Biophysical Chemistry 251 (2019) 106177

Contents lists available at ScienceDirect

Biophysical Chemistry

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

Quantifying length-dependent DNA end-binding by nucleoproteins

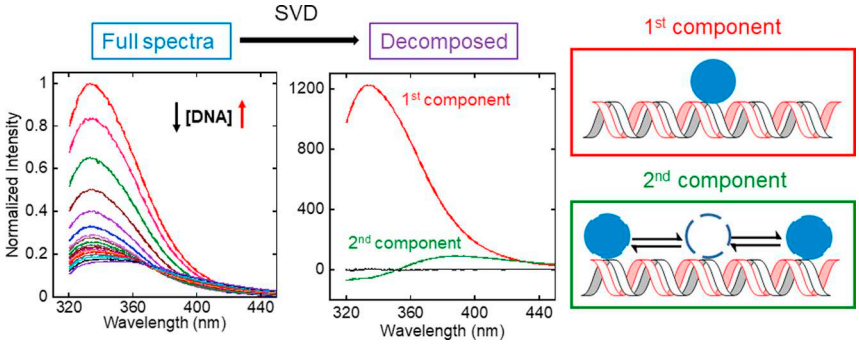
Tam Vo[a], Amanda V. Albrecht[a], W. David Wilson[a,b,*], Gregory M.K. Poon[a,b,*]

[a] Department of Chemistry, Georgia State University, Atlanta, GA 30303, United States of America
[b] Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, GA 30303, United States of America

HIGHLIGHTS

- Synthetic oligonucleotides are ubiquitously used to study ligand-nucleic acids interactions.
- Ends and other structural discontinuities such as hairpins present opportunistic binding sites.
- We dissect the fluorescence signatures of end and interior DNA binding by the model protein ETV6.
- End-binding is a non-electrostatic binding mode realized by translocation from interior sites.

GRAPHICAL ABSTRACT



ABSTRACT

The ends of nucleic acids oligomers alter the statistics of interior nonspecific ligand binding and act as binding sites with altered properties. While the former aspect of "end effects" has received much theoretical attention in the literature, the physical nature of end-binding, and hence its potential impact on a wide range of studies with oligomers, remains poorly known. Here, we report for the first time end-binding to DNA using a model helix-turn-helix motif, the DNA-binding domain of ETV6, as a function of DNA sequence length. Spectral analysis of ETV6 intrinsic tryptophan fluorescence by singular value decomposition showed that end-binding to nonspecific fragments was negligible at > 0.2 kbp and accumulated to 8% of total binding to 23-bp oligomers. The affinity for end-binding was insensitive to salt but tracked the affinity of interior binding, suggesting translocation from interior sites rather than free solution as its mechanism. As the presence of a cognate site in the 23-bp oligomer suppressed end-binding, neglect of end-binding to the short cognate DNA does not introduce significant error. However, the same applies to nonspecific DNA only if longer fragments (> 0.2 kbp) are used.

Sequence-specific nucleoproteins such as transcription factors function in a context of excess nonspecific nucleic acids. Nonspecific nucleic acids binding represents one application of a general class of ligand-lattice interactions, which include the well-known Ising model. The most influential formulation of nonspecific binding to nucleic acids, the McGhee-von Hippel (MvH) model, was formulated for an infinite lattice in which boundary conditions, or "end effects" do not exist [1,2]. The question of what lattice length is "effectively infinite" was taken up in Epstein's explicit treatment of binding to a finite-length lattice [3]. At a given level of lattice saturation, convergence with the infinite lattice approximation depends on the intrinsic binding affinity, cooperativity of binding, and the ligand site size [3,4].

Beyond modifying the statistics of interior binding, there is no general agreement on how end-binding is physically manifest. Epstein's model does not intrinsically differentiate end- from interior binding (same affinity, same site size requirement). A more recent treatment by Winzor and coworkers permit the end-binding site size to vary [5], while a modified MvH equation by Record and coworkers excludes end-binding in which the ligand overhangs the ends of the lattice [6]. The latter choice was motivated in part by binding studies of the model

* Corresponding authors at: P.O. Box 3965, Atlanta, GA 30302-3965, United States of America.
E-mail addresses: wdw@gsu.edu (W.D. Wilson), gpoon@gsu.edu (G.M.K. Poon).

https://doi.org/10.1016/j.bpc.2019.106177
Received 22 March 2019; Received in revised form 26 April 2019; Accepted 27 April 2019
Available online 30 April 2019
0301-4622/

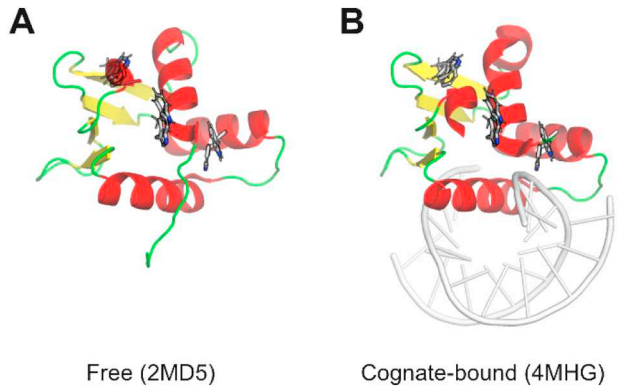


Fig. 1. Experimental structures of the ETS domain of murine ETV6. In the free state (A) and bound to cognate duplex DNA (B). Shown are residues 331 to 426, corresponding to the construct used in this study. The three Trp residues (338, 356, and 376) are shown as sticks.

oligocationic peptide KWK_6NH_2 to single-stranded dT oligomers in which electrostatic effects promote binding to interior over terminal sites [6,7].

The lack of agreement on end-binding reflects a dearth of empirical data on the nature of end binding. Several proteins, such as the DNA repair protein Ku, polynucleotide kinases/phosphatases and telomere-processing proteins derive their function *via* interactions with DNA ends [8–13]. Much more pervasively, end-binding represents a feature of empirical work on ligand binding to short oligomeric DNA that is unavoidable yet is seldom addressed. To what extent do proteins *actually* bind DNA ends? What impact do interior sites have on end-binding? What about other discontinuous structures, such as hairpins? To address these questions, we examined the binding modes of a representative DNA-binding domain, the ETS domain of ETV6. ETS proteins comprise a major family of eukaryotic transcription factors [14] and are structurally related to the ubiquitous helix-turn-helix motif [15]. ETV6 is also a good model because its structures in both free and DNA-bound states have been solved [16]. Both cognate and nonspecific binding are known to be mediated by the same contact surface [16], and three spatially separated Trp residues in the DNA-binding domain facilitate fluorescence detection of different binding modes [Fig. 1].

While quench of intrinsic Trp fluorescence has been used to monitor DNA binding, typical analyses at single wavelengths preclude unambiguous identification of multiple binding modes. To access these binding modes, singular value decomposition (SVD) offers a model-independent approach to dissecting complete spectra into orthogonal features. The decomposed data provide detailed, unbiased (in the least-square sense) information on the spectral properties of each binding mode, their contribution to overall binding, as well as their concentration dependence which may then, if desired, be fitted with specific models.

We titrated highly purified ETS domain ETV6 [Fig. S1A, Supporting Information] with 10 nM to 10 μM bp of nonspecific (NS) duplex DNA ranging from 23 bp to 0.9 kbp in length (see Methods in Supporting Information). The 23 bp duplex was annealed from oligonucleotides that contain no consensus sequence (5′-GGAA-3′) for the protein. Mixed-base salmon sperm DNA was sonicated to average lengths of 0.06, 0.2, 0.3, 0.5, and 0.9 kbp as sized by agarose electrophoresis (Fig. S1B). In agreement with previous observations [17], titration of ETV6

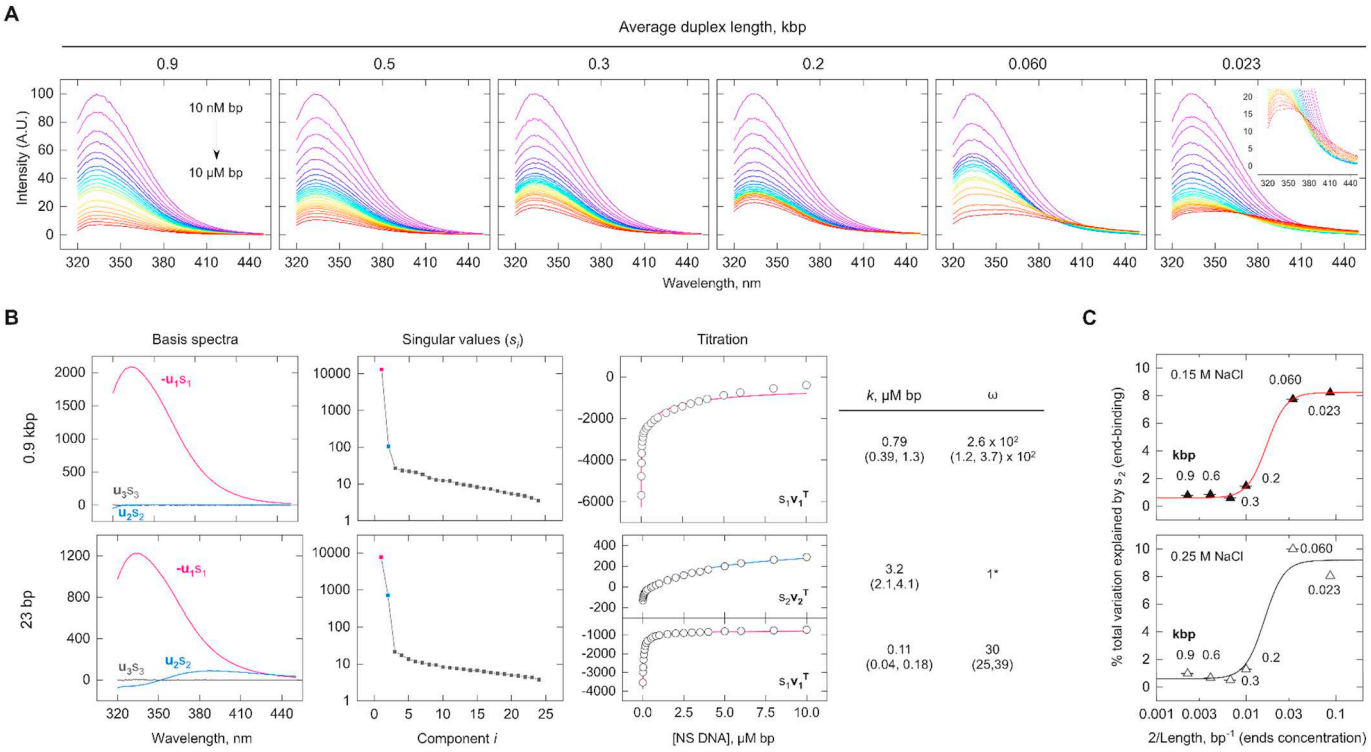


Fig. 2. DNA end-binding by the ETS domain of the transcription factor ETV6. A, Blank-subtracted emission spectra of ETV6 with increasing nonspecific DNA sizes. B, Representative SVD of the 0.9 kbp and 23 bp titrations. *Left*, the first three u vectors, scaled by their corresponding singular values to show their significant contribution to the observed spectra. *Middle*, Logarithmic "Scree plot" of the singular values. The emergence in significance of end-binding (component 2) for short DNA is reflected in the relative rise of s_2 (blue). *Right*, titration of interior (red) and end-binding (blue), fitted to the MvH equation (0.9 kbp) or Record's extension for finite-length lattices [6]. Parametric values are given with the 95% C.I. in parenthesis. The asterisk denotes fixing of the cooperativity parameter due to no statistical improvement when left floating (by Fisher's *F*-test for the sum of squares). C, Contribution of end-binding to total binding, taken as the fractional magnitude of s_2 over the sum of all singular values in the set, in 0.15 and 0.25 M total Na^+. Twice the reciprocal length is taken as a proxy for ends concentration. A Hill-type curve was fitted to suggest a correlation. The error bars are propagated from the standard deviations of the length distributions of the sheared salmon sperm DNA (c.f. Fig. S1B, Supporting Information). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

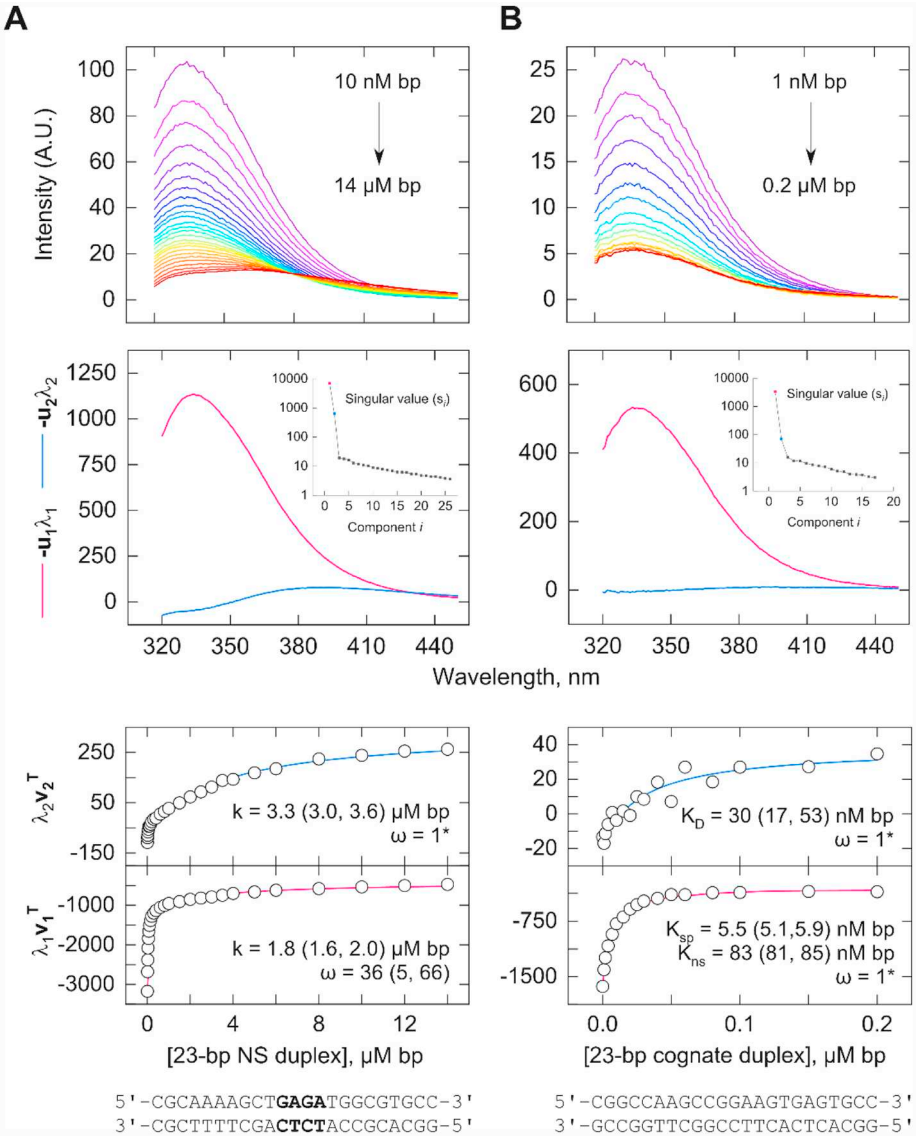


Fig. 3. Probing the electrostatic contribution and presence of a cognate site to DNA end-binding. SVD for the titration of ETV6 with a 23 bp oligomer harboring fully nonspecific DNA (A) or an ETS consensus site 5′-GGAA-3′ (B) in 0.25 M total Na^+. The initial ETV6 concentrations were 0.2 μM for the NS DNA titration, and 0.05 μM for the cognate DNA titration. Scaled basis vectors of the first two significant components are shown. The titrations were fitted with Record's extensions of the MvH equation [6]. The cooperative parameter was fixed at unity (asterisks) as described in Fig. 2. Parametric values are given with the 95% C.I. in parenthesis.

with DNA quenched the intrinsic Trp fluorescence of the protein. The peak at ~330 nm decreased monotonically in step with DNA addition. At < 0.3 kbp, the spectra revealed a secondary feature at higher DNA concentrations (> 1 μM), becoming more prominent with decreasing sequence length [Fig. 2A]. The feature exhibited an isosbestic point at ~370 nm, but only with spectra acquired at similarly high DNA concentrations.

Since DNA binding sites *increased* in abundance relative to protein over the course of the titrations, the secondary feature could not represent higher-order binding or aggregation of protein on the DNA. It also could not be due to supramolecular binding, as NMR studies had established only a single DNA contact surface on ETV6 [16], and the feature was more prominent at lengths below the persistence length of B-form DNA at 0.15 M NaCl (> 200 bp) [18]. Moreover, the secondary feature represented *de novo* emission at significantly red-shifted wavelengths, and therefore could not arise from a lack of binding. The possibility of contaminants in the protein preparation were excluded because the secondary feature was sensitive to the presence of an interior cognate site (*vide infra*). The evidence therefore pointed to the origin of the primary quench in interior binding, while the secondary emission represented binding at or near the ends.

To quantify the extent of end-binding, we applied SVD to extract orthogonal features from the observed spectra in decreasing order of significance (*i.e.*, magnitudes of the singular values). Analysis of binding to the 0.9 kbp polymer yielded the primary quench from interior binding as the only significant feature [Fig. 2B]. The second component contributed negligibly at 0.8%. Identical distributions between interior and end-binding were observed with decreasing average DNA length down to 0.3 kbp. Below this length, the relative contribution of the secondary feature corresponding to end-binding started to increase [Fig. 2C]. At 23 bp, the secondary feature accounted for 8.2% of the total variation in the data. As shown for the 0.9 kbp and 23 bp titrations, model fitting revealed positively cooperative binding to interior sites, suggesting favorable high-density loading (as observed with the O^6-alkylguanine-DNA transferase) [19], with dissociation constants in the 10^{-7} M range. In contrast, end-binding to the 23 bp oligomer was weaker by ~10 fold and non-cooperative, consistent with the disconnected ends in the oligomer. Thus, end-binding to *nonspecific* DNA

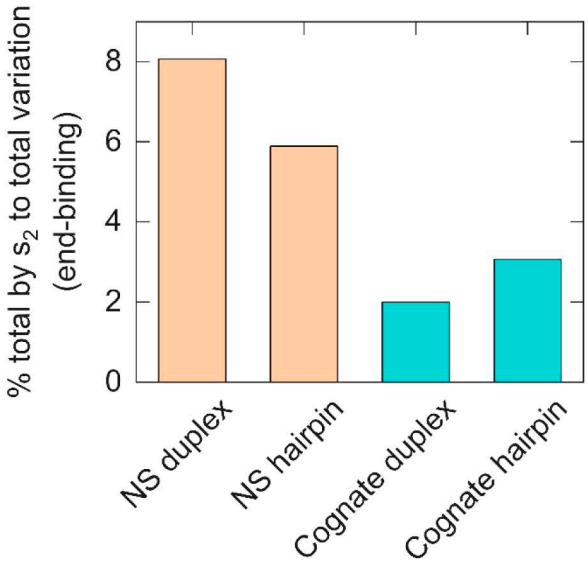


Fig. 4. Contribution of end-binding to 23 bp nonspecific and cognate oligomers as duplexes and hairpins. The duplexes present two ends for binding while the hairpins present one end and one hairpin.

by ETV6 was negligible down to 0.2 kbp and constituted ~8% of the total bound population for a 23 bp oligomer (Fig. 2C).

To probe the driving force of end-binding, we evaluated binding in 0.25 M total Na^+. Relative to 0.15 M Na^+, no significant difference in the distribution of the spectra components was detected (Fig. 2C). The titrations showed little change in end-binding affinity between the two salt conditions, while interior binding was ~10-fold weaker at 0.25 M Na^+ [Fig. 3A]. Since counter-ion densities on DNA backbone phosphates diminish towards the ends, this result implied that the end-bound protein interacted within < 3 to 5 bp from the termini, where the electrostatic potential drops significantly relative to interior sequences [20,21] and would contribute negligible electrostatically to binding.

Next, we investigated the impact of an interior cognate site on end-binding relative to nonspecific DNA. We compared binding of ETV6 to the 23 bp oligomer with another duplex harboring an ETS consensus (5′-GGAA-3′) of the same length. Since cognate binding occurs with ~100-fold higher affinity, the titrations were started with 50 nM protein at 0.25 M Na^+. End-binding to the 23 bp cognate duplex comprised only 2% of total binding, compared with 8% for its nonspecific counterpart, indicating a competition between interior and terminal binding by the protein [Fig. 3B]. This observation also excluded the possibility that the secondary binding was due to unforeseen contaminants which, if present, exhibited specific DNA preferences exactly as expected for *bona fide* ETV6.

Remarkably, while the apparent affinity for end-binding on the cognate oligomer was ~10-fold weaker than for interior binding, it was much stronger than on the nonspecific oligomer (Fig. 3B). Since the terminal segments for both oligomers were nearly identical, end-binding appeared to be coupled to interior binding. If association from free solution were significant, the apparent end-binding affinity should not depend on interior sequences, contrary to the strong dependence of end-binding on interior site identity. This observation therefore implied that association at the ends occurred primarily from short-range diffusion from the sequence interior. Mechanistically, we propose that the end-bound state begins with a pre-equilibrium at interior DNA sites with subsequent diffusion to the termini. An interior pre-equilibrium would account for the observed concentration dependence in end-binding of short oligomers, which harbor a limited reservoir of interior-bound complexes relative to polymeric DNA. Such a model would also imply that the end-reaching kinetics by linear diffusion are comparable or fast relative to dissociation from the DNA.

Finally, we wished to understand whether binding at ends was like hairpins, which are common substitutions for duplexes. We tested hairpin analogs of the same 23 bp duplexes harboring fully nonspecific DNA or a single cognate site (hairpin at the right end of sequences shown in Fig. 3). At 0.25 M Na^+, the data showed minor variations in secondary binding regardless the presence of cognate site: ± 2% of total binding for fully nonspecific and ± 1% in the existence of a single cognate site [Fig. 4]. Thus, ends and hairpins represented similar targets for ETV6 with respect to binding free energy, and the nature of structural discontinuity was not significant in impact relative to the presence of a cognate interior site.

In conclusion, we have defined the physical parameters of end-binding by the DNA-binding domain of ETV6, a representative helix-turn-helix motif variant among DNA-binding proteins. A plateau end-binding level of ~8% was observed in fully nonspecific 23-bp DNA but was suppressed to ~2% in the presence of a single cognate site. Moreover, end-binding was non-electrostatically mediated and realized *via* diffusion from interior sites. The red-shift in the emission associated with end-binding suggested loss of folded structure involving the three Trp residues [22]. For proteins with similar characteristics as ETV6 interaction with short cognate DNA, covering a single helical turn, explicit treatment of end-binding is not needed. Models that exclude binding at ends [6] are therefore appropriate so long as additional discontinuities are absent in the duplex. Parenthetically, end-binding to the 0.06-kbp sheared salmon sperm DNA was elevated relative to longer fragments despite similar statistical likelihoods of harboring fortuitous ETS consensus, thus underscoring the primacy of DNA length in determining end-binding of nonspecific DNA.

Acknowledgement

T.V. was supported by a GSU Molecular Basis of Disease Fellowship. This investigation was funded by NSF grant MCB 15451600 to G.M.K.P and NIH grant GM111749 to W.D.W.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bpc.2019.106177.

References

[1] J.D. McGhee, P.H. von Hippel, Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice, J. Mol. Biol. 86 (1974) 469–489.
[2] E. Di Cera, Y. Kong, Theory of multivalent binding in one and two-dimensional lattices, Biophys. Chem. 61 (1996) 107–124.
[3] I.R. Epstein, Cooperative and non-cooperative binding of large ligands to a finite one-dimensional lattice. A model for ligand-oligonucleotide interactions, Biophys. Chem. 8 (1978) 327–339.
[4] S.C. Kowalczykowski, L.S. Paul, N. Lonberg, J.W. Newport, J.A. McSwiggen, P.H. von Hippel, Cooperative and noncooperative binding of protein ligands to nucleic acid lattices: experimental approaches to the determination of thermodynamic parameters, Biochemistry 25 (1986) 1226–1240.
[5] P.D. Munro, C.M. Jackson, D.J. Winzor, Consequences of the non-specific binding of a protein to a linear polymer: reconciliation of stoichiometric and equilibrium titration data for the thrombin-heparin interaction, J. Theor. Biol. 203 (2000) 407–418.
[6] O.V. Tsodikov, J.A. Holbrook, I.A. Shkel, M.T. Record Jr., Analytic binding isotherms describing competitive interactions of a protein ligand with specific and nonspecific sites on the same DNA oligomer, Biophys. J. 81 (2001) 1960–1969.
[7] W. Zhang, H. Ni, M.W. Capp, C.F. Anderson, T.M. Lohman, M.T. Record Jr., The importance of coulombic end effects: experimental characterization of the effects of oligonucleotide flanking charges on the strength and salt dependence of oligocation (L8+) binding to single-stranded DNA oligomers, Biophys. J. 76 (1999) 1008–1017.
[8] M. Ono, P.W. Tucker, J.D. Capra, Ku is a general inhibitor of DNA-protein complex formation and transcription, Mol. Immunol. 33 (1996) 787–796.
[9] K.K. Khanna, S.P. Jackson, DNA double-strand breaks: signaling, repair and the cancer connection, Nat. Genet. 27 (2001) 247–254.
[10] M. O'Driscoll, Diseases associated with defective responses to DNA damage, Cold

Spring Harb. Perspect. Biol. 4 (2012) a012773.
[11] G.B. Morin, The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats, Cell 59 (1989) 521–529.
[12] C.W. Greider, E.H. Blackburn, A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis, Nature 337 (1989) 331–337.
[13] Z. Arslan, R. Wurm, O. Brener, P. Ellinger, L. Nagel-Steger, F. Oesterhelt, L. Schmitt, D. Willbold, R. Wagner, H. Gohlke, S.H. Smits, U. Pul, Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2, Nucleic Acids Res. 41 (2013) 6347–6359.
[14] J.M. Vaquerizas, S.K. Kummerfeld, S.A. Teichmann, N.M. Luscombe, A census of human transcription factors: function, expression and evolution, Nat. Rev. Genet. 10 (2009) 252–263.
[15] S.C. Harrison, A.K. Aggarwal, DNA recognition by proteins with the helix-turn-helix motif, Annu. Rev. Biochem. 59 (1990) 933–969.
[16] S. De, A.C. Chan, H.J. Coyne 3rd, N. Bhachech, U. Hermsdorf, M. Okon, M.E. Murphy, B.J. Graves, L.P. McIntosh, Steric mechanism of auto-inhibitory regulation of specific and non-specific DNA binding by the ETS transcriptional repressor ETV6, J. Mol. Biol. 426 (2014) 1390–1406.
[17] T. Vo, S. Wang, G.M.K. Poon, W.D. Wilson, Electrostatic control of DNA intersegmental translocation by the ETS transcription factor ETV6, J. Biol. Chem. 292 (2017) 13187–13196.
[18] H.B. Gray, V.A. Bloomfield, J.E. Hearst, Sedimentation coefficients of linear and cyclic wormlike coils with excluded-volume effects, J. Chem. Phys. 46 (1967) 1493–1498.
[19] M. Melikishvili, J.J. Rasimas, A.E. Pegg, M.G. Fried, Interactions of human O(6)-alkylguanine-DNA alkyltransferase (AGT) with short double-stranded DNAs, Biochemistry 47 (2008) 13754–13763.
[20] G.S. Manning, Approximate solutions to some problems in polyelectrolyte theory involving nonuniform charge distributions, Macromolecules 41 (2008) 6217–6227.
[21] S.A. Allison, End effects in electrostatic potentials of cylinders: models for DNA fragments, J. Phys. Chem. 98 (1994) 12091–12096.
[22] C.A. Royer, Probing protein folding and conformational transitions with fluorescence, Chem. Rev. 106 (2006) 1769–1784.