Salt Dependent Association of Novel Mutants of TATA-Binding Proteins to DNA: Predictions from Theory and Experiments

Johan H. Bredenberg[1] and Marcia O. Fenley[1,2,*]

[1] *Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA.*

[2] *Department of Physics, Florida State University, Tallahassee, FL 32306, USA.*

Abstract. The nonlinear Poisson-Boltzmann predictions of the salt-dependent association of proteins to DNA, SK_{pred}, are fairly insensitive to the choice of atomic charges, radii, interior dielectric constant and treatment of the boundary between a biomolecule and the solvent. In this study we show that the SK_{pred} is highly correlated with the conformational adaptability of the partners involved in the biomolecular binding process. This is demonstrated for the wild-type and mutant forms of the archaeon *Pyrococcus woesi* TATA-binding protein (*Pw*TBP) in complex with DNA, on which we performed molecular mechanics energy minimizations with different protocols, and molecular dynamics simulations and then computed the SK_{pred} on the resulting structures. It was found that the inter-molecular non bonded force field energy between the DNA and protein correlates linearly and significantly well with the SK_{pred}. This correlation encompasses the wild-type and mutant variants of the *Pw*TBP and provides us with a quick way to estimate the SK_{pred} from a large ensemble of structures generated with Molecular Dynamics or Monte Carlo simulations. The corresponding experimental SK_{obs} should also correlate with the inter-molecular non bonded force field energy between the protein and DNA, given that the underlying mechanisms in binding and salt-dependent effects are in fact the main contributors in the association of proteins/peptides to nucleic acids. We show that it is possible to fit experiments versus the inter-molecular non bonded force field energy between the protein and DNA, and use this relation to predict the SK_{obs} in absolute numbers. Thus, we present two novel approaches to estimate both the SK_{pred} and the SK_{obs} for *in silico* modelled *Pw*TBP novel mutants and even for TBPs from other organisms. This is a simple but powerful tool to suggest new experiments on the TBP-DNA type of macromolecular assemblies. We conclude by suggesting some mutants and a possible biological interpretation of how changes in solvent salinity affect the binding of proteins to DNA.

Corresponding author. Email addresses: `jbreden@sb.fsu.edu` (J. H. Bredenberg), `mfenley@sb.fsu.edu` (M. O. Fenley)

---

1 Introduction

TATA-binding proteins (TBP) [1] are involved in the primary transcription machinery, where they recognize and bind to variant combinations of thymine (T) and adenine (A) clustered in an eight base-pair DNA stretch, which is called the TATA-element [2]. Because of the seminal role in the transcription process [3], the TBP is present in a broad range of organisms that have adapted to live in different – and some cases extreme – environments [4–6]. Yet, all of the presently known TBPs are highly homologous and similar in the folded state, as revealed from their three-dimensional crystal structures [7–9]. They have many features in common, *e.g.* two conserved Phe residues whose side-chains intercalate the DNA and cause it to kink about 80 degrees, and an overall α-β-α-β fold. However, they differ in the overall charge distribution, total net charge and hydrophobic packing, which most likely is due to the adaptation to the environment in which the organism lives. For instance, the archaeon *Pyrococcus woesi* (*Pw*) is an organism that has adapted to a life in highly saline environments [10], where it grows optimally at non-physiological elevated temperatures (around the boiling point of water) and is therefore categorized as a hyperthermophilic halophilic organism [4]. The *Pw*TBP differ substantially in charge distribution and overall net charge compared to the mesophilic human *Homo sapiens* (*Hs*) and the yeast *Saccharomyces cerevisiae* (*Sc*) TBPs. The latter two are highly positively charged ($+16e$ and $+12e$, respectively) while the *Pw*TBP has an overall net charge of zero with 25 positively and 25 negatively charged protein residues distributed over the entire protein (Fig. 1). Some of these residues are located at the DNA-binding interface or within 6 Å to any nearest DNA atom (Fig. 1). Oddly, a large portion of these interfacial residues are acidic in character, *i.e.* negatively charged. Thus, one would expect these residues having an unfavorable repulsive effect in DNA binding since they are negatively charged, as are the phosphate groups in the DNA backbone.

Recently, we reported an exhaustive comparison between the *Hs*-, *Sc*-, and the *Pw*-TBP bound to DNA using a combined Molecular Mechanics/Poisson-Boltzmann (MM/PB) computational approach (Bredenberg, Russo and Fenley – Biophysical J., in press). This was done in order to investigate salt-mediated association effects when TBP binds to DNA. In particular, we focused on the *Pw*TBP and some of its mutants that have thermodynamic experimental data reported in the literature ([11] and references therein). Our results were qualitatively in agreement with thermodynamic isothermal titration calorimetry (ITC) experiments which measure the binding constant, K_{obs}, for the formation of a complex (*i.e.* when TBP binds to DNA) at different salt concentrations. The logarithmic of the K_{obs} is then plotted against the logarithmic of the salt concentration

A)

$^+$H$_3$N-MVDMSKVKLRIENIVASVDLFAQLDLEKVLDLCPNSKYNPEEFPGIIC

HLDDPKVALLIFSSGKLVVTGAKSVQDIERAVAKLAQKLKSIGVKFKRAPQI

DVQNMVFSGDIGREFNLDVVALTLPNCEYEPEQFPGVIYRVKEPKSVILLFS

SGKIVCSGAKSEADAWEAVKLLRELDKY-COO$^-$

B)
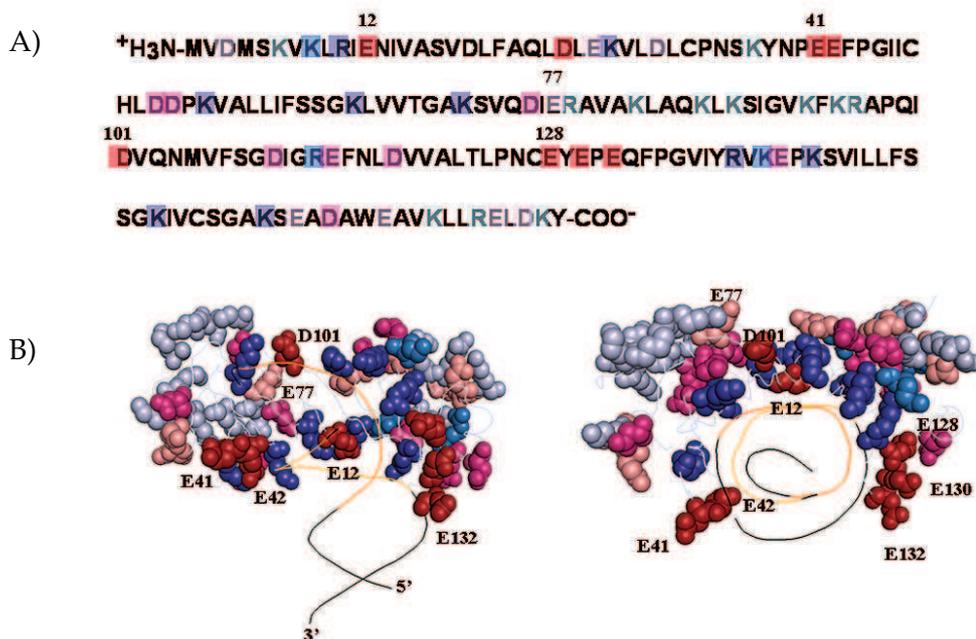
Figure 1: Color coded A) primary and B) tertiary sequence for the *Pyrococcus woesi* TATA-binding protein in complex with DNA (the specific DNA TATA-box is shown in orange). The protein backbone is traced in light-blue ribbons. All positively charged amino acids Lysine (K) and Arginine (R) are colored in blue and negatively charged amino acids Aspartate (D) and Glutamate (E) are colored in red. Darkest color means residues within 6 Å to the DNA, medium color represents residues at distances 6 to 12 Å from the DNA and the lightest coloring represents residues located > 12 Å from the DNA. In B) two views of the TBP-TATA complex is shown; a side view and the same orientation rotated 90 degrees in the z-plane. For reference, residues E12, E41, E42, E77, D101 E128, 130 and 132 are labeled. The graphical representation of the structure was generated with the PYMOL program (http://pymol.sourceforge.net/).

[c]. The slope, which is the salt dependence in binding of TBP to DNA, of this curve is $SK_{obs} = d\log K_{obs}/d\log[c]$ [12]. Assuming that electrostatics is the main driving force for the salt dependent TBP-DNA complexation and using Gibb's Law to substitute - ΔG_{elec} for $kT \cdot \ln K_{obs}$, one can compute the SK_{obs}, which we from hereon shall call $SK_{pred} = -d\Delta G_{elec}/2.303 \cdot kT \cdot d\log[c]$.

Our interpretation of SK_{pred} (and SK_{obs}) is that if $SK_{pred} < 0$, the binding between protein and DNA weakens with increasing salt concentration, and if $SK_{pred} > 0$, the binding becomes stronger with increased salt concentration.

Interestingly, we noted a clear correlation between the minimum distance between DNA phosphate-oxygens (O_p) to the acidic Glutamate (E) and Aspartate (D) side chain carboxylate-oxygens, O_δ or O_ε, respectively (for short this distance will be called R_{O-O}) in the wild-type PwTBP and the SK_{pred} (and SK_{obs}) for the mutation of the same E/D residue to an alanine (Fig. 2). However, while experiments indicate that an Alanine mutation of residues not involved in the DNA binding but still residing rather close to the DNA (See Fig. 1), renders the SK_{obs} unaltered, we got different results (Bredenberg, Russo
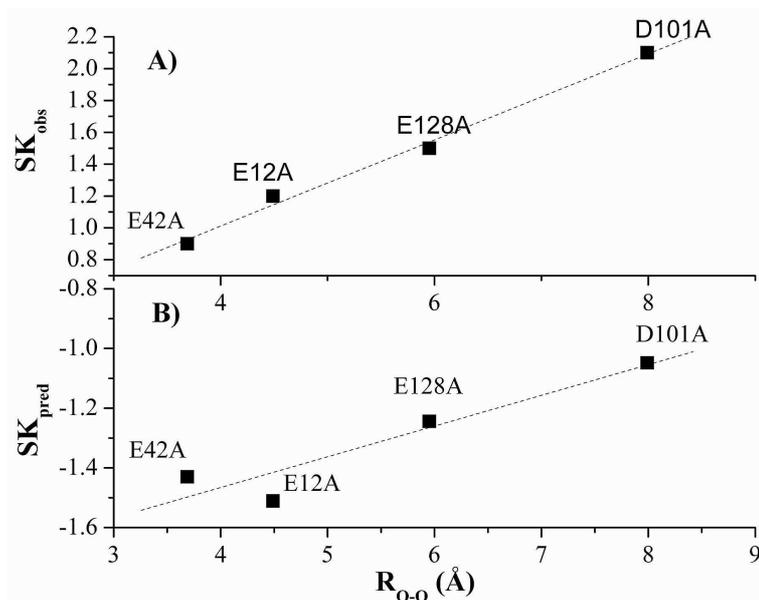
Figure 2: Linear fit of SK versus the minimum distance between the Asp/Glu side chain carboxylate-oxygen, O_δ or O_ε, to the DNA phosphate-oxygen, O_p, (R_{O-O}). All minimum distances are computed from the minimized wild-type PwTBP. SK_{obs} is obtained from references [11,17] and SK_{pred} is computed with the non-linear solution to the Poisson-Boltzmann Equation. Only PwTBP mutants that had experimental SK_{obs} are used in the fitting.

and Fenley – Biophysical J., in press). This was observed for the PwTBP D101A mutant. Furthermore, experiments reported that this mutant has no impact whatsoever on the SK_{obs} (although the error bar for this mutant is orders of magnitude larger than for the other mutants) [11].

From a theoretical point of view, any residue interacts with its environment; changing the charge of a particular (charged acidic) residue by neutralizing or reversing its charge, will change the SK_{pred}. When the R_{O-O} increases in the wild-type, the impact of the SK_{pred} from the mutant will be less and approach the SK_{pred} for the wild-type. This reasoning should hold for the experimental SK_{obs} too, given that the folding and DNA-bound state of the mutant protein is probably similar to that of the wild-type. We therefore decided to look further into the linear relationship between the mutant SK_{pred} for a certain E/D residue and the corresponding R_{O-O} for the same residue in the wild-type (Fig. 2).

An important feature in this paper is that the predictive power to a significant extent depends on the molecular mechanics protocol that is used for structure optimization prior to the electrostatic Poisson-Boltzmann (PB) computations. This optimization relies on an empirical potential energy function [13] that consists of two parts; interactions involving atoms separated with 2 or 3 bonds in sequence that are excluded from any Coulomb or van der Waals (vdW) energy. These interactions are called bonded terms. For any other atoms, the so-called non-bonded interaction energy prevails, and is composed of Coulomb and vdW energies. The latter energy term is short ranged and usually
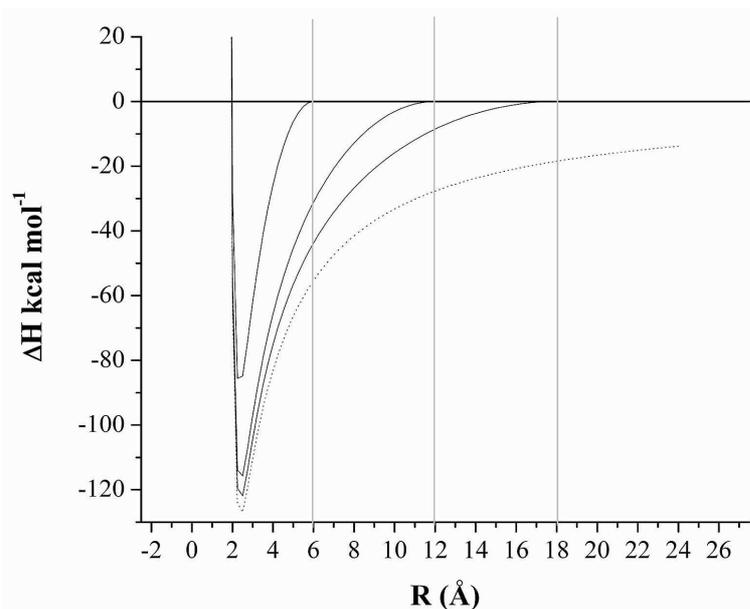
Figure 3: Non-bonded energy between two oppositely unit-charged particles as a function of inter-molecular distance (R). This energy is represented by Lennard-Jones 6-12 interaction energies and Coulombic energies; the L-J term has an attractive part that scales as R^{-6} and a repulsive part that scales as R^{-12}; the Coulomb energy scales as R^{-1} and follows a modified Coulomb's law with a shifting function. The cut-off limit is indicated with gray bars (where the energy is zero) and the "infinite" cutoff is shown by the dashed line.

modelled as Lennard-Jones 6-12 attractive/repulsive interactions and thus scales as r^{-6} and r^{-12} for the attractive and repulsive terms, respectively. The long range Coulombic interaction scale as $1/r$ and thus affects the molecule largely.

For minimization/molecular dynamics purposes, usually a 10-12-14 Å cut-off is applied for the Coulomb energy between atom pairs. This means that any atom within 10 Å from each other interacts according to, or with some clever modification in Coulomb's law, such that the entire charge-charge interaction energy is changed to zero between 10 and 12 Å. Finally, there is a look-up table that keeps track of atoms within 12 and 14 Å from each other, but do not effectively account for their interactions in the force field. This listing is updated either at fixed intervals or with some atomic displacement criteria. However, as illustrated in Fig. 3, the energy from the long range Coulombic term is significant for particles separated by distances greater than 12 Å, and thus a great portion of the Coulomb energy between pairs of atoms is neglected beyond the cut-off limit.

If one takes into account biological macromolecules that contains several thousands of atoms, and are highly charged (*e.g.* nucleic acids), there will a substantial contribution from the Coulomb potential. In many cases, modifications of the structure must be done prior to the computational act. Such modifications are for example: removal of co-binding proteins, water molecules, ions or modelling of one or more protein side chains to construct either missing residues or a mutant variant from the coordinates of the wild-

type structure, the addition of hydrogen atoms (if the structure was solved from a crystal that diffracted above 1 Å in the highest resolution shell). Since the interaction network within the solute matrix (*e.g.* a protein in complex with DNA) is multidimensional with respect to internal salt-bridges, inter- and intra-molecular hydrogen bonds and residue packing, then propagation or local rearrangements of amino acids is bound to occur upon such modifications of the initial coordinates.

Of course, if the mutant and wild-type structure of the protein already is known, and no other modifications need to be done, one can do theoretical analysis on the structure without the need of pre-mature modelling. This is usually not the case though, so in order to reduce the computational cost when optimizing structures that have been altered in some way, the molecular mechanics protocol use the approach of cut-offs in non-bonded energy during the minimizations (or molecular dynamics simulations). However, when solving the PBE, all atoms are included in the calculation: the cut-off is "infinite". Atoms that previously were oblivious to each other because they were beyond the distance cut-off in Coulomb energy during the molecular mechanics computation will suddenly experience the potential from other atoms that previously was outside the specified interaction range. Therefore, the calculation of the SK_{pred} (and other theoretical issues that are appropriately addressed with PB electrostatics approaches) is sensitive to this conformational adaptability of the solute (peptide/protein/nucleic acid). This conformational adaptability is the result from the minimized structure, which will look different depending on the minimization protocol and the type of force field used and the extent of modifications introduced in the original structure.

Steinbach and Brooks reported a solid comparison between different algorithms and cut-offs for handling the cut-off in non-bonded energies for use in molecular mechanics calculations on biomolecular structures and the structural impact for different methods/cut-offs [14]. We use only the default algorithm as implemented in the CHARMM program [15] and compare the wild-type and mutant *Pw*TBP structures obtained from minimizations performed in explicit solvent with varying cut-offs in the electrostatic non-bonded interaction energies and forces. We also carried out molecular dynamics (MD) simulations on these structures in explicit solvent (with the default settings of 10-12-14 Å). Finally, we carried out minimizations on these structures with the Generalized Born Molecular Volume (GBMVII) method [16]. This is an implicit solvent formulation which accounts for the free energy of solvation for the solute. For these minimizations, an "infinite" non-bonded cut-off was used, meaning that all atom pairs in the system were included in the Coulomb energy. All resulting structures from any protocol of energy-minimizations or molecular dynamics simulations were then subjected to PBE calculations at various salt concentrations in order to calculate the SK_{pred}.

We show that an extended cut-off in the Coulomb energy increases the precision in the estimate of the SK_{pred} and that a 4-6-8 Å scheme causes the linearity of the SK_{pred} versus the R_{O-O} to vanish. With this in mind, our findings demonstrate the importance of including preferentially all solute atoms in the structure optimization, and that even minor perturbations in the structure of the solute will claim its toll in the solution of the

PBE. Assuming that the linearity of the SK_{pred} versus the minimum distance holds we should be able to estimate the SK_{pred} for a specific R_{O-O}.

While the linearity of the change in SK_{pred} can be viewed in terms of distances, it can also be regarded as changes in the non bonded Coulomb and van der Waals inter-molecular energies between the protein and DNA. Thus, these energies would govern the SK_{pred} and be correlated with the force field energy: A change in the non bonded energy would lead to a change in the SK_{pred}. We show that this is indeed the case and that this information can be used for predicting the SK_{pred} from a large ensemble of structures and thus significantly reduce the computational burden for calculating the SK_{pred}.

If the three-dimensional structure of a binary complex is known and if the experimental data for the salt-dependence in binding are available, we found that the Coulomb and van der Waals energies computed from the force field are highly correlated with the experimental SK_{obs} for the wild-type (wt) and mutant variants of the *Pw*TBP [11,17]. This simple hybrid between experiments and theory, SK_{hyb}, predicted reasonably well the absolute numbers for the salt-dependence in binding of DNA to known single and multiple *Pw*TBP mutants. We also found that the experimental SK_{obs} for the *Sc*TBP [18] was well predicted from this relationship. In principle, one should be able to estimate the actual SK_{obs} (give or take) for any other single or multiple *Pw*TBP mutant or any other TBP from other organisms in complex with DNA. That is – if the overall fold of the TBPs are similar for different species. Our findings suggest that the salt-dependence in binding for different TBPs to DNA in general obeys the same physical law rather than binding to the DNA with different underlying mechanisms.

In this work, we present two novel and unique approaches in computing the SK_{pred} and SK_{obs}. To the best of our knowledge, none of these two methods have been reported anywhere in the literature before:

1. First we demonstrate how the SK_{pred} can be linearly fitted versus the force field inter-molecular protein-DNA energies from a set of structures and how this can be applied to estimate the SK_{pred} from a large ensemble of conformations sampled with molecular dynamics or Monte Carlo simulations. Thus, we can tie the information of conformational changes in the protein-DNA complex with the behavior of the salt-dependence in binding from a microscopic point of view.

2. We then show that the experimental SK_{obs} is highly linearly correlated with the inter-molecular force field energy between the protein and DNA in several variant *Pw*TBP bound to the same DNA. Thus, we can combine experiments and theory and use this as a guide to predict – in absolute numbers – what the experimental SK_{obs} would be. There will be some underlying system-dependence, but this dependence is internal: If the same model-system is used for the theoretical part, then the experimental conditions do not have to be similar to the theoretical model, given that the experimental results are accurate. Thus, we can use a macroscopic property and apply this property on a single structure in order to predict what the expected macroscopic property – in this case the SK_{obs} for the *Pw*TBP-DNA system – would be, and use this information to suggest new experiments.

2 Methods

The exact details of the protocol can be found elsewhere, but for easy reference a brief description of the protocols used in this study is given below.

2.1 Molecular mechanics

The CHARMM program [15] and the CHARMM 22/27 [19, 20] all atom topologies and force-field parameters was used in all modelling, minimizations and molecular dynamics simulations.

The initial coordinates for the archaeon PwTBP were retrieved from the RCSB Protein Data Bank (PDB code: 1AIS). The co-binding Transcription Factor IIB and crystal oxygen atoms located further away than 6 Å from the TBP/DNA-complex (in short: solute) were removed. Two 5-iodo-uracil bases were replaced by thymine and hydrogen atoms were added with HBUILD [21]. We used the recommended default CHARMM-potential for handling the electrostatic and van der Waals interaction energies and forces. While keeping all non-hydrogen atoms fixed, the hydrogen atoms were minimized with 200 steps of Steepest Descent (SD) followed by Adopted Basis Newton Raphson (ABNR) until the default convergence criteria of 0.0000 in the root mean square (rms) gradient of the potential energy was achieved. The solute was centered in a TIP3P [22] water-sphere of radius 47 Å. While keeping the solute fixed, all water molecules were subjected to 10 ps Langevin Dynamics at 300 K under the influence of stochastic boundary force that maintained the spherical shape of the solvent [23]. Thirty-six Na^+ and four Cl^- ions were then added to the solute-solvent system to render overall charge neutrality of the PwTBP-DNA system, and the solvent (*i.e.* water and added ions) was subjected to 20 ps of Stochastic Boundary Molecular Dynamics (SBMD) [24], where atoms within 44 Å from the sphere center were treated as Newton particles and the remaining atoms as Langevin particles. The final set of equilibrated solvent and initial solute coordinates served as the starting set for modelling mutant PwTBPs. We chose the following mutants that have the experimental SK_{obs} reported in the literature: E12A, E42A, D101A, E128A and we also chose one mutant that is located far from the binding interface (*i.e.*, that has an R_{O-O} of about 19.5 Å) – E77A. This mutant has – to the best of our knowledge – no available SK_{obs} data.

The mutants were constructed by remodelling the side chain atoms (if any) from the CHARMM-topology (*i.e.* in an extended conformation). For each mutation, this means that the target residue side chain was either truncated at the C_β-atom and had one hydrogen atom added – for substitution to an alanine – or missing atoms were built with the CHARMM22/27 topology files in conjunction with existing Cartesian coordinates. If the PwTBP mutant(s) resulted in non-zero overall net charge, the corresponding number of ion(s) – sodium or chloride – in the solvent were transformed into a TIP3P water oxygen (and adding its two hydrogen atoms) to reset charge neutrality. In cases were water molecules overlapped the extended side-chain (*i.e.* when new atoms were added), the water molecule(s) were removed. The mutated side-chain was then minimized while

keeping all other atoms fixed.

The resulting structures (*Pw*TBP and mutant variants) served as initial coordinates for the following three protocols: 1. Minimization in explicit solvent using 6, 12 and 18 Å for the truncation of all electrostatic energies and forces that were smoothly shifted to zero beyond these cut-offs [14]. 2. Molecular dynamics simulations using standard settings in the cut-off (the default of 12 Å). 3. Minimization with the Generalized Born Molecular Volume (II) implicit solvent model and setting the cut-off to "infinity", 99999 Å, for electrostatic energies and forces. For any of the protocols, SHAKE [25] was used to constrain all explicit bonds between hydrogen atoms and non-hydrogen atoms.

2.2 Minimization in explicit solvent

The solvent was fixed, and the solute was minimized for 5,000 steps of SD followed by ABNR until the default in function tolerance (0.0000) was reached and the minimization terminated. Then, a second round of minimization was carried out, allowing *all* atoms to move, and the system was minimized with 5,000 steps of SD, followed by ABNR until the default convergence criteria was reached.

2.3 Molecular Dynamics simulations in explicit solvent

Ten simulations in the NVT ensemble were carried out on each system by means of SBMD. The simulation details are the same as for the equilibration of ions in the preparation of the original system, as described previously, except that all atoms are allowed to move and that the initial coordinate set for the solute was not minimized. A seed for assigning Gaussian-distributed velocities for each atom, was randomly picked in each simulation, and the system was equilibrated for 10 ps at 298.15 K, followed by 20 ps of production run, and then quenched for 10 ps, targeting a temperature of zero K. Typically, the final quenched temperature in each simulation was about 75-85 K. The time-step was 2 femto-seconds and a non-bonded list was updated with a heuristic test, such that the list was reconstructed every time any atom had moved 1 Å since the last update. Altogether 400 ps were sampled for each system, but only the final structure from each simulation was selected for subsequent calculations using the non-linear Poisson-Boltzmann Equation (NLPBE).

2.4 GBMVII

All solvent atoms were stripped off, leaving the protein and DNA naked in space. The cut-off in all non-bonded electrostatic energies and forces were set to "infinity", *i.e.* 99999 Å. The parameters for the GBMVII were set as recommended in the literature [16] and all atoms were allowed to move instantly and the system was minimized with 5,000 steps of SD, followed by ABNR until the default convergence criteria in function tolerance was reached (0.0000). For this minimization an iterative tolerance threshold of 500 was set in order to prevent numerical underflow and ensure that the convergence was fully reached.

2.5 Poisson-Boltzmann calculations

The energy minimized/quenched MD solute coordinates (protein and DNA) were transformed into the Protein-Charge-Radii (PQR) format [26] with an in-house written program for use in the NLPBE calculation. For any atom, the charge and radii was adapted from the CHARMM 22/27 parameter files [19, 20]. All calculations were carried out at neutral pH (7.0) by assuming that Histidine residues were neutral and that all other charged residues were in their standard protonation states at room temperature (298 K) which is to say that each of the Glutamate and Aspartate residues carry an overall net charge of $-1e$ and that each of the Arginine and Lysine residues carry an overall net charge of $+1e$. We varied the 1:1 salt (NaCl) concentration from 0.1 to 0.4 M. The solute and solvent dielectric constants were set to $\varepsilon_{in}=2$ and $\varepsilon_{out}=80$, respectively. The solvent excluded molecular surface [27], which was based on a water probe radius of 1.4 Å, was used to define the dielectric interface that separates the solute and solvent regions. No ion exclusion region was considered. The total extent of the 3D grid was set to three times the largest dimension of the molecule in order to reduce outer boundary condition errors. Special outer boundary and energy corrections were enforced in order to obtain very accurate salt-dependent electrostatic free energies [28]. We choose a finest grid spacing of 0.3 Å for all PBE calculations for a compromise between accuracy and speed. All other default PBE code parameters were employed.

ΔG_{elec} was computed as the difference between the electrostatic free energy of the complex and the electrostatic free energy of the individual binding partners in their docked state at a fixed 1:1 salt concentration. The calculation of ΔG_{elec} involved three separate PBE calculations of G_{elec}: One for the complex, one for the isolated protein and one for the isolated DNA:

$$\Delta G_{elec} = G_{elec}(\text{complex}) - G_{elec}(\text{protein}) - G_{elec}(\text{DNA}). \tag{2.1}$$

We computed ΔG_{elec} at eight salt concentrations ranging from 0.1 M to 0.4 M, for a total 24 of PBE computations for each complex. If we assume that the long range electrostatic interaction energy predominates in the salt dependence of the TBP-DNA binding, we can approximate ΔG as ΔG_{elec} and use the relation $\Delta G = -kT \cdot \ln K$ (K is the equilibrium binding constant) to get the predicted salt-dependence in association, SK_{pred}:

$$SK_{pred} = -\frac{1}{2.3kT} \times \frac{d\Delta G_{elec}}{d\log[c]}. \tag{2.2}$$

SK_{pred} is the salt derivative of the electrostatic binding free energy, ΔG_{elec}, which consist of four energy terms [29]: the Coulomb, the reaction field, the dielectric stress, and the osmotic pressure energy (given in units of kT). k is the Boltzmann constant, T is the absolute temperature (here taken as 298K) and $[c]$ is the 1:1 salt concentration (*i.e.* NaCl). SK_{pred} is obtained as the slope of a least square linear fit of $-\Delta G_{elec}$ versus the logarithmic of salt concentration $[c]$. For more details about the different Poisson-Boltzmann energy terms the reader is referred to elsewhere [28, 29].

Once the SK_{pred} for mutant PwTBPs was computed (from structures generated with different protocols), it was plotted against the minimum distance between the corresponding wild-type side-chain carboxylate oxygen (O_δ for Asp and O_ε for Glu) and the nearest DNA phosphate oxygen (O_p). The same particular protocol used for computing a certain SK_{pred} was used in the minimum distance computation. For short this distance is called $R_{O\text{-}O}$.

With experiments, one can use the equilibrium binding constant, K_{obs}, to compute the corresponding salt-dependence in binding [12]:

$$SK_{obs} = \frac{d \log K_{obs}}{d \log [c]}. \qquad (2.3)$$

K_{obs} can be measured from thermodynamic experiments, *e.g.* Isothermal Titration Calorimetry (ITC), at various salt concentrations. The slope, SK_{obs}, can then be obtained by a linear fit of the logarithmic of K_{obs} versus the logarithmic of salt concentration $[c]$.

3   Results and discussion

First, we show that a charge-neutralization (*i.e.* mutation to an Ala) of an acidic Asp or Glu residue in a protein (or peptide in general) that binds to DNA (or nucleic acids in general) exhibit a linear behavior with respect to the minimum distance between the carboxylate-group of the amino-acid and the DNA phosphate-oxygen ($R_{O\text{-}O}$). This relates to the SK_{pred} of the alanine-mutant residue of that certain Asp or Glu residue. We demonstrate that there will be a substantial impact of the computed PBE result depending explicitly on the non bonded cut-off settings in the molecular mechanics (MM) protocol.

Second, we show that the SK_{pred} is highly correlated with the non bonded Coulomb and van der Waals inter-molecular interaction energies between the PwTBP and the DNA regardless if the protein is the wild-type or a mutant form. This relation can be used to extract SK_{pred} from ensembles of structures generated with molecular dynamics simulations or Monte Carlo methods.

Third, we show that there is an intrinsic and high degree of correlation between the experimental SK_{obs} and the inter-molecular force field energies between the DNA and the wild-type, single, or multiple mutant PwTBPs. This correlation can be used to estimate an expected SK_{obs} for other TBPs and novel PwTBP mutants. We call this mix of experimental and theoretical methods the SK_{hyb}.

Fourth, we address some biological interpretations of the PwTBP-DNA salt dependence in binding, and suggest some novel mutants that may lead to further interesting experiments that may be conducted on this specific system.

SK_{pred} from minimizations in explicit solvent with various non-bonded cut-offs and minimizations in implicit solvent at no-cutoff

Fig. 4 shows the extrapolation of a linear fit for the SK_{pred} of mutant residues E12A, E42A, D101A and E128A versus the R_{O-O} for the very same residues (E12, E42, D101 and E128) in the wild-type *Pw*TBP. These four mutants and the wild-type *Pw*TBP all have experimental SK_{obs} reported in the literature [11, 17]. Then, the SK_{pred} for the remotely positioned (some 19.5 Å) E77A was computed and plotted (but not included in the fit). Any of these SK_{pred} and distances was determined from the resulting coordinates used within the same particular protocol for consistency (see methods section). It is clear that the fitting of the 4-6-8 Å protocol do not predict the SK_{pred} for the E77A mutant in the graph (Fig. 4A). However, the wild-type and the E77A mutant have very similar SK_{pred} which should be expected (Table 1). In fact, the ΔSK_{pred} between them is zero for the 4-6-8 Å protocol, but since this protocol also predicts the wrong order of salt-dependence in binding (the coefficient is slightly negative), this is most likely an artifact because of the short cut-off in non-bonded force field energies. When the cut-off is increased to 10-12-14 Å, the linearity becomes more prominent and E77A appears to respond to this and fall reasonable within the linearity of the extrapolation in the graph (Fig. 4B) which now is in reasonable agreement with the computed SK_{pred} for the E77A mutant (Table 1). The extension of the cut-off to 16-18-20 improves this relation further (Fig. 4C) and the GBMVII approach reproduces the trend as well (Fig. 4D).

Three striking trends are observed in Table 1: 1. The SK_{pred} generally become less negative upon increasing the cut-off (which essentially means including more atoms-pairs in the Coulomb energy) until it for the GBMV case is greater than zero (Table 1). 2. The SK_{pred} increases with the distance and should in principle be very close to the wild-type at the most remotely positioned (with respect to R_{O-O} in the wild-type) acidic residue. 3. The ΔSK_{pred} for the wild-type and the mutant increases when the cut-off in non-bonded force field energy is increased, but it appears that the 16-18-20 cut-off is similar to the GBMV protocol implying that the upper limit for ΔSK_{pred} between the wild-type and E77A is about 0.3. This enhances the notion that there will always be a difference in the SK_{pred} if a charged residue is neutralized (or charge-reversed for that matter) relative the wild-type, regardless the location of that residue. This reasoning is well in line with a study of the distance and charge-dependence of the SK_{pred} that we recently reported in a study of the salt-mediated effects in the binding of aminoacyl tRNA-synthetases to their cognate tRNAs (Bredenberg, Boschitsch and Fenley – in this special Poisson-Boltzmann issue of the Commun. Comput. Phys.).

The fact that the minimization with the GBMVII resulted in $SK_{pred} > 0$ for the *Pw*TBP wild-type and the *Pw*TBP E77A mutant in this work (Table 1, Fig. 4D) implies that this mutation exhibit a similar SK_{pred} relative the wild-type, which should be expected since the R_{O-O} for this residue is about 19.5 Å. It appears that the linearity between SK_{pred} and R_{O-O} is improved when the non-bonded cut-off in the force field is extended (Fig. 4), or when all atoms are included in the calculations, as is the case for the GBMVII minimiza-
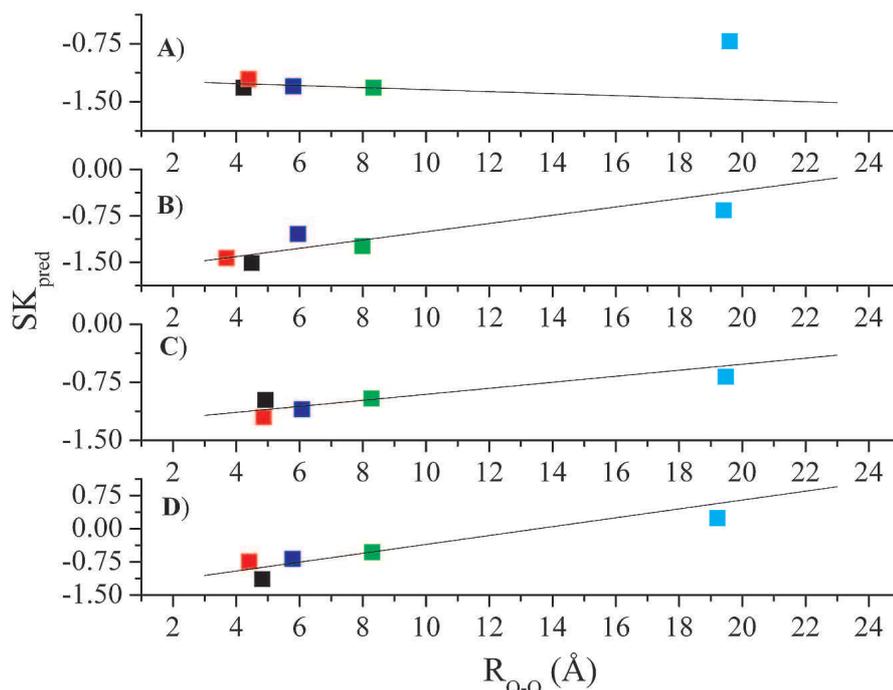
Figure 4: Linear fit of mutant SK_{pred} versus the minimum distance between the PwTBP wild-type Asp/Glu side chain carboxylate-oxygen (O_δ or O_ε) to the DNA phosphate-oxygen (O_p) as described in Fig. 2. Only, this time different protocols have been employed for the energy-minimization. The fitting and interpolation is done on the distances obtained for the wild-type structure using a particular protocol versus the corresponding SK_{pred} for mutant residues E12A (black square), E42A (red), D101A (blue) and E128A (green). These mutants have available experimental SK_{obs} [11, 17]. The light blue square is the SK_{pred} for residue E77A (which was not included in the fitting). For minimization in explicit solvent: A) 6 Å cut-off, B) 12 Å cut-off (default) and C) 18 Å cut-off. Minimization with the Generalized Born Molecular Volume II implicit model without any cut-off in non bonded Coulomb energy is shown in D). The SK_{pred} is computed with the non-linear solution to the Poisson-Boltzmann Equation.

tion protocol.

Also, since GBMV is a "simplified" approach towards the more rigorous Poisson-Boltzmann solution to the electrostatics in solvated molecules, it already circumvents some issues related to the choice of atomic parameters such as charges and radii. Our results in this work points to that one should include all atoms in the potential for geometry optimizations of structures that will be subjected to subsequent PB calculations. Such approach was also successfully reported recently by Honig and co-workers [30]. They used the Generalized Born Solvent Accessible Surface Area (GB/SA) [31] approach in the minimization of several protein-protein complexes in order to compute protein-protein binding energies with the NLPBE [30].

However, when using the GBMV [16] method, another frame of reference is used, such that no explicit solvent is present in the minimization procedure together with the

Table 1: Computed SK_{pred} on structures obtained with minimizations in explicit solvent at different cut-offs[a] in Coulomb energy, Molecular Dynamics at default cut-off (12 Å), and minimization with the Generalized Born Molecular Volume II [16] implicit solvent model at 99999 Å cut-off. The experimental SK_{obs} is provided for comparison.

	SK (kcal mol^{-2})					SK_{obs}
	SK_{pred}					
Method:	6	12	18	MD	GBMVII [16]	Exp [11, 17]
Protein[b]						
Wt	-0.73	-0.56	-0.33	-0.60 ± 0.22	0.53	2.1 ± 0.1
E12A	-1.32	-1.51	-0.98	-1.52 ± 0.11	-1.14	1.3 ± 0.3
E42A	-1.21	-1.43	-1.20	-0.91 ± 0.07	-0.74	0.9 ± 0.8
D101A	-1.32	-1.24	-0.96	-0.94 ± 0.32	-0.53	2.1 ± 1.1
E128A	-1.30	-1.04	-1.10	-1.02 ± 0.22	-0.68	1.6 ± 0.3
E77A	-0.72	-0.66	-0.68	-0.52 ± 0.14	0.24	n/a

[a] The boundary at which the interaction between two atoms becomes zero.
[b] Wild-type archaeon *Pyrococcus woesi* TATA-binding protein (*Pw*TBP, PDB code 1AIS) in complex with DNA. The mutants are expressed in single-letter abbreviations for negatively charged amino-acids Glutamate (E) and Aspartate (D) mutated to the neutral residue Alanine (A).

inclusion of all pair-wise Coulomb energies between the solute atoms. The latter will cause deviation from the original CHARMM-potential which has the default non-bonded interaction parameters set to 12 Å and also has based the force field development on these settings (see the CHARMM documentation and parameter-files). In addition to this, the CHARMM parameters were designed for molecular mechanics/dynamics in explicit solvent and thus parameterized against the TIP3P [22] water model. Still, for minimization purposes regarding CPU-hungry systems, the GBMV method offers a very reasonable alternative instead of solvating the biomolecule in explicit water.

Molecular dynamics simulations imply that the SK_{pred} is sensitive to the conformational adaptability

Since all structures were minimized from the same initial coordinates and since only one single-point calculation was carried out, we cannot completely depict the relation between the SK_{pred} and structural reformation from minimizations alone. Further, no error-bars can be obtained on one single structure (other than the dependence of grid-spacing and other parameters or settings in the PBE protocol). Therefore, we ran a series of MD simulations on each system, hunting for statistical quarry. We observed that the trend for E42A is similar to what is seen for all other protocols (except for the 20 Å cut-off protocol). This is probably because of the default non-bonded settings in the MD-protocol (Table 1) and may also be due to the limited simulation time.
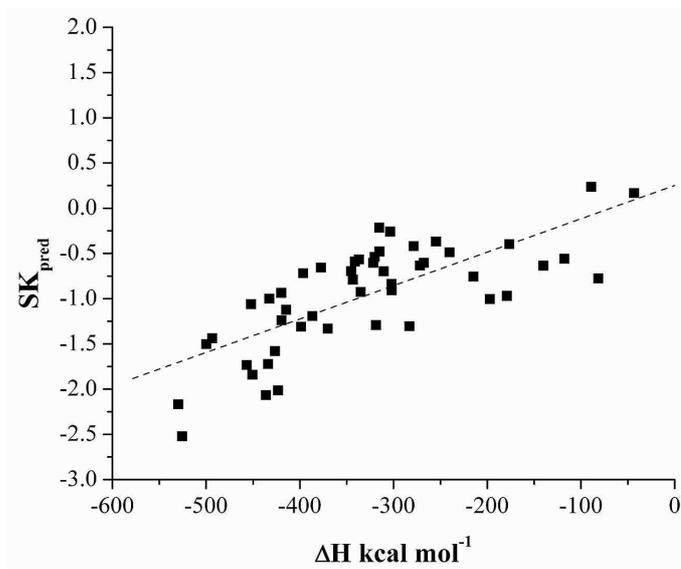
Figure 5: SK_{pred} versus the non bonded interaction energy between TBP and DNA (ΔH). The ΔH term consist of Lennard-Jones 6-12 (van der Waals) energies and Coulomb energies. Each ensemble has ten structures which are the final structure of ten independent Molecular Dynamics simulations using the default 10-12-14 Å non-bonded cut-off. All mutants indicated in Table 1 and the wild-type PwTBP are included in the plot. The correlation between the SK_{pred} and the force field enthalpy is 0.7943 and shown by the dotted line.

The potential felt by the DNA depends only on the number and distribution of protein charges, and thus the salt-dependence will be fairly linear with the interaction energy between the protein and the DNA in complex. In other words: the DNA does not care if the bound protein is altered, but will only respond to changes in the Coulomb and vdW interaction potential. This allows us to estimate a simple linear relationship between the two bound molecules and thus reduces the computational burden when computing the SK_{pred} for a large set of structures that have been generated from molecular simulations (*i.e.* from molecular dynamics or Monte Carlo). The only requirement is a series of computed SK_{pred} done on a structure bound to the very same DNA for consistency.

It is clear that the non-bonded force-field energies for the ensemble of structures generated from MD simulations correlates with the SK_{pred} (Fig. 5). Notice that this Coulomb energy is truncated at 12 Å during the molecular dynamics simulations and thus the structures are adapted accordingly to these circumstances, but that the single-point energy is computed including any of the solute atoms (*i.e.* no cut-off in Coulomb energy between any pairs of atoms) and setting the dielectric constant, ε, to 2 instead of 1. The reason for changing the ε value is to be consistent with the PBE setting, which has $\varepsilon_{in} = 2$.

We do observe low energies in the non-bonded force-field energies between TBP-DNA that has a weaker correlation with the corresponding SK_{pred}, but for *any one* of these outliers, at least one data-point was corrupt (and removed) or the variation in linearity between the data-points for the slope was larger. Nonetheless the relation between

conformational adaptability and the changes in SK_{pred} is significant, simple and linear, which provides us with a quick way to assess the SK_{pred} from a generated trajectory (of the same system) that may contain several thousands of configurations. Processing such a vast amount of configurations by means of PBE analysis is not very computationally efficient. For example, a "complete" SK_{pred} calculation with the NLPBE for one structure that takes some 15-30 minutes (which is fast and significantly efficient), would require 6000-12000 hours for a trajectory of 10 nano-seconds in length and some 24000 structures. Thus, the approach of computing just a few single-point SK_{pred} from a set of structures (preferentially independently generated or else arbitrarily picked from a production phase molecular dynamics simulation) with the NLPBE and then use the linear fitting of these SK_{pred} and the corresponding non bonded inter-molecular energy between the protein and DNA reduces the computational time substantially (about $\sim$1000 times faster). This approach opens up several possibilities to learn more about the actual biology behind the salt-dependence in binding for proteins/peptides bound to nucleic acids, and provides alternatives to connect SK_{pred} with structural properties in the complex.

Combination of experimental SK_{obs} and inter-molecular protein-DNA molecular mechanics energies provides a hybrid prediction of the salt-dependence in binding, SK_{hyb}

Inspired by the good correlation between the SK_{pred} versus the non-bonded energy obtained in the MD simulations (Fig. 5) that we described in the previous section, we examined the corresponding relation between the experimental SK_{obs} and the non-bonded energy for the minimized wild-type and mutant *Pw*TBP structures that we recently reported (Bredenberg, Russo, Fenley -Biophys. J, in press). Note that in this case we compute the energy for one single and minimized structure of each representative of the variant *Pw*TBPs (Fig. 6). We found that the Coulomb and Lennard-Jones energies between the protein and the DNA correlate well with the experimental SK_{obs} (Fig. 6). In this fitting, we computed the full non-bonded inter-molecular energy between the protein and DNA using a dielectric constant, $\varepsilon = 2$, just as we did for the SK_{pred} versus the non bonded force field energy in the previous section (using the standard $\varepsilon = 1$ would not change anything else than an offset of the data-points). The correlation between experimental SK_{obs} and non bonded interaction energies between the *Pw*TBP and DNA is good even for the cases of multiple mutants (Fig. 6), and thus provides us a direct way to predict what the experimental SK_{obs} would be for some selected mutants. We shall call this the SK_{hyb}, since it is calculated from a linear fit between experiments and molecular mechanics energies.

Interestingly, the interpolation even predicts the SK_{hyb} for TBPs from the mesophilic *Sc*TBP too (Fig. 6), being very close to the experimental SK_{obs}, which is -4.6. The latter have a comparable number of DNA base-pairs in their two crystal structures as do the *Pw*TBP, and we can see that the numbers of base-pairs need not to be exactly the same, nor does the bases need be identical to give a reasonable estimate of the SK_{hyb}. While this may be somewhat incidental, we also checked other *Pw*TBP-DNA structural vari-
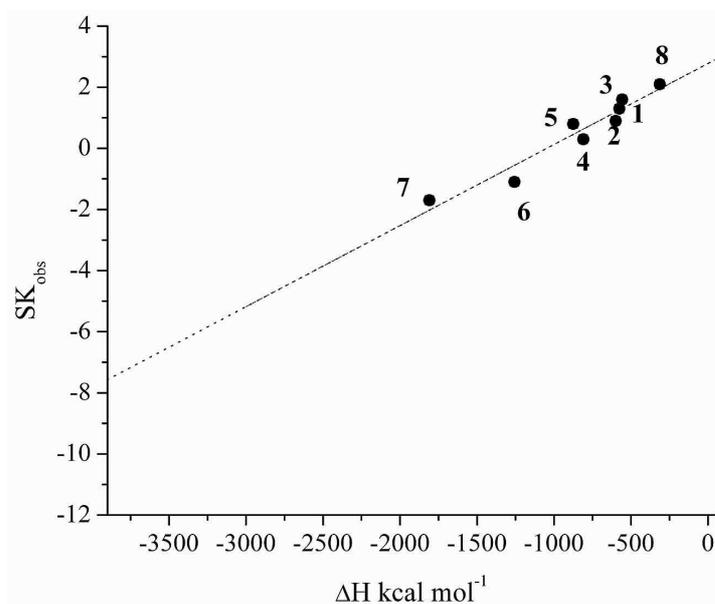
Figure 6: Experimental SK_{obs} versus interaction enthalpies between TBP and DNA in the wild-type and mini-mized PwTBP-DNA in explicit solvent (with the default cut-off). The enthalpy (ΔH) includes all atoms (with an "infinite" cut-off), and even double-, triple-, and the quadruple mutants was included in the fit. All struc-tures with available SK_{obs} used for fitting are labeled 1 to 8: E12A (1), E42A (2), E128A (3), E12AE128A (4), E12AE42A (5), E12AE42KE128A (6), E12AE41KE42KE128A (7) and the wild-type (8). Their respective SK_{obs} are adapted from references [11,17]. The correlation coefficient R is 0.96523 and the linear equation is $SK_{hyb} = 2.77932 + 0.00265 \cdot \Delta H \pm 0.28$, shown by the dotted line.

ants and found excellent agreement with experiments when exactly the same number of base-pairs were used in the structure (from another PwTBP crystal structure – PDB code 1d3u that has 23 base-pair in each strand which were truncated to 17 base-pairs/strand), and an offset of about 0.75 units when the number of base-pairs was increased from 17 to 23 which is not surprising since there will be a system-dependency on the fitting (Ta-ble 2). For any of the SK_{hyb}, we minimized the structures accordingly to the 10-12-14 Å explicit solvent minimization. We are aware that this is in a way system-dependent, but argue that the dependence is internal and due to the class of proteins and the num-ber of DNA-bases present in the modelling. Thus, we need not worry about mimicking the exact experimental conditions in the fitting between SK_{obs} and protein-DNA force field interaction energies, and plan to report a follow-up in the SK_{hyb} approach for other proteins or peptides in complex with DNA/RNA (Bredenberg and Fenley, preprint).

Importantly – however – is that the fitting of experimental versus a model case ap-pears in this case to hold for different TBPs and that we can use this method and arbitrar-ily choose the number of bases present in the structure for the fitting against experiments. This assures us that the force-field is doing a good job provided that the modelling is ac-curate and it also assures us that the experiments are consistent despite being conducted in independent laboratories [11,17,18,32].

Table 2: Suggested mutants from a fitting of known SK_{obs} versus the non-bonded interaction energy between wild-type or mutant variants of *Pyrococcus woesi* TATA-binding protein and DNA. Some experimental SK_{obs} for the *Pw*TBP that was not used in the fitting and from *Saccharomyces cerevisiae* TATA-binding protein to DNA is provided for comparison.

| Method: | SK_{pred} | | SK_{hyb} | SK_{obs} |
	12	GBMVII [16]	Exp. + MM	Exp. [11,17]
Protein				
Pw				
wt(2)	0.73		1.25	2.1
wt(3)	-0.04		2.12	2.1
D3A	-0.54	0.47	1.54	
*E12K	-2.26		0.57	1.2
D19A	-1.28	-0.40	1.38	
E27A	-0.60	-0.12	1.49	
K37E	0.15		3.26	
H49R	-1.82		1.11	
D51A	-0.79	-0.31	1.42	
D51K	-1.03		0.91	
D52A	-0.49	-0.08	1.52	
D52K	-1.03		1.02	
E77A	-0.67	0.24	1.50	
*D101A	-1.24	-0.53	1.38	2.1
*Q103A	-0.47		2.01	2.3
D110A	-0.87	-0.24	1.40	
E130A	-0.76	-0.81	1.35	
E130K	-0.97		0.78	
E128AE130A			0.67	
E128KE130K	-2.94		-0.53	
E41AE42A			0.66	
E41KE42K	-2.98		-0.67	
D51AD52A	-0.98		0.95	
D51KD52K	-1.11		~0	
E12AE41KE128A	-2.40		-0.55	
Sc				
wt(1)	-5.94		-4.59	-4.6
wt(2)	-5.92		-4.77	-4.6

Pw: wt(2), PDB code: 1D3U, with 23 base-pairs and wt(3), PDB code: 1D3U, with number of base-pairs reduced to 17. *Sc*: wt (1), PDB code: 1YTB and wt (2), PDB code: 1YTF. Mutants labeled with a * have experimental SK_{obs} available, but was not used in the fitting of the SK_{hyb}.

Biological implications in salt-dependent association of DNA and novel *Pw*TBP mutants suggested from SK$_{pred}$ or SK$_{hyb}$

Now, if SK >0 then the binding should increase with increasing salt concentration. It is clear from Fig. 6 that the enthalpy (*i.e.* non-bonded energy) of the solute increases with increasing SK$_{obs}$ and becomes less favorable. This implies that the entropy of the solute must be more favorable than the loss in enthalpy at higher SK and that the solvent (water and mobile ions) itself must gain in free energy since G$_{tot}$ = G$_{solute}$ + G$_{solvent}$. For the D51KD52K mutant the SK$_{hyb}$ is close to zero (Table 2) and if this holds true, then this mutant should be less sensible or even insensible in DNA-binding relative to the salt concentration in the solvent. This actually supports the idea that two ions are incorporated in the wild-type *Pw*TBP-DNA complex [32] (that is the experimental interpretation of SK$_{obs}$ >0) and may hint that the binding site of these two ions are close to residues D51 and D52. When these two acidic residues are substituted for two positively charged lysines, the protein no longer needs prosthetic ions from the solvent and thus will be insensitive to the salt-concentration. This does not contradict the possibility that the salt dependence in the binding of *Pw*TBP to DNA is accompanied by an increase in reactivity of the *Pw*TBP when the salinity is increased (Bredenberg, Russo and Fenley – Biophys. J., in press). On the contrary, when internal protein salt-bridges are broken, ions from the solvent may compensate by binding to the unbound partner of the disrupted ion pair in the protein. More theoretical and experimental studies are needed to unravel this relation fully.

The K37E mutant is predicted to have a SK$_{hyb}$ larger than the wild-type (Table 2). This residue (K37) is close to the DNA (Fig. 1) and forms hydrogen-bonds to the phosphate-backbone. When introducing an acidic residue at this position, the salt-dependent character of the binding will increase (provided that the mutant structure has a functional fold that binds to DNA). For the specific case of the K37E mutant, the force-field non bonded energy between the TBP and DNA would be >0 (Fig. 6), and thus it is questionable if other contributions to the free energy in binding (*e.g.* increasing the salt concentration in the solute) would suffice to allow the complex formation between this mutant and the DNA.

4   Conclusions

The use of default non-bonded settings in the force field for geometry optimizations of a structure should be fine as long no residues are located beyond the truncated zone of pair-wise Coulomb energies. For residue mutations that are remote from a binding co-partner (*i.e.* nucleic acid or another protein/peptide) the results from a subsequent PBE prediction of the binding, salt-dependence and other computed properties that relates to structure-energy relationships should be closely examined and caution should be made about conclusions drawn from such cases. Using a larger cut-off in the non-bonded scheme improves the linearity between the SK$_{pred}$ and R$_{O-O}$. Thus, ideally one

would want to include all atoms in the minimization/molecular dynamics simulations, but such an increase in non-bonded listing is also computationally more demanding.

We show that the non bonded inter-molecular energy between the DNA and the protein is highly correlated with the SK_{pred} within an ensemble of structures that are either identical in sequence or mutant forms of the same protein binding to the same stretch of DNA. We also show that the correlation between the enthalpic component in the force-field energies against experimentally determined SK_{obs} is highly linear and this suggests that the binding of TBP to its DNA-TATA recognition element follows the same physical laws and do not bind with different mechanisms. Yet simple, this hybrid – SK_{hyb} – from molecular mechanics non-bonded energies and experimental SK_{obs} may prove useful in extracting structural information coupled to changes in the salt dependence of peptides and proteins associating to nucleic acids. Thus, we are currently exploring several other systems for which three-dimensional structures together with thermodynamic binding affinity data at various salt concentrations are available in the literature. We have a series of studies with regard to the SK_{hyb} approach and its very diverse applications on other biomolecular systems are underway (Bredenberg and Fenley, in preparation).

Our results reported in this work enables us specifically to suggest novel mutants for the *Pw*TBP and – in general – may be a useful auxiliary tool in understanding the character of macromolecular assemblies of peptides/proteins and nucleic acids by using multi-disciplinary methods that are based on both theory and experiments. Thus, we expect that this paper will encourage and enhance collaborations between scientists from both theoretical and experimental labs focusing their research on nucleic acid association with other biologically important molecules: proteins, peptides or substrates for drug design.

Acknowledgments

References

[1] T. Rowlands, P. Baumann and S. P. Jackson, The TATA-binding protein: A general transcription factor in eukaryotes and archaebacteria, Science, 264 (1994), 1326-1329.
[2] S. K. Burley, The TATA box binding protein, Curr. Opinion Struct. Biol., 6 (1996), 69-75.
[3] S. K. Burley and K. Kamada, Transcription factor complexes, Curr. Opinion Struct. Biol., 12 (2002), 225-230.
[4] L. J. Rothschild and R. L. Mancinelli, Life in extreme environments, Nature, 409 (2001), 1092-1101.
[5] G. A. Patikoglou, J. L. Kim, L. Sun, S-H. Yang, T. Kodadek and S. K. Burley, TATA element recognition by the TATA box-binding protein has been conserved throughout evolution, Genes Dev., 13 (1999), 3217-3230.

[6] H. Koike, Y. Kawashima-Ohya, T. Yamasaki, L. Clowney, Y. Katsuya and M. Suzuki, Origins of protein stability revealed by comparing crystal structures of TATA binding proteins, Structure, 12 (2004), 157-168.

[7] D. B. Nikolov, H. Chen, E. D. Halay, A. Hoffmann, R. G. Roeder and S. K. Burley, Crystal structure of a human TATA box-binding protein/TATA element complex, PNAS, 93 (1996), 4862-4867.

[8] B. S. DeDecker, R. O'Brien, P. J. Fleming, J. H. Geiger, S. P. Jackson and P. B. Sigler, The crystal structure of a hyperthermophilic archaeal TATA-box binding protein, J. Mol. Biol., 264 (1996), 1072-1084.

[9] O. Littlefield, Y. Korkhin and P. B. Sigler, The structural basis for the oriented assembly of a TBP/TFB/promoter complex, PNAS, 96 (1999), 13668-13673.

[10] W. Zillig, H-P. Klenk, J. Trent, S. Wunderl, D. Janekovic, E. Imsel and B. Haas, Pyrococcus woesei, sp. nov., an ultra-thermophilic marinarchaebacterium, representing a novel order, Thermococcales. Syst. Appl. Microbiol., 9 (1987), 62-70.

[11] S. Bergqvist, M. A. Williams, R. O'Brien, J. E. Ladbury, Reversal of halophilicity in a protein-DNA interaction by limited mutation strategy, Structure, 10 (2002), 629-637.

[12] M. T. J. Record, T. M. Lohman and P. deHaseth, Ion effects on ligand-nucleic acid interactions, J. Mol. Biol., 107 (1976), 145-158.

[13] A. D. MacKerell Jr., B. R. Brooks, C. L. Brooks III, L. Nilsson, B. Roux, Y. Won and M. Karplus, CHARMM: The energy function and its parameterization with an overview of the program, in: P. v. R. Schleyer, P. R. Schreiner, N. L. Allinger, T. Clark, J. Gasteiger, P. Kollman and H. F. Schaefer III (Eds), The Encyclopedia of Computational Chemistry, Vol 1, Chichester: John Wiley & Sons, 1998, pp. 271-277.

[14] P. J. Steinbach and B. R. Brooks, New spherical-cutoff methods for long-range forces in macromolecular simulation, J. Comp. Chem., 15 (1994), 667-683.

[15] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, J. Comp. Chem., 4 (1983), 187-217.

[16] M. S. Lee, J. Salsbury, R. Freddie and C. L. Brooks III, Novel generalized Born methods, J. Chem. Phys., 116 (2002), 10606-10614.

[17] S. Bergqvist, The Role of Ions in an Archaeal Protein-DNA Interaction, London: University of London, 2002.

[18] S. Khrapunov and M. Brenowitz, Influence of the N-terminal domain and divalent cations on self-association and DNA binding by the Saccharomyces cerevisiae TATA binding protein, Biochemistry, 46 (2007), 4876-4887.

[19] A. D. MacKerell Jr, D. Bashford, M. Belott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B, 102 (1998), 3586-3616.

[20] N. Foloppe and A. D. MacKerell Jr, All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, J. Comp. Chem., 21 (2000), 86-104.

[21] A. T. Brünger and M. Karplus, Polar hydrogen positions in proteins: Empirical energy placement and neutron diffraction comparison, Proteins, 4 (1988), 148-156.

[22] W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. W. Impey and M. L. Klein, Comparison of

simple potential functions for simulating liquid water, J. Chem. Phys., 79 (1983), 926-935.

[23] C. L. Brooks III and M. Karplus, Deformable stochastic boundaries in molecular dynamics, J. Chem. Phys., 79 (1983), 6312-6325.

[24] C. L. Brooks III, A. T. Brünger and M. Karplus, Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach, Biopolymers, 24 (1985), 843-865.

[25] J-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, J. Comp. Phys., 23 (1977), 327-341.

[26] T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe and N. A. Baker, PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations, Nucl. Acids. Res., 35(suppl 2) (2007), W522-525.

[27] T. J. Richmond, Solvent accessible surface area and excluded volume in proteins: Analytical equations for overlapping spheres and implications for the hydrophobic effect, J. Mol. Biol., 178 (1984), 63-89.

[28] A. H. Boschitsch and M. O. Fenley, A new outer boundary formulation and energy corrections for the nonlinear Poisson-Boltzmann equation, J. Comp. Chem., 28 (2007), 909-921.

[29] A. H. Boschitsch and M. O. Fenley, Hybrid boundary element and finite difference method for solving the nonlinear Poisson-Boltzmann equation, J. Comp. Chem., 25 (2004), 935-955.

[30] C. Bertonati, B. Honig and E. Alexov, Poisson-Boltzmann calculations of nonspecific salt effects on protein-protein binding free energies, Biophys. J., 92 (2007), 1891-1899.

[31] C. W. Still, A. Tempczyrk, R. C. Hawley and T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, J. Am. Chem. Soc., 112 (1990), 6127-6129.

[32] S. Bergqvist, M. A. Williams, R. O'Brien and J. E. Ladbury, Heat capacity effects of water molecules and ions at a protein-DNA interface, J. Mol. Biol., 336 (2004), 829-842.