ELSEVIER

Biophysical Chemistry 117 (2005) 239 – 254

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Review

Thermodynamic calculations in biological systems

L. Mario Amzel *, Xavier Siebert [1], Anthony Armstrong [1], German Pabon [1]

Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205, United States

Received 3 May 2005; received in revised form 3 May 2005; accepted 3 May 2005
Available online 24 June 2005

Abstract

The ability to compute intra- and inter-molecular interactions provides the opportunity to gain a deeper understanding of previously intractable problems in biochemistry and biophysics. This review presents three examples in which molecular dynamics calculations were used to gain insight into the atomic detail underlying important experimental observations. The three examples are the following: (1) Entropic contribution to rate acceleration that results from conformational constraints imposed on the reactants; (2) Mechanism of force unfolding of a small protein molecule by the application of a force that separates its N- and C-terminals; and (3) Loss of translational entropy experienced by small molecules when they bind to proteins.

Keywords: Molecular dynamics; Translational entropy; Rate acceleration; Extension by force

Contents

1. Example 1: entropic contributions to reaction rates . . . 240
   1.1. Methods . . . 240
   1.2. Results and discussion . . . 242
2. Example 2: unfolding proteins by force . . . 243
   2.1. Methods . . . 243
   2.2. Results . . . 243
   2.3. Discussion . . . 246
3. Example 3: translational entropy of binding in solution . . . 246
   3.1. Introduction . . . 246
   3.2. Methods . . . 247
      3.2.1. Molecular dynamics simulation . . . 247
      3.2.2. Normal mode analysis . . . 248
   3.3. Results and discussion . . . 249
      3.3.1. Solute ξ in water . . . 249
      3.3.2. The case of pure water . . . 250
      3.3.3. Benzene binding to the TCM T4-lysozyme . . . 251
      3.3.4. Parameterization . . . 252
Acknowledgments . . . 253
References . . . 253

Experimental biophysical studies continue to provide a wealth of information about the properties of proteins, their interactions with each other and with small molecules. Recently, even the behavior of individual molecules has

* Corresponding author. Tel.: +1 410 955 3955; fax: +1 410 955 0637.
E-mail address: mario@neruda.med.jhmi.edu (L.M. Amzel).
[1] These authors contributed equally to this work.

0301-4622/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2005.05.004

been studied using atomic force microscopes and optical tweezers. Most of these experiments, however, can benefit from atomic level interpretation of the results. Molecular Mechanics/Dynamics calculations can be used to try to gain insight into the molecular events that give rise to the experimental observations. This paper reviews three independent examples in which calculations of this type can contribute to our understanding of important molecular events: (1) reaction rates of small molecule models for enzyme catalysis, (2) atomic details of single molecule force unfolding, and (3) entropy contributions critical for prediction of drug binding to proteins.

1. Example 1: entropic contributions to reaction rates

Rates of enzymatic reactions have been interpreted in terms of entropic and enthalpic contributions to catalysis. The difficulty in analyzing reaction kinetics lies in how to treat the transition state, a transient high energy species having distorted molecular orbital geometries and force constants only accessible through time consuming quantum-mechanical calculations. Efforts in the past to determine which factors contribute most to enzymatic reactions have focused on simpler systems, small molecules undergoing intramolecular reaction [1–6]. In 1960 Bruice and Pandit determined the rates of cyclization of a series of monoesters of dicarboxylic acids differing in the number of rotatable bonds between reactive groups, and in the location and the identity of geminal substituents placed along the carbon backbone of the molecules [7,8] (Fig. 1). The data showed an increase in reaction rate when torsions between reactive groups were either eliminated or effectively frozen. Substituents placed at the α- and β-carbons also increased the reaction rate, with the more bulky substituents having a greater effect. The interpretation of these data in terms of entropic and enthalpic contributions, however, has remained a topic of debate.

Lightstone and Buice [1], using stochastic search methods, estimated the activation enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) in this series and found a strong correlation between $H^{\ddagger}$ and reaction rate but no correlation with $S^{\ddagger}$. More recently, Armstrong and Amzel [9] carried out different estimates of $S^{\ddagger}$ that showed a strong correlation with reaction rates. Their results will be reviewed here.

1.1. Methods

In the study by Armstrong and Amzel [9], two methods of estimating relative activation entropies were employed. They differed in the way the transition state was treated: as an ensemble of species restricted to a single ground state conformational well (Method I), or as a single conformer (Method II). Both methods discretely sampled the energy landscape of each compound. The structure of each compound was energy minimized and the dihedrals between the attacking carboxylate and the ester carbonyl were successively incremented through 360°. The first and last torsions were incremented using steps of 10°; the remaining torsions were incremented using a step size of 20°. After each increment the energy was minimized, constraining the torsions at their grid values (adiabatic mapping). (These calculations were carried out using the program CHARMM [10,11]). The probability of a conformation is given by:

$$p_i = \frac{e^{-E_i/kT}}{\sum_i e^{-E_i/kT}} \tag{1}$$

where the sum is taken over all points of the energy landscape. Fig. 2 shows the adiabatic energy surface and corresponding probability surface for structure 36ETHP calculated using dielectric constants of 1 and 80.

To calculate the entropy of activation ($S^{\ddagger}$), the Near Attack Conformation (NAC) was used as a surrogate of the reaction transition state (TS), since the energetics of the TS cannot be evaluated using molecular mechanics semiempirical potentials. The NAC concept, central to the calculations of Lightstone and Bruice [1], is defined as a ground state conformation having a geometry in which molecular orbitals are aligned for reaction (within a 30° cone) and are in close proximity (between 2.8 and 3.2 Å).

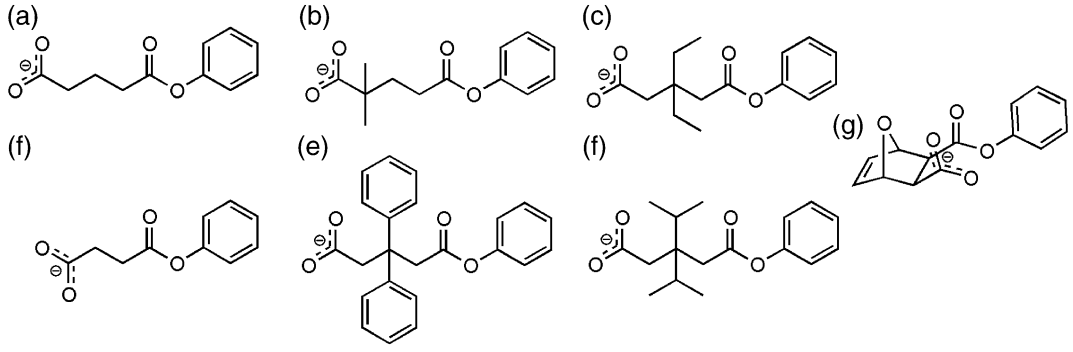

Fig. 1. The seven compound considered in this study. They are the monophenyl esters of (A) glutarate (GLUT), (B) α-gem-dimethylglutarate (αGDMG), (C) β-gem-diethylglutarate (βGDEG), (D) succinate (SUCC), (E) β-gem-diphenylglutarate (βGDPG), (F) β-gem-diisopropylglutarate (βGDIG), and (G) 3,6-endoxo-Δ4-tetrahydrophthalate (36ETHP).

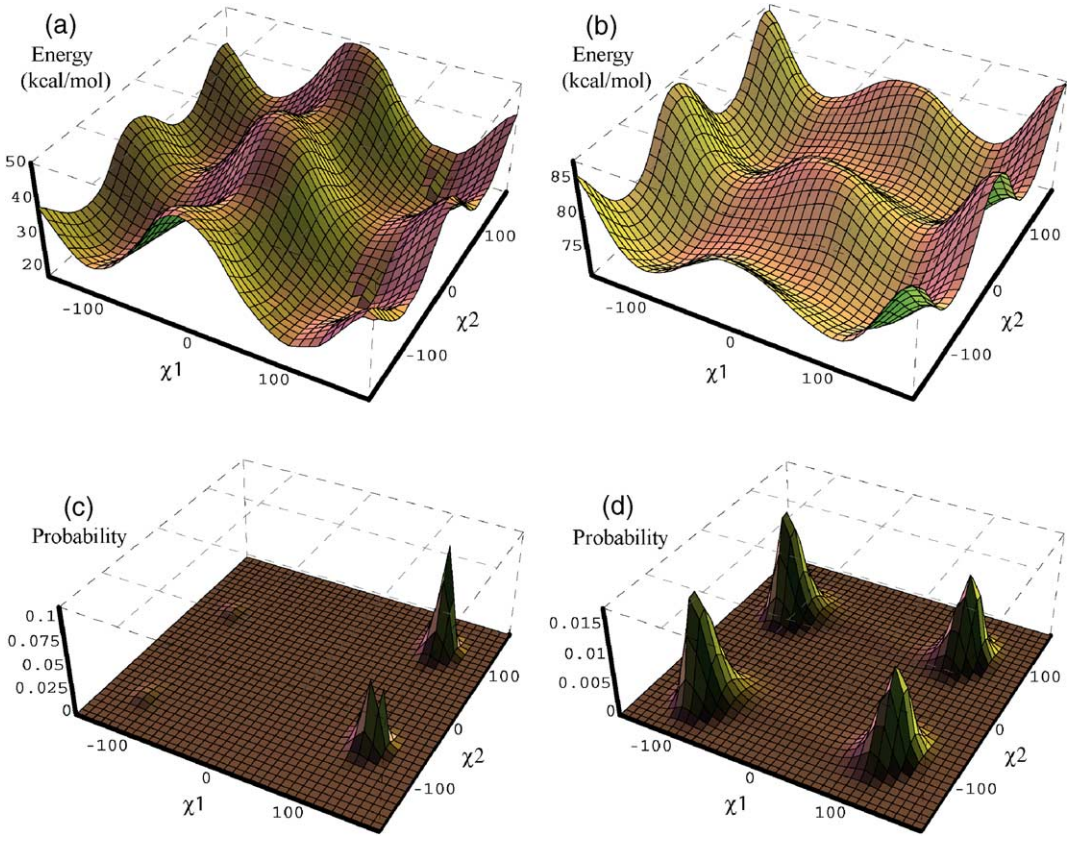

Fig. 2. The adiabatic energy and probability surfaces for structure 36ETHP. Panels A and C are the result of calculations performed using a gas phase dielectric constant. A dielectric constant of 80 was used in the calculations resulting in panels B and D.

In Method I, the activation entropy was considered to be composed of two parts: the change in entropy associated with restricting all ground state conformations (state S) to a single conformer identified as a near-attack conformer (NAC), $\Delta S_{\mathrm{NAC,S}}$, and the change in vibrational entropy upon conversion of the NAC to the transition state, $\Delta S_{\mathrm{TS,NAC}}$.

$$\Delta S^{\ddagger} = S_{\mathrm{NAC,S}} + S_{\mathrm{TS,NAC}} \tag{2}$$

Since in this series of compounds, all NACs and all TSs have similar geometries and charge distributions, the assumption was made that $\Delta S_{\mathrm{TS,NAC}}$ is either a constant or is proportional to $\Delta S_{\mathrm{NAC,S}}$. In either case, any correlation between the activation entropy and the logarithm of the rate constant would be observable as a correlation between $\Delta S_{\mathrm{NAC,S}}$ and the logarithm of the rate constant. Therefore, the quantity $\Delta S_{\mathrm{NAC,S}}$ was estimated as

$$S_{\mathrm{NAC,S}} = S_{\mathrm{NAC}} - S_{\mathrm{S}} \tag{3}$$

with

$$S_{\mathrm{S}} = -k \sum_i p_i \ln p_i \tag{4}$$

where the sum was carried out over all points of the energy landscape. These values are obviously not absolute entropies, but rather entropic terms dependent on the sampling interval chosen. Thus, to allow a comparison of $\Delta S_{\mathrm{NAC,S}}$ among the structures differing in the number of rotatable bonds, all sampling intervals were normalized to 10°.

To calculate S_{NAC}, the energy well comprising the largest number of conformations having NAC geometry was used for each compound. Slices through the energy minimum were taken, allowing one torsion to vary at a time. For each slice, a quadratic fit to the energy basin was determined and probabilities were calculated for points in a grid equivalent to that used for the ground state. A term S_{slice} was calculated using Eq. (4) and summing over the points of the slice. S_{NAC} was then computed as:

$$S_{\mathrm{NAC}} = \sum_i S_{\mathrm{Slice},i} \tag{5}$$

where the summation was carried out over all slices.

In Method II, the energy landscape was divided into bins corresponding to canonical conformers. Bin probabilities were calculated as the sum of the probabilities of all grid points within each bin. Bins were defined as follows: Rotations about a bond between two sp^3 carbons were divided into 3 equal parts corresponding to gauche$^+$, gauche$^-$, and *trans* rotamers. Rotations about bonds

between sp^2 and sp^3 atoms were divided into 6 bins. In this treatment of the activation entropy the transition state was modeled as a structure confined to a single rotameric bin, therefore having a conformational entropy of 0 (ln 1). The relative activation entropies of the members of the monoester series were then estimated as:

$$\Delta S^{\ddagger}_{\text{Conform.}} = \Delta S^{\text{Conform.}}_{\text{NAC,S}} = k \sum_i p_i \ln p_i \quad (6)$$

In this case the summation was not over individual conformations but rather over all conformers.

1.2. Results and discussion

According to absolute rate theory, reaction rates depend on the difference in free energy between a low energy basin (ground state) and the saddle point along the reaction coordinate (transition state) [12]. These two states are the only states relevant to the rate of reaction. In the calculations of Armstrong and Amzel [9], NACs were used as surrogates of the transition state. Entropies of activation were estimated by two different methods (Table 1), both employing adiabatic mapping of the energy landscape. Adiabatic mapping simplifies the computationally intractable problem of computing the energy landscape by isolating the significant degrees of freedom and minimizing with respect to the remainder. Low energy conformations contribute most to the entropy; therefore, points on the calculated surface represent the conformations with the specified torsion angles that contribute most to the entropy.

Method I estimates $\Delta S_{\text{NAC,S}}$, the entropic differences between all ground state conformers and a single ground state conformer (NAC) differing in entropy from the transition state by a value $\Delta S_{\text{TS,NAC}}$ assumed to be constant or proportional to $\Delta S_{\text{NAC,S}}$. In either case, a correlation between rate and the activation entropy would result in a correlation between rate and $\Delta S_{\text{TS,NAC}}$. The assumption that $\Delta S_{\text{TS,NAC}}$ is either constant or proportional to $\Delta S_{\text{NAC,S}}$ is

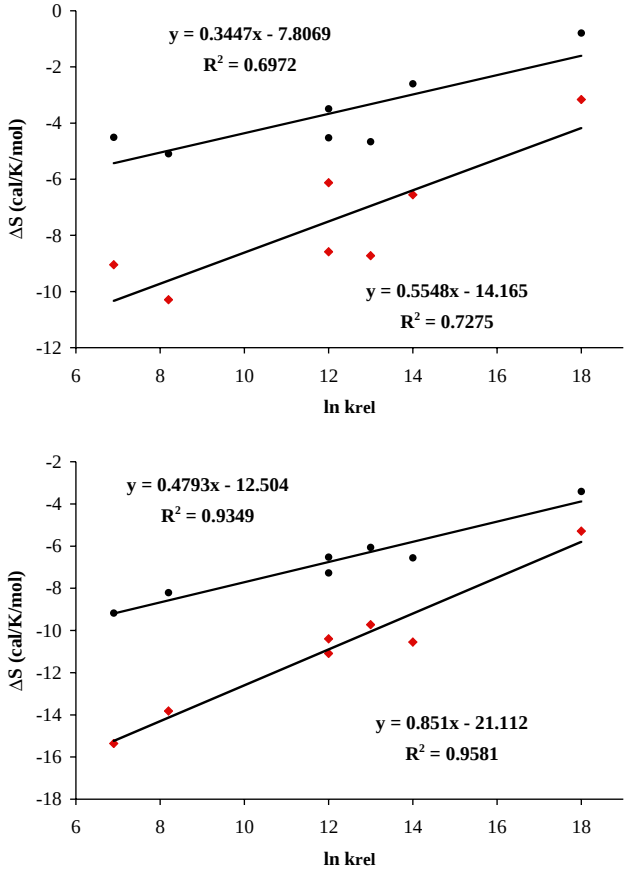

Fig. 3. Plots of calculated entropic terms against the logarithm of the experimentally determined relative rate constants. The top represents calculations performed using a gas phase dielectric; the bottom represents calculation performed with constant dielectric of 80. Method I is indicated by red diamonds; Method II is indicated by black circles.

Table 1
Experimentally determined relative rates of reaction and calculated entropic values for the seven structures considered

			EPS 1		EPS 80		ΔS_{Vib} (cal/K/mol)
			ΔS (cal/K/mol)		ΔS (cal/K/mol)		
			$-T\Delta S$ (kcal/mol)		$-T\Delta S$ (kcal/mol)		
	krel	ln krel	Method I	Method II	Method I	Method II	
GLUT	1.0E+03	6.9	−9.048	−4.505	−15.355	−9.180	−2.160
			2.698	1.343	4.578	2.737	
αGDMG	3.6E+03	8.2	−10.292	−5.090	−13.813	−8.210	−2.234
			3.069	1.518	4.118	2.448	
βGDEG	1.8E+05	12	−8.582	−4.523	−10.395	−6.526	−1.858
			2.559	1.349	3.099	1.946	
SUCC	2.3E+05	12	−6.125	−3.494	−11.090	−7.274	−1.117
			1.826	1.042	3.306	2.169	
βGDPG	2.7E+05	13	−8.727	−4.659	−9.725	−6.060	−1.709
			2.602	1.389	2.900	1.807	
βGDIG	1.3E+06	14	−6.555	−2.597	−10.545	−6.559	−2.487
			1.954	0.774	3.144	1.956	
36ETHP	8.0E+07	18	−3.163	−0.793	−5.294	−3.413	−0.352

All values are symmetry corrected using a factor of Rln2 due to the indistinguishable nature of OT1 and OT2 of carboxylate.

because molecular geometries [13] and charge distributions in the transition states are similar for all structures in this molecular series. In Method I, the term $\Delta S_{NAC,S}$ is a partial configurational entropy containing conformational as well as vibrational (torsional librations) components.

Strong correlations were found between the estimates of the relative activation entropies and the logarithm of the experimentally determined relative rate constants using either Method I or II (Fig. 3). Method I results in a better correlation, probably because it accounts for the details in the shapes of the conformational wells. For both methods, the correlations are stronger for calculations done using a dielectric constant of 80 than for those done using a dielectric constant of 1. The experimental rates were determined in a solution of 50% dioxane which has a dielectric value of between 26.8 and 43.9, the values reported for 60% and 40% solutions, respectively, [14]. The effect of using different dielectric constants is most easily seen in the differences between the two probability surfaces constructed for 36ETHP (Fig. 2). Four conformers, nearly equally populated, are observed when the calculations are done using a dielectric constant of 80, whereas calculations performed with a dielectric constant of 1 are dominated by exaggerated electrostatic interactions that result in only two significantly populated conformers (Fig. 2).

Calculations employing Method I using a dielectric of 80 yield values for $-T\Delta S^{\ddagger}$ with a maximum difference of 3.0 kcal/mol (Table 1), corresponding to a 160-fold rate enhancement. This is smaller than the experimentally determined range of rate enhancements of 8×10^4 and suggests that entropic considerations alone cannot fully account for the increased reaction rate. Lightstone and Bruice [1] estimated an enthalpy range ($\Delta(\Delta H)$) of 4.5 kcal/mol corresponding to a 2200-fold enhancement in rate for the same series of small molecules. Combination of these enthalpy and entropy estimates gives a range of enhancements of 3.5×10^5, in close agreement with the experimentally determined range.

2. Example 2: unfolding proteins by force

Single molecule experiments in which proteins are stretched by mechanical forces reveal a wealth of information about their mechanical properties [15–20]. Atomistic-level computer simulation [21,22,25–28] performed in conjunction with these experiments may reveal details of the mechanism of a protein's response to tensile loading. A combination of experimental [15,20,21,24,27,28] and computational [22,23,26,27] studies have revealed the structural changes experienced by TI I27, the 27 th immunoglobulin (Ig) domain of the I band of human muscle titin, as it unfolds when the ends of the polypeptide chain are pulled apart. Steered molecular dynamics (SMD) used in the computational studies involves the application of time-dependent external forces to single molecules [36] and has been applied to investigate the molecular events that take place when proteins are subjected to deformation, as in AFM or in optical tweezer experiments.

A direct comparison of pulling experiments with MD simulations is, however, often impossible because of the disparity between their time scales. At the much faster rate typical of MD simulations, proteins may unfold via mechanisms different from those occurring in experimental studies [29]. In experiments, the mean domain unfolding force usually exhibits a weak logarithmic dependence on the pulling speed. This dependence is generally expected when unfolding is driven by thermally activated barrier crossing [30].

Several approaches have been developed recently to extrapolate single-molecule stretching data to slower pulling speeds. Jarzynski [31,32] and subsequently Hummer and Szabo [33] showed that the equilibrium free energy dependence can, in principle, be reconstructed from single molecule experiments or simulations even when those are performed far from equilibrium. This approach has been tested experimentally [35] and by calculations [34], but since it involves averaging over multiple single-molecule trajectories, it is computationally expensive.

To overcome these difficulties, Pabon and Amzel [38] investigated the stretching and unfolding of TI I27 using an explicit solvent model with an alternative, speed-independent stretching simulation protocol.

2.1. Methods

All calculations were carried out with the program CHARMM [10,11]. The starting model was the energy-minimized average NMR structure of the Ig domain, TI I27 of the I-band of cardiac titin (PDB accession code 1TIT; [39]). The structure of TI I27 was solvated with 4680 pre-equilibrated TIP3P water molecules [40] and periodic boundary conditions were used. All atoms, including waters and hydrogens, were modeled explicitly resulting in a system containing 13 926 atoms.

Stretching of TI I27 was accomplished by fixing the Cα of the N-terminal leucine and applying an external force to the Cα of the C-terminal leucine. The force was applied by restraining harmonically the pulled end to a fixed restraint-point. After each full equilibration, the data were analyzed and the restraint-point was moved an additional 4 Å along in the line connecting the N- and C-terminal atoms. This process was repeated until a final extension of 17 Å was reached. At each extension, the calculation was repeated with the spring constant associated with the restraining force set to five different values between 0.5 and 5 kcal/mol Å to explore the dynamical response of the structure.

2.2. Results

The force versus extension curve (Fig. 4) shows a peak of about 1440 pN at an extension of 15.5 Å. This peak

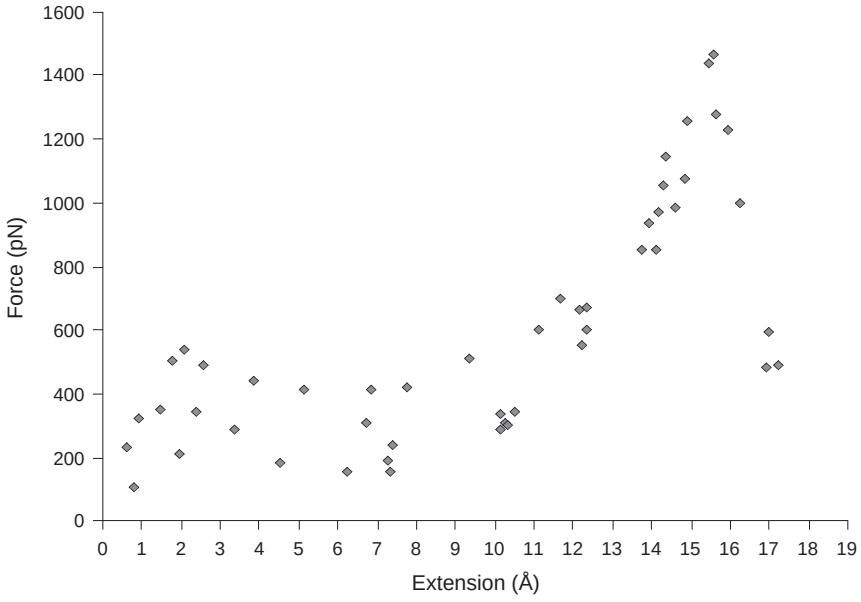

Fig. 4. Force extension profile of the I27 domain observed in our simulation. The stretching process was stopped at 16.8 Å.

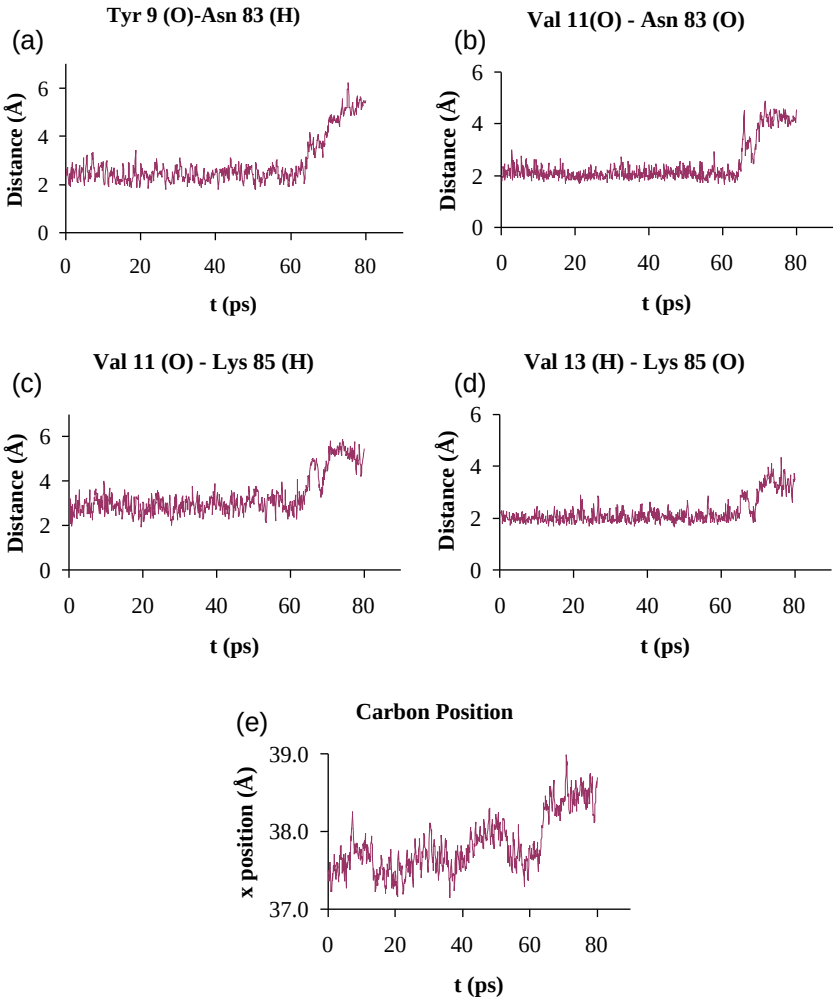

Fig. 5. Hydrogen bonding distances of β-strands A′, G of I27. (A) Tyr 9 (O)–Asn 83 (H); (B) Val 11 (H)–Asn 83 (O); (C) Val 11 (O)–Lys 85 (H); (D) Val 13 (H)–Lys 85 (O); (E) The *x* position of Leu 89 Cα recorded during the first 80 ps of equilibration.

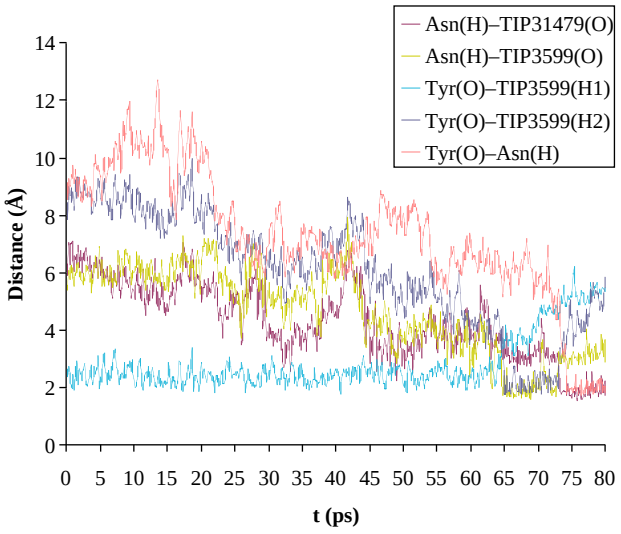

Fig. 6. Distances behavior between Tyr 9 (O), Asn 83 (H), TIP3599, and TIP31479 water molecules during the first 80 ps of simulation.

corresponds to the main energy barrier separating the folded and the unfolded states of the domain. After the barrier is overcome, the force decreases rapidly to average values around 300 pN. The maximal extension reached during the procedure was 16.8 Å.

An analysis of the trajectory corresponding to the equilibration phase at the extension of the maximal force reveals that, four hydrogen bonds between β-strands A′and G break concurrently, as pointed out in [25]. The O–H distances of the backbone H-bonds between β-strands A′and G during this event are shown in (Fig. 5). The remaining H-bonds between β-strands A′and G and between β-strands A and B break at an early step (not shown).

A full analysis of the H-bond behavior shows that they are stable during the process, although they break and reform often. Of the six H-bonds between A′ and G, the middle four bonds are very stable, but the two at the ends of this β-sheet are more flexible.

Figs. 6 and 7 show a breaking event for the H-bond between Tyr 9 (O) and Asn 83 (H). Until 65 ps of the equilibration run, this backbone H-bond is still formed; the closest water molecule is at least 3 Å away from either of the H-bond partners (Fig. 7A). At 65 ps, a water molecule approaches and forms H-bonds with Tyr 9 (O) and Asn 83 (H) simultaneously (Figs. 6 and 7B) and immediately after, the backbone H-bond between Tyr 9 (O) and Asn 83 (H) is broken. Although it remains permanently broken, the O–N distance remains at just over 4 Å to allow both groups to make H-bonds to the same water molecule. At 75 ps, a second water approaches and forms a H-bond with Asn 83 (H) (Fig.

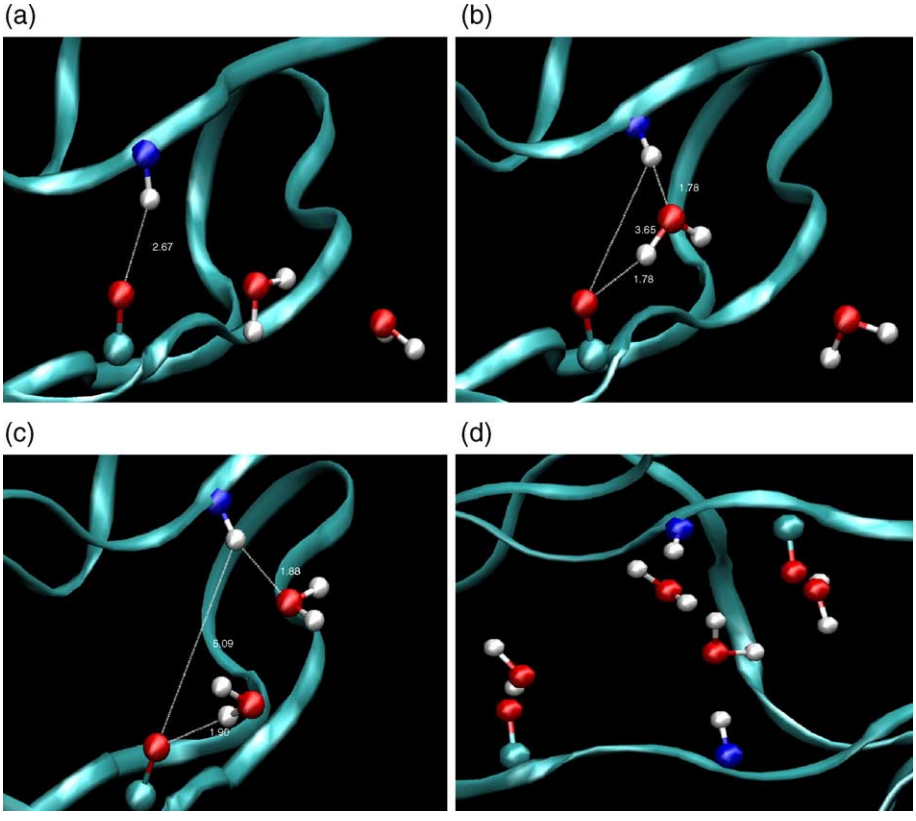

Fig. 7. (A)–(C) Snapshots of the H-bond behavior between Tyr 98 (O) and Asn 83 (H). Two water molecules are involucred in the H-bond breaking event (see text for explication). (D) Snapshot of the H-bond behavior between Val 11 (H) and Asn 83 (O). These figures were created with VMD [37].

7C). Several picoseconds later the bonds are still formed and the backbone O–N distance increases beyond 5 Å. Similar behavior is shown by the H-bond between Val 11 (H) and Asn 83 (O): two water molecules are involved in keeping it permanently broken (Fig. 7D).

2.3. Discussion

AFM and optical tweezers have been used to successfully monitor stretching of single proteins; however, understanding of the molecular mechanism of unfolding by stretching remains elusive. Molecular dynamics simulations that are consistent with observations can provide hints about the mechanism of this unfolding. However, because of the short simulation times imposed by limitations of computer resources, SMD simulations must stretch Ig domains at speeds six to eight orders of magnitude higher than in AFM and optical tweezer experiments. As a result, SMD simulations yield force peaks that are one order of magnitude greater than the experimental values. This is because SMD simulations operate in regime where forces are large enough to pull the system above energy barriers, whereas AFM experiments operate in a regime in which forces are low but there is enough time to allow the system to thermally equilibrate, thus lowering the barriers and promoting crossing within the millisecond experimental time scale [26]. Pabon and Amzel [38] carried out calculations that reflected more closely the conditions of the experimental work, namely, short extensions followed by long equilibrations.

The simulations of Pabon and Amzel [38] show features of the unfolding process and force profiles similar to those reported by others [25,26], but with significant differences in the force peak value and in the details of the timing of the events. Monitoring of hydrogen bond participants reveals that resistance of TI I27 to unfolding under external forces can be attributed to inter-β-sheet backbone hydrogen bond breaking, with water molecules playing a key role. Water molecules continuously form and break H-bonds with backbone oxygens and NH-hydrogens. During the equilibration (Fig. 5), backbone H-bonds fluctuate between formed and broken states due to attacks by water molecules but remain intact for most of the time. Under stretching forces, the protein has a tendency to slightly elongate and backbone H-bonds become less stable. Under these conditions, water molecules have a better chance of approaching and forming H bonds with backbone atoms. Each backbone–backbone H-bond breaks only when one or two water molecules have made H-bonds to both donor and acceptor of the original bond. Full separation of the strands require that all NH and CO group Hydrogen bond to water molecules. Therefore, with long equilibration times water molecules do not replace backbone H-bonds *after* they break instead they must hydrogen bond to the backbone groups *before* the backbone H-bonds can break.

3. Example 3: translational entropy of binding in solution

The ability to predict binding affinities of molecules in solution is essential to structure-based drug design [41]. Enthalpic (ΔH) and entropic (ΔS) effects both contribute to affinity and it is desirable to estimate them from first principles, using only the structures of the molecules participating in the reaction and an adequate force-field. ΔH has been recently studied and parameterized using the polarity of the molecular surfaces at play [42]. It is important to develop simple and efficient frameworks to estimate binding entropies, usually decomposed into translational, rotational, vibrational and conformational contributions. Among these, the *translational* entropy is uniquely dependent on solute concentration and on complex interactions between solute and solvent. This latter feature complicates considerably the evaluation of ΔS and has led us to propose a statistical–mechanical framework [43,44] that allows efficient predictions of translational entropies of solutes.

This section reviews our previous results for test cases (pure water, solutes in water) in which we used molecular dynamics (MD) simulations to compute the translational entropy. These results led us to suggest a correlation between the translational entropy of a solute and its polarity and size that can be used for routine calculations, in the same spirit as the correlation developed for binding enthalpies [45]. Calculation of the translational entropy loss upon binding of benzene to a cavity-mutant of T4-lysozyme will be discussed as an example of a small molecule binding to a protein. We also use a normal-mode analysis to demonstrate that the motion of benzene within the binding site of lysozyme can indeed be treated as a translation and compare these results with a recent estimate based on a body-restraint algorithm for MD simulation [46].

3.1. Introduction

The loss of translational entropy when two molecules A and B participate in the reaction $A+B \rightleftharpoons AB$ in solution (e.g., A=ligand, B=protein, AB=complex) is given by $\Delta S^{\mathrm{trsl}}=S_{AB}^{\mathrm{trs}l}-S_A^{\mathrm{trsl}}-S_B^{\mathrm{trsl}}$. The notation S_ξ^{trsl} ($\xi=A$, B or AB) represents the entropy of the species ξ *in solution* at a 1 M standard concentration. Estimates of ΔS^{trsl} require adequate estimates of each term of the type S_ξ^{trsl}. From classical statistical mechanics we have [12]

$$S_\xi^{\mathrm{trsl}} = k\left(\ln Q_\xi^{\mathrm{trsl}} + T\frac{\delta \ln Q_\xi^{\mathrm{trsl}}}{\delta T}|_V\right) \tag{7}$$

where T is the temperature, and k Boltzmann's constant. Q_ξ^{trsl} is the translational partition function of the N_ξ solute molecules and can be expressed in terms of q_ξ^{trsl}, the translational partition function of one solute molecule, as

$Q_\xi^{\mathrm{trsl}}=(q_\xi^{\mathrm{trsl}})N_\xi/N_\xi!$ for indistinguishable molecules (e.g. gas or liquid state) or $Q_\xi^{\mathrm{trsl}}=(q_\xi^{\mathrm{trsl}})N_\xi/N_\xi^{N_\xi}$ otherwise (e.g. crystalline solid state or bound molecule). Therefore, even at the same density, the entropy of liquid molecules is higher than that of a solid by a term kN_ξ, sometimes called "communal entropy" [47]. This reflects the ability of molecules in a liquid to occupy any position in a given volume, whereas those in a solid are fixed [12]. Eq. (7) assumes that the different contributions (translational, rotational, vibrational) are separable, a common assumption for many-body calculations [12]. Moreover, the total partition function of the system also contains terms associated with degrees of freedom of the solvent, contributing to hydration terms that are parameterized independently and were therefore excluded here from the translational entropy term. The partition function of a solute molecule is given by

$$q_\xi^{\mathrm{trsl}}=\frac{1}{\Lambda_\xi^3}\int_V e^{-\beta U(\mathbf{q}_\xi,\{\mathbf{q}_\mathbf{w}\})}\mathrm{d}\mathbf{q}_\xi \tag{8}$$

where $\Lambda_\xi=\frac{h}{\sqrt{2\pi m_\xi kT}}$ is the De Broglie wavelength [12] and V is the total volume. $\mathbf{q}_\xi$ and $\{\mathbf{q}_\mathbf{w}\}$ are the positions of the solute and the set of water molecules, respectively. Evaluation of Eq. (7) requires evaluation of the configurational integral in Eq. (8), which can only be carried out using approximations. Assuming that the intermolecular potential $U(\mathbf{q}_\xi,\ \{\mathbf{q}_\mathbf{w}\})=0$ leads to the gas-phase "Sackur–Tetrode" (ST) formula, sometimes used as a crude estimate of the translational entropy in solution [48–50].

$$S_{\xi,\mathrm{ST}}^{\mathrm{trsl}}=k\left(N_\xi\ln\frac{V}{N_\xi\Lambda_\xi^3}+\frac{5}{2}N_\xi\right) \tag{9}$$

Siebert and Amzel recently proposed [43,44] a framework based on the cell model (CM) for condensed phases [47,51,52] to approximate simply but efficiently the configurational integral. CM considers that the center of mass of each molecule (solute or solvent) is confined in a cell delimited by the field of the surrounding solvent molecules leading to the concept of "free volume"

$$v_{\mathrm{f},\xi}=\int_{\frac{V}{N_\xi+N_w}} e^{-\beta U\left(\mathbf{q}_\xi,\{\mathbf{q}_\mathbf{w}\}\right)}\mathrm{d}\mathbf{q}_\xi \tag{10}$$

and to the translational entropy of the N_ξ solute molecules immersed in N_w solvent molecules [44]

$$S_{\xi,\mathrm{CM}}^{\mathrm{trsl}}=k\left(N_\xi\ln\frac{(N_w+N_\xi)v_{\mathrm{f},\xi}}{N_\xi\Lambda_\xi^3}+\frac{5}{2}N_\xi\right) \tag{11}$$

In the limiting case of an ideal gas of N_ξ solute molecules without solvent ($N_w=0$), the free volume is $v_{\mathrm{f},\xi}=V/N_\xi$, and we recover the ST expression in Eq. (9).

The problem of evaluating the translational entropy is now reduced to the evaluation of the free volume using Eq. (10). As the intermolecular interactions in solution are dominated by the short-range repulsive part of the potential $U(\mathbf{q}_\xi,\ \{\mathbf{q}_\mathbf{w}\})$, the exponential factor in Eq. (8) vanishes except on a small portion of the space, which makes the calculation tractable. With this framework, one can estimate the terms S_A^{trsl} and S_B^{trsl} directly using Eq. (11). The term S_{AB}^{trsl} takes into account the free volumes of A and B in the complex (A_c and B_c), which can be calculated using Eq. (11) if the integral giving the free volume of B_c is restricted to the portion of space where B_c is bound to A_c [44]. This obliterates the "communal entropy" of B_c, lowering its entropy by a factor kNB_c.

Thus loss of translational entropy upon binding, expressed at a standard state of 1 M of solute in 55 M of water (N_av being Avogadro's number), is given by

$$\Delta S_{\mathrm{CM}}^{\mathrm{trsl}}=kN_\mathrm{av}\left(-\ln 56+\ln\frac{v_{\mathrm{f},A_\mathrm{c}}v_{\mathrm{f},B_\mathrm{c}}}{v_{\mathrm{f},A}v_{\mathrm{f},B}}\right)-kN_\mathrm{av} \tag{12}$$

Conservation of mass imposes cancelation of the terms corresponding to the integration over the momenta [53,54].

3.2. Methods

3.2.1. Molecular dynamics simulation

Molecular dynamics (MD) simulations are used to evaluate the energy of interaction of the solute with the solvent, $U(\mathbf{q}_\xi,\ \{\mathbf{q}_\mathbf{w}\})$ in Eq. (10). The CHARMM28 force field [10,11] and the SHAKE algorithm [55] were used with an integration time step of 2 fs, saving snapshots of configurations every 2 ps for analysis. The temperature

Table 2
Comparison of free volumes (in Å^3) for various solutes in water, for a molecule of pure water, and for benzene in the TCM T4-lysozyme

Molecule i (by decreasing size)	$\mathrm{ASA_T}$ [Å^2]	$\mathrm{ASA_P}$ [Å^2]	$\bar{v_\mathrm{f}}\pm\sigma(v_\mathrm{f})$ [Å^3]
4-methyl, 1-hydroxyl naphthalene	322.8	47.17	0.053±0.030
Naphthylhydroquinone	314.4	94.71	0.030±0.013
Naphthylquinone	312.8	90.96	0.049±0.024
Naphthalene	286.7	0.000	0.146±0.084
Quinone	255.5	102.2	0.059±0.031
Glycerol	238.2	154.6	0.030±0.013
Benzene	225.0	0.000	0.245±0.147
Methylacetate	218.8	64.46	0.120±0.059
Methanol	148.9	61.54	0.129±0.073
Pure water	113.1	113.1	0.099±0.088
Benzene in TCM T4-lysozyme			0.095±0.036

$\bar{v_\mathrm{f}}=\frac{1}{N}\sum_{i=1}^{N}v_\mathrm{f}^i$ is the average free volume over all snapshots and $\sigma(v_\mathrm{f})$ $\left(\frac{1}{N}\sum_{i=1}^{N}\left(v_\mathrm{f}^i-\bar{v_\mathrm{f}}\right)^2\right)^{1/2}$ its standard deviation. $\mathrm{ASA^T}$ and $\mathrm{ASA^P}$ are the total accessible surface area and the polar accessible surface area of the molecule, respectively.

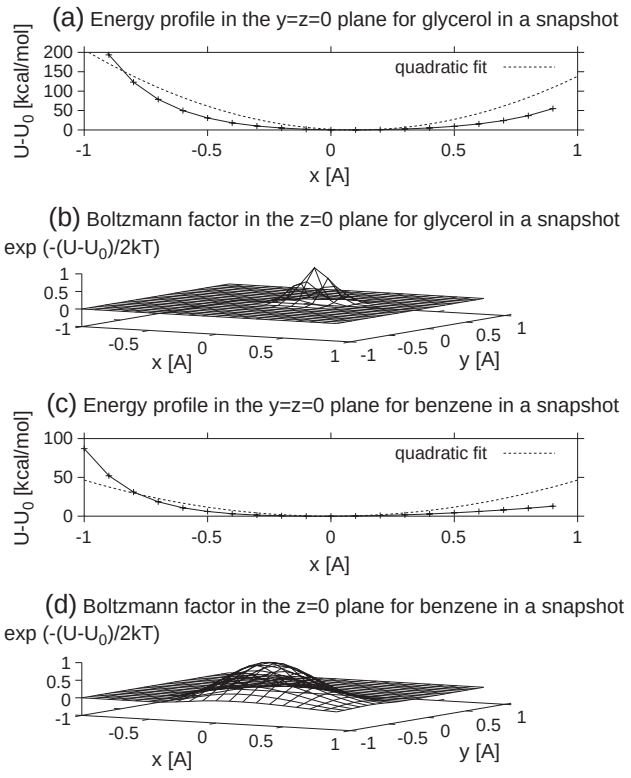

Fig. 8. Illustrations of the restriction of the motion of glycerol (A, B) and benzene (C, D) in the field of the water molecules: (A) One-dimensional energy profile in the plane $y=z=0$, for displacements of the glycerol molecule along the x-axis. A quadratic fit to the data illustrates the anharmonic behavior of the energy. (B) Two-dimensional profile of the Boltzmann factor $\exp\left[-\frac{\Delta U(x,y,0,\{\mathbf{q}_w\})}{kT}\right]$ for glycerol moving along x and y in the plane $z=0$. (C)–(D) The corresponding plots for benzene.

was maintained at 300 K, and a cutoff of 12 Å was used for non-bonded interactions.

3.2.1.1. Solute in water. Trajectories for 512 water molecules (TIP3 model) placed in a cubic box of side length = 24.8346 Å were run for 800 ps. The solute (see Table 2 for the list of solutes) was then inserted, removing the water molecules that overlapped with it. As the 12 Å cutoff is slightly less than half the box side length, solute molecules do not interact with their images. This mimics an ideal, infinitely dilute solution [48,53,56,57]. The standard state concentration of 1 M solute is obtained by choosing adequates N_ξ and N_w in Eq. (11). Trajectories of 800 ps were run after equilibration.

3.2.1.2. Benzene binding to a mutant T4-lysozyme. A mutant T4-lysozyme L99A, C54T, C97A (or TCM, for Triple Cavity Mutant) was designed to create a hydrophobic cavity large enough to accommodate a benzene molecule [58]. The 1.9 Å crystal structure (PDB entry: 1L84, [58]) was used in the simulation. Small harmonic constraints (2.4 kcal/mol Å^2 force constant) were imposed on the Cα atoms to avoid unfolding of the protein but still allow the ligand to move in the cavity during 800 ps of production.

3.2.2. Normal mode analysis

The motion of benzene bound to TCM T4 lysozyme was also analyzed using the normal modes analysis facility (Vibran) of the CHARMM28 program [10,11]. The TCM T4-lysozyme was kept fixed, and the normal modes were

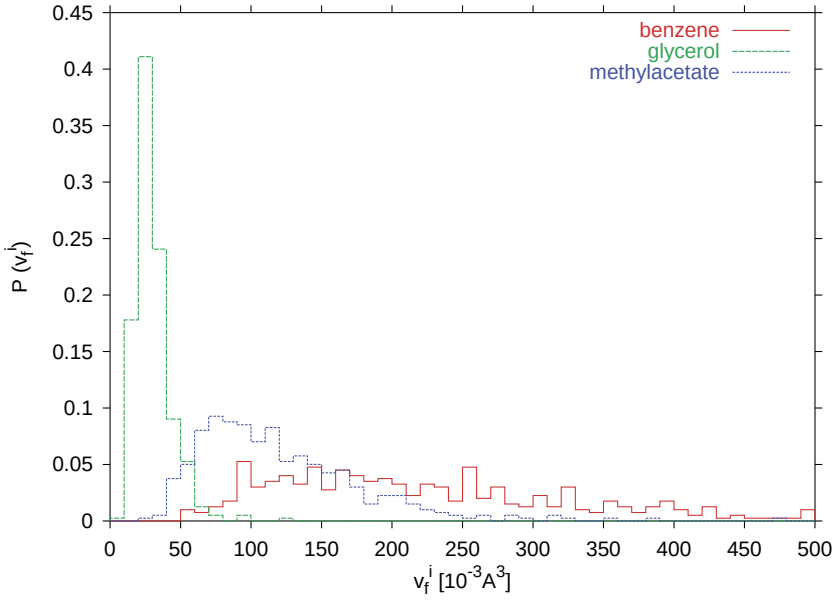

Fig. 9. Normalized probability distributions of free volume for three solutes of similar size in water: Glycerol (green), Methylacetate (blue) and Benzene (red). v_f^i data are binned in 10 Å^3 increments to generate the histograms. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

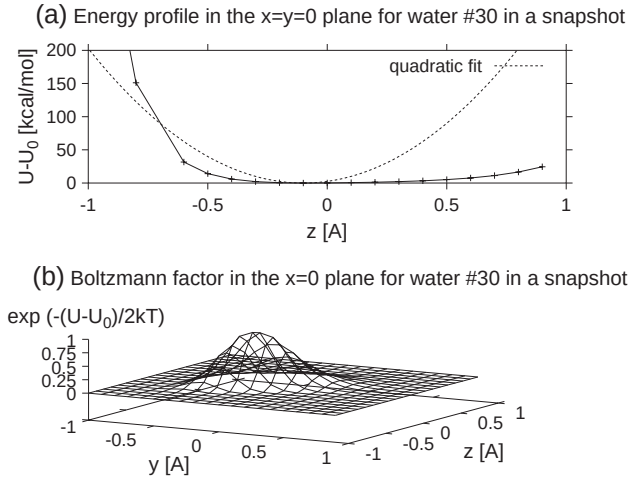

Fig. 10. Illustrations of the restriction of the motion of a water molecule in the field of the other molecules: (A) One-dimensional energy profile in the plane $y=z=0$, for displacements of a water molecule (number #30) along the x-axis. A quadratic fit to the data illustrates the anharmonic behavior of the energy. (B) Two-dimensional profile of the Boltzmann factor $\exp\left[\frac{\Delta U(x*,y*,z*,\{\mathbf{qw}\neq\mathbf{w}^*\})}{2kT}\right]$ in the plane $z=0$, for the same water molecule.

calculated in the reduced basis corresponding to the motion of the benzene in the field created by the fixed lysozyme.

3.3. Results and discussion

3.3.1. Solute ξ in water

The various solutes used in this analysis (Table 2) were chosen for their different sizes and patterns of surface polarities. The free volume of a solute ξ in water reflects restriction of the motion of its center of mass and depends on the interaction energy $U(\mathbf{q}_\xi, \{\mathbf{q}_\mathbf{w}\})$ between the solute and the field created by the surrounding water molecules. This interaction energy was calculated for each snapshot on a three-dimensional grid (spacing $\delta_x=\delta_y=\delta_z=0.1$ Å) centered around the equilibrium position of the solute [44]. Fig. 8A and C show one-dimensional slices (along the x axis, in the $y=z=0$ grid plane) of the interaction energy $U(x, 0, 0, \{\mathbf{q}_\mathbf{w}\})$, as well as a quadratic fit to the energy profile. The figure emphasizes the anharmonic character of the energy profile for solutes as different as glycerol (Fig. 8A) and benzene (Fig. 8C). Fig. 8B and D show the 2D-profile of the exponential factor $\exp\left[-\frac{\Delta U(x,y,0,\{\mathbf{q}_\mathbf{w}\})}{kT}\right]$ and illustrate that the translation of the solute molecule in the field created by the surrounding water molecules is restricted to less than 1 Å in each direction, at which point the Boltzmann factor falls to zero. Due to the high number of H-bonding groups, the motion of glycerol in water is more restricted than that of the nonpolar benzene.

The corresponding energy profiles for methylacetate (data not shown) are of the same order of magnitude and present features intermediate those of glycerol and benzene, consistent with the respective surface polarities of these molecules. The free volumes v_f^i, computed for each snapshot i, follow the distributions plotted in Fig. 9 for three selected solutes of comparable size (glycerol, methylacetate, benzene).

The distribution for glycerol (Fig. 9) shows a sharp peak at low v_f^i, reflecting the tight packing of water molecules hydrogen-bonded to the glycerol, and a small tail at large v_f^i that extends to about 0.1 Å^3. In contrast, the histogram for benzene is broader, with a much smoother cutoff at low v_f^i and a wider tail at large v_f^i. The average free volume for benzene shifts towards higher values (see Table 2). This reflects the looser packing of the clathrate-

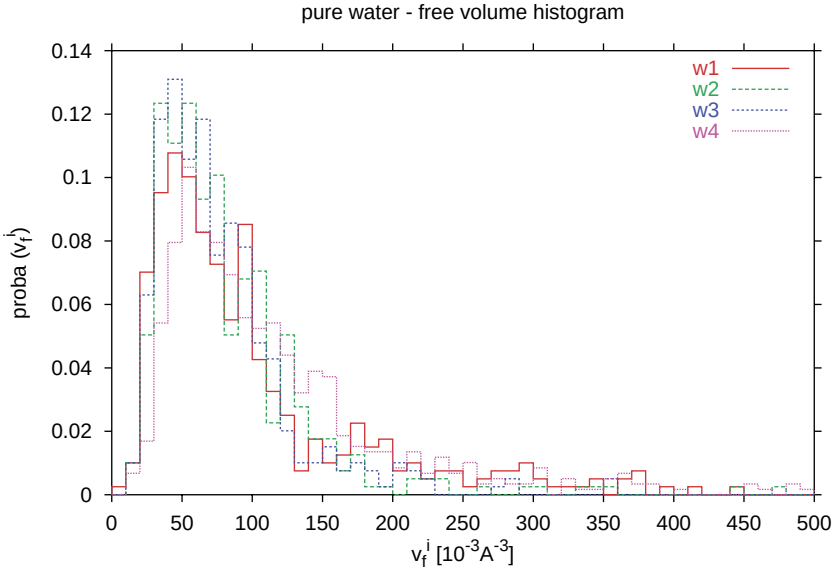

Fig. 11. Normalized probability distributions of free volumes for 4 different water molecules, calculated for snapshots taken every 2 ps during the 800 ps simulation. v_f^i data are binned in 10 Å^3 increments to generate the histograms.

Table 3
Comparison of the experimental values (EXP) with our cell model (CM) and with the Sackur–Tetrode (ST) estimates

	ST	CM	EXP
S_w^{total}	25.15→30.8	13.75→18.4	16.7
ΔS_{vap}^{total}	14.27→20.16	21.02→26.57	28.29

S_w^{total} is the entropy of liquid water at 298K and ΔS_{vap}^{total} is the entropy of vaporization of water at 298 K. The arrows indicate boundaries calculated using lower and higher bounds of the rotational entropy S_w^{rot} as in [44]. Values are in entropic units (e.u.=cal/mol K).

forming water molecules around the hydrophobic benzene molecule. For methylacetate, the situation is intermediate that for glycerol and benzene because of the presence of both hydrophobic (methyl) and hydrophilic (acetate) groups.

For each solute considered in this study (Table 2), the free volume is very small compared to the size of the solute itself, because it corresponds to the volume defined by the movement of the *center of mass* of the solute molecule. The large standard deviations could be thought to arise from using a limited number of steps in the simulations. However, the histograms after 400 ps and 800 ps are very similar (data not shown), indicating that these deviations are a reflection of the local variations in density in the liquid state. Similar calculations for pure water show the same phenomenon.

3.3.2. The case of pure water

As detailed in [44], CM calculations for pure water require a slight adaptation of Eq. (10) since in this case the solute itself is a water molecule, which violates the assumption that the solute molecules do not interact with each other (ideal solution).

The energy profiles of Fig. 8 (glycerol or benzene in water) and Fig. 10 (pure water) reveal that the range of motion of the center of mass of a molecule of glycerol or benzene in water is similar to that of a water molecule (less than 1 Å in each direction). The distributions of $v_{f,w}^{*}$ for 4 different water molecules (labeled w1, w2, w3 and w4, respectively) are shown in Fig. 11.

All four histograms in Fig. 11 have similar widths and are skewed to the right. The sharp cutoff at low v_f^i corresponds to densely packed regions, and the tail at high v_f^i to low-density regions. As in the previous section, the width of the distribution reflects the local density variations in liquid water [60].

The average value of the free volume from the combination of these four histograms is 0.099 Å^3, with a standard deviation of 0.088 Å^3. With this value, the translational entropy of water is $S_{w,CM}^{trsl}=8.90\pm1.76$ e.u.. The ST estimate (Eq. (9)) at the density of liquid water gives $S_{w,ST}^{trsl}=20.3$ e.u.. These values are compared in Table 3 with the experimental values for the entropy of liquid water and for the entropy of vaporization.

Clearly, CM provides a much better approximation of the translational entropy of water than ST. This is not very surprising, considering that ST is derived from ideal gas statistical mechanics, neglecting the interactions between molecules, whereas CM takes the intermolecular interaction explicitly into account in calculating the free volume, as seen in Eq. (11).

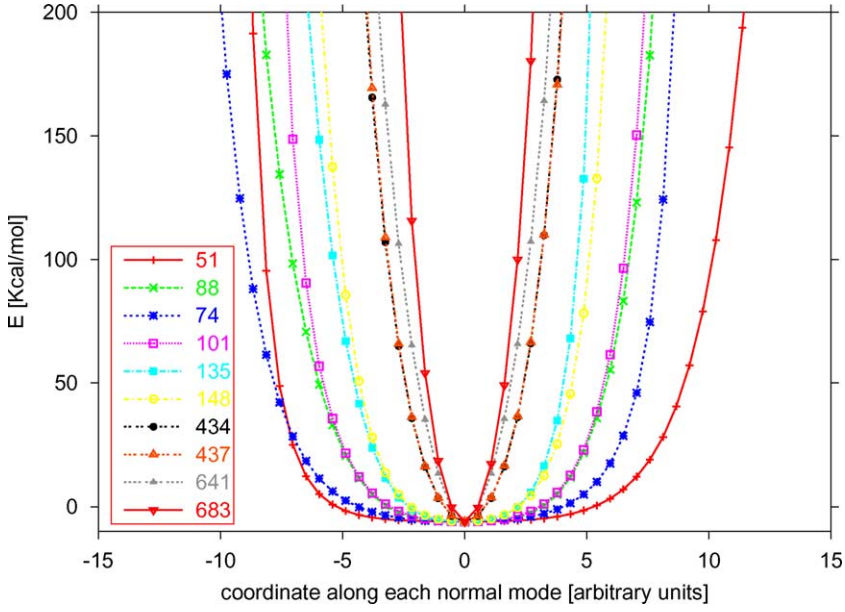

Fig. 12. Energy profile along the ten lowest-frequency modes. The frequency of each mode (in cm^{-1}) is listed in the figure. Modes 1 to 6 (51 to 148 cm^{-1}) exhibit a relatively flat energy minimum corresponding to translations/rotations, whereas modes 7 to 10 (434 to 683 cm^{-1}) correspond to higher frequency harmonic motions.

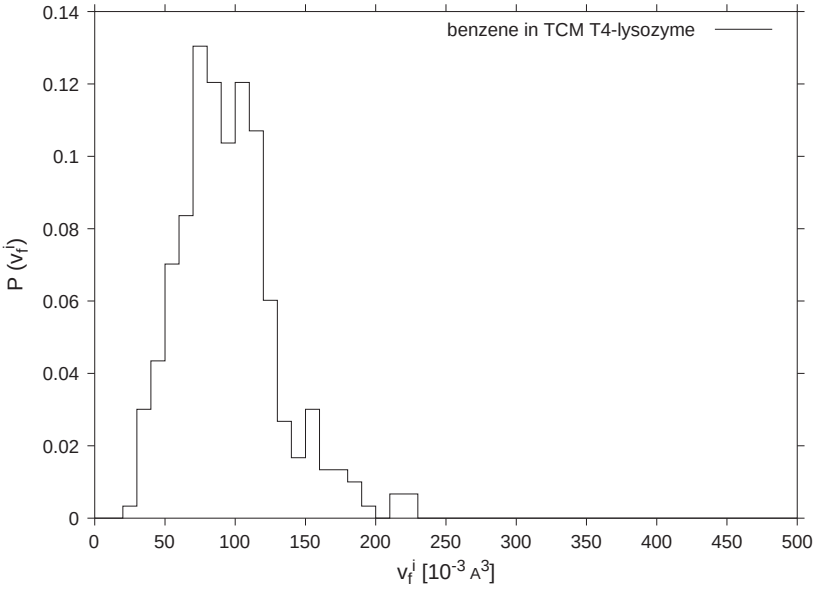

Fig. 13. Normalized probability distributions of free volume for benzene in TCM T4-lysozyme. v_f^i data are binned in 10 Å^3 increments to generate the histogram.

3.3.3. Benzene binding to the TCM T4-lysozyme

3.3.3.1. Normal mode analysis. Fig. 12 presents the energy profile for small displacements of the benzene molecule in the cavity of the TCM T4-lysozyme along each of the ten lowest-frequency modes. Modes 1 to 6 exhibit frequencies about 4 times lower than the next modes, and their energy profile has a relatively flat minimum corresponding to anharmonic translational/rotational motions (similar to a particle-in-a-box motion). Modes 7 to 10 correspond to low frequency internal modes, with a clear harmonic behavior. This analysis shows that the motion of the benzene within the binding site of the TCM T4-lysozyme can be best described as a combination of translations.

3.3.3.2. Free volume. Fig. 13 shows a histogram of the free volume of benzene in the hydrophobic cavity of

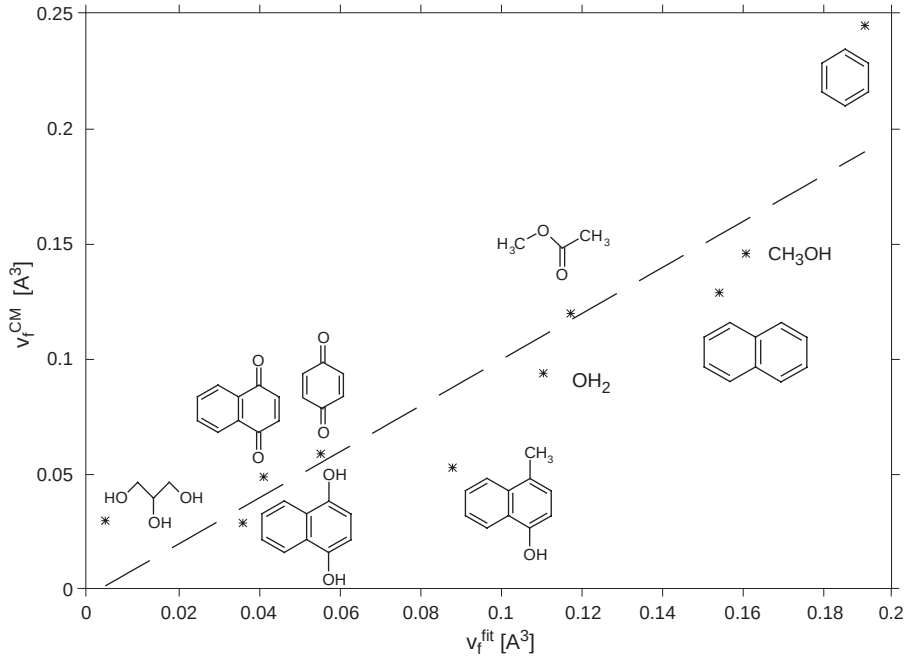

Fig. 14. Free volume calculated with the cell model (v_f^{CM}) vs. free volume calculated from a correlation with total and polar surface areas (v_f^{fit}) for the various solutes used in this study as well as for pure water. The perfect fit is given by the line ($v_f^{CM}=v_f^{fit}$).

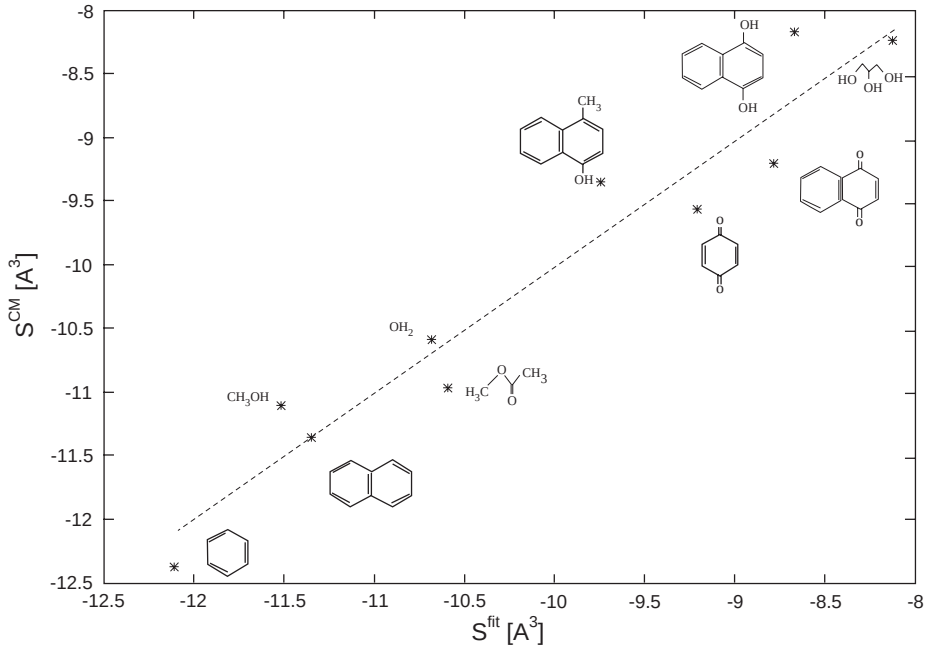

Fig. 15. Translational entropy loss: calculation with the cell model (S^{CM}_{trsl}) vs. calculation from a correlation with total and polar surface areas (S^{fit}_{trsl}) for the various solutes used in this study as well as for pure water. The perfect fit is given by the line ($S^{CM}_{trsl}=S^{fit}_{trsl}$).

TCM T4-lysozyme. A comparison of Figs. 13 and 9 indicates that the average free volume of benzene in the TCM T4-lysozyme is much smaller than that of benzene in water (see Table 2). Moreover, the general shape of the histogram in Fig. 13 is similar to that of glycerol in water. This reflects the tight encapsulation of the benzene by the hydrophobic cavity as well as the higher density of the protein interior (about 1.4 times that of water [61]).

Using Eq. (12) and considering that the free volume of the TCM T4-lysozyme is similar to that of the complex, we get $\Delta S^{trsl}_{CM} \approx 11.78$ e.u.. Hermans and Wang [46] estimated the rms displacements of the benzene in the binding pocket using MD simulations and obtained a value of 0.13 Å for the motion perpendicular to the benzene ring and 0.3 Å in two orthogonal directions in the plane of the ring corresponding to a volume of about $0.26\times0.6\times0.6=0.0936$ Å^3, in excellent agreement with our free volume calculations.

3.3.4. Parameterization

The results obtained for the free volumes calculated with CM (Table 2) suggest some degree of correlation between the free volume and the hydrogen-bonding ability or polar surface area of the solute calculated with the program Areaimol [59] using a probe radius of 1.4 Å. Fig. 14 shows a fit of the free volume to a linear combination of polar (ASA_P) and total (ASA_T) surface areas of the solute,

$$v^{fit}_f = a\ ASA_P + b\ ASA_T + c$$

The value of the fitting parameters are:

$$a = -1.18\times10^{-3}\ \text{Å}\quad b = -4.78\times10^{-4}\ \text{Å}\quad c = 0.2978\ \text{Å}^3$$

This indicates that for the range of solutes considered in this study, the maximum free volume is $c=0.2978$ Å^3. It also indicates that the free volume of a molecule is more influenced by its polarity than by its size (i.e. $a>b$). This correlation applies to molecules with ASA_T smaller than about 600 Å^2. Larger molecules, including proteins, are expected to have a free volume of approximately 0.03 Å^3 or smaller.

To transform the correlation between the free volume and the polarity of the solute into a correlation directly involving the entropy (Eq. (12)), we consider the fictitious case where each solute binds to the same protein of average density 1.4 g/cm^3. The correlation shown on Fig. 15 takes the form

$$\Delta S^{fit}_{trsl} = d\ ASA_P + e\ ASA_T + f$$

The parameters are:

$$d = 0.02\frac{\text{e.u.}}{\text{Å}^2}\quad e = 0.012\frac{\text{e.u.}}{\text{Å}^2}\quad f = -14.85\text{e.u.}$$

Although more data on several other solutes are necessary to establish this correlation more precisely, this promising result opens the possibility of obtaining estimates of the translational entropy of a solute in water, bypassing the MD simulations. We therefore propose the CM as a

simple method to evaluate the loss of translational entropy upon molecular association.

Acknowledgments

We acknowledge the National Cancer Institute for allocation of computing time and support staff at the Frederick Biomedical Supercomputing Center of the Frederick Cancer Research and Development Center. This work was supported by grants GM51362 and GM45540 from NIGMS. We thank Julio Fernandez for helpfull discussions. The work of AAA was supported by the Burroughs Wellcome Fund through its funding of the Johns Hopkins Program in Computational Biology.

References

[1] F.C. Lightstone, T.C. Bruice, Ground state conformations and entropic and enthalpic factors in the efficiency of intramolecular and enzymatic reactions: 1. Cyclic anhydride formation by substituted glutarates, succinate, and 3,6-endoxo-delta(4)-tetrahydrophthalate, J. Am. Chem. Soc. 118 (1996) 2595.

[2] T.C. Bruice, F.C. Lightstone, Ground state and transition state contributions to the rates of intramolecular and enzymatic reactions, Acc. Chem. Res. 32 (1999) 127.

[3] W.P. Jencks, Catalysis in Chemistry and Enzymology, McGraw-Hill, New York, 1969.

[4] P.G. Schultz, Catalytic antibodies, Acc. Chem. Res. 22 (1989) 287.

[5] T.C. Bruice, A. Brown, D.O. Harris, Concept of orbital steering in catalytic reactions, Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 658.

[6] M.I. Page, W.P. Jencks, Entropic contributions to rate accelerations in enzymic and intramolecular reactions and chelate effect, Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 1678.

[7] T.C. Bruice, U.K. Pandit, The effect of geminal substitution, ring size, and rotamer distribution on the intramolecular nucleophilic catalysis of the hydrolysis of monophenyl esters of dibasic acids and the solvolysis of the intermediate anhydrides, J. Am. Chem. Soc. 82 (1960) 5858.

[8] T.C. Bruice, W.C. Bradbury, Gem effect: 4. Activation parameters accompanying increased steric requirements of 3,5′ substituents in solvolysis of mono-*p*-bromophenyl glutarates, J. Am. Chem. Soc. 90 (1968) 3808.

[9] A.A. Armstrong, L.M. Amzel, Role of entropy in the increased rates of intramolecular reactions, J. Am. Chem. 47 (2003) 14596.

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

[11] A.D. MacKerell Jr., B. Brooks, C.L. Brooks III, L. Nilsson, B. Roux, Y. Won, M. Karplus, CHARMM: the energy function and its parameterization with an overview of the program, P.v.R. Schleyer , et al., (Eds.), The Encyclopedia of Computational Chemistry, vol. 1, John Wiley & Sons, Chichester, 1998, pp. 271–277.

[12] T.L. Hill, An Introduction to Statistical Thermodynamics, Dover Publications, Inc., New York, 1986.

[13] F.C. Lightstone, T.C. Bruice, Separation of ground state and transition state effects in intramolecular and enzymatic reactions: 2. A theoretical study of the formation of transition states in cyclic anhydride formation, J. Am. Chem. Soc. 119 (1997) 9103.

[14] Y.Y. Akhadov, Dielectric Properties of Binary Solutions, Pergamon Press, Oxford, 1981.

[15] R. Lavery, A. Lebrun, J.-F. Allemand, D. Bensimon, V. Croquette, Structure and mechanics of single biomolecules: experiment and simulation, J. Phys., Condens. Matter 14 (2002) R383.

[16] N. Becker, E. Oroudjev, S. Mutz, J.P. Clevel, P.K. Hansma, C.Y. Hayashi, D.E. Makarov, H.G. Hansma, Molecular nanosprings in spider capture-silk threads, Nat. Mater. 2 (2003) 278.

[17] R.B. Best, D.J. Brockwell, J.L. Toca-Herrera, A.W. Blake, D.A. Smith, S.E. Radford, J. Clarke, Force mode atomic force microscopy as a tool for protein folding studies, Anal. Chim. Acta 479 (2003) 87.

[18] D.J. Brockwell, G.S. Beddard, J. Clarkson, R.C. Zinober, A. Blake, J. Trinick, P.D. Olmsted, D.A. Smith, S.E. Radford, The effect of core destabilization on the mechanical resistance of I27, Biophys. J. 83 (2002) 458.

[19] T.E. Fisher, P.E. Marsalek, J.M. Fernandez, Stretching single molecules into novel conformations using the atomic force microscope, Nat. Struct. Biol. 7 (2000) 719.

[20] T.E. Fisher, A.F. Oberhauser, M.C. Vezquez, P.E. Marsalek, J. Fernandez, The study of protein mechanics with the atomic force microscope, Trends Biochem. Sci. 24 (1999) 379.

[21] R.B. Best, B. Li, A. Steward, V. Daggett, J. Clarke, Can non-mechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulation, Biophys. J. 81 (2001) 2344.

[22] P.E. Marsalek, H. Lu, H. Li, M. Carrion-Vazquez, A.F. Oberhauser, K. Schulten, J. Fernandez, Mechanical unfolding intermediates in titin modules, Nature (London) 402 (1999) 100.

[23] M. Rief, J.M. Fernandez, H.E. Gaub, Elastically coupled two-level systems as a model for biopolymer extensibility, Phys. Rev. Lett. 81 (1998) 4764.

[24] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM, Science 276 (1997) 1109.

[25] Z. Bryant, V.S. Pande, D.S. Rokhsar, Mechanical unfolding of a -hairpin using molecular dynamics, Biophys. J. 78 (2000) 584.

[26] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation, Biophys. J. 75 (1998) 662.

[27] H. Lu, K. Schulten, Steered molecular dynamics simulation of conformational changes of immunoglobulin domain I27 interpret atomic force microscopy observations, Chem. Phys. 247 (1999) 141.

[28] T. Shen, L.S. Canino, J.A. McCammon, Unfolding proteins under external forces: a solvable model under the self-consistent pair contact probability approximation, Phys. Rev. Lett. 89 (2002) 068103.

[29] D.E. Makarov, P.K. Hansma, H. Metiu, Kinetic Monte Carlo simulation of titin unfolding, J. Chem. Phys. 114 (2001) 9663.

[30] E. Evans, K. Ritchie, Strength of a weak bond connecting flexible polymer chains, Biophys. J. 76 (1999) 2439.

[31] C. Jarzynski, Nonequilibrium equality for free energy differences, Phys. Rev. Lett. 78 (1997) 2690.

[32] C. Jarzynski, How does a system respond when driven away from thermal equilibrium? Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 3636.

[33] G. Hummer, A. Szabo, From the cover: free energy reconstruction from nonequilibrium single-molecule pulling experiments, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 3658.

[34] J. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante, Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, Science 296 (2002) 1832.

[35] F. Ritort, C. Bustamante, I. Tinoco, A two-state kinetic model for the unfolding of single molecules by mechanical force, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 13544.

[36] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, W. F. Molnar, K. Wriggers, Steered molecular dynamics, in: P. Deuflhard, J. Hermans, B. Leimkuhler, S. A.E.Mark, R.D. Reich (Eds.), Computational Molecular Dynamics: Challenges, Methods, Ideas, Lecture Notes in Computational Science and Engineering, vol. 4, Springer, Berlin, p. 36.

[37] W.F. Humphrey, A. Dalke, K. Schulten, VMD — visual molecular dynamics, J. Mol. Graph. 14 (1996) 33.
[38] G. Pabon and L.M. Amzel. Personal communication.
[39] F.C. Bernstein, T.F. Koetzle, G.J. Williams, E.F. Meyer, M.D. Brice, J.R. Rogers, O. Kennard, T. Shimanouchi, M. Tasumi, The protein data bank: a computer-based archival file for macromolecular structures, J. Mol. Biol. 112 (1997) 535.
[40] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential function for simulating liquid water, J. Chem. Phys. 79 (1983) 926.
[41] L.M. Amzel, Structure-based drug design, Curr. Opin. Biol. 9 (1998) 366.
[42] V.J. Hilser, J. Gomez, E. Freire, The enthalpy change in protein folding and binding: refinement of parameters for structure-based calculations, Proteins 26 (1996) 123.
[43] L.M. Amzel, Loss of translational entropy in binding, folding, and catalysis, Proteins 28 (1997) 144.
[44] X. Siebert, L.M. Amzel, Loss of translational entropy in molecular associations, Proteins 54 (2004) 104.
[45] I. Luque, E. Freire, Structural parameterization of the binding enthalpy of small ligands, Proteins 49 (2002) 181.
[46] J. Hermans, L. Wang, Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme, J. Am. Chem. Soc. 119 (1997) 2707.
[47] J.O. Hirschfelder, D. Stevenson, H. Eyring, A theory of liquid structure, J. Chem. Phys. 57 (1937) 896.
[48] A.V. Finkelstein, J. Janin, The price of lost freedom: entropy of bimolecular complex formation, Protein Eng. 3 (1989) 1.
[49] G.P. Brady, K.A. Sharp, Entropy in protein folding and in protein—protein interactions, Curr. Opin. Struct. Biol. 7 (1997) 215.
[50] Y.B. Yu, P.L. Privalov, R.S. Hodges, Contribution of translational and rotational motions to molecular association in aqueous solution, Biophys. J. 81 (2001) 1632.
[51] H. Eyring, J.O. Hirschfelder, The theory of liquid state, J. Chem. Phys. 41 (1937) 249.
[52] J.G. Kirkwood, Critique of the free volume of the liquid state, J. Chem. Phys. 18 (1950) 380.
[53] M.K. Gilson, et al., The statistical–thermodynamic basis for computation of binding affinities: a critical review, Biophys. J. 72 (1997) 1047.
[54] R. Luo, M.K. Gilson, Synthetic adenine receptors: direct calculation of binding affinity and entropy, J. Am. Chem. Soc. 122 (2000) 2934.
[55] J.P. Ryckaert, G. Cicotti, H.J.C. Berendsen, Numerical integration of Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327.
[56] J. Janin, Elusive affinities, Proteins 21 (1995) 30.
[57] J. Janin, For Guldberg and Waage, with love and cratic entropy, Proteins 24 (1996) 1.
[58] A.E. Eriksson, et al., A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene, Nature 355 (1992) 371.
[59] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, J. Mol. Biol. 55 (1971) 379.
[60] D. Eisenberg, W. Kauzman, The Structure and Properties of Water, Oxford University Press, London, 1969.
[61] F.M. Richards, The interpretation of protein structures: total volume, group volume distributions and packing density, J. Mol. Biol. 82 (1974) 1.