein molar ratio (Fig. 3, *bottom panel*). The resulting thermogram showed a best-fit according to a model for a single binding site by using a non-linear least squares method. The thermodynamic parameters, indicated in Fig. 3, *bottom panel*, were determined from the fit. The experimental errors of the thermodynamic values shown comprise a standard deviation of the fit for three independent experiments and the accuracy in determination of reactant concentrations.

The ITC experiments revealed that *Yl*Tay1p binds to the double-stranded telomeric DNA with relatively high affinity (K_D^{ITC} = 80 nM) (Fig. 3). Furthermore, the position of the

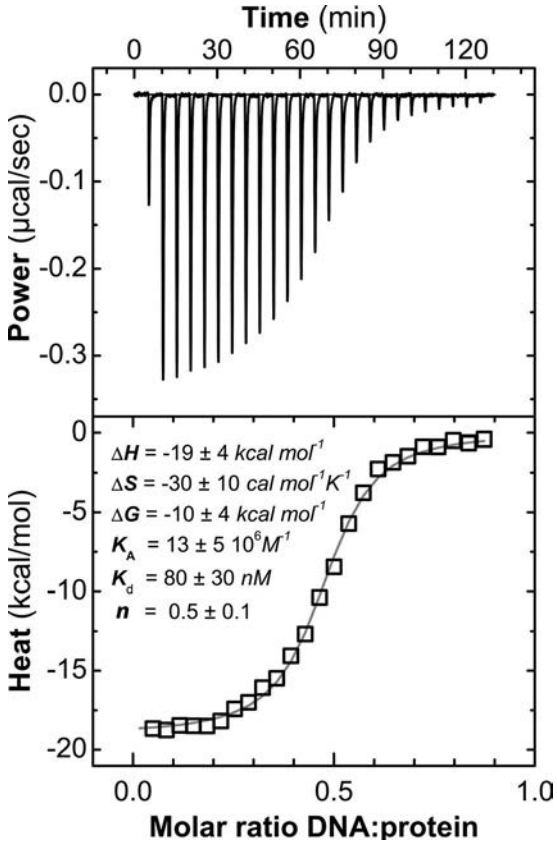


FIGURE 3. **Isothermal calorimetric titration of 15-bp telomeric DNA duplex into full-length *Yl*Tay1 protein.** The measurements were conducted at 25 °C in 50 mM sodium phosphate buffer (pH 7.0) containing 50 mM NaCl. The *top panel* shows the primary titration data; the *lower panel* shows the binding isotherm constructed from the primary data. The model curve (*gray*) represents the result of a nonlinear least squares fitting of the binding isotherm.

inflection point of the thermogram indicates that one molecule of the DNA duplex YlTR1.5 is bound by two molecules of *Yl*Tay1p (Fig. 3). The thermodynamic data show that the relatively high negative binding enthalpy ($\Delta H = -19$ kcal/mol) is partly compensated by entropy contribution ($\Delta S = -30$ cal/mol/K^{-1}; $-T\Delta S = 9$ kcal/mol). The negative binding enthalpy observed by isothermal titration calorimetry demonstrates the formation of hydrogen bonds and van der Waals interactions during protein-DNA complex formation. Additionally, the calculated negative entropy change corresponds to the constraint of the intramolecular vibrational flexibility and the reduction in the translational and rotational degrees of freedom after formation of the nucleoprotein complex (23). The enthalpy and entropy contributions result in a well pronounced free energy of binding ($\Delta G = \Delta H - T\Delta S = -10$ kcal/mol), which corresponds to the observed high binding affinity. From the thermodynamic parameters obtained, it is obvious that the DNA binding of *Yl*Tay1 protein is enthalpy-driven. Overall, the thermodynamic analysis of *Yl*Tay1p binding to *Y. lipolytica* telomeric DNA disclosed the high affinity and dimeric protein binding stoichiometry.

The High Affinity Binding of YlTay1p to Telomeric Repeats Is a Result of Synergic Effect of the Two Myb Domains—If both Myb domains of *Yl*Tay1p are able to bind preferentially to telomeric repeats, what is their contribution to the overall binding characteristics of the protein? To address this question, we employed fluorescent anisotropy (FA) and compared the binding affinity of WT along with the mutant versions of *Yl*Tay1p lacking Myb domains. The binding affinity of *Yl*Tay1p variants to double-stranded DNA was monitored by FA measurement. If the solution contains only free fluorescently labeled DNA molecules, FA is relatively low, due to the fast rotational rearrangement of DNA molecules. If the protein aliquots are added to the solution of labeled DNA, a bulky slower rotating protein-DNA complex is formed, and the anisotropy value increases. We used YlTR1.5 (labeled with Alexa Fluor 488) as a DNA substrate for *Yl*Tay1p binding assays (same substrate as the one used for ITC experiments). The anisotropy change described the extent of *Yl*Tay1p variants binding to telomeric DNA duplex. The equilibrium binding affinity was quantified by analysis of the recorded binding isotherms.

Importantly, both the association (K_a) and dissociation (K_D) constants obtained from the measurements of binding of WT *Yl*Tay1p to the DNA were in excellent agreement with these parameters calculated from the ITC data (Fig. 4, Table 2). The measurements of the binding of the mutant versions of *Yl*Tay1p to the telomeric oligonucleotide revealed that both single mutants (Δ1, Δ2) exhibit almost 4-fold lower affinity to the substrate than the WT protein (Fig. 4*A*; Table 2). The mutant version of *Yl*Tay1p lacking both Myb domains (ΔΔ), in agreement with the experiments described above, exhibited negligible binding to the DNA (Fig. 4*A*; Table 2). Thus, the equilibrium binding data support the cooperative contribution of both Myb domains to the overall high DNA binding affinity of *Yl*Tay1p.

YlTay1 Protein Exhibits Higher Binding Affinity to Both Y. lipolytica and Human DNA Telomeric Repeats Than Human TRF1 and TRF2—As the amino acid sequences of both Myb domains of *Yl*Tay1p are highly similar to the Myb domains of hTRF1 and hTRF2 (15), it was of interest to compare the binding affinity of all three proteins (*Yl*Tay1p, hTRF1, and hTRF2) to either *Y. lipolytica* (YlTR1.5) or human (HsTR2.0) telomeric DNA. The binding of *Yl*Tay1p variants to double-stranded telomeric oligonucleotides was monitored by fluorescence anisotropy. As one would assume, *Yl*Tay1p exhibited much higher (13–23-fold) affinity for the *Y. lipolytica* telomeric DNA (YlTR1.5) than either hTRF1 or hTRF2 (Fig. 4*B*, Table 2).

Surprisingly, the same result was observed when a human telomeric oligonucleotide was used for the binding reaction. Whereas the affinity of hTRF1 ($K_D = 114$ nM) and hTRF2 ($K_D = 135$ nM) for HsTR2.0 was similar to the affinity of *Yl*Tay1p for YlTR1.5 ($K_D = 81$ nM), WT *Yl*Tay1p bound human telomeric oligonucleotide with an ~8-fold higher affinity than either of the human telomeric proteins (Fig. 4*B*, Table 2). Further ITC analysis of WT *Yl*Tay1p binding to human telomeric DNA oligonucleotide HsTR2.0 has shown binding stoichiometry 2:1 (protein:DNA); *i.e.* two molecules of *Yl*Tay1p bind one dsDNA oligonucleotide. Moreover, the ITC measurement confirmed the higher binding affinity ($K_D^{ITC} = 18$ nM) of WT *Yl*Tay1p to human telomeric dsDNA HsTR2.0 (data not shown). Even one of the single mutants lacking the Myb1 domain (Δ1) exhibited more than a 2-fold higher affinity for HsTR2.0 than hTRF1 or hTRF2; similarly, the affinity of Δ2 was 2-fold higher than in the

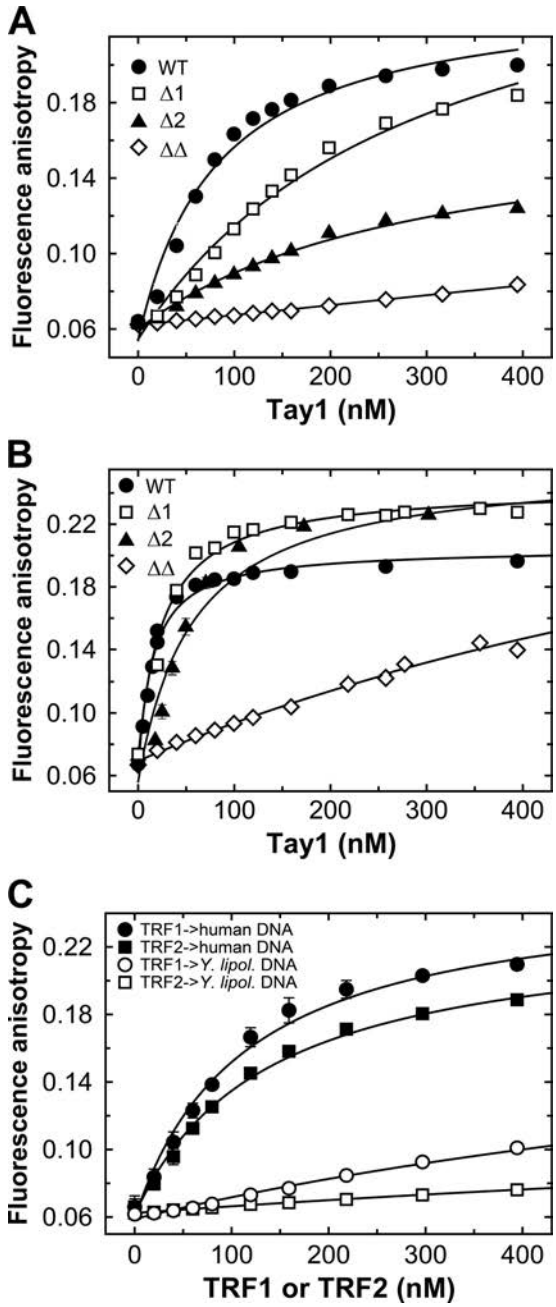


FIGURE 4. **Binding of *Yl*Tay1 protein to telomeric dsDNA.** *A*, shown is equilibrium binding of full-length *Yl*Tay1 (*WT*, *close circle*) protein and its deletion variants lacking Myb1 (Δ1, *open square*), Myb2 (Δ2, *close triangle*), and both Myb domains (ΔΔ, *open diamond*) to fluorescently labeled DNA duplex YlTR1.5 monitored by fluorescence anisotropy. The 15-bp telomeric dsDNA (10 nM) was titrated with protein solution at 25 °C in 50 mM sodium phosphate buffer (pH 7.0) containing 50 mM NaCl. The sequence of the DNA was the same as used in ITC studies. *B*, shown is equilibrium binding of full-length *Yl*Tay1 protein and its deletion variants to human telomeric dsDNA hTR2 monitored by fluorescence anisotropy. The conditions were the same as in *A*. *C*, shown is equilibrium binding of full-length TRF1 and TRF2 to human telomeric DNA HsTR2.0 (*close circle*, *close rectangle*, respectively) or to *Y. lipolytica* telomeric DNA duplex YlTR1.5 (*open circle*, *open rectangle*, respectively). The experimental conditions were the same as described in *A*.

case of the human proteins (Fig. 4*B*, Table 2). Altogether, comparison of the binding affinity of *Yl*Tay1p variants to *Y. lipolytica* and human telomeric DNA demonstrated that *Yl*Tay1p binds both telomeric sequences significantly tighter than human telomeric proteins TRF1 and TRF2.

TABLE 2

Parameters of binding of *Yl*Tay1p, hTRF1, and hTRF2 to *Y. lipolytica* (YlTR1.5) and human (HsTR2.0) telomeric oligonucleotides obtained by FA measurements

ND, not determined.

	YlTR1.5			HsTR2.0		
Protein	$K_a \times 10^{-6}$	K_D	Fit relative error	$K_a \times 10^{-6}$	K_D	Fit relative error
	M^{-1}	nM	%	M^{-1}	nM	%
WT	12.30	81	11	66.7	15	8
Δ1	3.30	302	12	22.7	44	12
Δ2	3.28	305	11	16.9	59	21
ΔΔ	<0.33	>3000	ND	0.58	1700	30
hTRF1	0.91	1100	18	8.77	114	16
hTRF2	0.53	1900	11	7.41	135	6

DISCUSSION

Telomere binding factors associated with either ss- or dsDNA regions of the chromosomal ends represent an evolutionarily highly divergent group of proteins (7, 8). Their only common feature is the presence of a conserved sequence and/or structural domain (OB-fold or Myb domain) mediating specific binding to telomeric repeats. However, the topology and the number of these DNA binding modules on the corresponding proteins greatly vary. Comparative analysis of the representatives of various types of telomere-binding proteins is the only approach enabling extraction of both specific and general characteristics of telomere protection in eukaryotes (12).

*Yl*Tay1p of *Y. lipolytica* together with its homologues in *S. pombe* (Mug152) and basidiomycetous fungi (15, 24) represent a unique group of telomeric proteins possessing two Myb domains that exhibit high similarity to the Myb domains of mammalian TRF1 and TRF2. The similarities between amino acid sequences of the Myb domains in some cases approaches 50% identity (Myb1 of *Yl*Tay1p and Myb of TRF1; (15)), which is a much higher value than that of the Myb domain of *Sp*Taz1, the most intensively studied fungal telomeric protein considered to be a functional orthologue of TRF1/TRF2. The main difference between *Yl*Tay1p and *Sp*Taz1p, TRF1, or TRF2 is the presence of two Myb domains within the single polypeptide of Tay1p. Although our previous data (gel filtration, electron microscopy) indicate that *Yl*Tay1p forms oligomers *in vitro*, it is possible that in contrast to *Sp*Taz1p, TRF1, and TRF2, where homo-oligomerization is a prerequisite for high affinity DNA binding, the tandem Myb domains of *Yl*Tay1p would allow a single monomer of the protein to bind effectively its target sequences.

The data supporting this hypothesis are derived from an artificial, yet highly informative experimental system. When we tried to test the ability of *Yl*Tay1p to bind *Y. lipolytica* telomeric repeats using the one-hybrid system, we were unable to transform *S. cerevisiae* with an expression plasmid carrying the *YlTAY1* gene under the control of a strong constitutive *ADH1* promoter.[3] When we placed *YlTAY1* under the inducible *GAL1* promoter, we demonstrated that the expression of this gene inhibits growth of *S. cerevisiae* (Fig. 1). We currently do not understand the reason for this inhibitory effect. EMSA assays using *Y. lipolytica*, *S. pombe*, and *S. cerevisiae* telomeric tracts as probes revealed that *Yl*Tay1p is able to bind to telomeres of distantly related yeast species *in vitro* (supplemental Fig. 1).

However, it is unlikely that the inhibition of growth of *S. cerevisiae* cells expressing the *YlTAY1* gene is caused by its binding to telomeres. First, we did not observe any differences in telomere length in the *S. cerevisiae*-expressing *YlTAY1* gene when the cells were transferred from glucose to galactose (supplemental Fig. 2). Second, *YlTAY1* seems to be an essential gene for *Y. lipolytica* (similarly to its *S. pombe* homologue Mug152 (15)), which is able to survive in the absence of telomeric repeats (14). It is more likely that in *S. cerevisiae Yl*Tay1p competes for the binding sites with essential DNA-binding proteins, possibly *Sc*Rap1p and/or *Sc*Tbf1p, both telomere-associated proteins that are indispensable not due to their telomeric functions but due to their binding to nontelomeric sites within the *S. cerevisiae* genome (25, 26). Displacement of such proteins by *Yl*Tay1p may thus interfere with their vital functions. Regardless of the reason, the ability of *Yl*Tay1p to inhibit growth in *S. cerevisiae* is dependent on its DNA binding activity (Fig. 1). The fact that mutant versions lacking either Myb1 or Myb2 domain do not affect growth of *S. cerevisiae* indicates that the high affinity binding is achieved by the concerted action of both Myb domains.

Experiments reconstituting the binding of *Yl*Tay1p and its mutant versions to telomeric sequences using EMSA and electron microscopy revealed that the individual Myb domains bind DNA (Fig. 2*B*) and exhibit preference for telomeric sequences (Fig. 2*C*). It is likely that the binding of the mutant proteins containing a single Myb domain results from oligomerization of the protein (see above). On one hand, this would suggest that the primary function of oligomerization for the wild-type protein is not to facilitate the specific binding to the DNA substrate but to mediate formation of telomere-telomere bridges as observed in our previous report (15). On the other hand, when the protein lacks a single Myb domain but retains its ability to form oligomers, it should still bind DNA with some specificity similar as in the case of the single hTRF1 Myb domain, which is able to specifically bind to human telomeres (27, 28). This line of argument together with data from *S. cerevisiae* (Fig. 1) predicts that even though single deletion mutants exhibit a preference for telomeric repeats, their affinity for the substrate should be substantially lower than in the case of wild-type protein.

The ITC and FA measurements point to exactly that direction (Figs. 3 and 4). Single deletion mutants lacking either Myb1 or Myb2 domain exhibited an almost 4-fold lower affinity for the *Y. lipolytica* telomeric sequence than the wild-type protein. This indicates that *Yl*Tay1p Myb domains act synergistically to set the affinity of the protein to the values observed by both ITC and FA ($K_D = 70 - 80$ nM). Similar K_D values have been obtained for binding of human TRF1 and TRF2 to human telomeric repeats (Fig. 4*B*, Table 2). These values are in good agreement with data obtained by Hanaoka *et al.* (21), although these authors observed that the DNA binding domain of hTRF1 has an almost 4-fold higher affinity for human telomeric repeats ($K_D \sim 200$ nM) compared with hTRF2 ($K_D \sim 750$ nM). The K_D values for hTRF1 DNA binding domain obtained by König *et al.* (27) were much lower (~3 nM). Furthermore, we compared the K_D values obtained for *Yl*Tay1p lacking either the Myb1 or Myb2 domain with those of plant proteins from the SMH family containing single Myb-like domains (29). We found that the affinities observed for *Yl*Tay1p variants with only one Myb domain fall into the interval observed for the SMH proteins ($K_D \sim 100 - 400$ nM).

The thermodynamic values determined by isothermal titration calorimetry are in good agreement with the values described in the single previous thermodynamic study showing equivalent c-Myb domain interaction with their cognate DNA sequence (30). In this study, Oda *et al.* (30) characterized the specific binding of c-Myb R2R3, containing two Myb domains, to its cognate DNA sequence using isothermal titration calorimetry. The value of binding enthalpy for c-Myb R2R3 at 25 °C ($\Delta H = -15$ kcal/mol) is in good agreement with the value −19 kcal/mol measured for the *Yl*Tay1p interaction with DNA in our study. Similarly, values of free energy of binding ΔG are comparable: −10 kcal/mol for *Yl*Tay1p and −12 kcal/mol for c-Myb R2R3; both values were measured at 25 °C. The comparable values of thermodynamic parameters for DNA binding of c-Myb R2R3 or *Yl*Tay1p, both containing two DNA binding domains, allow us to speculate that the binding is mediated by the same number of Myb domains. So, even though two molecules of *Yl*Tay1p containing a total of four DNA binding domains bind to the target sequence, only two DNA binding domains may take part in the interaction with DNA at any time. It is possible that the remaining two DNA binding domains may contribute to interactions leading to shaping of telomeric DNA into higher-order structures *in vivo*. The comparison of *Yl*Tay1p binding affinity to its cognate *Y. lipolytica* and human telomeric DNA reveals unexpected results. Intriguingly, *Yl*Tay1p exhibits about 8-fold higher affinity for human telomeric repeats than its mammalian counterparts (in fact, it binds to human telomeres with about 4-fold higher affinity than to its native *Y. lipolytica* telomeric substrate) (Fig. 4*B*, Table 2). Hanaoka *et al.* (21) observed that several substitutions at critical positions within the Myb domain substantially increased the affinity of binding of hTRF2, indicating that the natural selection does not necessarily lead to maximization of the binding affinity. In addition to differences in amino acid sequences, the higher affinity binding of *Yl*Tay1p to human telomeric repeats compared with hTRF1/hTRF2 may be caused by the presence of tandem Myb domains on a single polypeptide, making the binding more efficient than in the case of a single Myb domain-carrying protein that provides the two Myb domains in the form of homodimer. The observation that the affinity of Δ1 version of *Yl*Tay1p is within the same range as the TRF1/TRF2 goes in line with this argument (providing it still can form homo-oligomers observed for wild-type protein). Yet it is surprising that Δ1 has higher affinity for the human probe than Δ2, as Myb1 (retained in Δ2) is more similar to the Myb domain of TRF1/TRF2 than Myb2 (retained in Δ1), especially when both single mutants exhibit very similar affinity to *Y. lipolytica* telomeric repeats. At present we do not have an explanation for this discrepancy. Perhaps the presence of a lysine in Myb1, in the position in which serine or alanine mediate the contact with DNA (Fig. 1*A*, *arrow*; Ref 21), is responsible for a decreased binding of the Myb1 domain to the human telomeric sequence. As in the case of other telomere-binding proteins, like TRF1, TRF2, and *Sc*Rap1 (21, 27, 28, 31–33), more detailed structural

analysis of *Yl*Tay1p (or at least its Myb domains) would be highly instrumental in addressing these issues.

The observations that *Yl*Tay1p exhibits higher affinity for human telomeres than to *Y. lipolytica* telomeres and that it shows lower dissociation constants for binding to human telomeres than human telomeric proteins TRF1/TRF2 underline one important general evolutionary principle: natural selection does not necessarily lead to maximization of the affinity of a particular DNA-binding protein to its cognate DNA substrate. Perhaps the binding properties of a DNA-binding protein are tuned to inferior values, thus, enabling its dynamic association with the target DNA loci.

Comparative analyses, similar to that presented in our study, allow one to formulate hypotheses about evolutionary paths leading to contemporary telomeric binding factors. Even when limiting the comparison to fungi and mammals and disregarding proteins like Tbf1 able to bind telomeric repeats *in vitro* but binding to subtelomeric regions *in vivo* (25), the repertoire of Myb domain-containing proteins directly associated with double-stranded regions of telomeres is quite wide. Modern fungal or mammalian cells contain a diverse combination of three major players: (i) TRF-like proteins (*Sp*Taz1, TRF1, TRF2) possessing a single Myb domain and binding to telomeres as homooligomers (although a Myb domain of hTRF1 can bind to telomeric sequence (27)), (ii) *Yl*Tay1-like with two tandem Myb domains exhibiting high similarity to Myb domains of TRF-like proteins, and (iii) *Sc*Rap1-like proteins possessing two Myb domains exhibiting a weak similarity to the former two groups of proteins (32). The evolution of these three groups of proteins could have proceeded in two phases (supplemental Fig. 3). During the first phase, the ancestral genomes accumulated precursors of all three types of proteins possibly via the neutral evolutionary ratchet (34) involving gene and domain duplications. This might have generated a complex set of proteins that could have adapted to either general functions related to regulation of gene expression or more specialized functions at telomeres. In some cases the proteins started to play both telomere- and non-telomere-associated roles (*Sc*Rap1p, *Yl*Tay1p) making them essential components of the cell. In other cases, all three types of proteins were adopted for a specialized function(s), leading to loss of DNA binding activity (*Sp*Rap1, hRap1) and/or gene loss (there does not seem to be a *Yl*Tay1-like protein in mammalian cells and conventional yeast models like *S. cerevisiae*, *K. lactis*, *C. albicans*, or *S. pombe*). Are there any advantages of having two single Myb domain-containing proteins *versus* a protein containing two Myb domains within a single molecule? For example, the former is probably more dynamic in its effect (in response to its cellular level) as a certain minimum level of a protein is necessary to form homodimers with a reasonably high binding affinity. In addition, having two separate telomeric proteins forms a basis for higher levels of regulation that may be important for complex mammalian cells. On the other hand, the latter results in formation of more stable complexes independently of the instantaneous concentration of the protein and thus may be more tolerant to changes in environmental conditions as could be expected in single-cell organisms. Future studies on telomere-binding proteins as well as other components involved in telomere maintenance from various phylogenetic groups will fill the gaps in this incomplete picture.

Acknowledgments—We thank Ladislav Kovac (Comenius University, Bratislava, Slovak Republic) for inspiration and continuous support, members of our laboratories for discussions, Prof. Titia de Lange (Rockefeller University) for providing the hTRF2 protein construct, and Michal Zimmermann for preparation of the hTRF1 protein construct. This work was realized in part in the Central European Institute of Technology with research infrastructure supported by the project CZ.1.05/1.1.00/02.0068 financed from the European Regional Development Fund and the project of the Czech Science Foundation (P205/12/0550).

REFERENCES

1. de Lange, T. (2009) How telomeres solve the end-protection problem. *Science* **326,** 948–952
2. McEachern, M. J., Krauskopf, A., and Blackburn, E. H. (2000) Telomeres and their control. *Annu. Rev. Genet.* **34,** 331–358
3. Greider, C. W., and Blackburn, E. H. (1987) The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51,** 887–898
4. Theobald, D. L., Mitton-Fry, R. M., and Wuttke, D. S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32,** 115–133
5. Palm, W., and de Lange, T. (2008) How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* **42,** 301–334
6. de Lange, T. (2005) Shelterin. The protein complex that shapes and safeguards human telomeres. *Genes Dev.* **19,** 2100–2110
7. Linger, B. R., and Price, C. M. (2009) Conservation of telomere protein complexes. Shuffling through evolution. *Crit. Rev. Biochem. Mol. Biol.* **44,** 434–446
8. Lewis, K. A., and Wuttke, D. S. (2012) Telomerase and telomere-associated proteins. Structural insights into mechanism and evolution. *Structure* **20,** 28–39
9. Karamysheva, Z. N., Surovtseva, Y. V., Vespa, L., Shakirov, E. V., and Shippen, D. E. (2004) A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in *Arabidopsis*. *J. Biol. Chem.* **279,** 47799–47807
10. Kuchar, M., and Fajkus, J. (2004) Interactions of putative telomere-binding proteins in *Arabidopsis thaliana*. Identification of functional TRF2 homolog in plants. *FEBS Lett.* **578,** 311–315
11. Peška, V., Schrumpfová, P. P., and Fajkus, J. (2011) Using the telobox to search for plant telomere binding proteins. *Curr. Protein Pept. Sci.* **12,** 75–83
12. Lue, N. F. (2010) Plasticity of telomere maintenance mechanisms in yeast. *Trends Biochem. Sci.* **35,** 8–17
13. Dujon, B. (2010) Yeast evolutionary genomics. *Nat. Rev. Genet.* **11,** 512–524
14. Kinsky, S., Mihalikova, A., Kramara, J., Nosek, J., and Tomaska, L. (2010) Lack of the catalytic subunit of telomerase leads to growth defects accompanied by structural changes at the chromosomal ends in *Yarrowia lipolytica*. *Curr. Genet.* **56,** 413–425
15. Kramara, J., Willcox, S., Gunisova, S., Kinsky, S., Nosek, J., Griffith, J. D., and Tomaska, L. (2010) Tay1 protein, a novel telomere binding factor from *Yarrowia lipolytica*. *J. Biol. Chem.* **285,** 38078–38092
16. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning, A Laboratory Manual, 3rd Ed.,* Cold Spring Harbor Laboratory Press, Cold Pring Harbor, New York
17. Gietz, R. D., Schiestl, R. H., Willems, A. R., and Woods, R. A. (1995) Studies on the transformation of intact yeast cells by the LiAc/ss-DNA/PEG procedure. *Yeast* **11,** 355–360
18. Griffith, J. D., and Christiansen, G. (1978) Electron microscope visualization of chromatin and other DNA-protein complexes. *Annu. Rev. Biophys. Bioeng.* **7,** 19–35

19. Heyduk, T., and Lee, J. C. (1990) Application of fluorescence energy transfer and polarization to monitor *Escherichia coli* cAMP receptor protein and lac promoter interaction. *Proc. Natl. Acad. Sci. U.S.A.* **87,** 1744–1748
20. Kuzmic, P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data. Application to HIV proteinase. *Anal. Biochem.* **237,** 260–273
21. Hanaoka, S., Nagadoi, A., and Nishimura, Y. (2005) Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA binding activities. *Protein Sci.* **14,** 119–130
22. Konishi, A., and de Lange, T. (2008) Cell cycle control of telomere protection and NHEJ revealed by a ts mutation in the DNA binding domain of TRF2. *Genes Dev.* **22,** 1221–1230
23. Oda, M., and Nakamura, H. (2000) Thermodynamic and kinetic analyses for understanding sequence-specific DNA recognition. *Genes Cells* **5,** 319–326
24. Sánchez-Alonso, P., and Guzman, P. (2008) Predicted elements of telomere organization and function in *Ustilago maydis*. *Fungal Genet. Biol.* **45,** S54–S62
25. Koering, C. E., Fourel, G., Binet-Brasselet, E., Laroche, T., Klein, F., and Gilson, E. (2000) Identification of high affinity Tbf1p-binding sites within the budding yeast genome. *Nucleic Acids Res.* **28,** 2519–2526
26. Lieb, J. D., Liu, X., Botstein, D., and Brown, P. O. (2001) Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.* **28,** 327–334
27. König, P., Fairall, L., and Rhodes, D. (1998) Sequence-specific DNA recognition by the myb-like domain of the human telomere binding protein TRF1. A model for the protein-DNA complex. *Nucleic Acids Res.* **26,** 1731–1740
28. Nishikawa, T., Okamura, H., Nagadoi, A., König, P., Rhodes, D., and Nishimura, Y. (2001) Solution structure of a telomeric DNA complex of human TRF1. *Structure* **9,** 1237–1251
29. Hofr, C., Sultesová, P., Zimmermann, M., Mozgová, I., Procházková Schrumpfová, P., Wimmerová, M., and Fajkus, J. (2009) Single-Myb-histone proteins from *Arabidopsis thaliana.* A quantitative study of telomere binding specificity and kinetics. *Biochem. J.* **419,** 221–228
30. Oda, M., Furukawa, K., Ogata, K., Sarai, A., and Nakamura, H. (1998) Thermodynamics of specific and non-specific DNA binding by the c-Myb DNA binding domain. *J. Mol. Biol.* **276,** 571–590
31. Court, R., Chapman, L., Fairall, L., and Rhodes, D. (2005) How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA. A view from high resolution crystal structures. *EMBO Rep.* **6,** 39–45
32. Konig, P., Giraldo, R., Chapman, L., and Rhodes, D. (1996) The crystal structure of the DNA binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* **85,** 125–136
33. Nishikawa, T., Nagadoi, A., Yoshimura, S., Aimoto, S., and Nishimura, Y. (1998) Solution structure of the DNA binding domain of human telomeric protein, hTRF1. *Structure* **6,** 1057–1065
34. Lukeš, J., Archibald, J. M., Keeling, P. J., Doolittle, W. F., and Gray, M. W. (2011) How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life* **63,** 528–537

Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1

Karel Kubicek, Hana Cerna, Peter Holub, et al.

Genes Dev. 2012 26: 1891-1896 originally published online August 14, 2012
Access the most recent version at doi:10.1101/gad.192781.112

RESEARCH COMMUNICATION

Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1

Karel Kubicek,[1,3] Hana Cerna,[1,3] Peter Holub,[1] Josef Pasulka,[1] Dominika Hrossova,[1] Frank Loehr,[2] Ctirad Hofr,[1] Stepanka Vanacova,[1,4] and Richard Stefl[1,4]

[1]CEITEC-Central European Institute of Technology, Masaryk University, Brno, 62500, Czech Republic; [2]Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, Goethe-University Frankfurt, 60438 Frankfurt am Main, Germany

Recruitment of appropriate RNA processing factors to the site of transcription is controlled by post-translational modifications of the C-terminal domain (CTD) of RNA polymerase II (RNAP II). Here, we report the solution structure of the Ser5 phosphorylated (pSer5) CTD bound to Nrd1. The structure reveals a direct recognition of pSer5 by Nrd1 that requires the *cis* conformation of the upstream pSer5–Pro6 peptidyl-prolyl bond of the CTD. Mutations at the complex interface diminish binding affinity and impair processing or degradation of noncoding RNAs. These findings underpin the interplay between covalent and noncovalent changes in the CTD structure that constitute the CTD code.

Supplemental material is available for this article.

Received March 26, 2012; revised version accepted July 12, 2012.

The C-terminal domain (CTD) of the largest subunit of RNA polymerase II (RNAP II) consists of multiple tandem repeats of the heptapeptide consensus Tyr1–Ser2–Pro3–Thr4–Ser5–Pro6–Ser7 that is conserved from yeast to humans. The CTD is essential and forms a flexible tail of RNAP II. It serves as a binding platform for various cotranscriptional processing factors (Hirose and Manley 2000; Maniatis and Reed 2002; Meinhart et al. 2005). Phosphorylation and dephosphorylation of Ser2, Ser5, and Ser7 create a unique pattern in coordination with the transcription cycle (Komarnitsky et al. 2000; Meinhart et al. 2005; Phatnani and Greenleaf 2006; Kim et al. 2010; Mayer et al. 2010; Tietjen et al. 2010; Bataille et al. 2012). This phosphorylation pattern, often called the CTD code (Buratowski 2003, 2009; Chapman et al. 2008; Egloff and Murphy 2008), controls the recruitment, activation, and displacement of various factors involved in transcription and RNAP II transcript processing (Meinhart et al. 2005). It has been proposed that the CTD code is also affected by noncovalent changes in the CTD structure, such as peptidyl-prolyl bond isomerization (Buratowski 2003); however, the structural basis for such tuning of the CTD code and its role in recruitment of RNA processing factors are not yet fully understood. The importance of the CTD isomerization for Ser5 dephosphorylation has recently been demonstrated for the human and insect Ssu72 phosphatases (Xiang et al. 2010; Werner-Allen et al. 2011).

The pSer5 marks occur predominantly in the early elongation phase and are essential in 3′ end processing of short noncoding genes (Komarnitsky et al. 2000; Gudipati et al. 2008; Vasiljeva et al. 2008a). These marks are specifically recognized by Nrd1 and are required for the Nrd1-dependent termination pathway, used at small nuclear/nucleolar RNAs (sn/snoRNAs), cryptic unstable transcripts (CUTs), and other short RNAP II transcripts (Steinmetz et al. 2001; Arigo et al. 2006; Thiebaut et al. 2006; Vasiljeva et al. 2008a). Recent genome-wide studies in yeast demonstrated the co-occurrence of pSer5 and pSer7 marks at some genes in the early elongation phase (Kim et al. 2010; Mayer et al. 2010; Tietjen et al. 2010). In addition to specific phosphorylation, the Nrd1 termination pathway requires the Ess1 (Pin1 in humans) peptidyl-prolyl isomerase (Singh et al. 2009) that specifically isomerizes the pSer5–Pro6 peptidyl-prolyl bond in the CTD (Gemmill et al. 2005). In yeast, Ess1 stimulates dephosphorylation of pSer5–Pro6 in vivo (Singh et al. 2009), and therefore it has been hypothesized that it regulates the Nrd1 association with the CTD (Singh et al. 2009). To gain insights into the recruitment process of Nrd1 to the 5′ regions of genes, we determined the solution structure of the *Saccharomyces cerevisiae* Nrd1 CTD-interacting domain (CID) in complex with a CTD peptide phosphorylated at Ser5. The Nrd1 CID structure reveals a conserved CTD-binding site that engages the β-turn motif of the CTD formed by $Ser2_b$–$Pro3_b$–$Thr4_b$–$pSer5_b$ and a site recognizing selectively the upstream $pSer5_a$ and the *cis* conformation of the $pSer5_a$–$Pro6_a$ peptidyl-prolyl bond of the CTD. Furthermore, we show that the specific recognition of pSer5 CTD by Nrd1 CID is important for the processing and degradation of noncoding RNAs (ncRNAs) in vivo.

Results and Discussion

Affinity of Nrd1 to the CTD with 'early' phosphorylation marks

A previous study demonstrated that Nrd1 favors binding to the CTD with "early" pSer5 marks over the CTD with "late" pSer2 marks (Vasiljeva et al. 2008a). To test the effect of the unphosphorylated CTD and the CTD with the "early" pSer7 and pSer5 CTD marks on the affinity to Nrd1, we performed a quantitative solution-binding assay using fluorescence anisotropy (FA) experiments. We found that Nrd1 binds the pSer5 CTD with a significantly stronger affinity compared with the pSer7 CTD or unphosphorylated CTD (Fig. 1A). Nrd1 shows also only a slightly weaker binding to the doubly phosphorylated pSer5/7 CTD than to the pSer5 CTD (Fig. 1A),

[*Keywords*: RNA polymerase II; CTD code; phosphorylation; proline isomerization; RNA processing and degradation; NMR spectroscopy; structure]
[3]**These authors contributed equally to this work.**
[4]**Corresponding author**
E-mail richard.stefl@ceitec.muni.cz
E-mail vanacova@chemi.muni.cz
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.192781.112.

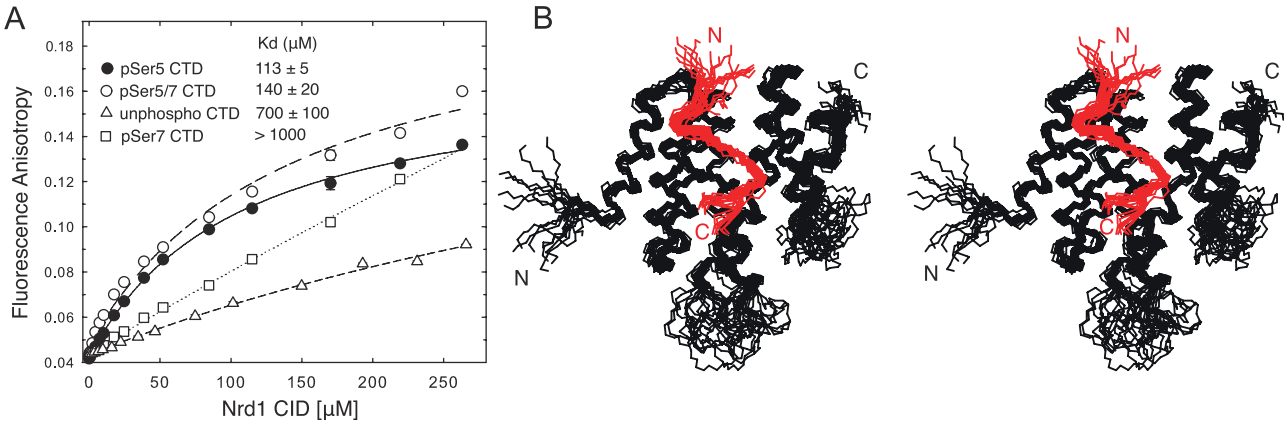


Figure 1. Structure of the Nrd1 CID–pSer5 CTD complex. (*A*) Equilibrium binding of the Nrd1 CID with differently phosphorylated CTD peptides monitored by FA. Binding isotherms and dissociation constants (K_d) are shown. (*B*) Overlay of the 20 lowest-energy structures of the Nrd1 CID–pSer5 CTD complex shown in stereo view. The backbone of the Nrd1 CID and pSer5 CTD is shown in black and red, respectively.

suggesting that the co-occurrence of the pSer5 and pSer7 marks may not impair the recruitment of the Nrd1 complex in the early elongation phase.

Structure of Nrd1 CID bound to the phosphorylated CTD

To understand how the pSer5 CTD is recognized by Nrd1, we determined the solution structure of a reconstituted complex consisting of the CID (residues 1–153) of Nrd1 and a 14-amino-acid peptide, the pSer5 CTD (two repeats of the heptapeptide CTD consensus phosphorylated at Ser5; $Tyr1_a$–$Ser2_a$–$Pro3_a$–$Thr4_a$–$pSer5_a$–$Pro6_a$–$Ser7_a$–$Tyr1_b$–$Ser2_b$–$Pro3_b$–$Thr4_b$–$pSer5_b$–$Pro6_b$–$Ser7_b$) (Fig. 1B; Supplemental Table S1; Supplemental Fig. S1). The ^1H, ^{13}C, and ^{15}N chemical shift assignments for the bound Nrd1 CID were obtained as described previously (Kubicek et al. 2011). The structure of the Nrd1 CID is formed by eight α helices in a right-handed superhelical arrangement (Fig. 1B) and is virtually identical to the structure of the Nrd1 CID in the free form (Vasiljeva et al. 2008a). The pSer5 CTD peptide contacts helices α2, α4, and α7 of the Nrd1 CID (Fig. 2A).

Recognition of the phosphorylated CTD by Nrd1

The CTD peptide adopts a β-turn conformation at $Ser2_b$–$Pro3_b$–$Thr4_b$–$pSer5_b$ and docks into a hydrophobic pocket of the Nrd1 CID that is formed by Ile29, Tyr67, Leu127, Ile130, and Met126 using $Tyr1_b$ and $Pro3_b$ residues (Fig. 2A,B). The hydroxyl group of $Tyr1_b$ forms a hydrogen bond with a conserved aspartate (Asp70) of Nrd1 (Fig. 2A,B). The binding mode of the peptide at the β-turn conformation resembles other previously determined structures of the CTD bound to CIDs of Pcf11, SCAF8, and Rtt103 (Supplemental Fig. S2; Meinhart and Cramer 2004; Becker et al. 2008; Lunde et al. 2010). However, in contrast to these CID–CTD complexes, Nrd1 binds more residues upstream of the pSer5 CTD via a conserved region at the N-terminal tip of helix α2 (Fig. 2A; Supplemental Fig. S3A). This unique region of Nrd1 is used to specifically recognize $pSer5_a$ via hydrogen bonding of Ser25 and Arg28 to the phosphate group of $pSer5_a$ (Fig. 2A,B). Another region that is more upstream in the α1–α2 loop has been previously suggested as the phosphoserine-binding site of Nrd1 based on a sulfate ion that was found in the crystal structure of the free Nrd1 CID (Vasiljeva et al. 2008a). The sulfate ion located ~8 Å away from

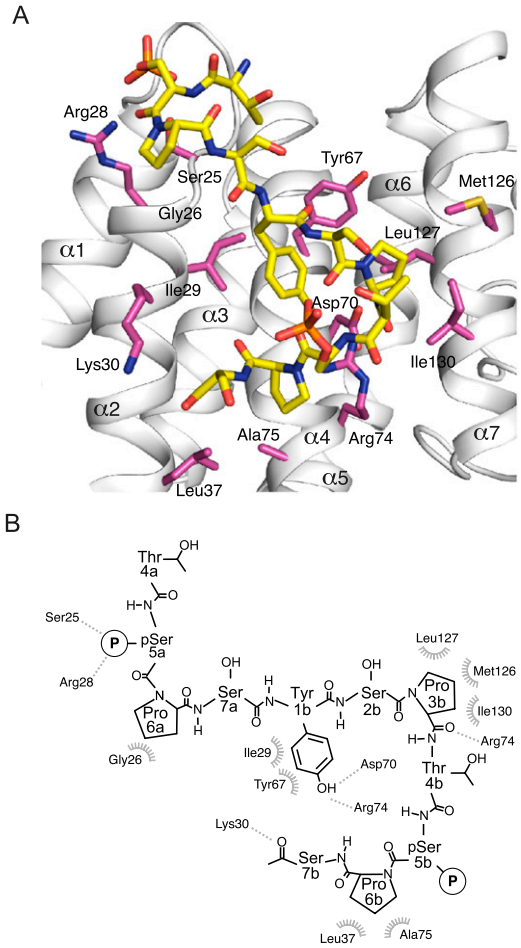


Figure 2. Recognition of the pSer5 CTD by Nrd1. (*A*) Scheme showing contacts between the Nrd1 CID and the CTD peptide. Protein residues that form hydrogen bonds and hydrophobic contacts to the CTD peptide are shown in white sticks. (*B*) Scheme showing contacts between the Nrd1 CID and the CTD peptide.

the phosphoserine location identified in our structure is coordinated using nonconserved amino acids. The structure of the Nrd1 CID–pSer5 CTD complex shows that the specific recognition of $pSer5_a$ is facilitated by the *cis* conformation of the $pSer5_a$–$Pro6_a$ peptidyl-prolyl bond. This *cis* conformation maximizes the intermolecular contacts and prevents the peptide from clashing with the α1–α2 loop of Nrd1. A conserved G26 in Nrd1 (all other CTD-containing proteins have a bulky and charged residues in this position) (Supplemental Fig. S3B) allows for loading of $pSer5_a$ into a highly electropositive pocket (Supplemental Fig. S3A). This is the first CID–CTD structure in which a phosphoserine–proline bond is observed in the *cis* conformation and in which a direct recognition of pSer5 is found. Importantly, both features are interconnected and thus required for the efficient binding. A similar conformation of the pSer5 CTD peptide was found in the crystal structure of the Ssu72–pSer5 CTD complex (Supplemental Fig. S4; Xiang et al. 2010; Werner-Allen et al. 2011).

Interaction between Nrd1 and the CTD is important for cell viability and the processing or degradation of ncRNAs

Specific association of the pSer5 CTD with the Nrd1 CID was further tested in a quantitative in vitro binding assay using FA. We titrated the wild-type and mutant Nrd1 CID against the fluorescently labeled pSer5 CTD. Alanine or aspartate (charge-swapping) substitutions at positions Ser25, Gly26, Arg28, Ile29, and Lys30 significantly decreased the binding affinity with the pSer5 CTD (Fig. 3A; Supplemental Fig. S6, control mutations of nonessential residues). In comparison, the effect of mutants at nonconserved positions Leu20, Lys21, and Ser22 in the region that was previously suggested to bind the phosphoserine (see above) is much smaller (Vasiljeva et al. 2008a). D70R and R74D variants of Nrd1 could not be assayed due to their instability at the high concentrations required for FA measurements. Next, we tested the effect of removal of phosphorylation in the downstream CTD repeat (at $Ser5_b$). We found that the CTD peptide with a single phosphorylation [pSer5(1P)] has the same affinity for Nrd1 as the CTD peptide phosphorylated at both Ser5s (Supplemental Fig. S5). Furthermore, mutations at $Pro6_a$ in the CTD peptide ($P6_aA$ CTD and $P6_aR$ CTD) have a larger negative effect on affinity to Nrd1 than mutations at $Pro6_b$ ($P6_bA$ CTD and $P6_bR$ CTD), confirming the requirement of the *cis* conformation at $pSer5_a$–$Pro6_a$ for the binding to Nrd1 (Supplemental Fig. S5). The effect of the $P6_aR$ mutation is larger than for $P6_aA$, as the bulkier side chain (the side chain at position 6_a is solvent-exposed) creates more unfavorable interactions with the adjacent side chain of pSer5, decreasing the stability of the *cis* conformation. This corroborates previous studies that showed that proline-to-alanine mutations do not necessarily alter the *cis* conformation if it is enforced by the structural context (Mayr et al. 1994; Xiong et al. 2000). Altogether, the FA data strongly support the phosphorylation-specific recognition ob-

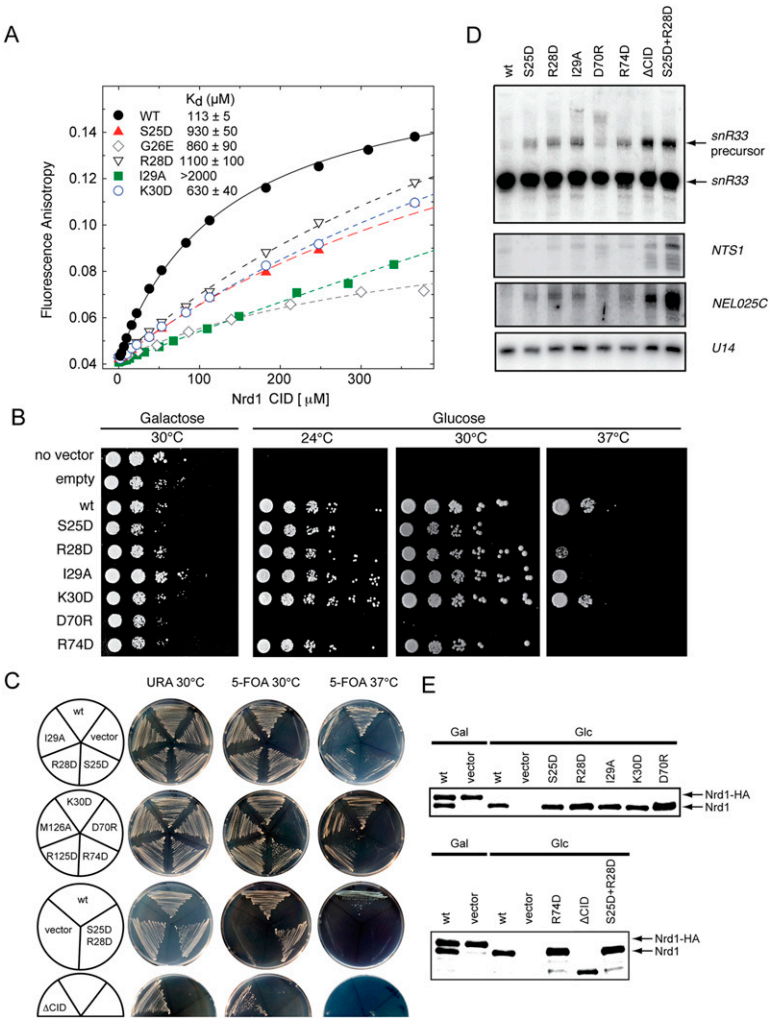


Figure 3. Critical residues of the Nrd1 CID that are required for CTD binding, cell viability, and RNA processing and degradation. (*A*) Equilibrium binding of the Nrd1 CID mutants with the pSer5 CTD peptide monitored by FA. Binding isotherms and dissociation constants (K_d) are shown for individual mutants. (*B*) Phenotypic analysis of the Nrd1 CID mutants. The mutants were expressed from *pRS415* plasmids in the *GAL1::NRD1* strain background. Growth on glucose-containing plates leads to the repression of *GAL1*-driven wild-type Nrd1 and thus shows the functionality of the different Nrd1 CID mutants. (*C*) Phenotypic analysis of the Nrd1 CID mutants. The mutants were expressed episomally from *pRS415* plasmids in the yeast strain where the endogenous *NRD1* was deleted and growth was supplemented with *NRD1* on the *URA3* plasmid (*pRS316*). To test the functionality of the different Nrd1 CID mutants, cells were grown on 5-FOA-containing plates for 3 d at the indicated temperatures. Wild-type *NRD1* was used as positive control, empty pRS415 plasmid was used as negative control, and $Nrd1_{\Delta 1-150}$ was the ΔCID. (*D*) *snR33* snoRNA processing efficiency and stability of *NTS1* and *NEL025C* CUTs analyzed by Northern blot analysis. Mature *U14* snoRNA represents a loading control. Total RNA was purified from cells expressing wild-type *NRD1* or the indicated mutants grown in glucose-containing medium. (*E*) Western blot analysis of expression levels of the wild-type and mutant Nrd1 originating from *pRS415* (faster-migrating band) in the yeast strain where the endogenous *NRD1* is under the galactose promoter. Protein extracts were prepared from the original *GAL1::NRD1* strain transformed with plasmids carrying wild-type and mutant *NRD1* grown in either galactose-containing medium (Gal) or glucose-containing medium (Glc). (*Top* band) The genomic *NRD1* copy contains a fusion HA tag, resulting in slower gel migration. Proteins were detected using specific antibodies against Nrd1p. Vectors with no insertion (vector) or containing the wild-type *NRD1* were used as negative and positive controls, respectively.

served in the structure of the Nrd1–pSer5 CTD complex and that the interaction relies on the presence of the invariant basic residues in the CID domain (Supplemental Fig. S3A).

To determine the importance of these individual residues for Nrd1 function in vivo, we monitored cell viability and ncRNA processing/stability in single-amino-acid mutants (Fig. 3B–D). The Nrd1 strain lacking the CID (Nrd1Δ1–150) was used as a reference for the CID-related function. Deletion of the CID is not lethal (Vasiljeva et al. 2008a); however, we observed that deletion of the CID led to inviability at 37°C (Fig. 3C). Similarly, to a lesser extent, a temperature-sensitive (ts) growth defect was observed in mutants of Ser25, Arg28, Ile29, and Arg74 (Fig. 3B,C). The double mutant in the residues contacting the phosphorylated serine (Ser25+Arg28) exhibited the same growth phenotype as ΔCID Nrd1 (Fig. 3C). Nrd1 CID deletion causes an accumulation of *snR33* precursors in vivo (Vasiljeva et al. 2008a). We observed that point mutants with the *ts* growth phenotype showed snoRNA processing and CUT degradation defects demonstrated by an accumulation of *pre-snR33* snoRNA and *NEL025c* and *NTS1* CUTs, respectively (Fig. 3D,E; Supplemental Fig. S7). None of the Nrd1 CID mutants tested displayed transcription termination defects that would be represented by readthrough product accumulation. Importantly, the Ser25+Arg28 double mutant showed processing and degradation defects comparable with those of the mutant lacking the entire Nrd1 CID. Taken together, these data demonstrate that specific recognition of the pSer5 CTD by the Nrd1 CID is important for the processing and degradation of ncRNAs in vivo.

The Nrd1 complex associates with the exosome and the TRAMP complex (Vasiljeva and Buratowski 2006). Based on our data, it is tempting to speculate that the Nrd1 CID mediates interaction with this processing and degradation apparatus. As the Nrd1 truncation lacking CID (Δ39–169) has no effect on the exosome copurification (demonstrated for Rrp6) (Vasiljeva and Buratowski 2006), it is likely that other RNA processing and degradation auxiliary factors, such as the TRAMP complex (LaCava et al. 2005; Vanacova et al. 2005), are recruited through the Nrd1 CID.

Previous chromatin immunoprecipitation (ChIP) experiments showed that Ess1 promotes the release of Nrd1 from terminator regions (Singh et al. 2009). This observation has been attributed to the indirect effect of Ess1 in which it stimulates dephosphorylation by Ssu72 (Singh et al. 2009). Ssu72 targets specifically the *cis* conformation of the peptidyl-prolyl bond of the pSer-Pro-containing peptides (Xiang et al. 2010; Werner-Allen et al. 2011). Here we suggest that Ess1 can also directly regulate the association of Nrd1 with the RNA Pol II CTD, as it specifically recognizes the pSer5–Pro6 CTD in the *cis* conformation, indicating that Ess1 may play a dual role in regulating the Nrd1 pathway. A detailed understanding of this mechanism will require further studies, but it is an exciting possibility, particularly in light of recent observations that suggested that Ssu72 may be a less "conformation-specific" phosphatase (acting also on the pSer7–Tyr1 CTD peptide that is unlikely to exist in the *cis* conformation) than previously expected (Bataille et al. 2012; Zhang et al. 2012).

Conclusions

The structure of Nrd1 CID–pSer5 CTD presented here reveals that the CTD recognition by Nrd1 requires both phosphorylation and isomerization of the RNAPII CTD. This suggests that the coupling of covalent and noncovalent changes in the CTD structure regulated by kinases/phosphatases and isomerases is crucial for the dynamical process of recruitment and displacement of appropriate processing factors during the transcriptional cycle. In addition, we show that specific recognition of the pSer5 CTD by the Nrd1 CID is important for the processing and degradation of ncRNAs in vivo, suggesting that these events occur cotranscriptionally.

Materials and methods

Protein expression and mutagenesis

The DNA encoding the *Saccharomyces cerevisiae* Nrd1 CID domain (residues 1–153) was amplified and cloned into a pET22b expression vector (Novagen) via NdeI and XhoI restriction sites. Details on cloning, expression, and purification of the Nrd1 CID construct have been described previously (Kubicek et al. 2011). Protein mutants were designed on the basis of the NMR structure of the Nrd1–pSer5 CTD complex and were prepared using the QuikChange site-directed mutagenesis kit (Stratagene).

The CTD of RNAP II

It has been established previously that the CTD mimic consisting of two repeats of the CTD canonical heptad yields the same binding affinity to Nrd1 CID as the CTD mimic of four repeats (Vasiljeva et al. 2008a). Thus, we used a 14-amino-acid peptide, the pSer5 CTD (two repeats of the heptapeptide CTD consensus phosphorylated at two Ser5s; $Tyr1_a$–$Ser2_a$–$Pro3_a$–$Thr4_a$–$pSer5_a$–$Pro6_a$–$Ser7_a$–$Tyr1_b$–$Ser2_b$–$Pro3_b$–$Thr4_b$–$pSer5_b$–$Pro6_b$–$Ser7_b$), in our study to mimic the CTD phosphorylated at Ser5. Similarly, the unphosphorylated CTD, the pSer7 CTD, and the doubly phosphorylated pSer5/7 CTD were used. The peptides were purchased from Clonestar Peptide Services.

NMR

All NMR spectra for the backbone and side chain assignments of 2.0 mM uniformly ^{15}N,^{13}C-labeled Nrd1 CID in 50 mM sodium phosphate buffer (pH 8.0), 100 mM NaCl, and 10 mM β-mercaptoethanol (90% H_2O/10% D_2O) were recorded on Bruker AVANCE 600- and 950-MHz spectrometers equipped with a cryoprobe at a sample temperature of 20°C. The spectra were processed using an NMRPipe package (Delaglio et al. 1995), and the protein resonances were assigned manually using Sparky software (T.G. Goddard and D.G. Kellner, University of California at San Francisco). The ^1H, ^{13}C, and ^{15}N chemical shifts of the bound form of the Nrd1 CID were assigned as described elsewhere (Kubicek et al. 2011). All distance constraints were derived from the three-dimensional (3D) ^{15}N- and ^{13}C-separated NOESYs and two-dimensional (2D) ^1H-^1H NOESY (with a mixing time of 80 msec) collected on a 950-MHz spectrometer. Intermolecular distance constraints were obtained from the 3D F_1-^{13}C/^{15}N-filtered NOESY-[^{13}C,^1H]-HSQC experiment (Zwahlen et al. 1997; Peterson et al. 2004), with a mixing time of 150 msec on a 950-MHz spectrometer. Intramolecular distance constraints of the bound CTD peptide (unlabeled) were derived from a 2D F_1,F_2-^{13}C/^{15}N-filtered [^1H,^1H]-NOESY (τ_m = 150 msec) (Zwahlen et al. 1997; Peterson et al. 2004). The NOEs were semi-quantitatively classified based on their intensities in the 2D and 3D NOESY spectra.

Structure calculations

The preliminary structure determinations of the Nrd1–pSer5 CTD complex were performed with the automated NOE assignment module implemented in the CYANA program (Guntert 2004). In the next step, CYANA-generated restraints along with manually assigned protein–CTD intermolecular restraints were used for further refinement of the preliminary structures with AMBER 10.0 software (Case et al. 2005). These calculations used a modified version (AMBER ff99SB) of the force field described by Cornell et al. (1995) using a protocol described previously (Stefl et al. 2010; Hobor et al. 2011). From 40 refined structures, the 20 conformers with the lowest AMBER energy were selected to form the final ensemble of structures. Molecular graphics were generated using MOLMOL (Koradi et al. 1996) and PyMOL (http://www.pymol.org). The

atomic coordinates and restraints for the Nrd1 CID–pSer5 CTD complex have been deposited in the Protein Data Bank under ID code 2lo6.

FA

The equilibrium binding of the Nrd1 CID to the differently phosphorylated CTD was analyzed by FA. The CTD peptides were N-terminally labeled with the 5,6-carboxyfluorescein (FAM). The measurements were conducted on a FluoroMax-4 spectrofluorometer (Horiba Jobin-Yvon). The instrument was equipped with a thermostatted cell holder with a Neslab RTE7 water bath (Thermo Scientific). Samples were excited with vertically polarized light at 477 nm, and both vertical and horizontal emissions were recorded at 525 nm. All measurements were conducted at 10°C in 50 mM phosphate buffer (pH 8.0) containing 100 mM NaCl and 10 mM β-mercaptoethanol. Each data point is an average of five measurements. The experimental binding isotherms were analyzed by nonlinear least-squares regression in SigmaPlot 11 software (Systat Software) using a single-site binding model according to Heyduk and Lee (1990).

Construction of yeast plasmids

The *pRS415* plasmid (*CEN*, *LEU2*) with insertion of the wild-type *NRD1* gene surrounded by the *NRD1* promoter and terminator (Vasiljeva et al. 2008a) was used as a template for QuikChange site-directed mutagenesis (Stratagene). See Supplemental Table S2 for primer sequences and Supplemental Table S3 for constructs generated in this study. The *NRD1 ΔCID* region was amplified as follows: Fragment 1: −340 nt up to +6 nt of *NRD1* with SVO F71 and SVO F72; Fragment 2: +453 nt up to TAA +300 nt from the 3′ untranslated region (UTR) with SVO F73 and SVO F74 primers. The two resulting PCR products were ligated together and inserted into the above-mentioned *pRS415*-based construct.

Yeast cultures and manipulation

Yeast were cultured under standard conditions in media with selective markers corresponding to particular strains and vectors. Yeast transformations were performed by the lithium acetate method.

Yeast growth test analysis

W303 (*GAL1::NRD1*)-derived strains (Supplemental Table S4) carrying appropriate mutant *NRD1* plasmids were grown in SD-LEU-HIS + 2% galactose at 30°C to an O.D. of 1.0. The cultures were serially diluted by a factor of 10 and spotted onto SD-LEU-HIS medium containing 2% glucose to repress the expression of the endogenous *NRD1* or control medium (SD-HIS + 2% galactose). Plates were incubated at 25°C, 30°C, and 37°C. EJS101-9d-derived strains (Supplemental Table S4) were grown for 3 d on SC-LEU plates, then spread on SC-LEU plates with or without 5-FOA and incubated at 25°C, 30°C, and 37°C.

Whole-cell protein extract preparation for Western blot analysis

Protein extracts were prepared from cultures grown on either galactose- or glucose-containing medium to an O.D. of 1.0. Five milliliters of culture was harvested and lyzed by 1.85 M NaOH for 15 min on ice, and proteins were subsequently precipitated with ice-cold trichloroacetic acid. Pellets were resuspended in 5% SDS and 8 M urea buffer prior to SDS-PAGE analysis. Proteins were resolved on a 12% SDS-PAGE gel, transferred to a nitrocellulose membrane by a semidry electroblotter (Bio-Rad), and probed for the presence of Nrd1p with the anti-Nrd1 sera kindly provided by David Brow (Steinmetz and Brow 1998).

RNA isolation and analysis

For RNA analysis, cells were inoculated to an O.D. of 0.1 in glucose- or galactose-containing SD medium and grown for 16 h at 30°C. RNA was isolated by hot phenol extraction and stored at −80°C. Five micrograms of total RNA was denatured in 25% formamide, separated on an 8% denaturing (8 M urea) polyacrylamide gel, and transferred to nylon membrane using semidry electro-transfer. RNA was cross-linked to the membrane by UV light (120 mJ/cm^2) and hybridized with a probe in Ultra-Hyb buffer (Ambion) according to the manufacturer's instructions.

Preparation of DNA probes for Northern blot analysis

The DNA probe for *snR33* was amplified from *S. cerevisiae* S288C genomic DNA with primers Forward, 5′-CGGAACGGTACATAAGAA TAGAAGAG-3′, and Reverse, 5′-TAAAGAAAACGATAAGAACTAA CCTC-3′. The *NTS1* probe 1 was prepared according to Vasiljeva et al. (2008b), by using primers Forward, 5′-TGAGTGCTTGTATAAGTTTA GAGAATTGA-3′, and Reverse, 5′-TTAATACTTTCCTCTTCGTCTTTT TCTAC-3′. The *NEL025c* probe was amplified with primers Forward, 5′-CCTGTTGACATTGCAGACAA-3′, and Reverse, 5′-GCAAAGATCTG TATGAAAGG-3′. The resulting PCR products were used as templates for random primed labeling using [α-^{32}P]dATP and the commercial kit (Roche). To detect *U14* snoRNA, the oligonucleotide 5′-TCACTCAGACATCC TAGG-3′ was 5′-phosphate-labeled by T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP.

Acknowledgments

We thank Anton Meinhart for helpful advice. We also thank Domenico Libri and David Brow for yeast strains, constructs, and antibodies. This work was supported by the project "CEITEC-Central European Institute of Technology" (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund, Czech Science Foundation (P305/12/G034; K.K. was supported by P305/10/1490, D.H. was supported by P305/11/1095, and C.H. was supported by P205/12/0550), Wellcome Trust 084316/Z/07/Z, and EMBO Installation Grant 1642. The NOESY spectra were obtained at the BMRZ NMR facility supported by the EU-NMR program (RII3-026145). P.H. is in receipt of the Brno City Municipality Scholarship for Talented PhD Students.

References

Arigo JT, Eyler DE, Carroll KL, Corden JL. 2006. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell* **23:** 841–851.

Bataille AR, Jeronimo C, Jacques PE, Laramee L, Fortin ME, Forest A, Bergeron M, Hanes SD, Robert F. 2012. A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Mol Cell* **45:** 158–170.

Becker R, Loll B, Meinhart A. 2008. Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **283:** 22659–22669.

Buratowski S. 2003. The CTD code. *Nat Struct Biol* **10:** 679–680.

Buratowski S. 2009. Progression through the RNA polymerase II CTD cycle. *Mol Cell* **36:** 541–546.

Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. *J Comput Chem* **26:** 1668–1688.

Chapman RD, Heidemann M, Hintermair C, Eick D. 2008. Molecular evolution of the RNA polymerase II CTD. *Trends Genet* **24:** 289–296.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J Am Chem Soc* **117:** 5179–5197.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biol NMR* **6:** 277–293.

Egloff S, Murphy S. 2008. Cracking the RNA polymerase II CTD code. *Trends Genet* **24:** 280–288.

Gemmill TR, Wu X, Hanes SD. 2005. Vanishingly low levels of Ess1 prolyl-isomerase activity are sufficient for growth in *Saccharomyces cerevisiae*. *J Biol Chem* **280:** 15510–15517.

Gudipati RK, Villa T, Boulay J, Libri D. 2008. Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat Struct Mol Biol* **15:** 786–794.

Guntert P. 2004. Automated NMR structure calculation with CYANA. *Methods Mol Biol* **278:** 353–378.
Heyduk T, Lee JC. 1990. Application of fluorescence energy transfer and polarization to monitor *Escherichia coli* cAMP receptor protein and *lac* promoter interaction. *Proc Natl Acad Sci* **87:** 1744–1748.
Hirose Y, Manley JL. 2000. RNA polymerase II and the integration of nuclear events. *Genes Dev* **14:** 1415–1429.
Hobor F, Pergoli R, Kubicek K, Hrossova D, Bacikova V, Zimmermann M, Pasulka J, Hofr C, Vanacova S, Stefl R. 2011. Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (NAB) 3 protein. *J Biol Chem* **286:** 3645–3657.
Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock DD, Megee PC, Bentley DL. 2010. Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat Struct Mol Biol* **17:** 1279–1286.
Komarnitsky P, Cho EJ, Buratowski S. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* **14:** 2452–2460.
Koradi R, Billeter M, Wuthrich K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graph* **14:** 51–55.
Kubicek K, Pasulka J, Cerna H, Lohr F, Stefl R. 2011. ^{1}H, ^{13}C, and ^{15}N resonance assignments for the CTD-interacting domain of Nrd1 bound to Ser5-phosphorylated CTD of RNA polymerase II. *Biomol NMR Assign* **5:** 203–205.
LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121:** 713–724.
Lunde BM, Reichow SL, Kim M, Suh H, Leeper TC, Yang F, Mutschler H, Buratowski S, Meinhart A, Varani G. 2010. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **17:** 1195–1201.
Maniatis T, Reed R. 2002. An extensive network of coupling among gene expression machines. *Nature* **416:** 499–506.
Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, Cramer P. 2010. Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* **17:** 1272–1278.
Mayr LM, Willbold D, Rösch P, Schmid FX. 1994. Generation of a non-prolyl *cis* peptide bond in ribonuclease T1. *J Mol Biol* **240:** 288–293.
Meinhart A, Cramer P. 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. *Nature* **430:** 223–226.
Meinhart A, Kamenski T, Hoeppner S, Baumli S, Cramer P. 2005. A structural perspective of CTD function. *Genes Dev* **19:** 1401–1415.
Peterson RD, Theimer CA, Wu H, Feigon J. 2004. New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA–protein complexes. *J Biol NMR* **28:** 59–67.
Phatnani HP, Greenleaf AL. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20:** 2922–2936.
Singh N, Ma Z, Gemmill T, Wu X, Defiglio H, Rossettini A, Rabeler C, Beane O, Morse RH, Palumbo MJ, et al. 2009. The Ess1 prolyl isomerase is required for transcription termination of small non-coding RNAs via the Nrd1 pathway. *Mol Cell* **36:** 255–266.
Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB, et al. 2010. The solution structure of the ADAR2 dsRBM–RNA complex reveals a sequence-specific readout of the minor groove. *Cell* **143:** 225–237.
Steinmetz EJ, Brow DA. 1998. Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. *Proc Natl Acad Sci* **95:** 6699–6704.
Steinmetz EJ, Conrad NK, Brow DA, Corden JL. 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. *Nature* **413:** 327–331.
Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D. 2006. Transcription termination and nuclear degradation of cryptic unstable transcripts: A role for the nrd1–nab3 pathway in genome surveillance. *Mol Cell* **23:** 853–864.
Tietjen JR, Zhang DW, Rodriguez-Molina JB, White BE, Akhtar MS, Heidemann M, Li X, Chapman RD, Shokat K, Keles S, et al. 2010. Chemical-genomic dissection of the CTD code. *Nat Struct Mol Biol* **17:** 1154–1161.
Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W. 2005. A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol* **3:** e189. doi: 10.1371/journal.pbio.0030189.
Vasiljeva L, Buratowski S. 2006. Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. *Mol Cell* **21:** 239–248.
Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A. 2008a. The Nrd1–Nab3–Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **15:** 795–804.
Vasiljeva L, Kim M, Terzi N, Soares LM, Buratowski S. 2008b. Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. *Mol Cell* **29:** 313–323.
Werner-Allen JW, Lee CJ, Liu P, Nicely NI, Wang S, Greenleaf AL, Zhou P. 2011. *Cis*-proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *J Biol Chem* **286:** 5717–5726.
Xiang K, Nagaike T, Xiang S, Kilic T, Behav MM, Manley JL, Tong L. 2010. Crystal structure of the human symplekin–Ssu72–CTD phosphopeptide complex. *Nature* **467:** 729–733.
Xiong Y, Juminaga D, Swapna GV, Wedemeyer WJ, Scheraga HA, Montelione GT. 2000. Solution NMR evidence for a *cis* Tyr–Ala peptide group in the structure of [Pro93Ala] bovine pancreatic ribonuclease A. *Protein Sci* **9:** 421–426.
Zhang DW, Mosley AL, Ramisetty SR, Rodríguez-Molina JB, Washburn MP, Ansari AZ. 2012. Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA polymerase II C-terminal domain is essential for viability and transcription termination. *J Biol Chem* **287:** 8541–8551.
Zwahlen C, Legault P, Vincent SJF, Greenblatt J, Konrat R, Kay LE. 1997. Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage λ N-peptide/boxB RNA complex. *J Am Chem Soc* **119:** 6711–6721.

Bioorganic & Medicinal Chemistry 21 (2013) 1078–1081

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

New PAH derivatives functionalized by cyclic nitrone framework: Synthetic design, anti-proliferative activity and interaction with DNA

Marian Buchlovič [a], Zdeněk Kříž [b], Ctirad Hofr [c], Milan Potáček [a,*]

[a] Department of Chemistry, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic
[b] National Centre for Biomolecular Research, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic
[c] Department of Experimental Biology, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

ARTICLE INFO

Article history:
Available online 11 January 2013

Keywords:
Polycyclic aromatic hydrocarbon (PAH)
Nitrones
DNA
Intercalation
Docking

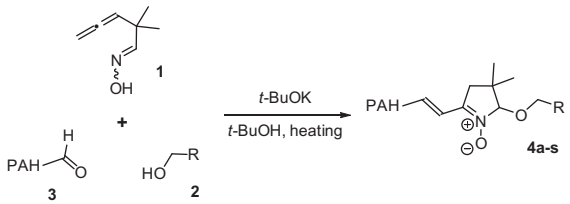
ABSTRACT

Novel approach to functionalized polycyclic aromatic hydrocarbons (PAHs) is presented. Incorporation of cyclic nitrone framework into the structure of PAHs was studied with respect to their anti-proliferative activities and interaction with double stranded DNA. Theoretical docking studies and UV titration methods were used for preliminary evaluation of binding of new PAH derivatives to DNA structure.

Polycyclic aromatic hydrocarbons (PAHs) are widely studied class of organic compounds from both chemical and biological point of view. Well known carcinogenic properties of PAHs are believed to be a result of preliminary metabolic activation.[1] In sharp contrast to the carcinogenic effects of PAHs, functionalized PAHs were recently identified as important candidates for cancer therapy.[2] Mechanism of action is often simplified and studied at the DNA level. A lot of discussion has been held about functional groups being introduced to the structure of PAHs and their contribution to the anticancer properties. It seems that polar side chains containing OH, NH_2, NHR, NR_2 and other similar groups may be capable of interaction with DNA backbone by non-covalent interactions (hydrogen bonding) and thereby provide the enhanced stability of the intercalation complex.[3] Limited number of studies have shown a relatively good correlation between antitumor properties of DNA intercalators and strength of DNA binding,[4] however plenty of other examples contradict this simplification and suggest a rather more complex[5a] or completely different mechanism of action.[5b–e] Furthermore, formation of free-radical species and reactive oxygen species, especially in the proximity of DNA molecule, was connected to many successful drug candidates including those that are well known DNA interacalators.[5e,6] Investigation of free-radical species involvement in the mechanism of action of anticancer agents and/or DNA intercalators therefore requires a new synthetic targets that will hold all of the desired features. In the following study we have selected a cyclic nitrone framework linked to PAH moiety as new model of functionalized PAH derivatives with a potential for DNA intercalation and resultant anticancer properties. Our selection of nitrone moiety as a key building block was motivated by the well known reactivity of nitrones with biologically interesting radicals[7] and broad experimental evidence dealing with biological activity of simple structured nitrones.[8] Interaction of target molecule with DNA and the evaluation of anticancer properties in vitro are the first choice methods for the study of new functionalized PAH derivatives and will be presented in this report.[9]

The synthetic design of new PAH derivatives was achieved by a multicomponent reaction of compounds **1**–**3** (Scheme 1). All relevant details of the presented nitrone chemistry were reported previously by our research group and can be found in the precedent work.[10]

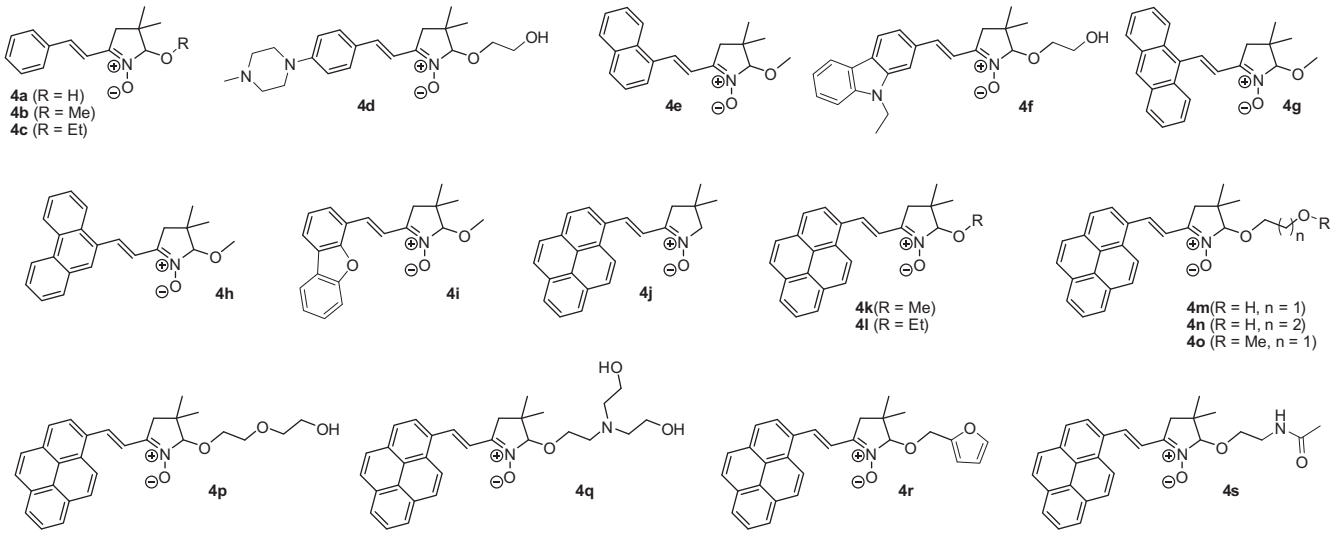
The facile synthetic method (Scheme 1) yielded nitrones **4a–s** that are characterized by both of the desired structural frameworks; (i) PAH moiety containing **1**–**4** fused aromatic cycles and (ii) nitrone moiety bearing variously functionalized polar side chains (see Scheme 2 for full list of prepared compounds).[11]



Scheme 1.

* Corresponding author. Tel.: +420 549496615; fax: +420 549492688.
E-mail address: potacek@chemi.muni.cz (M. Potáček).

0968-0896/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2013.01.004



Scheme 2.

All prepared nitrones **4a–s** were subjected to in vitro anticancer screening on five cancer cell lines.[12] Obtained results are summarized in Table 1. Roscovitine and Imatinib were used as reference compounds with well known antiproliferative activity associated with CDK inhibition.[13]

According to results shown in Table 1, compounds **4a–e** containing less than three fused aromatic cycles were found completely inatictive and shown none (**4a–d**) or modest (**4e**) antiproliferative activity in vitro. On the other hand, significantly lower IC_{50} values were obtained for compounds **4g–s**. In particular, medium to good antiproliferative activities (on micromolar level) that were comparable to reference compounds, were found for nitrones **4j–s**. As expected, compounds that contain 1-pyrenyl moiety showed significantly higher activities than their analogues with lower number of cycles in the aromatic conjugate (compare activities for compounds **4b,e,g,h,k**). It seems that nature and size of PAH moiety is crucial for the antitumor activity. These observations are in good agreement with the known low binding ability of mono and bicyclic aromates to DNA when compared to aromatic polycycles.[4] Moreover, a significantly lower in vitro and in vivo antitumor activity of aromates with less than three fused benzene rings was also observed in this regard.[4,5a] To verify the hypothesis in the case of compounds **4**, we have tested computational docking[14] of selected derivatives **4** (see Table 2) on a model of DNA duplex containing 10 base pairs. Two model sequences[15] $(AAGAATTCTT)_2$ and $(GGAAGCTTCC)_2$ were used in our calculations as simulations of AT and GC intercalation sites, respectively. According to the used method, interaction of ligand (compound **4**) and DNA duplex takes place implicitly at central couple of base pairs.

Calculated scoring function values of best complexes (Table 2) showed expected trend in stability of intercalation complexes. When number of cycles in the PAH moiety was increased (from monocycic **4b** to polycyclic **4g**), most significant energy gap was found between PAHs that contained two and three fused benzene rings (compare values that was calculated for compounds **4e** and **4g**). Similar relative difference in scoring function values was observed when the structure of side chains was altered while the type of PAH was kept unchanged. It seems that type and structure of functionalized nitrone moiety can alter the relative affinity to DNA duplex. In all studied examples, the lowest values of scoring function values were found for arrangement that allowed nitrone

Table 1
In vitro antiproliferative activity of compounds **4a–s** on human cancer cell lines[12]

Compound	IC_{50} (μM)				
	K562	MCF7	CEM	HCT116	BJ
4a	>100	>100	>100	n.a.	>100
4b	>100	>100	>100	n.a.	>100
4c	>100	>100	>100	n.a.	>100
4d	>100	>100	n.a.	n.a.	n.a.
4e	69.7	>100	40.0	>100	n.a.
4f	>100	>100	n.a.	n.a.	n.a.
4g	>25	>25	>25	n.a.	n.a.
4h	>12.5	>12.5	>12.5	n.a.	>12.5
4i	>12.5	>12.5	>12.5	n.a.	>12.5
4j	>10	>10	n.a.	n.a.	n.a.
4k	8.9	>50	13.5	14.5	n.a.
4l	>10	>10	n.a.	n.a.	n.a.
4m	11.2	>25	13.9	>25	11.1
4n	6.8	10.9	8.6	7.5	6.8
4o	>20	>20	n.a.	n.a.	n.a.
4p	>20	>20	n.a.	n.a.	n.a.
4q	7.1	9.6	n.a.	n.a.	n.a.
4r	>20	>20	n.a.	n.a.	n.a.
4s	>20	>20	n.a.	n.a.	n.a.
Roskovotin	42	11	n.a.	n.a.	n.a.
Imatinib	0.5	>10	n.a.	n.a.	n.a.

Tissue origin—K562 (blood), MCF7 (breast), CEM (blood), HCT116 (colon), BJ (skin), n.a. = data were not analyzed.

Table 2
Docking study of compounds **4b**,**e**,**g**,**h**,**k**–**n**,**p**,**q** and corresponding scoring function values of compound **4**/DNA complex

Compound	Scoring function value (kcal/mol)	
	$(AAGAATTCTT)_2$	$(GGAAGCTTCC)_2$
4b	−37.4	−32.1
4e	−40.4	−40.3
4g	−48.9	−47.8
4h	−49.9	−47.6
4k	−47.8	−46.9
4l	−50.4	−51.4
4m	−52.1	−53.2
4n	−53.1	−51.4
4p	−54.5	−55.6
4q	−60.3	−59.1

The scoring function values represent the lowest values for the orientation (arrangement) of compound **4** and DNA duplex.[14] All nitrones were used as neutral molecules for the docking study, no protonation of nitrone moiety or other functional groups in the structure of nitrones were considered in the calculation.

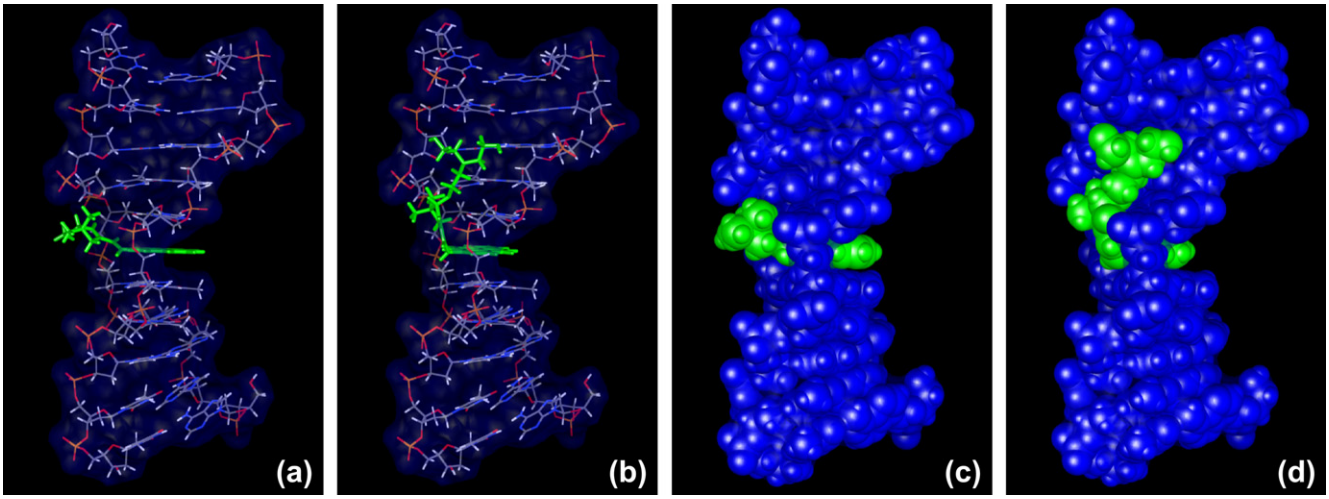


Figure 1. Visualization of compounds **4k** (a), (c) and compound **4q** (b), (d) intercalated into $(AAGAATTCTT)_2$ duplex. Complexes are shown in wireframe and spherical representation, respectively.

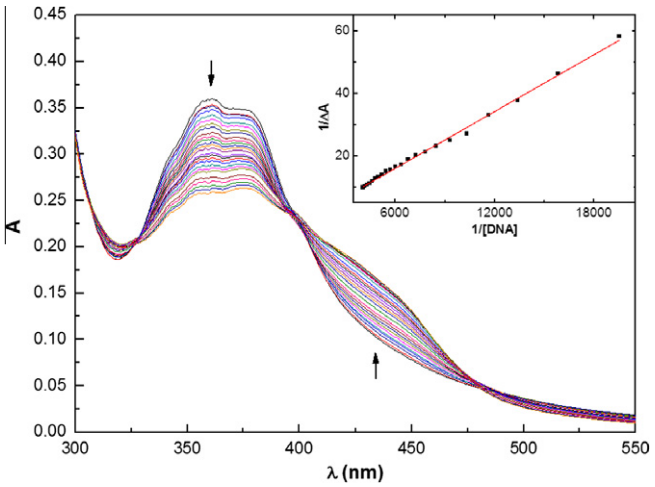


Figure 2. UV absorption profile of compound **4n** (2×10^{-5} M) in the presence of increasing amounts of ct-DNA (0–2.5×10^{-4} M). The titration was done in 5×10^{-3} M phosphate buffer at pH 7.0. The graph inset shows double-reciprocal plot binding analysis at 360 nm (ct-DNA concentration 5.1×10^{-5}–2.5×10^{-4} M).[16]

moiety and corresponding side chain to sit in the minor groove region of DNA duplex (Table 2). Thus the arrangement of compound **4** and DNA duplex complexes was analogous for simply structured nitrones (**4k**) as well as for longer or branched molecular side chain containing nitrones (**4p**). For selected visualizations of modeled structures see Figure 1. On the other hand, no significant trend and/or difference in scoring function value was found when intercalation at AT and GC site was compared for calculated complexes.

To study the interaction of compounds **4** and double stranded DNA experimentally, spectrophotometric titration was used. Nitrone **4n** was selected for this purpose, because this showed promising anti-proliferative activity (Table 1). In addition, relatively simple and linear side chain in compound **4n** should provide a well defined mode of binding to DNA. Figure 2 shows UV profile of compound **4n** in the presence of increasing concentration of ct-DNA. As expected, UV absorbance maxima of compound **4n** at 360 nm shows significant hypochromism (~28%) and new absorbance maxima appears around 435 nm. Three well defined isosbestic points were observed at 327, 399 and 480 nm.

The results obtained for compound **4n** are in good accordance with the behavior of substituted pyrenes intercalated to double stranded DNA structure. Both red shift and strong hypochromism of UV absorption maxima were observed for pyrenes coupled with five-membered heterocycles[9] or with other similar polar (hydrophilic) functional groups.[16] UV absorption data of compound **4n** at 360 nm were analyzed according to known methods[17] (see graph inset in Fig. 2) to give estimation of association constant value $K_a = 3.2 \times 10^3$ M^{-1}. The analysis offered a clear indication of binding of nitrone **4n** to ct-DNA that is comparable to related pyrene derivatives.[9] Due to the presence of unique nitrone functionality in the structure of compounds **4** that has not precedent example in the literature, further investigation is needed for full confirmation of the binding mode.[18]

In conclusion, we have presented a novel synthetic design of the first PAH derivatives coupled to cyclic nitrones. According to our preliminary studies presented in this Letter, the proposed methodology offers target molecules that are capable of interaction with double stranded DNA structure. Examples of 19 new synthetic derivatives of PAHs were tested for their antiproliferative activity against five tumor cell lines. Nitrones that contain PAH with more than two cycles showed promising activity on micromolar concentration level. In particular, substituted pyrenes **4k,n,q** offered the best results from tested compounds. Functionalized pyrenes also showed strongest binding to DNA duplexes according to calculated relative energies of corresponding intercalation complexes. Interaction of compound **4n** with ct-DNA showed UV characteristics that are typical for intercalative mode of binding.

Acknowledgments

The authors thank Vladimír Kryštof and co-workers from Laboratory of Growth Regulators of the Faculty of Science, Palacký University, for evaluation of antiproliferative activity. The research was supported by Grant Agency of the Czech Republic, Grant No. 203/09/1345.

References and notes

1. Luch, A. *The Carcinogenic Effects of Polycyclic Aromatic Hydrocarbons*; Imperial Collage Press: London, 2005.
2. (a) Banik, B. K.; Becker, F. F. *Curr. Med. Chem.* **2001**, *8*, 1513. and references cited therein; (b) Banik, B. K.; Becker, F. F. *Bioorg. Med. Chem.* **2001**, *9*, 593; (c) Rescifina, A.; Varrica, M. G.; Carnovale, C.; Romeo, G.; Chiacchio, U. *Eur. J. Med. Chem.* **2012**, *51*, 163.
3. Wunz, T. P.; Craven, M. T.; Karol, M. D.; Hill, C. G.; Remers, W. A. *J. Med. Chem.* **1990**, *33*, 1549.
4. For examples see: Ohara, K.; Smietana, M.; Restouin, A.; Mollard, S.; Borg, J. P.; Collette, Y.; Vasseur, J. J. *J. Med. Chem.* **2007**, *50*, 6465.

5. (a) Bair, K. W.; Andrews, C. W.; Tuttle, R. L.; Knick, V. C.; Cory, M.; McKee, D. D. *J. Med. Chem.* **1983**, *1991*, 34; (b) Graves, D. E.; Velea, L. M. *Curr. Org. Chem.* **2000**, *4*, 915. and references cited therein; (c) Snyder, R. D. *Mutat. Res.* **2007**, *623*, 72; (d) Denny, W. A.; Rewcastle, G. W.; Baguley, B. C. *J. Med. Chem.* **1990**, *33*, 814; (e) Baguley, B. C.; Wakelin, L. P. G.; Jacintho, J. D.; Kovacic, P. *Curr. Med. Chem.* **2003**, *10*, 2643.
6. Kovacic, P.; Osuna, J. A. *Curr. Pharm. Des.* **2000**, *6*, 277.
7. (a) Berliner, L. J. *Appl. Magn. Reson.* **2009**, *36*, 157. and references cited therein; (b) Gomez-Mejiba, S. E.; Zhai, Z.; Akram, H.; Deterding, L. J.; Hensley, K.; Smith, N.; Towner, R. A.; Tomer, K. B.; Mason, R. P.; Ramirez, D. C. *Free Radical Biol. Med.* **2009**, *46*, 853.
8. (a) Floyd, R. A.; Hensley, K.; Foster, M. J.; Kelleher-Andersson, J. A.; Wood, P. L. *Mech. Ageing Dev.* **2002**, *123*, 1021; (b) Floyd, R. A.; Kopke, R. D.; Choi, Ch. H.; Foster, S. B.; Doblas, S.; Towner, R. A. *Free Radical Biol. Med.* **2008**, *45*, 1361.
9. For examples of related studies on intercalation of pyrene derivatives, see: Rescifina, A.; Chiacchio, U.; Piperno, A.; Sortino, S. *New J. Chem.* **2006**, *30*, 554.
10. For the chemistry of nitrones **4**, see: (a) Buchlovič, M.; Man, S.; Potáček, M. *Synthesis* **2012**, *44*, 973; (b) Buchlovič, M.; Man, S.; Kislitson, K.; Mathot, Ch.; Potáček, M. *Tetrahedron* **1821**, *2010*, 66; (c) Buchlovič, M.; Man, S.; Potáček, M. *Tetrahedron* **2008**, *64*, 9953; (d) Man, S.; Buchlovič, M.; Potáček, M. *Tetrahedron Lett.* **2006**, *47*, 6961; (e) Buchlovič, M.; Man, S.; Potáček, M. *Tetrahedron Lett.* **2010**, *51*, 5801; (f) Buchlovič, M.; Hebánová, S.; Potáček, M. *Tetrahedron* **2012**, *68*, 3117.
11. For multicomponent reaction of allenyloxime **1** and various alcohols **2** and aldehydes **3** that was used for synthesis of nitrones **4a–i**,**k–s**, see Ref. 10a Synthesis of nitrone **4j** was achieved by reductive cyclization of oxime **1** in the presence of $NaBH_3CN$ (see Ref. 10c) and subsequent condensation with 1-pyrene carboxaldehyde according to conditions used in Ref. 10a
12. In vitro antiproliferative activity was evaluated by the same method as described in: Kryštof, V.; McNae, I. W.; Walkinshaw, M. D.; Fischer, P. M.; Müller, P.; Vojtěšek, B.; Orság, M.; Havlíček, L.; Strnad, M. *Cell. Mol. Life Sci.* **2005**, *62*, 1763.
13. (a) Huber, R. J.; O'Day, D. H. *J. Cell. Biochem.* **2012**, *113*, 868; (b) Ranza, E.; Mazzini, G.; Facoetti, A.; Nano, R. *J. Neurooncol.* **2010**, *96*, 349.
14. The double stranded DNA with sequence $(AAGAATTCTT)_2$ and $(GGAAGCTTCC)_2$ has been prepared using nucleic acid builder and leap modules of the Amber software package.[19] The canonical B-DNA secondary structure was prepared. Obtained structures have been optimized using simulated annealing protocol with application of Watson–Crick base pair restraints and distance restraints between central couple of base pairs according to distances of intercalated ethidium chloride between pairs DNA from X-ray structure from Cambridge Structural Database (CSD). The system was heated to 600 K in first 5 ns, then was slowly cooled to 100 K for 16 ns and at the end it was cooled to 0 K for 4 ns. At the end the MD simulation with fully solvated structure in explicit solvent were ran for 2 ns with applicate distance restraints for couple of central base pairs. The TIP3P water model and FF99SB Amber force field were used for MD simulation. The structures of compounds **4** were modeled using the PC-Model software.[20] The 3D structure was first minimized using molecular mechanics method with application of MMX force field. The obtained structures were then optimized using ab initio HF method with 6-31G* basis set implemented in the GAUSSIAN 03[21] and RESP charges[22] were calculated. The ligand molecules in required format were prepared using antechamber module of Amber software package.[19] The optimized structure of DNA duplex from the end of MD simulation in explicit solvent was used as model of receptor for docking studies. The flexible anchor and growth approach implemented in Dock 6.4 software[23] were used. For the docking study following parameters were used: maximal number of orientation was set to 50 millions, only ligand was set as flexible for docking, the 6–12 van der Waals potential was used for scoring and maximally 100,000 iterations was used for anchor and also for grow optimizations. The internal ligand energy was also used for scoring function. Grid step was set 0.3 Å for electrostatic and van der Waals grids. All atom model was used for scoring function evaluations. All structures were performed clustering analysis with RMS threshold of 2.0 Å and 10 best clusters were saved.
15. We used rather short DNA sequences in our calculations to make the system as simple as possible; note that shorter DNA duplexes (less than 10 base pairs) were found unstable when subjected to our MD simulation (see Ref. 14 for method details).
16. Wolfe, A.; Shimer, G. H.; Meehan, T. *Biochemistry* **1987**, *26*, 6392.
17. UV spectra were recorded on Shimadzu UV 1601 spectrometer. Titration was done directly in 1.0 cm quartz cells. Concentration of compound **4n** was kept the same during the titration. Calf thymus DNA (Sigma–Aldrich #D1501) solution in phosphate buffer and solution of **4n** in phosphate buffer were prepared and used immediately before use. Concentration of DNA in base pair was determined using the extinction coefficient 6600 M^{-1} cm^{-1} at 260 nm. Binding constant K_B was determined according to published method[16] from double reciprocal plot of the change in apparent UV absorption of compound **4n**: $1/\Delta\varepsilon = 1/\Delta\varepsilon_{max} + 1/(K_B\Delta\varepsilon_{max}[DNA])$
Here $\Delta\varepsilon$ corresponds to apparent decrease in A_{360} of **4n**, $\Delta\varepsilon_{max}$ is decrease in A_{360} of totally bound **4n** and [DNA] is actual concentration of calf thymus DNA in base pair.
18. For examples of binding mode evaluation, see: Palchaudhuri, R.; Hergenrother, P. J. *Curr. Opin. Biotechnol.* **2007**, *18*, 497.
19. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, 2008.
20. PC-Model v. 9.0, Serena Software, Bloomington, 2009, IN 47402-3076.
21. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T., Jr.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. GAUSSIAN *03, Revision E.01*; GAUSSIAN, Inc.: Wallingford, 2004.
22. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.
23. (a) Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. T.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA* **2009**, *15*, 1219; (b) Graves, A. P.; Shivakumar, D. M.; Boyce, S. E.; Jacobson, M. P.; Case, D. A.; Shoichet, B. K. *J. Mol. Biol.* **2008**, *377*, 914.