Single-myb histone proteins from Arabidopsis thaliana: Quantitative study of telomere binding specificity and kinetics

Ctirad Hofr, Pavla Šultesová, Michal Zimmermann, Iva Mozgová, Petra Procházková Schrumpfová, Michaela Wimmerová, Jiří Fajkus

Single-Myb Histone Proteins from *Arabidopsis thaliana*: 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*

[‡]*Laboratory of DNA-Molecular Complexes, Institute of Biophysics, Czech Academy of Sciences, CZ-61265 Brno, Czech Republic*

[1]*Corresponding authors:*

Tel. +420-549494003, Fax. +420-549492654, E-mail:fajkus@sci.muni.cz (J. Fajkus)

Tel. +420-549495952, Fax. +420-549492654, E-mail:hofr@sci.muni.cz (C. Hofr)

Short title: DNA binding kinetics of SMH proteins

SYNOPSIS

Proteins that bind telomeric DNA modulate the structure of chromosome ends and control telomere function and maintenance. It has been shown that *Arabidopsis thaliana* AtTRB proteins from the single-myb-histone (SMH) family selectively bind double-stranded telomeric DNA and interact with the telomeric protein AtPOT1b, which is involved in telomere capping. Here 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 analyzed 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: telomere protein-DNA interaction; single-myb-histone; fluorescence anisotropy; kinetics; surface plasmon resonance

Abbreviations used: SMH, single-myb-histone; FA, fluorescence anisotropy; EMSA, electrophoretic mobility-shift assay; SPR, surface plasmon resonance

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 a specific capping structures as e.g., 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 (telomeric repeat-binding factors), 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) 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 TRF proteins 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 TRF proteins 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 homeostasis was demonstrated [10, 11]. Interestingly, besides the TRF-like proteins, a plant-specific family of other telobox proteins was described [12]. This group of proteins, termed the Single-Myb-Histone (SMH) 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*, five SMH proteins were identified (AtTRB1-5) [12] and three of them were characterized recently [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 TRF proteins), 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 AtTRB proteins 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 domain are available [18]. Similarly, kinetic studies are limited to the interaction of a Myb-domain bearing fragment with a short telomeric DNA oligonucleotide (13bp) [18] and a non-equilibrium technique was used to describe binding kinetics of telomere repeat binding factors 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 labeled DNA may be monitored by fluorescence anisotropy (FA). 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 due 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 labeled DNA duplexes, after each addition of protein into solution, describes the extent of protein-DNA binding [20, 21].
Here we report a detailed study to reveal stoichiometry and kinetics of AtTRB1 and AtTRB3 proteins binding to telomeric DNA. Proteins AtTRB1 and AtTRB3 have been chosen for these functional assays because they showed the highest structural stability from the AtTRB family of proteins. Interactions of full-length proteins with telomeric DNA are analyzed by a combination of electrophoretic mobility shift assay (EMSA) and quantitative biophysical methods employing FA and surface plasmon resonance (SPR). Kinetic analyses at different salt conditions enable us to estimate the electrostatic component of binding and explain different affinities of the two AtTRB proteins 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. Based on 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 proteins

The cDNA sequence of AtTRB1 (locus At1g49950) was obtained by RT-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 E.coli-C41(DE3) cells [14]. The cells were grown on Luria-Bertani (LB) medium with ampicilin (100 µg/mL) at 37°C overnight. The next day cells were diluted 20 fold into ZYM 5052 complex autoinducing media containing ampicilin [22]. The cells were incubated at 37°C for 5 hours. 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 hours. At OD_{600} = 0.6, the overexpression of AtTRB3 was induced by the addition of Isopropyl-1-thio-β-D-galactopyranoside to concentration 1mM. After lowering the incubation temperature to 25°C, the growth went on for an additional 3 hours.
The following extraction and purification steps were the same for the both recombinant proteins. After harvesting by centrifugation, the pellet was dissolved in buffer containing 50 mM sodium phosphate pH 8.0 with 300 mM NaCl and 10 mM imidazole and sonicated for 5 minutes. The sonicated cell extract was cleared by centrifugation (30 000g, 1h, 4°C) and subsequent filtration (0.45 µm filter). Affinity purification was performed on a column filled with a TALON-Metal-Affinity Resin (BD Biosciences, Clontech). Protein was eluted at 80 mM imidazole concentration. The eluent was loaded on a Heparin HiTrap column (GE Healthcare). A concentration gradient of NaCl from 0.4 to 1M 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 liter of bacterial culture. The concentration of purified protein was determined by Bradford assay [23].

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 C6 aminoalkyl linker and labeled with Rhodamine Red-X (RedX) (Molecular Probes, Invitrogen) 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 extinction coefficients of the single strands were estimated with the employment of phosphate assay [24]. Extinction coefficients were 281 000 (RedX labeled strand in R4), 278 000 (complementary strand in R4), 148 000 (RedX labeled strand in R2), 140 000 (complementary strand in R2), 284 000 (RedX labeled strand in N), and 265 000 M^{-1} cm^{-1} (complementary strand in R2) for DNA oligonucleotides shown in Figure 1.

Electrophoretic mobility shift assay

Protein-DNA binding reactions were performed in 10 µL volume containing the same amount of labeled DNA duplex (30 pmol) and various concentration 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 additional 90 min at 3 V/cm. Gels were analyzed 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 spectrofluorometer (Horiba, Jobin-Yvon) 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 quartz-glass cuvette 10 x 4 mm 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 DNA solution (20nM) 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 is 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 protein binding to DNA duplexes was performed with the assumption of a non-cooperative 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 sodium chloride (Figure 4, 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 equation:

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

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 about 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 salt-dependent electrostatic part.[28, 29]. When the linear dependence of log K_a is extrapolated to the salt concentration 1 M, the electrostatic term in equation (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 ΔG_a=-2.3·RT·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 1M and the overall energy of binding is given only by the non-electrostatic term, $\Delta G_a = \Delta G_a^{nel}$ = -2.3·RT·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 Supplementary material.

RESULTS

Stoichiometry of protein-DNA complexes

In order to estimate the binding ratio of AtTRB proteins 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 this study (Figure 1).
/Figure 1 location/

Both AtTRB1 and 3 proteins bind telomeric DNA with the stoichiometry of one protein monomer per one telomeric repeat

The binding stoichiometry was analyzed by EMSA with samples containing a variable protein to 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 labeled 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 further characterize interaction stoichiometry of AtTRB proteins with telomeric DNA, 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 two-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). Based on these data we could rephrase our initial statement regarding stoichiometry to the following form: one dimer of AtTRB protein 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 analyzed 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 2a.
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 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.
/Figure 2 location/

Binding kinetics

The binding affinity of AtTRB variants to double-stranded DNA was further analyzed by FA measurements. In these measurements, protein aliquots were added to the solution of labeled 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 R4 and N are shown in Figure 3a. 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 90 nM 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. Protein 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 nM and 2900 nM, respectively. Protein AtTRB3 shows more then 7-fold higher affinity to telomeric DNA duplex than to non-telomeric DNA.

The absolute value of the dissociation constant was verified by surface plasmon resonance

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 protein 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 previously determined value with a factor of 2 at similar salt concentration (See more in Supplementary material).

AtTRB1 and AtTRB3 proteins show reduced binding affinity to R2 when compared to 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 N duplex. 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 drop in the binding affinity of both AtTRB1 and AtTRB3 proteins.

Electrostatic contribution to binding affinity

The binding of AtTRB1 or AtTRB3 to duplex R4 containing two putative binding sites induces the formation of 4 or 3 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 versus the salt concentration (Figure 4, 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, approximately four ion pairs are formed upon binding of AtTRB1 and approximately three ion pairs upon binding of AtTRB3 to the telomeric DNA.

/Figure 4 location/

The binding energy is provided mainly by a non-electrostatic component in the case of both AtTRB proteins

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 $kJmol^{-1}$ for AtTRB3, it can be concluded that the non-electrostatic interactions contribute to the total energy of binding by approximately 60% for AtTRB1 and about 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 favorable overall binding energy of AtTRB1 compared to 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 due to 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

Experiments presented here show 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 to 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 AtTRB proteins form tightly bound homo- and heterodimers and multimers [14, 15]. Relatively strong mutual interactions of AtTRB proteins were also verified independently using SPR (data not shown) and their dimerization ability was demonstrated by gel chromatography (Suplementary material). Therefore, the feasibility of protein dimerization and stoichiometric data presented here support the assumption that AtTRB proteins bind to DNA in dimeric form. In this respect, the AtTRB proteins behave similar to human TRF1 and TRF2 [30-32], with the exception that TRF proteins do not form heterodimers.

Surprisingly, the affinity of AtTRB1 to telomeric substrate R4 is four-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 here for AtTRB1 and AtTRB3 correspond very well to K_D values obtained for 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 four-fold higher affinity than TRF2.

In order to explain potential reasons for 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 here at a corresponding NaCl concentration, the magnitude of K_D for AtTRB3 is slightly lower with value 0.9 μM (see Table 2), 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 AtTRB proteins 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 the 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 alpha from *Oxytricha nova* induces the formation of 2 ion pairs upon binding to DNA and the electrostatic contribution to the free energy of binding is about 15% [25]. On the other hand, proteins containing a strongly positively charged scissor-grip motif for DNA recognition induce formation of 6 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 4 and 3 ion pairs upon AtTRB1 or AtTRB3 binding to R4 and the corresponding electrostatic contribution to the total free energy of binding is 40% and 30 %, respectively. This correlates well with data available for electrostatic interactions of other DNA binding proteins. The DNA binding event of AtTRB proteins is driven mainly by non-electrostatic interactions. On the whole, our results show that AtTRB proteins 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 AtTRB proteins with telomeric DNA.

In general, one might suppose that the same binding preferences to telomeric DNA is 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-cooperative 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 to 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 AtTRB proteins 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 about 13 bp [18]. If AtTRB proteins 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 drop 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 AtTRB proteins 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 here 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 AtTRB proteins, a speculative protein arrangement on telomeric DNA could be considered (Figure 5).

/Figure 5 location/

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 drop of binding activity when the DNA substrate is 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 less 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 heteromultimers 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, as e.g. t-loops, in a similar way to human TRF proteins [31, 32].

Thus, on the basis of results presented here and the available data, we suggest that interactions of the two AtTRB proteins 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 AtTRB proteins. 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 considerable lack of general knowledge on 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 behavior of protein would not be markedly changed in the cellular environment. The access of AtTRB proteins to their telomeric target sites is restricted in both spatial and temporal way 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].

AtTRB proteins might be first recruited by a weak non-specific binding to multiple chromosome regions. Then, once the specific target sites become accessible, high 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, AtTRB proteins 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 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 are to be determined, these results demonstrate the advantage of the approach employed in this work 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. This research paper demonstrates that the detailed quantification of protein-DNA interactions may provide new insights into the structural dynamics of telomeres.

ACKNOWLEDGEMENTS

We are grateful for the critical reading of manuscript by M. Chester. This work was supported by the Grant Agency of the Czech Republic (521/08/P452 and 204/08/H054), Czech Ministry of Education (LC06004), Grant Agency of the Czech Academy of Sciences (IAA500040801) and the institutional support (MSM0021622415, MSM0021622413, AV0Z50040702 and AV0Z50040507).

REFERENCES

1 Zakian, V. A. (1995) Telomeres: beginning to understand the end. Science **270**, 1601-7

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**, 503-14

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**, 271-9

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**, 1659-68

6 Baumann, P. and Cech, T. R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. Science **292**, 1171-5

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**, 2667-78

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**, 1294-303

9 Kuchar, M. (2006) Plant telomere-binding proteins. Biologia Plantarum **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**, 3370-85

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**, 1333-42

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**, 1336-50

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**, 316-24

14 Mozgova, I., Prochazkova Schrumpfova, P., Hofr, C., Fajkus, J. (2008) Functional characterisation of domains in AtTRB1, a putative telomere-binding protein in Arabidopsis thaliana. Phytochemistry **69**, 1814-9

15 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**, 1400-6

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**, 311-5

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**, 7725-33

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**, 1744-8

21 LeTilly, V. and Royer, C. A. (1993) Fluorescence anisotropy assays implicate protein-protein interactions in regulating trp repressor DNA binding. Biochemistry **32**, 7753-8

22 Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. **41**, 207-34

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**, 248-54

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**, 273-82

25 Buczek, P. and Horvath, M. P. (2006) Thermodynamic characterization of binding Oxytricha nova single strand telomere DNA with the alpha protein N-terminal domain. J. Mol. Biol. **359**, 1217-34

26 Kuzmic, P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. Anal. Biochem. **237**, 260-73

27 Olmsted, M. C., Bond, J. P., Anderson, C. F. and Record, M. T., Jr. (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**, 634-47

28 Record, M. T., Jr., 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**, 5192-7

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**, 1785-94

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**, 351-61

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**, 119-30

34 Revzin, A. (1990) The Biology of Nonspecific DNA-protein Interactions, CRC Press, Boston

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**, 163-9

37 Fajkus, J. and Trifonov, E. N. (2001) Columnar packing of telomeric nucleosomes. Biochem. Biophys. Res. Commun. **280**, 961-3

Table 1 Dissociation and association constants for binding of AtTRB1 and AtTRB3 to DNA*

	R4		N		R2	
	K_D†	K_a‡	K_D†	K_a‡	K_D†	K_a‡
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

*Means ± standard errors of three independent experiments in 50 mM Sodium Phosphate pH 7.5, 100 mM NaCl measured at 25°C; † K_D in 10^{-9} M; ‡ K_a in 10^6 M^{-1}.

Table 2 Salt concentration dependence of association constants for binding of AtTRB1 and AtTRB3 to R4

	[NaCl] (mM)	log K_a	$\delta\log K_a/\delta\log[NaCl]^*$	log K_{nel} *	Z
AtTRB1	100	7.04			
	119	7.08			
	141	6.67	2.8 ± 0.2	4.31 ± 0.2	4.4
	167	6.54			
	200	6.25			
AtTRB3	100	6.41			
	119	6.31			
	141	6.06	2.0 ± 0.1	4.4 ± 0.1	3.2
	167	6.00			
	200	5.81			

*Mean values are presented with the corresponding standard errors.

FIGURE LEGENDS

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 position 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].

Figure 2 Non-denaturing electromobility shift assay
(a) AtTRB1 binding to fluorescently labeled 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 AtTRB proteins were incubated with increasing amounts of protein. The numbers under electrophoretic lanes denote protein/DNA stoichiometric ratio. The protein/DNA ratio corresponding to binding saturation is indicated with a gray line.

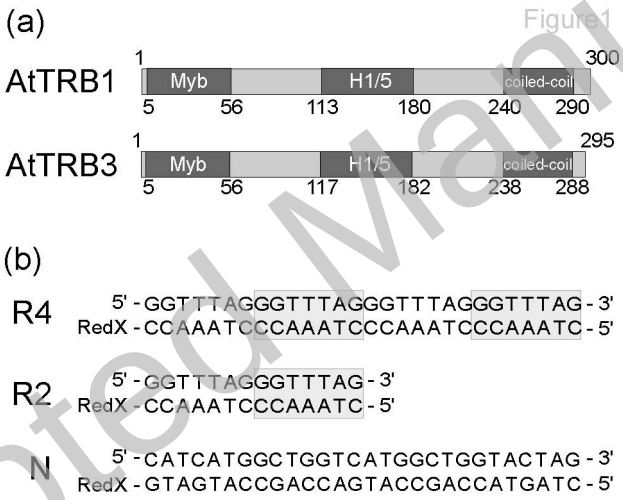
Figure 3 Binding of AtTRB1 and AtTRB3 to DNA duplexes
(a) Fluorescence anisotropy measurements of binding of AtTRB1 to telomeric duplex R4 or non-telomeric duplex N. The binding at 20 nM DNA occurred in buffer containing 50 mM sodium phosphate pH 7.5 and 100 mM NaCl (b) Fluorescence anisotropy measurements of binding of AtTRB3 to R4 or N duplex. Binding conditions are the same as in part a. (c) Binding isotherms of AtTRB1 and AtTRB3 with telomeric duplex R2 measured by fluorescence anisotropy. Binding conditions are the same as in part a.

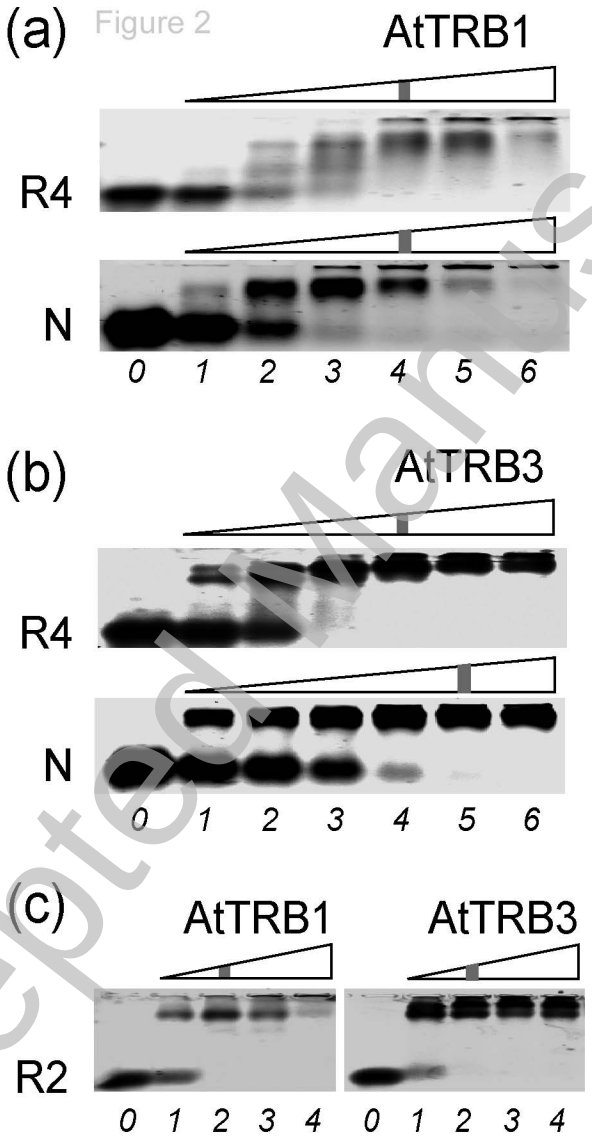
Figure 4 Dependence of the association constants for binding of AtTRB1 and AtTRB3 to substrate R4 on NaCl concentration
Inset: The electrostatic and non-electrostatic components of the free energy of association of AtTRB1 or AtTRB3 with substrate R4.

Figure 5 Speculative model of interaction of AtTRB proteins with telomeric DNA
Both homo- and heterodimers of AtTRB may participate in the interaction with telomeric DNA.

Biochemical Journal Immediate Publication. Published on 22 Dec 2008 as manuscript BJ20082195

Figure1





Biochemical Journal Immediate Publication. Published on 22 Dec 2008 as manuscript BJ20082195

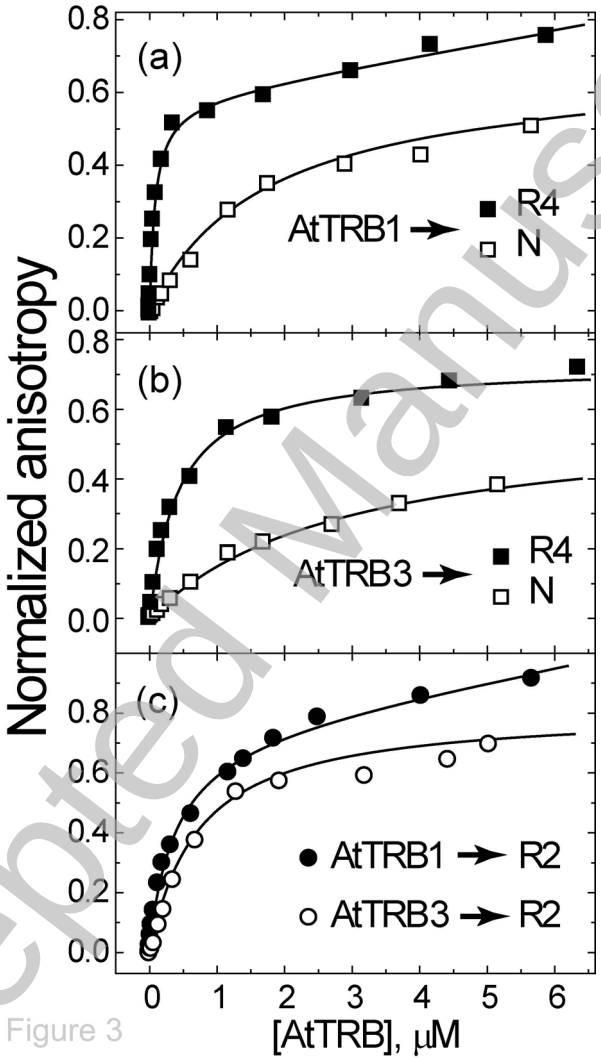


Figure 3

Biochemical Journal Immediate Publication. Published on 22 Dec 2008 as manuscript BJ20082195

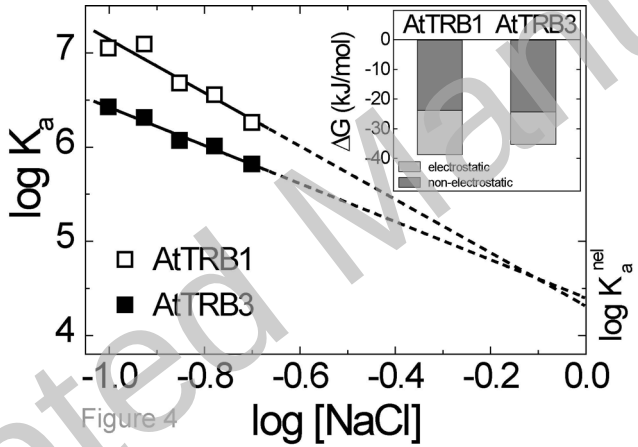


Figure 4

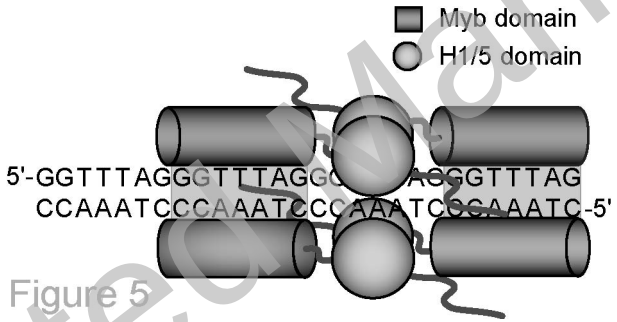


Figure 5