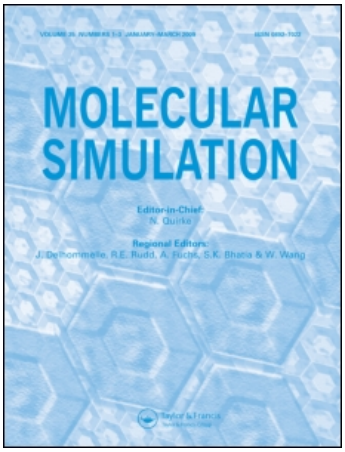
This article was downloaded by:
On: *14 January 2011*
Access details: *Access Details: Free Access*
Publisher *Taylor & Francis*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Simulation

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Calculation of the Affinity of the λ Repressor-Operator Complex Based on Free Energy Component Analysis

David C. Kombo[a]; B. Jayaram[a]; Kevin J. McConnell[a]; David L. Beveridge[a]
[a] Department of Chemistry and Molecular Biophysics Program, Hall-Atwater Laboratory, Wesleyan University, Middletown, CT

Online publication date: 26 October 2010

To cite this Article Kombo, David C. , Jayaram, B. , McConnell, Kevin J. and Beveridge, David L.(2002) 'Calculation of the Affinity of the λ Repressor-Operator Complex Based on Free Energy Component Analysis', Molecular Simulation, 28: 1, 187 — 211
To link to this Article: DOI: 10.1080/08927020211970
URL: http://dx.doi.org/10.1080/08927020211970

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Molecular Simulation, 2002, Vol. 28(1–2), pp. 187–211

Taylor & Francis
Taylor & Francis Group

CALCULATION OF THE AFFINITY OF THE λ REPRESSOR-OPERATOR COMPLEX BASED ON FREE ENERGY COMPONENT ANALYSIS

DAVID C. KOMBO[a], B. JAYARAM[b], KEVIN J. McCONNELL[a] and DAVID L. BEVERIDGE[a,*]

[a]*Department of Chemistry and Molecular Biophysics Program, Hall-Atwater Laboratory, Wesleyan University, Middletown, CT 06457;*
[b]*Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi-110016, India*

Downloaded At: 18:49 14 January 2011

(Received September 2000; accepted January 2001)

A calculation of the binding free energy of the λ repressor-operator complex is described based on free energy component analysis. The calculations are based on a thermodynamic cycle of seven steps decomposed into a total of 24 individual components. The values of these terms are estimated using a combination of empirical potential functions from AMBER, generalized Born – solvent accessibility calculations, elementary statistical mechanics and semiempirical physicochemical properties. Two alternative approaches are compared, one based on the crystal structure of the complex and the other based on the molecular dynamics simulation of the λ repressor-operator complex. The calculated affinity is −19.7 kcal/mol from the crystal structure calculation and −17.9 kcal/mol from the MD method. The corresponding experimental affinity of the complex is about −12.6 kcal/mol, indicating reasonable agreement between theory and experiment, considering the approximations involved in the computational methodology. The results are analyzed in terms of contributions from electrostatics, van der Waal interactions, the hydrophobic effect and solvent release. The capabilities and limitations of free energy component methodology are assessed and discussed on the basis of these results.

Keywords: Free energy calculations; Component analysis; Molecular dynamics; Protein; DNA interactions; Binding affinity; Bacteriophage λ system

*Corresponding author. Tel.: (860)-685-2575, Fax: (860)-685-2211, e-mail: dbeveridge@mail.wesleyan.edu

ISSN: 0892-7022 © 2002 Taylor & Francis Ltd DOI: 10.1080/08927020290004458

1. INTRODUCTION

Free Energy Component Analysis (ΔGCA) is a widely used, albeit highly approximate, method for the calculation of the free energy of complexation of biological molecules. The ΔGCA approach is tractable for large complex systems, and useful for relating structural characteristics of a system to thermodynamic stability. However, due to the uncertainties in the evaluation of individual terms and propagation of errors, the reliability of ΔGCA results is questionable. To address this issue, we have formulated a ΔGCA methodology with a more exhaustive enumeration of terms than previously used, and are proceeding to apply this to a series of systems as rigorously as possible. Our objective is to determine as fully as possible the capabilities and limitations of the method and develop guidelines for informed utilization of the procedures. In previous articles, a general theory and methodology was described, and applied to a calculation of the affinity of the EcoRI endonuclease-DNA complex [1] and used to examine the specificity of a site specific mutation in the U1A-RNA complex [2]. To extend our characterization and assessment, we describe herein a ΔGCA study of the λ repressor operator complex. The λ repressor-operator complex is a propitious choice for a case study – as the prototype case for the recognition process in genetic expression systems, it is well characterized in both structure determination and biochemistry.

Downloaded At: 18:49 14 January 2011

2. BACKGROUND

The calculation of affinities (ΔG), specificities (ΔΔG) and cooperativities (ΔΔΔG) is important for the understanding of protein-DNA binding processes. However, the reliable theoretical calculation of theses quantities has turned out to be a challenging theoretical undertaking since the systems are complex and the dimensionality of the problem, especially with water and mobile ions included, is very large [3]. The net affinity and related properties depend on a sensitive balance of contributions such as effects of shape and electronic complementarity, electrostatics, hydration, counterion release, conformational flexibility and solvent reorganization upon complexation. A well-defined theoretical approach to this problem is free energy *via* molecular simulation (ΔGMS), *i.e.*, the determination of the free energy of complexation directly using Monte Carlo (MC) or molecular dynamics (MD) simulation combined with thermodynamic integration or the perturbation method [4–9]. The most popular ΔGMS methods are the

thermodynamic cycle perturbation method and the Brownian dynamics. In the perturbation method, the free energy problem is cast in terms of a determination of relative free energy difference, between two well-defined states. Sen and Nillson have recently applied this approach in a computational study of the wild-type EcoRI-DNA complex and some of its mutants [10]. They found that the predicted free energy differences qualitatively agreed well with experiments. Free energy calculation by the Brownian reactive dynamics method has been used to predict the rate of initial diffusional encounter between reactant molecules in solution. This method was applied to study the nonspecific λ Cro repressor protein-DNA complex formation [11]. Favorable electrostatic interactions were found to be partially offset by a loss of entropy, as the incoming protein dimer orientations become increasingly restrictive. For protein-DNA complexes, only a few examples of ΔGMS have been reported so far [10, 12], and this has not yet become a practical approach for studies over a series of systems, or a series of mutants of a single system.

Downloaded At: 18:49 14 January 2011

At the opposite end of the computational spectrum is "free energy component analysis" in which the free energy determination is reduced to sums of terms independently calculated within a well-defined thermodynamic cycle. Such calculations come in a variety of flavors [13]. The methodology has shown particular recent success in the treatment of protein–protein binding energetics by Janin and coworkers [3, 14, 15] and by others [16, 17]. We note particularly the recent calculations of Honig and coworkers on the binding free energies of MHC Class I protein–peptide interactions using continuum electrostatics [18]. They found that the net electrostatics oppose formation and non-polar interactions favor complexation. Gilson and coworkers have applied this extensively to drug–DNA interactions [19, 20]. In the formulation applied herein, we utilize empirical energy functions, crystallographic and MD information combined with continuum treatments of solvation by means of GBSA models of hydration [21–26], Debye-Huckel (DH) theory of added salt effects, and semiempirical estimates of physicochemical effects. Basically, the following phenomena are quantified; the structural adaptation of the protein and the DNA molecules upon binding, the electrostatic effects of desolvating the macromolecule and the counterions, the van der Waals interactions with the solvent, the elimination of the solvent cavity in which the macromolecule is accommodated, the change in the added salt effects, the electrostatic and van der Waals interactions between partners in the complex, and the entropy loss due to decreased translational and rotational degrees of freedom in the complex relative to the unbound species.

The scientific applications of free energy component analysis are best advanced in the context of "case studies", in which calculations are formulated and performed on protein nucleic acid complexes in the most theoretically rigorous manner possible and the results are examined and assessed alongside corresponding experimental data. In a previous study, we applied an expanded version of this method to the EcoRI endonuclease-DNA complex, and found that the calculated standard free energy of formation obtained, -11.5 kcal/mol, agreed with experiment to within 5 kcal/mol [1]. Electrostatic effects were found to be destabilizing when solvent was included, a counter-intuitive result with precedents in protein-protein complexes [18]. In a parallel treatment of the U1A-RNA complex [2], the methodology revealed that subtle aspects of the specificity were linked to changes in the coupling of consensus and nonconsensus elements of the protein structure. In summary, the ΔGCA approach is tractable, readily interpretable and amenable to successive improvements. However, there are disadvantages as well – the additivity of free energy terms is assumed, and there are difficulties and uncertainties in estimating individual terms which propagate in the calculation of absolute affinities. These problems enter in different forms in the determination of relative free energies of specificity and cooperativity, but are nonetheless important.

Downloaded At: 18:49 14 January 2011

The λ repressor is one of the proteins that control gene regulation and expression in the phage λ through self-assembly, site-specific recognition and cooperativity (for review, see Ptashne [27]). The N-terminal domain of the λ repressor (residues 1 to 92) also has the ability to form dimers that then bind to DNA [28]. Its crystal structures, both in the free and DNA-bound forms, are known [29–32]; as are its corresponding solution structures [33–36]. It consists of five α-helices interspaced by loops (Fig. 1). The second and third α-helices and the loop between them constitute the helix-turn-helix DNA-recognition unit, while the fifth helices mediate dimerization by virtue of packing interactions between each other. The third helix of each monomer fits into the major groove and is termed the "recognition helix". Besides, a flexible N-terminal arm contributes to the recognition by wrapping around the DNA operator site. In addition to structural studies, extensive genetic and biochemical studies (for review see Sauer *et al.* [37]) as well as studies of energetic contributions of the λ repressor-DNA association [38–43] have provided insight into the mechanism of the coupled equilibrium dimerization-DNA binding process of the λ repressor-operator complex. A review article on protein-DNA calculations in general, using the λ system as a case study, will be reported elsewhere [44].

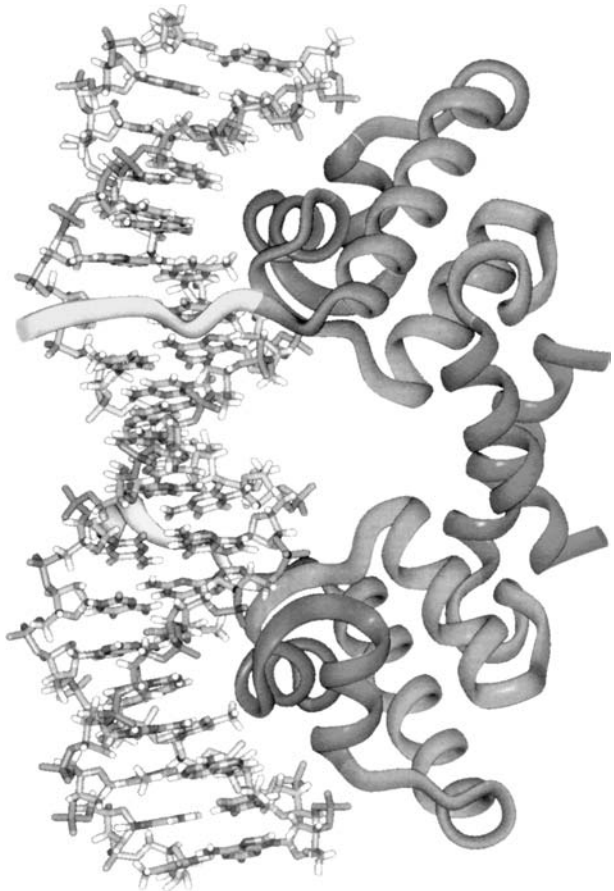


FIGURE 1 Structure of the N-terminal domain of the λ repressor-OL1 DNA complex. The N-terminal arms (residues 1–6) are shown in yellow. The helix-turn-helix unit in both monomers is shown in magenta. The helix-5 and helix-5′, which mediate the dimerization process, are shown in red.

Building on the studies described above, we report herein ΔGCA computations on the λ repressor-operator complex, considering two possible ways of formulating the problem: (a) using only the crystal structure data, and (b) using information from a molecular dynamics simulation on the λ repressor-operator complex. The results form a basis for further understanding of the problems and prospects of component analysis methodology, and reveal as well some interesting features of protein DNA complexation accessible only with computational modeling. MD studies on the λ repressor-OL1 operator complex, probing sequence-dependent

Downloaded At: 18:49 14 January 2011

geometrical futures of the DNA operator, are reported elsewhere [45]. Our results show that in the MD analysis, electrostatic interactions and water release favor complexation, while van der Waals interactions, considering both intramolecular and solvation effects, prove slightly unfavorable. The free energy component analysis accurately predicts that the number of ions released upon complexation of the N-terminal domain of the repressor with OL1 is equal to 3.9, which agrees very well with the experimentally-determined value of 3.7 ion equivalents released upon binding [38]. The calculated standard free energy of formation agrees with experiment to within 5 kcal/mol. Other theoretical studies of protein-DNA complexes are increasingly being carried out, thanks to the considerable advances in digital computer power [46–49].

Downloaded At: 18:49 14 January 2011

3. METHODS AND CALCULATIONS

3.1. Free Energy Cycle

Our calculations are referred to the thermodynamic cycle for protein–DNA binding in solution which is shown in Figure 2. The net binding process is decomposed into seven steps. Step I describes the process of converting uncomplexed DNA, denoted "D", and counterions to the form D^*, in which the DNA has adapted its structure to that of the bound form. It is important to note that some of the counterions are considered "condensed" on the DNA and are treated explicitly.

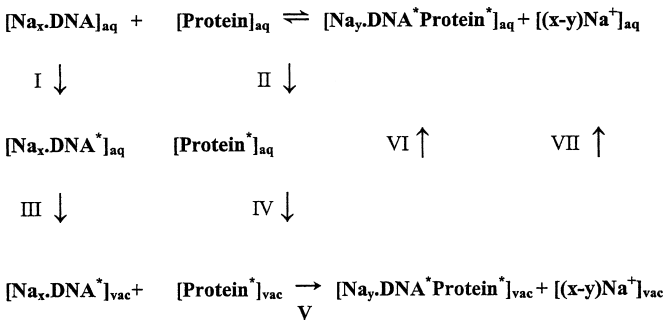


FIGURE 2 The thermodynamic cycle used for the analysis of binding free energies of protein-DNA complexes.

3.2. Free Energy Components

The contributions of the various chemical forces to the complexation process are determined by properly summing up the terms defined in Table I and Ref. [1]. Specifically, the contribution of structural adaptation to free energy can be written as

$$\Delta G^{adpt} = \Delta G_1^{adpt.D} + \Delta G_2^{adpt.P} \tag{1}$$

In this study, we have assumed that both terms of Eq. (1), *i.e.*, the adaptation energy of the DNA molecule and the protein molecule to be practically equal to zero. In fact, crystal and NMR solution structures of the λ repressor-OL1 complex and the free protein, have shown that the complexation process does not alter significantly the structures of the free partners [29–36].

The contribution of electrostatics (excluding the small ion effects) to the free energy result can be expressed as

$$\Delta G^{el} = \Delta G_3^{el.D} + \Delta G_8^{el.P} + \Delta H_{13}^{el.C} + \Delta G_{19}^{el.C} \tag{2}$$

The van der Waals interactions, effectively the net energetics of shape complementarity, is reflected in the sum,

$$\Delta G^{vdW} = \Delta G_5^{vdW.D} + \Delta G_{10}^{vdW.P} + \Delta H_{15}^{vdW.C} + \Delta G_{21}^{vdW.C} \tag{3}$$

The total contribution of cavitation effects to the binding is

$$\Delta G^{cav} = \Delta G_6^{cav.D} + \Delta G_{11}^{cav.P} + \Delta G_{22}^{cav.C} \tag{4}$$

These nonelectrostatic contributions to the standard free energy (vdW and cavity terms) are defined as a linear function of the solvent accessible surface area with an empirical coefficient, defining the proportionality [21]. However, as pointed out by Hummer, Chandler and Pratt, and their coworkers, we acknowledge that exposed surface area alone does not fully characterize hydrophobicity, a collective and complex phenomenon [70, 71]. Nevertheless, our use of a surface area approximation is based on a large literature, empirical parametrization and testings.

The entropy change on complexation is described by the combination

$$\Delta G^{trvc} = -T\Delta S_{17}^{tr\,\&\,rot} - T\Delta S_{18}^{vib\,\&\,conf} \tag{5}$$

These terms, which relate to the structures in the gas phase, are based on ideal gas statistical mechanics and other considerations [50]. The calculation

Downloaded At: 18:49 14 January 2011

Downloaded At: 18:49 14 January 2011

TABLE I Calculated values of the various contributions to the standard free energy of binding for the λ Repressor-OL1 Operator complex at 298 K

Step	Term	Component	Value (kcal/mol) Crystal	Value (kcal/mol) MD	Method[1]
Step-I*: *Structural Adaptation of DNA					
1	$\Delta G_1^{adpt.D}$	Free energy change for the process $D \rightarrow D^*$	~0.0	~0.0	EX
Step-II*: *Structural Adaptation of Protein					
2	$\Delta G_2^{adpt.P}$	Free energy change for the process $P \rightarrow P^*$	~0.0	~0.0	EX
Step-III*: *Desolvation of DNA					
3	$\Delta G_3^{el.D}$	Electrostatic component of DNA desolvation	13762.2	12817.9	GB
4	$\Delta G_4^{el.ci.D}$	Counterion effect on D^* desolvation	−7981.4	−7369.3	GB
5	$\Delta G_5^{vdW.D}$	vdW component of D^* desolvation	311.2	296.5	SA
6	$\Delta G_6^{cav.D}$	Cavity component of D^* desolvation	−367.5	−350.1	SA
7	$\Delta G_7^{DH.D}$	Loss of added salt interactions with $Na \cdot D^*$	30.4	28.4	DH
Step-IV*: *Desolvation of Protein					
8	$\Delta G_8^{el.P}$	Electrostatic component of P^* desolvation	3154.7	3207.7	GB
9	$\Delta G_9^{elci.P}$	Counterion effect on P^* desolvation	~0.0	~0.0	GB
10	$\Delta G_{10}^{vdW.P}$	vdW component of P^* desolvation	421.0	444.6	SA
11	$\Delta G_{11}^{cav.P}$	Cavity component of P^* desolvation	−497.2	−525.0	SA
12	$\Delta G_{12}^{DH.P}$	Loss of added salt interactions with P^*	19.9	20.0	DH
Step-V*: *Complex formation in vacuo					
13	$\Delta H_{13}^{el.C}$	Electrostatic interactions of P^*D^*	−3358.8	−3714.2	FF
14	$\Delta H_{14}^{ci.C}$	Change in counterion interactions on P^*D^* binding	1924.7	2099.2	FF
15	$\Delta H_{15}^{vdw.C}$	vdW interactions of P^*D^*	−204.4	−148.5	FF
16	$-T\Delta S_{16}^{ci.C}$	Entropy of (complex-DNA-protein) counterions	−13.5	−11.1	
17	$-T\Delta S_{17}^{tr\&rot}$	Rotational and translational entropy change	32.4	32.4	PF
18	$-T\Delta S_{18}^{v\&cnf}$	Vibrational and configurational entropy change	25.6	24.6	

Downloaded At: 18:49 14 January 2011

Step-VI*: *Solvation of Complex					
19	$\Delta G_{19}^{el.C}$	Electrostatic component of complex solvation	−13535.9	−12318.6	GB
20	$\Delta G_{20}^{el.ci.C}$	Counterion effect on complex solvation	6364.3	5603.7	GB
21	$\Delta G_{21}^{vdW.C}$	vdW component of complex solvation	−579.4	−581.1	SA
22	$\Delta G_{22}^{cav.C}$	Cavity component of complex solvation	684.2	686.2	SA
23	$\Delta G_{23}^{DH.C}$	Added salt interactions with complex	−41.9	−39.5	DH
Step-VII*: *Solvation of freed counterions					
24	$\Delta G_{24}^{r.ci}$	Solvation free energy of released counterions	−170.3	−221.6	GB
	ΔG^{0}**net**	**Net binding free energy**	−19.7	−17.9	

[1] The method used to determine each term is indicated with the abbreviations: DH: Debye Huckel, Exp: Experimental, FF: Force Field, GB: Generalized Born, SA: Surface Area.

of vibrational and configurational entropy change on complexation follows Janin's procedure [3, 14], which includes contributions due to the motional restriction of amino acid side chains at the contact surface and from the additional low-frequency vibrational modes which arise on complex formation. The number of amino acids contacting the DNA are computed and an entropy loss of [Rln3] for each amino acid is taken into account. The underlying assumption is that each amino acid side chain can exhibit three conformations and because of the contacts with DNA, this degree of freedom is frozen. A better method is to use the simulations on the free protein, the DNA and the complex and carry out a quasiharmonic analysis of entropies on the lines of Karplus and Kushick's theory [69]. However, a common problem encountered in these studies is the dimensionality of the covariance matrix. The DNA is manageable but not the protein and certainly not the complex. Some approximations have to be introduced in estimating the determinants for such large systems. These are under investigation.

Downloaded At: 18:49 14 January 2011

Small ion effects on free energy, due to both explicit ions and added salt in the model, can be summed as

$$\Delta G^{\text{ions}} = \Delta G_4^{\text{el.ci.D}} + \Delta G_7^{\text{DH.D}} + \Delta G_9^{\text{el.ci.P}} + \Delta G_{12}^{\text{DH.P}} + \Delta H_{14}^{\text{ci.C}} - T\Delta S_{16}^{\text{ci.C}} + \Delta G_{20}^{\text{el.ci.C}} + \Delta G_{23}^{\text{DH.C}} + \Delta G_{24}^{\text{r.ci}} \quad (6)$$

Here, the value for the corresponding solvation free energy is estimated from experimental data; each unit of charge gains a solvation free energy of -98.3 kcal/mol upon transfer to bulk [51]. We assume a model of the ion atmosphere of the DNA in which counterions neutralize the Manning fraction of the DNA charge. In this model, the discrete counterions with Na^+-sized solvatons bear an effective charge of 0.76, with a further correction for oligonucleotides of finite length, as described by Jayaram *et al.* [1]. Consequently, the solvation free energy of each ion released will be equal to -56.78 kcal/ion released.

The counterion treatment is phenomenological in nature and in a way a true atomic level reproduction of Manning theory and the binding process. The solvation energies of ions are based on Born treatment and match with experimental values. The ion – DNA interactions are based on AMBER force field which is now well tested *via* a series of MD simulations on DNA. The only weak link in the ion energy treatment is the entropy loss of a condensed ion (2 kcal) and this was taken based on the literature values of entropies of sodium ions in water and alcohols [51, 68]. However, in the post facto analysis of MD simulations, the ion treatment is much more rigorous. No

assumptions with regard to placement of the ions or their charge is made. It must be noted that the energetics of ion distribution around DNA in explicit solvent simulations is a potential of mean force problem and any estimate of free energy of ions condensed vis-a-vis released, based on a single simulation is an approximation.

The sum of all these terms equals the net standard free energy of binding, *viz*.

$$\Delta G^0 = \Delta G^{adpt} + \Delta G^{el} + \Delta G^{vdW} + \Delta G^{nel} + \Delta G^{ions} + \Delta G^{trvc} \qquad (7)$$

The terms of the Eqs. (1) through (6) are defined in Table I. The mathematical and physico-chemical formalism of the individual contributions has been extensively documented in a previous analogous study which did set the stage for the present work [1]. The electrostatic contributions to solvation (Eq. (2)) computed using the GB equation employed the effective radii parameters derived by Jayaram *et al*. [25] based on AMBER charges. This allows the calculation of both the intramolecular and solvation electrostatics based on a single set of charges, eliminating a possible inconsistency in the model. Note that usage of $\varepsilon = 1$ in the computation of direct electrostatic interactions between protein and DNA is consistent with the GB methodology for solvation. The molecular surface area calculations required for the non-electrostatic contribution to the solvation energy (Eqs. (3)–(4)) were performed using the ACCESS program for solvent accessibility based on the algorithm of Lee and Richards [52] and AMBER parm94 vdW radii. The added salt concentration employed in the Debye-Huckel factor (Eq. (6)) was 0.18M.

Downloaded At: 18:49 14 January 2011

3.3. Crystal Structure Analysis

The atomic coordinates of the X-ray crystal structure of the λ repressor-OL1 DNA complex, used in this study, were obtained from the Nucleic acid data base (code : pdr010). First, explicit hydrogen atoms were added to the crystal structure. Next, the protonation state of ionizable groups was set at that corresponding to $pH = 7$ and assumed to be constant. Then, energy minimization of the protein-DNA complex was performed using the Sander module of the Amber 4.1 molecular modeling package [53] employing the most recent parameterization of the AMBER empirical energy function, the "parm94" force field [54]. In the energy minimization, we seek only to relieve any unfavorable clashes in the crystal structure and prepare the system for further study. Here 500 steps of minimization restraining heavy

atoms {50 steps of steepest descent (SD) followed by 450 steps of conjugate gradient (CG)}, followed by a further 250 steps (50 SD + 200 CG) of free minimization are carried out to a tolerance of 0.5 kcal/mol Å. The structure obtained at this stage is still very close to the crystal structure, and forms the basis for further analysis of the binding process.

3.4. Molecular Dynamics Simulation

The X-ray coordinates sets of the λ repressor-OL1 operator complex were the same as in the above-mentioned analysis. All water molecules were removed from the crystal structure of the complex. A computer model for the N-terminal arm in the non-consensus protein monomer was built as described elsewhere [55]. Using the Amber Edit module, 36 Na^+ ions were added to neutralize the DNA charges. To provide electroneutrality for the protein molecule at PH 7.0, 13 Cl^- and 12 Na^+ ions were added per protein monomer, in the vicinity of the side-chain atoms of Lys and Arg and Glu and Asp, respectively. Additional 27 Na^+ and 27 Cl^- ions were added to the total system protein-DNA. Thus, the total number of ions in the system was 140. A Monte Carlo simulation of the system of the protein, DNA, ions and co-ions enclosed in a cylindrical box of 74 Å and 33 Å radius was subsequently carried out using a distance dependent dielectric function. All solute atoms were kept fixed throughout the simulation whereas all the ions were allowed to move. Geometric and energy parameters from the force field of Cornell *et al.* [54] were used to determine the total configurational energy in the MC simulation. The lowest-energy configuration of the system was then solvated with 7833 TIP3P water molecules using the Amber edit module and immersed in a box of $84.0 \times 65 \times 59$ Å^3. Six rounds of energy minimization were carried out on the system composed of the protein-DNA complex, ions and water molecules, prior to an MD simulation, using the protocol described by Cheatham *et al.* [56]; Young *et al.* [57]. We have used the force field of Cornell *et al.* [54], as implemented in the Amber 4.1 computer package developed by Pearlman *et al.* [53], with the Particle Mesh Ewald method to treat the electrostatics, constant pressure, SHAKE [58] on hydrogens, a 2 fs time step, and a temperature of 300 K with Berendsen coupling, and a 9 Å cutoff. MD simulations in the nanosecond regime were then performed on the Cray Y-MP machine at the Pittsburgh Supercomputing center. The analysis of the MD trajectory was carried out using the Molecular Dynamics Tool Chest (MDTC2.0) analysis software program [59] and the results will be reported elsewhere [45].

Downloaded At: 18:49 14 January 2011

3.5. Protein-DNA and Protein – Protein Interaction Energies

An interpretation of the Protein-DNA interaction energies *per se* was carried out using three terms of the potential function: van der Waals, electrostatic and hydrophobic. A similar calculation of protein – protein interaction energies between the protein monomers was carried out for residues located at the dimer interface. The magnitude of the hydrophobic potential energy was assigned to be proportional to apolar solvent accessible surface area alone with coefficients of 25 cal/A^2 for aliphatic carbon atoms and 16 cal/A^2 for aromatic carbon atoms [60]. In all the present study, the solvent accessible surface area is calculated using a probe radius of 1.4 Å. A sigmoidal distance-dependent dielectric function was used to model solvation effects. Computations were performed on structures of the protein-DNA complex extracted from the MD trajectory. Water molecules and ions were removed from the MD-derived structures of the complex and thus were not explicitly included in the calculation procedure.

Downloaded At: 18:49 14 January 2011

4. RESULTS

The calculated contributions to the standard free energy of binding for the λ repressor protein to the OL1 DNA operator are shown in Table I. Negative values of energy are favorable and positive values are unfavorable to binding. We find that 10 out of the 24 components listed here are favorable to binding. In particular, the value obtained for the solvation energy of released counterions (term 24) allows us to predict that 3.9 counterions are released upon complexation. An analysis of the results of Table I based on the contributions as defined in Eqs. (1) – (7) is presented schematically in Figure 3a for the MD simulation, and in Figure 3b, for the crystal structure of the λ repressor-operator OL1 complex. Figure 3a shows that the differential effect of electrostatics and of cavity formation upon complexation are seen to be favorable to binding. All other terms, including direct van der Waals interactions between the protein and DNA, are seen to be unfavorable to binding. These results corroborate the widespread emphasis on hydrophobic contacts, hydrogen bonds, phosphate contacts, found in the protein-DNA complex literature. That the small ions effect is unfavorable to complexation is consistent with the results of the calculations based on finite-difference solutions to the nonlinear Poisson-Boltzmann (NLPB) equation carried out by Misra *et al.*, on the λ repressor-operator complex [61]. They found that the contribution of the ion atmosphere to the binding

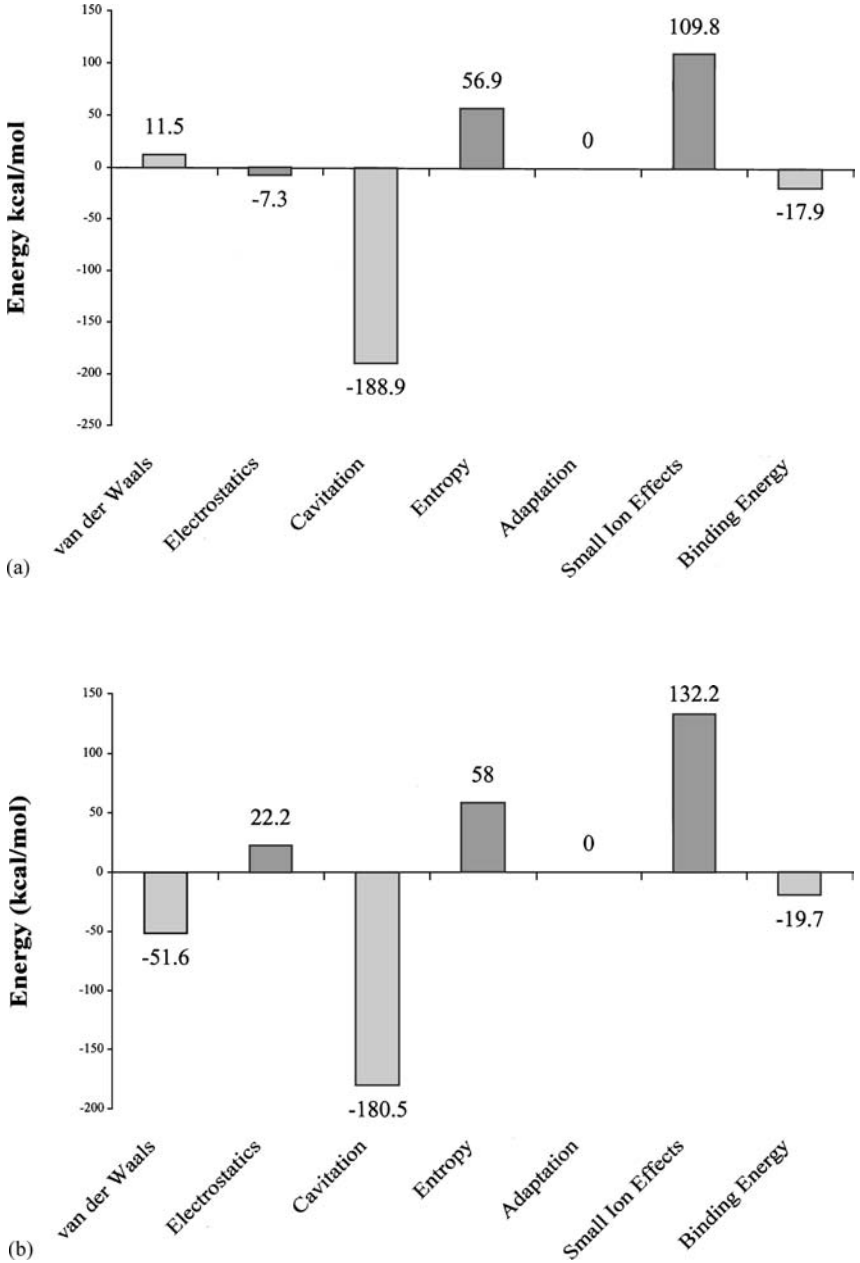


FIGURE 3 A histogram view of the calculated contributions to the binding free energy of protein-DNA complexes. The reference state is separated protein and DNA, with negative values favorable to binding and positive values unfavorable. (a) Results for the crystal structure analysis; (b) results the MD-derived structures.

Downloaded At: 18:49 14 January 2011

free energy substantially destabilizes the λ repressor-OL1 operator DNA complex.

The results shown in Figure 3b, obtained in the crystal structure analysis, show that electrostatics are unfavorable to binding, while van der Waals interactions favor complexation. This difference between the MD results and the single crystal structure analysis presumably results from the effects of explicit water molecules and explicit ions, as well as the conformational flexibility taken into account in the molecular dynamics simulation. Nevertheless, compensation effects between various terms result in very close values in the net binding energy in both cases; -19.7 kcal/mol and -17.9 kcal/mol, for the crystal structure and the MD simulation, respectively.

As described in the method section, we have computed the interaction energies between the protein and the OL1 DNA, using the ensemble of structures generated in the MD trajectory, after removing explicit water molecules and explicit ions. Figure 4 shows the total per-residue contribution of the flexible N-terminal arms to the total protein–DNA interaction energies. Interestingly, we notice that the immutable Lys-4 exhibits

Downloaded At: 18:49 14 January 2011

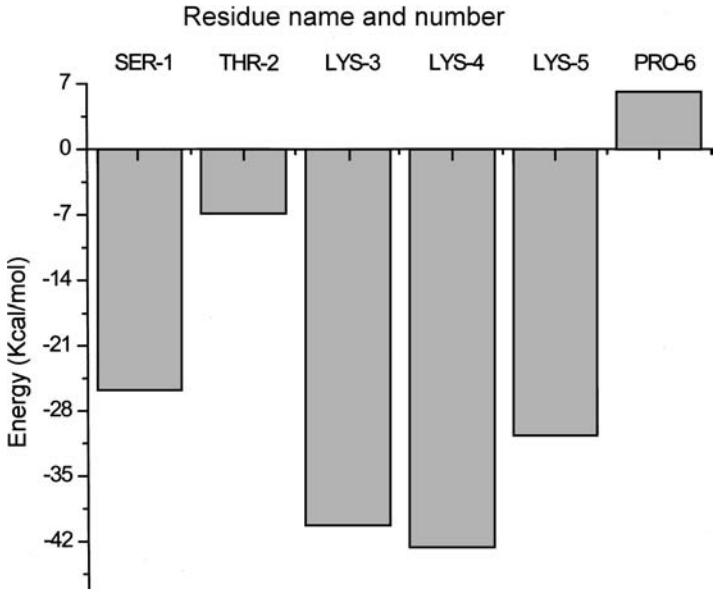


FIGURE 4 Histogram of interaction energies of P* with D* partitioned with respect to amino acids of the N-terminal flexible arm of the λ repressor. The presentation is for the total sum of interaction energies of each protein monomer with its cognate operator half-site.

the highest value in favorable interaction energy, followed by Lys-3 and Lys-5, the two residues which only tolerate one substitution with Arg [32]. Similarly, Ser-1, Thr-2, and Pro-6, which exhibit lower if not unfavorable interaction energy, were also found to tolerate any substitution, in the same cassette-directed mutagenesis experiment.

The histogram of interaction energies of P^* and D^* partitioned with respect to amino acids of the helix-turn-helix unit is show in Figure 5. We notice that Gln-33, Gln-44 and Ser-45 exhibit the strongest contributions to the interaction energies. This result is consistent with the contacts found in the crystal structure of the complex, where Gln-33 was found to contact the sugar phosphate backbone, while Gln-44 and Ser-45 make contact with DNA bases, in both operator half-sites [31]. The negatively charged amino acid Glu-34 and Asp-38, exhibit the most unfavorable contributions, presumably because of their repulsive electrostatic interactions with the negatively charged DNA phosphate groups.

Binding of the N-terminal domain of λ repressor to DNA is coupled to dimerization. Hydrophobic interactions between helix-5 and helix-5′ drive the packing at the dimer interface. We have carried out calculations of both protein–protein interaction energies between the protein monomers and

Downloaded At: 18:49 14 January 2011

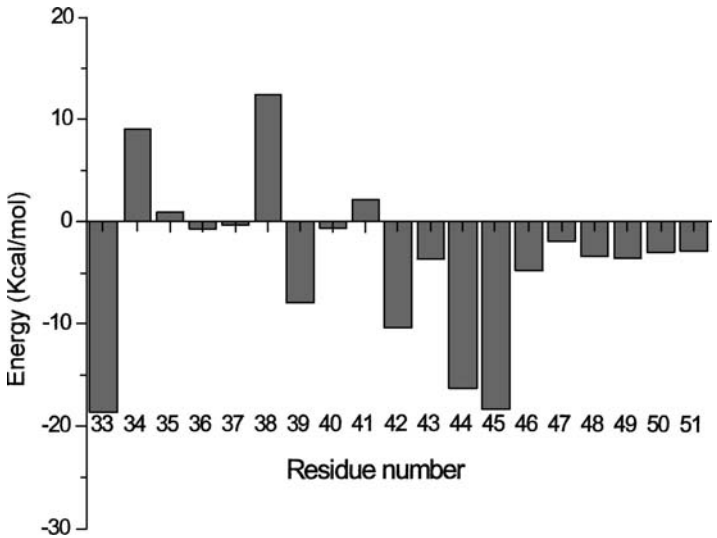


FIGURE 5 Histogram of interaction energies of P^* with D^* partitioned with respect to amino acids of the helix-turn-helix unit of the λ repressor. The presentation is for the total sum of interaction energies of each protein monomer with its cognate operator half-site.

calculations of protein–DNA interaction energies. The total sum of these 2 types of interaction energies, partitioned with respect to amino acids residues at the dimer interface, is shown in Figure 6. We notice that the following residues exhibit the strongest contributions to the interaction energies: Ile-84, Met-87, Tyr-88, Ala-90, and Val-91, while the most unfavorable contributions are mediated through Glu-75, Pro-78, Glu-83, and Ser-92.

A corresponding analysis referenced to the nucleotides of the 17-bp OL1 DNA operator is shown in Figure 7. Here, we clearly notice the asymmetric pattern in the contacts made by each operator half-site with the corresponding protein monomer. In particular, the GC-rich central region, which makes contacts with the N-terminal flexible arms and Ala-56, Asn-58, Asn-55 and Asn-61, in the crystal structure, generally exhibits stronger contributions than do the outer-edge region, with the exceptions of both terminal AT dinucleotide steps. The strong contributions observed in these two steps corroborate the direct contacts between Lys-19, Tyr-22, Gln-33, and Asn-52 with the sugar-phosphate backbone of the DNA at these positions, which have been observed by X-ray crystallography [31].

Downloaded At: 18:49 14 January 2011

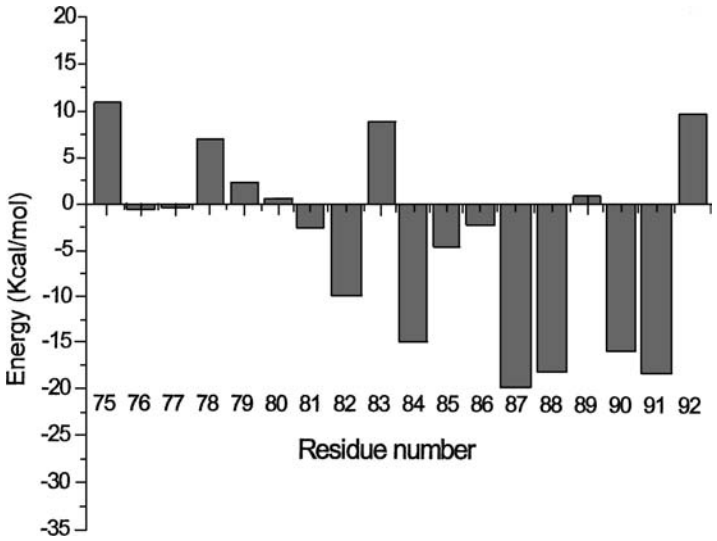


FIGURE 6 Histogram of interaction energies of P* with D* partitioned with respect to amino acids of the N-terminal flexible arm of the λ repressor. The presentation is for the total sum of interaction energies of each protein monomer with its cognate operator half-site.

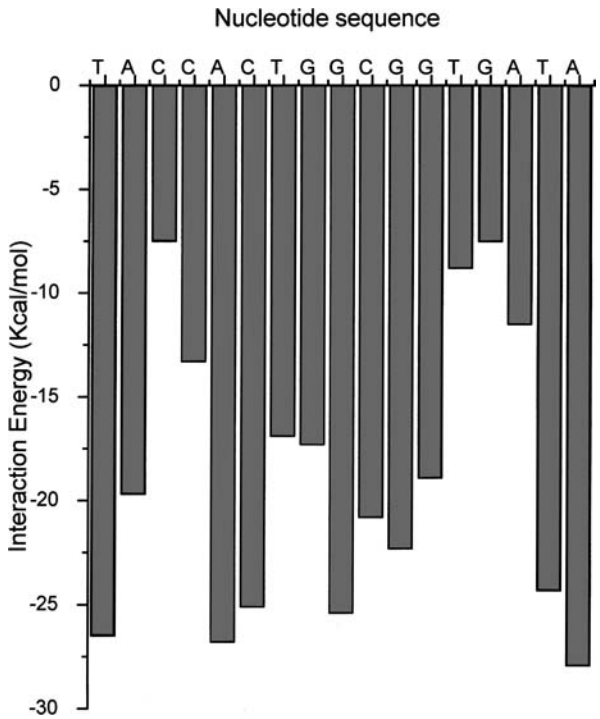


FIGURE 7 Histogram of interaction energies of D* with P* partitioned with respect to nucleotides of the DNA. The presentation is for both strands of the OL1 operator DNA duplex.

Downloaded At: 18:49 14 January 2011

5. DISCUSSION

Our free energy component analysis of the binding of the N-terminal domain of the λ repressor protein to its cognate OL1 operator has resulted in a net binding free energy of -17.9 kcal/mol for the MD-derived structures, and of -19.7 kcal/mol for the crystal structure. These values are close to the experimental values of -15.6 to -12.6 kcal/mol observed in the salt concentration range of 50 mM to 200 mM KCl [38]. In the latter experiment, if the entire salt effect is attributed to KCl-linked interactions, the slopes of the lines obtained in the study of the salt dependence on the site specific binding of the N-terminal domain to the DNA operator right (OR) region, translated to 3.7 ion equivalents released upon binding [38]. Interestingly, our MD-based computational study predicts a value of 3.9 ion equivalents released upon binding at the OL1 site, which is very close to the value experimentally-determined. One should recall that OR1 and

OL1 only differ at one base-pair position and almost bind with the same affinity.

The trends in the partition of interactions energies involving the arms (Fig. 4) are in agreement with the tolerance for amino acid substitution in these regions of the protein. The stronger the interaction energies of a particular amino acid residue, *i.e.*, the more negative the values, the lower the mutability of that residue. For example, in the flexible N-terminal arm region, Lys-4 exhibits the strongest contribution to the protein–DNA interaction energy, followed by Lys-3 and Lys-5. These results are in agreement with the experimental findings as to which Lys is absolutely required at position 3 for proper DNA-binding of the protein, while Lys-4 and Lys-5 can only tolerate Arg. The three other residues, Ser-1, Thr-2 and Pro-6, which exhibit the three lowest contributions to protein–DNA interaction energy, were found to tolerate any substitution, in the same experiment.

Downloaded At: 18:49 14 January 2011

Reidhaar-Olson and Sauer found a rough correlation between the tolerance for amino acid substitution in helix-5 and the loop between helix-4 and helix-5, and the solvent accessibility surface area of the side-chains [62, 63]. However, some exceptions to the general rule were found. For example, Pro-78, a highly surface-exposed residue, was found to be conserved, while Ala-90, a buried residue, was found to tolerate many functional substitutions. In the present study, we find that, in general, the stronger the total contributions to both protein–protein and protein–DNA interactions (Fig. 6), the lower the tolerance for amino-acid substitution. For example, Met-87, Tyr-88, Val-91, Ala-90, and Ile-84 exhibit the strongest total contributions to the sum of the protein–protein and protein–DNA interaction energies, ranging from -14.5 to -20.0 kcal/mol. Random mutagenesis experiments have also shown that these five residues do tolerate only one to 9 functionally acceptable residues. Residues such as Ser-79, Arg-82, Glu-83, Tyr-85, Glu-86, and Glu-89 exhibit very low contributions to the interaction energies, ranging from -9.5 kcal to $+8.5$ kcal. Random mutagenesis experiments have also shown that these five residues do in fact tolerate a high number of functionally acceptable substitutions, ranging from 12 to 13. However, about 6 residues show exceptions to this correlation. Specifically, Glu-75, Phe-76, Ser-77, Pro-78, Ile-80, and Ala-81, show relatively very low interaction energies but still, these residues only tolerate 1 to 6 functional substitutions. In the exceptional case of Pro-78, further experimental studies aimed at investigating the requirement for proline at this position, showed that this residue was essential in the λ repressor protein and was required for resistance to intracellular proteolysis [64].

Figure 7, which shows the protein-DNA interaction energy partitioned with respect to nucleotides, emphasizes that the λ repressor dimer, made of identical subunits, recognizes the nearly palindromic operator sequence asymmetrically. This finding is consistent with systematic base-substitution experiments [65] and X-ray crystallographic studies as well [29–32].

The most striking difference between the results obtained for the crystal structure analysis and the MD-derived ensemble of structures lies in the magnitude and the sign of the contributions of electrostatics and van der Waals interactions to the free energy of binding (Figs. 3a and 3b). It looks as though there is a compensation effect between these two contributions when one goes from the crystal structure study to the MD study. In fact, while the contribution of electrostatics tends to favor complexation, the contribution of van der Waals interactions tends to become unfavorable, once explicit solvent and ions, as well as conformational flexibility are taken into account through MD simulation. Our finding that electrostatics are favorable to complexation in the present case, is consistent with the important role played by the Lys-rich flexible N-terminal arm in the affinity of the λ repressor protein with its cognate OL1 DNA site. In fact, Eliason *et al.*, have shown that removal of the whole arm from the N-terminal domain of the λ repressor reduces the affinity of the protein by at least 3 orders of magnitude [66]. In addition, it is widely believed that the arrangement of helical dipoles in the helix-turn-helix unit is very important for recognition [37]. The closeness of the N-terminal end of each helix of the helix-turn-helix unit to the DNA results in a favorable electrostatic interaction between the DNA and the partial positive charge at the N-terminus of the helix. Moreover, H-bonding interactions, which are electrostatic in nature, contribute significantly in the contacts mediated by the helix-turn-helix units and the N-terminal arms.

Misra *et al.*, reported a Poisson-Boltzmann study on the pH dependence of binding constants of the N-terminal domain of the λ repressor, using the crystal structure of the complex. Their study captured the experimentally-observed trends quite well. In a further dissection of the electrostatic effects in the λ system, the authors demonstrated that desolvation electrostatics opposes direct favorable electrostatic interactions between the protein and the DNA to such an extent that the electrostatics is net unfavorable to the overall binding free energies. They proposed that water release and enhanced packing interactions could be the major driving forces in protein-DNA complexation. Similar to their study, we also found that the desolvation free energy of DNA is the largest unfavorable contribution to the electrostatic binding free energy, while the desolvation of charged groups on the protein also opposes binding, but to a smaller extent than in

Downloaded At: 18:49 14 January 2011

the case of DNA. Moreover, the signs of the contributions to the free energy obtained by Misra *et al.*, are the same compared with our results on the crystal structure. However, the magnitude of the absolute values obtained in our study are very large compared to theirs. Lower differences are also observed. For example, they found that nonpolar contributions amount to -150 kcal/mol. We found a corresponding value of -180.5 and -188.9 kcal/mol in the crystal structure and the MD analysis, respectively.

With respect to the net binding free energy, it appears that the result obtained at 180 mM added salt concentration in the MD analysis (-17.9 kcal/mol) is closer to the range of -15.6 to -12.6 kcal/mol, which was experimentally observed in the salt concentration range of 50 mM to 200 mM KCl value, than is the corresponding value derived from the crystal structure analysis. The MD-derived free energy components were averaged over an ensemble of structures sampled during the simulation, while the crystal structure analysis used a single structure. Moreover, it is interesting to notice that the MD analysis predicts a number of ions released equal to 3.9, which is closer to the experimental value of 3.7 than does the value of 3.0, obtained in the crystal structure analysis. The MD analysis presents also the advantage of taking into account the conformational flexibility, a property akin to macromolecular recognition and mimic fairly well the solution environment; two properties that are lacking in the crystal structure analysis.

Downloaded At: 18:49 14 January 2011

6. SUMMARY AND CONCLUSIONS

We have learned that the free energy component analysis described in this article is capable of predicting reasonably the binding free energy of a protein to DNA, and the number of ions released upon complexation. The calculated standard free energy of formation, derived from MD analysis, agrees with experiment to within 5 kcal/mol. This result is reminiscent of the value obtained with the crystal structure analysis of the EcoRi endonuclease-DNA complex [1]. In both studies, hydrophobic contributions favor complexation, while entropic and small ion effects do not. Nevertheless, some errors and uncertainties are akin to the process of free energy component analysis applied to a complex system. Examples include the small differences between large numbers problem, qualitative approximations in some of the theoretical methods used to estimate the various terms, and the nonuniqueness of the partitioning into components. Even if biology obeys the laws of chemistry and physics, can we expect molecular bioenergetics to be explained quantitatively by these physical sciences or

should we consider the results in a qualitative manner? To make matters worse, agreements with experiments is not *a condition sine qua non* for the accuracy of the computational procedure. Studies aimed at improving the methods for calculating the added salt effects and the free energy of adaptation are currently underway in our laboratory. Analysis of individual contributions to the standard free energy of complexation from amino acid residues of the protein have shown that some contact interactions disfavor complexation and that, tolerance for amino acid substitution in a DNA-binding protein can be predicted and explained as well.

One could argue that the approximate agreement with experiment for the overall free energy of binding of the N-terminal domain of the λ repressor might be fortuitous. But it could not have been fortuitous for several other systems which have also been examined, in addition to the EcoRI endonuclease-DNA complex [1]. These include some protein-drug complexes as well. While the method is not yet quantitative, it has the merits of exhaustively enumerating all the molecular level phenomena associated with a binding reaction and providing estimates albeit semi-quantitative/qualitative which is new knowledge in itself.

Downloaded At: 18:49 14 January 2011

Acknowledgments

This work was supported by a Patrick and Catherine Donaghue Foundation Biomedical Research postdoctoral fellowship to D. C. Kombo. B. Jayaram acknowledges the financial support and encouragement received from the Indian Department of Science and Technology, the Council of Scientific and Industrial Research, and the Indo-French Centre for the Promotion of Advanced Research. K. J. McConnell was supported by a traineeship in molecular biophysics from the NIH Training Grant GM-08271. Support was also received from Grant GM-37909 from the National Institute of Health awarded to D. L. Beveridge. Molecular dynamics simulations were carried out using the Pittsburgh Supercomputer center facility and the National Center for Supercomputer Applications at the University of Illinois.

References

[1] Jayaram, B., McConnel, K. J., Dixit, S. B. and Beveridge, D. L. (1999). "Free energy analysis of protein–DNA binding: the EcoRI endonuclease-DNA complex", *J. Comp. Phys.*, **151**, 333.

[2] Blakaj, D. M., McConnell, K. J., Beveridge, D. L. and Baranger, A. M. (2000). "Molecular dynamics and thermodynamics of a protein-RNA complex: mutation of a

conserved aromatic residue modifies stacking interactions and structural adaptation in the U1A-stem loop 2 RNA complex", *J. Amer. Chem. Soc.*, submitted.
[3] Janin, J. (1995). "Elusive affinities", *Proteins*, **21**, 30.
[4] Beveridge, D. L. and DiCapua, F. M. (1989). "Free energy *via* molecular simulation: applications to chemical and biomolecular systems", *Ann. Rev. Biophys. Chem.*, **18**, 431.
[5] Jorgensen, W. L. (1989). "Free energy calculations: A breakthrough for modeling organic chemistry in solution", *Acc. Chem. Res.*, **22**, 184.
[6] McCammon, J. A. (1991). "Free energy from simulations", *Curr. Opin. Struct. Biol.*, **1**, 196.
[7] Kollman, P. A. (1993). "Free energy calculations: Applications to chemical and biochemical phenomena", *Chem. Rev.*, **93**, 2395.
[8] Simon, E. J. (1996). *Dissertation*, Harvard Univ.
[9] Van Gunsteren, W. F. and Berendsen, H. J. C. (1990). "Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry", *Agnew. Chem. Int. Ed. Engl.*, **29**, 992.
[10] Sen, S. and Nilsson, L. (1999). "Free energy calculations and molecular dynamics simulations of wild-type and variants of the DNA-*ECORI* complex", *Biophys. J.*, **77**, 1801.
[11] Thomasson, K. A., Ouporov, I. V., Baumgartner, T. *et al.* (1998). "Free energy of nonspecific binding of Cro repressor protein to DNA", *J. Phys. Chem. B*, **101**, 9127.
[12] Sen, S. and Nilsson, L. (1999). "Structure, interaction, dynamics and solvent effects on the DNA-EcoRI complex in aqueous solution from molecular dynamics simulation", *Biophys. J.*, **77**, 1782.
[13] Ajay and Murcko, M. A. (1995). "Computational methods to predict binding free energy in ligand-receptor complexes", *J. Med. Chem.*, **38**, 4953.
[14] Janin, J. (1995). "Protein–protein recognition", *Prog. Biophys. Mol. Biol.*, **64**, 145.
[15] Finkelstein, A. V. and Janin, J. (1989). "The price of lost freedom: Entropy of biomolecular complex formation", *Protein Eng.*, **3**, 1.
[16] Lazardis, T., Archontis, G. and Karplus, M. (1995). "Enthalpic contribution to protein stability: Insights from atom-based calculations and statistical mechanics", *Adv. Protein Chem.*, **47**, 231.
[17] Novotny, R. E., Bruccoleri, M. Davis *et al.* (1997). "Empirical free energy calculations: a blind test and further improvements to the method", *J. Mol. Biol.*, **268**, 401.
[18] Proloff, N., Windemuth, A. and Honig, B. (1997). On the Calculation of Binding Free Energies using Continuum Methods: Application to MHC Class I Protein–Peptide Interactions, *Protein Science*, **6**, 1293 (1997).
[19] Gilson, M. K., Given, J. A., Bush, B. L. *et al.* (1997). "The statistical thermodynamic basis for computation of binding affinities: a critical review", *Biophys. J.*, **72**, 1047.
[20] Forsyth, W. R., Gilson, M. K., Antosiewicz, J. *et al.* (1998). Theoretical and experiment analysis of ionization equilibria in ovomucoid third domain, *Biochemistry*, **37**, 8643.
[21] Still, W. C., Tempczyk, R. C., Hawley, R. C. *et al.* (1990). "Semianalytical treatment of solvation for molecular mechanics and dynamics", *J. Am. Chem. Soc.*, **112**, 6127.
[22] Hawkins, G. D., Cramer, C. J. and Truhlar, D. G. (1996). Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium, *J. Phys. Chem.*, **100**, 19824.
[23] Hawkins, G. D., Cramer, C. J. and Truhlar, D. G. (1995). "Pairwise solute descreening of solute charges from a dielectric medium, *Chem. Phys. Lett.*, **246**, 122.
[24] Jayaram, B., Liu, Y. and Beveridge, D. L. (1998). "A modification of the generalized Born theory for improved estimates of solvation energies and pKa Shifts", *J. Phys. Chem.*, **109**, 1465.
[25] Jayaram, B., Sprous, D. and Beveridge, D. L. (1998). "Solvation free energy of biomacromolecules: parameters for a modified generalized Born model consistent with the AMBER force field", *J. Phys. Chem. B*, **102**, 9571.
[26] Reddy, M. R., Erion, M. D., Agarwal, A. *et al.* (1998). "Solvation free energies calculated using the GB/SA model", *J. Comp. Chem.*, **19**, 769.
[27] Ptashne, M. (1992). "A genetic switch. Second edition. Phage λ and higher organisms", Blackwell scientific publications and Cell Press, Cambridge.
[28] Pabo, C. O., Sauer, R. T., Sturtevart, J., Ptashne, M. (1979). "The λ repressor contains two domains", *Proc. Natl. Acad. Sci. USA*, **76**, 1608.

Downloaded At: 18:49 14 January 2011

[29] Pabo, C. O. and Lewis, M. (1982). "The operator-binding domain of λ repressor: structure and DNA recognition", *Nature*, **298**, 443.
[30] Jordan, S. R. and Pabo, C. O. (1988). "Structure of the lambda complex at 2.5 Å resolution: details of the repressor–operator interactions", *Science*, **242**, 893.
[31] Beamer, L. J. and Pabo, C. O. (1992). "Refined 1.8 crystal structure of the λ repressor-operator complex", *J. Mol. Biol.*, **227**, 177.
[32] Clarke, N. D., Beamer, L. J., Goldberg, H. R., Berkower, C. and Pabo, C. O. (1991). "The DNA binding arm of λ repressor: critical contacts from a flexible region", *Science*, **254**, 267.
[33] Weiss, M. A., Pabo, C. O., Karplus, M. and Sauer, R. T. (1987a). "Dimerization of the operator binding domain of phage λ repressor", *Biochemistry*, **26**, 897.
[34] Weiss, M. A., Karplus, M. and Sauer, R. T. (1987b). "Quaternary structure and function in phage λ repressor: 1H-NMR studies of genetically altered proteins", *J. Biomolec. Struc. Dyn.*, **5**, 539.
[35] Benevides, J. M., Weiss, M. A. and Thomas, G. J. Jr. (1991). "DNA recognition by the helix-turn-helix motif: investigation by laser Raman spectroscopy of the phage λ repressor and its interaction with operator sites O_L1 and O_R3", *Biochemistry*, **30**, 5955.
[36] Benevides, J. M., Weiss, M. A. and Thomas, G. J. Jr. (1994). "An altered specificity mutation in the λ repressor induces global reorganization of the protein–DNA interface", *J. Biol. Chem.*, **269**, 10869.
[37] Sauer, R. T., Jordan, S. R. and Pabo, C. O. (1990). "A model system for understanding protein–DNA interactions and protein stability", *Adv. Prot. Chem.*, **40**, 1.
[38] Bain, D. L. and Ackers, G. K. (1994). Self-association and DNA binding of λ cI repressor N-terminal domains reveal linkage between sequence-specific binding and the C-terminal cooperativity domain. *Biochemistry*, **33**, 14679.
[39] Koblan, K. S. and Ackers, G. K. (1991). "Energetics of subunit dimerization in bacteriophage λ cI repressor: linkage to protons, temperature, and KCL", *Biochemistry*, **30**, 7817.
[40] Koblan, K. S. and Ackers, G. K. (1992). "Site-specific enthalpic regulation of DNA transcription at bacteriophage λ O_R", *Biochemistry*, **31**, 57.
[41] Senear, D. F. and Batey, R. (1991). "Comparison of operator-specific and nonspecific DNA binding of the lambda cI repressor: [KCl] and pH effects", *Biochemistry*, **30**, 6677.
[42] Senear, D. F. and Ackers, G. K. (1990). "Proton-linked contributions to site-specific interactions of λ cI repressor and O_R", *Biochemistry*, **29**, 6568.
[43] Nelson, H. C. M. and Sauer, R. T. (1985). "Lambda repressor mutations that increase the affinity and specificity of operator binding", *Cell*, **42**, 549.
[44] Kombo, D. C., Jayaram, B. and Beveridge, D. L. (2001). "Computational studies of protein-DNA interactions: lessons from the bacteriophage λ repressor-OL1 DNA operator complex", Manuscript in preparation.
[45] Kombo, D. C., McConnel, K. J., Young, M. A. and Beveridge, D. L. (2001). "Molecular dynamics simulation reveals sequence-intrinsic and protein-induced geometrical features of the OL1 DNA operator", *Biopolymers*, in the press.
[46] Harris, L. F., Sullivan, M. R., Popken-Harris, P. D. *et al.* (1994). "Molecular dynamics simulations in solvent of the estrogen receptor protein in complex with a glucocorticoid response element DNA sequence", *J. Biomol. Struct. Dyn.*, **12**, 249.
[47] Harris, L. F., Sullivan, M. R. and Popken-Harris, P. D. (1999). "Molecular dynamics simulation in solvent of the bacteriophage 434 cI repressor protein DNA binding domain amino acids (R1-69) in complex with its cognate operator (OR1) DNA sequence", *J. Biomol. Struct. Dyn.*, **17**, 1.
[48] Roxstroem, G., Velazquez, I., Paulino, M. *et al.* (1998). "Molecular dynamics simulation of a Zif268-DNA complex in water: spatial patterns and fluctuations sensed from a nanosecond trajectory", *J. Phys. Chem. B*, **102**, 1828.
[49] Duan, Y., Wilkosz, P. and Rosenberg, J. M. (1996). "Dynamics contributions to the DNA binding entropy of the EcoRI and EcoRV restriction endonucleases", *J. Mol. Biol.*, **264**, 546.
[50] Atkins, P. W., *Physical Chemistry* (Freeman, NY, 1990).
[51] Friedman, H. L. and Krishnan, C. V. (1973). In: *Water*: *A Comprehensive Treatise*, Edited by Franks, F., Plenum, New York, **3**.

Downloaded At: 18:49 14 January 2011

[52] Lee, B. K. and Richards, F. M. (1971). "The Interpretation of Protein Structures: Estimation of Static Accessibility", *J. Mol. Biol.*, **55**, 379.
[53] Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham III, T. E., Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. A. and Kollman, P. A. (1995). *Ambe 4.1. UCSF*, San Francisco, CA.
[54] Cornell, W. D., Cieplack, P., Bayly, C. I. *et al.* (1995). "A second generation force field for the simulation of proteins, nucleic acids and organic molecules", *J. Am. Chem. Soc.*, **7**, 5179.
[55] Kombo, D. C., Young, M. A. and Beveridge, D. L. (2000). "One nanosecond molecular dynamics simulation of the N-terminal domain of the λ repressor", *Biopolymers*, **53**, 596.
[56] Cheatham III, T. E., Miller, J. L., Fox, T., Darden, T. A. and Kollman, P. A. (1995). "Molecular dynamics simulation on solvated biomolecular systems: The particle Mesh Ewald method leads to stable trajectories of DNA, RNA and proteins", *J. Amer. Chem. Soc.*, **117**, 4193–4194.
[57] Young, M. A., Jayaram, B. and Beveridge, D. L. (1997). "Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: fractional occupancy of electronegative pockets", *J. Am. Chem. Soc.*, **119**, 59.
[58] Ryckaert, J.-P., Cicotti, B. and Berendsen, H. J. C. (1977). "Numerical integration of the cartesian equations of motions of a system with constraints: molecular dynamics of n-alkanes", *J. Comp. Phys.*, **23**, 327.
[59] Ravishanker, G. and Beveridge, D. L. (1997). *MD Tool Chest*. Wesleyan University, Middletown, CT.
[60] Karplus, P. A. (1997). "Hydrophobicity regained", *Protein Sci.*, **6**, 1302.
[61] Misra, V. K., Hecht, J. L., Sharp, K. A., Friedman, R. A. and Honig, B. (1994). "Salt effects on protein–DNA interactions. The λ cI repressor and *Eco*RI endonuclease", *J. M. B.*, **238**, 264.
[62] Reidhaar-Olson, J. F. and Sauer, R. T. (1988). "Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences", *Science*, **241**, 53.
[63] Reidhaar-Olson, J. F. and Sauer, R. T. (1990). "Functionally acceptable substitutions in two α-helical regions of λ repressor", *Prot. Struct. Funct. Genet.*, **7**, 306.
[64] Reidhaar-Olson, J. F., Parsell, D. A. and Sauer, R. T. (1990). "An essential proline in λ repressor is required for resistance to intracellular proteolysis", *Biochemistry*, **29**, 7563.
[65] Sarai, A. and Takeda, Y. (1989). "λ repressor recognizes the approximately 2-fold symmetric half-operator sequences asymmetrically", *Proc. Natl. Acad. Sci. USA*, **86**, 6513.
[66] Eliason, J. L., Weiss, M. A. and Ptashne, M. (1985). "NH_2-terminal arm of phage λ repressor contributes energy and specificity to repressor binding and determines the effects of operator binding". *Proc. Nat. Acad. Sci., USA*, **82**, 2339.
[67] Misra, V. K., Hecht, J. L., Yang, A. S. and Honig, B. (1998). "Electrostatic contributions to the binding free energy of the λ cI repressor to DNA". *Biophys. J.*, **75**, 2262.
[68] Krestov, G. A., *Thermodynamics of solvation* (Ellis-Horwood, New york, 1991).
[69] Karplus, M. and Kushick, J. N. (1981). "Method for estimating the configurational entropy of macromolecules". *Macromolecules*, **14**, 325.
[70] Lum, K., Chandler, D. and Weeks, J. D. (1999). "Hydrophobicity at small and large length scales". *J. Phys. Chem. B*, **103**, 457.
[71] Hummer, G., Garde, S., Garcia, A. E. and Pratt, L. R. (1996). "An information theory model of hydrophobic interactions". *Proc. Natl. Acad. Sci. USA*, **93**, 8951.

Downloaded At: 18:49 14 January 2011