Protein Science (2000), *9*:252–264. Cambridge University Press. Printed in the USA.

Structure-based thermodynamic analysis of the dissociation of protein phosphatase-1 catalytic subunit and microcystin-LR docked complexes

PIERRE LAVIGNE,[1] JOHN R. BAGU,[2] ROBERT BOYKO,[2] LEIGH WILLARD,[1] CHARLES F.B. HOLMES,[2] AND BRIAN D. SYKES[1]

[1]Department of Biochemistry and The Protein Engineering Network of Centres of Excellence, University of Alberta, Edmonton, Alberta T6G 2S2, Canada
[2]Department of Biochemistry and MRC Group in Protein Structure and Function, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

(RECEIVED June 2, 1999; FINAL REVISION October 18, 1999; ACCEPTED October 30, 1999)

Abstract

The relationship between the structure of a free ligand in solution and the structure of its bound form in a complex is of great importance to the understanding of the energetics and mechanism of molecular recognition and complex formation. In this study, we use a structure-based thermodynamic approach to study the dissociation of the complex between the toxin microcystin-LR (MLR) and the catalytic domain of protein phosphatase-1 (PP-1c) for which the crystal structure of the complex is known. We have calculated the thermodynamic parameters (enthalpy, entropy, heat capacity, and free energy) for the dissociation of the complex from its X-ray structure and found the calculated dissociation constant (4.0×10^{-11}) to be in excellent agreement with the reported inhibitory constant (3.9×10^{-11}). We have also calculated the thermodynamic parameters for the dissociation of 47 PP-1c:MLR complexes generated by docking an ensemble of NMR solution structures of MLR onto the crystal structure of PP-1c. In general, we observe that the lower the root-mean-square deviation (RMSD) of the docked complex (compared to the X-ray complex) the closer its free energy of dissociation (ΔG_d°) is to that calculated from the X-ray complex. On the other hand, we note a significant scatter between the ΔG_d° and the RMSD of the docked complexes. We have identified a group of seven docked complexes with ΔG_d° values very close to the one calculated from the X-ray complex but with significantly dissimilar structures. The analysis of the corresponding enthalpy and entropy of dissociation shows a compensation effect suggesting that MLR molecules with significant structural variability can bind PP-1c and that substantial conformational flexibility in the PP-1c:MLR complex may exist in solution.

Keywords: complex dissociation; docking; microcystin-LR; NMR; protein phosphatase-1; structure-based thermodynamics

Protein phosphorylation is a general mechanism for the regulation of many important cellular processes (Cohen, 1989; Shenolikar, 1994; Barford, 1995). Due to the reversible nature of phosphorylation, there is generally an antagonistic relationship between activation of cellular processes achieved by protein kinases and deactivation of cell signals by protein phosphatases. Dephosphorylation of serine and threonine is mainly accomplished by four subgroups of phosphatases: protein phosphatase-1 (PP-1c), -2a, -2b (calcineurin), and -2c (Cohen, 1989). Some of these classes (PP-1c and PP-2a) are inhibited by metabolites of cyanobacteria (e.g., microcystins and nodularins), dinoflagellates (e.g., okadaic acid), and compounds isolated in sponges (e.g., calyculin A). These metabolites are liver toxins and have powerful tumor promotion activity linked to morphological changes in hepatocytes (Ohta et al., 1992). The microcystins are cyclic 7 amino acid peptides containing several unusual amino acids. There are several varieties of microcystins with the differences normally localized to changes in the two variable amino acids or alterations in the methylaspartic acid and/or N-methyldehydroalanine residues (Craig et al., 1993). Two examples are microcystin-LR (MLR) and microcystin-LL

Reprint requests to: Dr. Brian D. Sykes, Department of Biochemistry and The Protein Engineering Network of Centres of Excellence, University of Alberta, Edmonton, Alberta T6G 2S2, Canada; e-mail: brian.sykes@ualberta.ca.

Abbreviations: ASA, accessible surface area; ΔG_d°, standard Gibbs free energy of dissociation; ΔH_d°, standard enthalpy of dissociation; ΔS_d°, standard entropy of dissociation; $\Delta G_{d,corr}^\circ$, standard corrected Gibbs free energy of dissociation; $\Delta H_{d,corr}^\circ$, standard corrected enthalpy of dissociation; $\Delta S_{d,corr}^\circ$, standard corrected entropy of dissociation, ΔG_{mlr}°, standard relative Gibbs free energy of individual MLR conformers; ΔH_{mlr}°, standard relative enthalpy of individual MLR conformers; ΔS_{mlr}°, standard relative entropy of individual MLR conformers; ΔG_{comp}°, standard relative Gibbs free energy of individual docked complexes; ΔH_{comp}°, standard relative enthalpy of individual docked complexes; ΔS_{comp}°, standard relative entropy of individual docked complexes; MLL, microcystin-LL; MLR, microcystin-LR; RMSD, root-mean-square deviation; PP-1c, protein phosphatase-1 catalytic subunit.

(MLL), wherein Arg has been substituted for Leu. Microcystins are able to covalently link to PP-1c with the N-methyldehydroalanine residue joining with Cys273 (MacKintosh et al., 1995; Runnegar et al., 1995). This covalent linkage is time dependent and has no impact on the initial inhibition of PP-1c or PP-2a. The consequence of the covalent linkage is to irreversibly inhibit the phosphatase preventing any further activity.

The NMR solution structure of MLR has been determined (Bagu et al., 1995) as well as the X-ray structure of PP-1c covalently complexed with MLR (Goldberg et al., 1995). The free and bound forms of MLR were found to have similar overall structures and, most strikingly, the conformation of the cyclic backbone of the solution structure of MLR is almost identical to the structure in the complex. The relationship between the structure of a ligand in its bound form(s) and in its free form(s) is of particular interest to the field of drug design. For example, if one could be able to discover the active conformation(s) of a peptide in a family of NMR structures, this would accelerate drug discovery. One way of addressing this problem is to develop approaches capable of calculating reliable energetic or thermodynamic parameters for the association (dissociation) of complexes generated by docking solution structures onto a target of known structure. Thus, the solution structures of the ligands present in complexes that give calculated thermodynamic parameters that agree well with measured ones should be close to the active form(s) of the ligand or at least competent binding conformations.

In a previous attempt to understand the relationship between the free form and the bound form of MLR in the inhibition process of PP-1c, we first generated a model for the PP-1c:MLR complex by a rigid body docking procedure using the average solution structure of MLR (Bagu et al., 1997). The success of docking the average NMR solution structure to PP-1c in the same position as the bound crystal MLR allowed for further successful dockings of microcystin-LL, motuporin, okadaic acid, and calyculin A. These toxins all had similar three-dimensional structures despite significant primary structural differences and were proposed to bind to the same site as MLR. The quality of the models was assessed on the basis of surface complementarity and potential energy obtained from a molecular mechanics force field (Bagu et al., 1997).

While molecular mechanics force fields are useful to maintain proper noncovalent and covalent stereochemistry, they seem not to be able to discriminate the correct fold within clusters of docked complexes with minimal potential energies. This problem, sometimes referred to as the "docking problem," involves the discrimination between the correct answer and the "false positives" that have similar potential energy but incorrect structures (Cummings et al., 1995). Recently, recourse to solvation free energy corrections has been shown to partially solve that problem (Cummings et al., 1995). On the other hand, empirical free energy functions have been successfully used to reliably calculate the binding free energies (ΔG_{bind}) or relative ΔG_{bind} from the structure of complexes (reviewed in Vajda et al., 1997). Therefore, it has been proposed (Abagyan & Totrov, 1994; Jackson & Sternberg, 1995) that the minimization of accurate or realistic empirical free energy functions might alleviate the docking problem. The development and the use of accurate free energy functions are of the utmost importance both in the protein folding problem and molecular recognition fields. Ultimately, what is needed is an empirical free energy function(s) that is realistic (faithful) and accurate enough to reproduce both the experimental configuration and the corresponding thermodynamics for the folding and binding processes of polypeptide chains.

Binding free energies (ΔG_{bind}) can be computed from first principles using statistical mechanical approaches (reviewed in Gilson et al., 1997). Although these calculations rely on exact results of statistical mechanics, they are not easily tractable. Empirical free energy functions can be obtained from structural database statistics or from linear regression fitting of different empirical free energy parameters (scaling with molecular surface) with experimental ΔG_{bind} (Moult, 1997; Vajda et al., 1997). Both approaches have been used with some success, although the latter approach seems to be restricted to homologous systems (Vajda et al., 1997). A third approach that is transferable to other systems and has been successful to predict the experimental ΔG_{bind} (or $\Delta\Delta G_{bind}$ for mutants) separates the total ΔG_{bind} into hydrophobic force (per Å^2 of surface), electrostatic interactions (Poisson–Boltzmann equation), conformational entropy and overall rotational and translational entropy contributions (Weng et al., 1996; Novotny et al., 1997). A fourth approach can calculate the free energy of binding from a parameterization (per Å^2 of polar and nonpolar ASA) of the heat capacity, enthalpy, and solvation entropy obtained from a global fit of structural and thermodynamic database of globular proteins (Murphy & Freire, 1992; Baker & Murphy, 1998; Luque & Freire, 1998). This method has the advantage of being tractable and of separating the ΔG into enthalpic and entropic contributions. This approach has been successfully used to calculate from the crystal structure of protein–peptide, protein–ligand, and protein–protein complexes the energetics of dissociation (enthalpy (ΔH), entropy (ΔS), and heat capacity change (ΔC_p) of complexes that agree well with experimentally determined values (Baker & Murphy, 1998; Luque & Freire, 1998). This approach has also been utilized in order to validate the model structure of a complex (Baker & Murphy, 1997).

In this study, we use a slightly modified version of the structure based approach described above to address the relationship between the bound conformation and the free solution structure of MLR. More specifically, we calculate the thermodynamic parameters for the dissociation of 47 complexes generated by docking an ensemble of NMR solution structures of MLR on the crystal structure of the PP-1c. We compare these values with the ones calculated from the crystal structure of the complex and explore the relationship between structural diversity and the energetics of the different docked complexes.

Results and discussion

Calculation of the thermodynamics of dissociation of MLR and PP-1c from the X-ray complex

We present in Figure 1A the X-ray structure of MLR when bound to PP-1c (Goldberg et al., 1995), referred to as X-ray MLR in the rest of the text, and in Figure 1B the ensemble of NMR solution structures of MLR determined by NMR (Bagu et al., 1995) and referred to as NMR MLR throughout. Most of the residues in the seven amino acid cyclic peptide are unique or altered amino acids. Starting at D-alanine (D-Ala) the sequence is L-leucine (Leu), β-linked D-erythro-β-methylaspartic acid (Masp), L-arginine (Arg), β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda), γ-linked D-glutamic acid (D-Glu), and N-methyldehydro-alanine (Mdha). The cyclic backbone is saddle shaped (Bagu et al., 1995) with the Arg pointing above the saddle, the large hydrophobic side chain Adda pointing

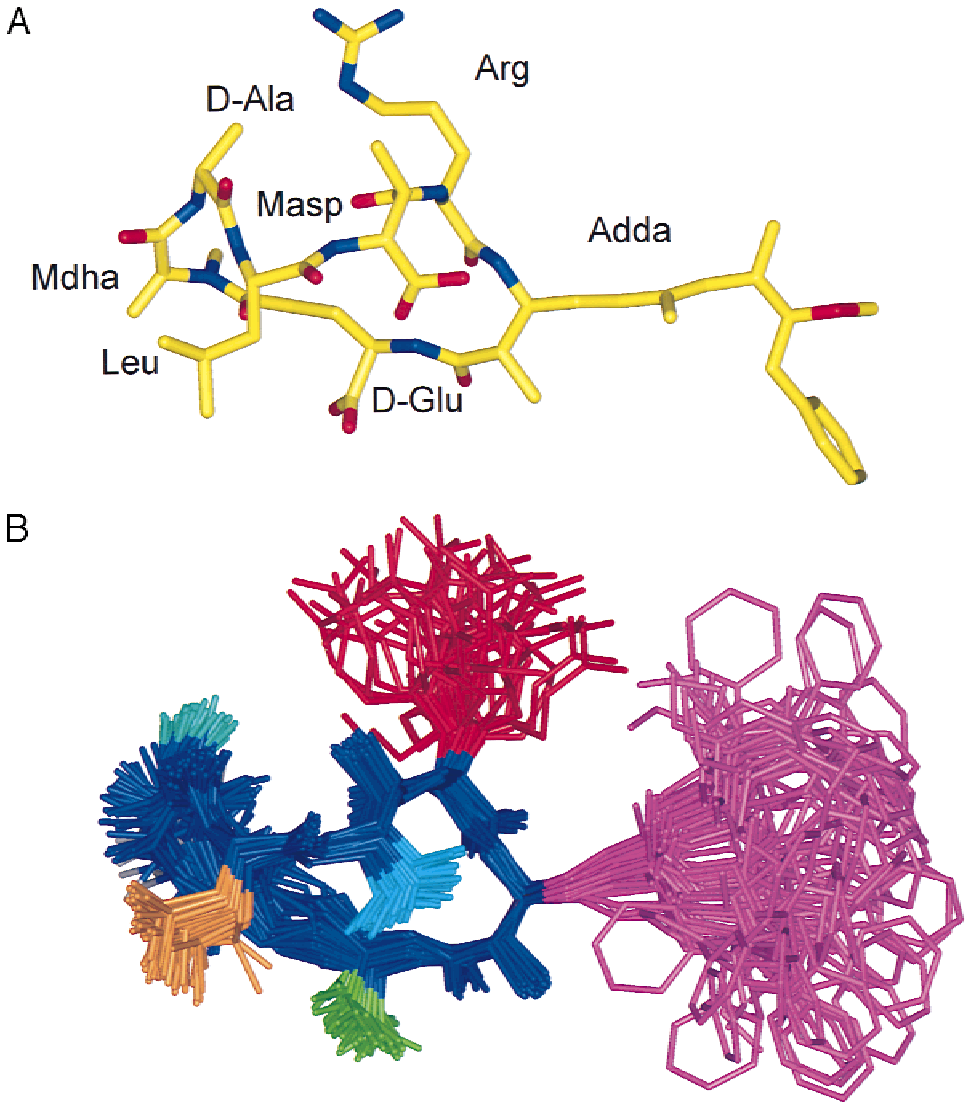

Fig. 1. The structure of MLR when bound to the PP-1C and the ensemble of structures calculated from solution state NMR for MLR free in solution. **A**: Bound X-ray structure of MLR (Goldberg et al., 1995). The residues are D-alanine (D-Ala), L-leucine (Leu), β-linked D-erythro-β-methylaspartic acid (Masp), L-arginine (Arg), β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda), γ-linked D-glutamic acid (D-Glu), and N-methyldehydro-alanine (Mdha). The cyclic backbone is saddle shaped (Bagu et al., 1995). **B:** Ensemble of 46 calculated solution structures and average minimized solution structure of MLR (Bagu et al., 1995). The Leu side chain is brown, Masp light blue, Arg red, Adda purple, D-Glu green, D-Ala cyan, and the cyclic backbone is dark blue. When compared with the average free solution structure, the backbone is almost identical to the bound X-ray form (backbone RMSD is 0.65 Å). For clarity hydrogens are not shown.

behind the saddle, and the negatively charged carboyxl groups located underneath the saddle (as observed in the orientation of Fig. 1). The Mdha residue, which covalently links with PP-1c, is located at the top, front of the saddle. The Leu side chain is brown, Masp is yellow, Arg is red, Adda is purple, D-Glu is green, and the cyclic backbone is in blue. The backbone of the average solution structure is almost identical to that of the bound form with RMSD of 0.65 Å. It is notable that in the free solution structure the Arg and Adda side chains are highly flexible adopting multiple conformations. Conversely, the side chains of the D-Glu and the Masp residues are constrained by the backbone as they have only one rotable bound.

The X-ray complex of MLR bound to PP-1c is represented in Figure 2 (Goldberg et al., 1995) where only PP-1c residues within 4 Å of MLR are labeled. As discussed elsewhere (Goldberg et al., 1995), specific salt-bridges or H-bonds between the MLR and PP-1c involve Masp (Arg96 and Tyr134) and D-Glu (Arg96 and Arg221). The Arg side chain of the toxin is also found to lie between the carboxylates of Asp220 and Glu275 at the surface the enzyme with distances $<$6.0 Å giving rise to potential solvent exposed salt-bridges and/or H-bonds. The hydrophobic Adda side chain interacts with hydrophobic residues lining the hydrophobic groove on the enzyme as described before (Goldberg et al., 1995). Finally, the toxin is found to be covalently linked to Cys273 (Sγ) through the Cβ of the Mdha. This covalent attachment has been observed to be slow (hours) and not to impair or affect the initial inhibitory action of the toxin.

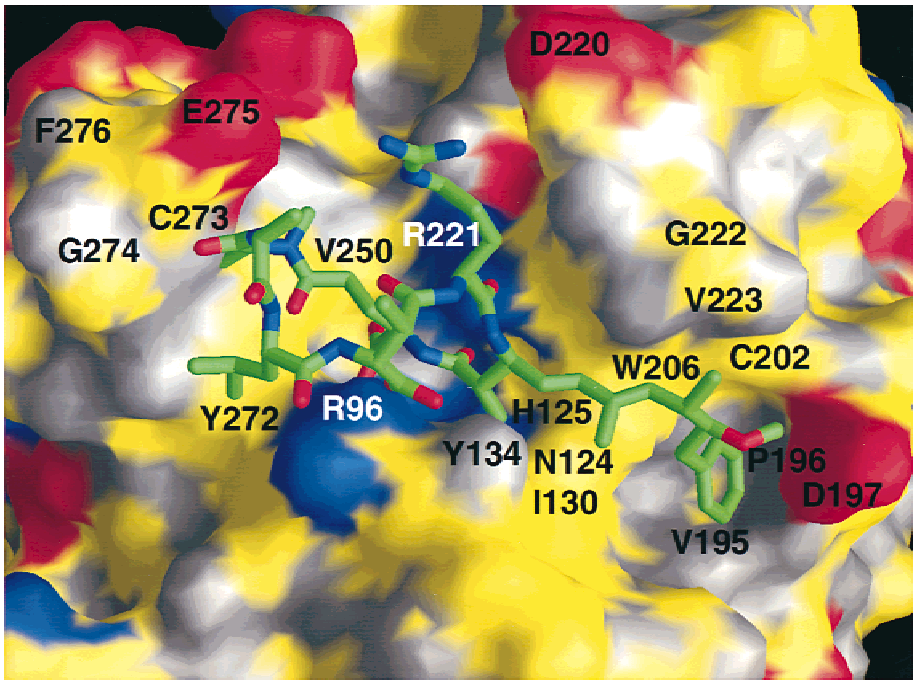

Fig. 2. The X-ray complex of MLR and PP-1c (Goldberg et al., 1995) displayed using the program Grasp (Nicholls et al., 1991). Only PP-1c residues within 4 Å of MLR are labeled. Red are surfaces that are potentially negatively charged, blue are potentially positively charged, yellow represents hydrophobic surfaces, and white are polar surfaces.

Figure 3 (top equilibrium) sketches the process for the structure-based thermodynamic calculations using the X-ray complex. First, the ASA of every atom of the X-ray complex, the dissociated PP-1c, and MLR are calculated. Second, the differences in ASA (ΔASA) for every atom of the PP-1c and MLR in both forms are obtained. Finally, the corresponding ΔG_d° are computed on a per residue basis as described in Materials and methods. The changes in ASA per residue (nonpolar and polar) upon dissociation of the X-ray complex are displayed in Figure 4A and the corresponding ΔG_d° on Figure 4B. Overall, MLR exposes 542 and 150 Å^2 of nonpolar and polar ASA, respectively, while PP-1c exposes 282 Å^2 of nonpolar and 296 Å^2 polar ASA upon dissociation. Listed in Table 1 are the thermodynamic parameters obtained for the total and per MLR residue ΔASA values for the dissociation of the X-ray complex.

As one can notice (Table 1), the unfavorable dissociation free energy calculated: $\Delta G_d^\circ(25) = 14.1$ kcal·mol^{-1} is dominated by an overall unfavorable dissociation entropy ($-T\cdot\Delta S_d^\circ(25) = 16.0$ kcal·mol^{-1}) with a slightly favorable dissociation enthalpy ($\Delta H_d^\circ(25) = -1.87$ kcal·mol^{-1}). It is evident that most of the affinity comes from the unfavorable change in solvation entropy ($-T\cdot\Delta S_{sol}(25) = 21.9$ kcal·mol^{-1}) upon dissociation of the X-ray complex. The Adda residue that buries an extensive amount of nonpolar surface, roughly 40% (317 Å^2) of the total ΔASA_{np}, is a major contributor to the affinity (Fig. 4B) through the hydrophobic effect, i.e., decrease in the solvent entropy upon dissociation ($-T\cdot\Delta S_{sol}(25) = 10.7$ kcal·mol^{-1}, see Table 1). It can be seen that the $T\cdot\Delta S_{sol}(25)$ of the Adda is offset by a favorable $\Delta H_d^\circ(25)$ and $-T\cdot\Delta S_{conf}$ (Table 1). This side chain has been hypothesized to be critical for the binding of MLR to PP-1c. Indeed, a structural isomer of the Adda side chain was determined to inhibit PP-1c activity 100 times more weakly than the maternal MLR (Nishiwaki-Matsushima et al., 1991). The second largest $\Delta G_d^\circ(25)$ on MLR is D-Glu, which is dominated by an unfavorable $\Delta H_d^\circ(25)$ indicative of favorable interactions with Arg96 and Arg221 (Table 1). Interestingly, it has been shown that esterification of the D-Glu side chain has a significant reduction in toxicity suggesting that it is indeed important for the binding of MLR to PP-1c (Stotts et al., 1993). It is noted in Figure 4B that residues Arg96, Arg221 also have a high individual $\Delta G_d^\circ(25)$ (>1 kcal·mol^{-1}) indicating that they are contributing substantially to the affinity of the complex. Interestingly, mutagenesis studies have shown that the replacement of Arg221 (Arg221Ser) and Arg96 (Arg96Ala) resulted in drastic reduction in K_i by MLR (Huang et al., 1997) supporting the high $\Delta G_d^\circ(25)$ values calculated here.

As discussed in Materials and methods and stated above, the calculated gain in conformational entropy results solely from side chains becoming exposed. It is worth pointing out, though, that if the MLR was unfolding (linear) upon dissociation, then the gain in entropy of the backbone would make the affinity of MLR to PP-1c much lower. Assuming a median value for the conformational entropy change for the backbone of $\Delta S_{bb} \sim 6$ cal·K^{-1}·mol^{-1} per residue (Brady & Sharp, 1997), an additional gain in ΔS_{conf} of ~ 42 cal·K^{-1}·mol^{-1} can be estimated ($-T\cdot\Delta S_{conf}(25) \sim -1.5$ kcal·mol^{-1} per residue), which would lower the $\Delta G_d^\circ(25)$ from 14.1 kcal·mol^{-1} to practically 0 kcal·mol^{-1}. This is a rough and probably underestimated value since most of the residues of the toxin have more than two rotable bonds and reinforces the idea that the cyclic and folded nature of the backbone of the toxin contributes a lot to its high affinity.

The overall $\Delta H_d^\circ(25)$ of -1.87 kcal·mol^{-1} indicates that the disruption of the favorable noncovalent interactions at the interface of the complex is compensated by an almost equally favorable solvation enthalpy of the overall molecular surfaces exposed upon

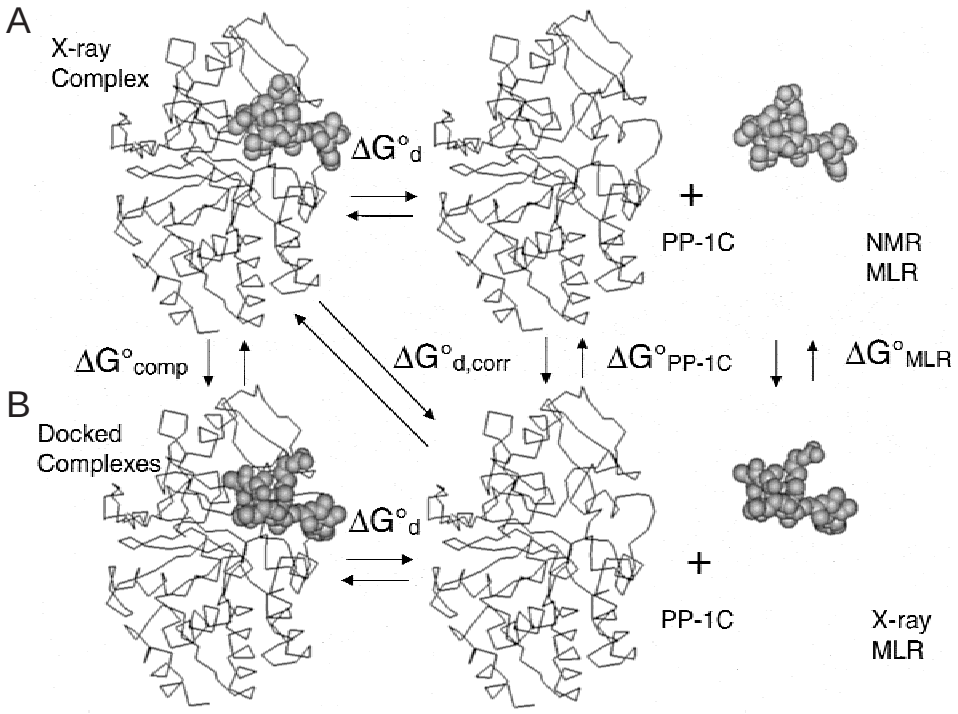

Fig. 3. Schematic representation of the (**A**) dissociation equilibrium for the X-ray complex and (**B**) the different docked complexes. Since the different docked complexes are different from the X-ray complex, a relative difference in free energy (ΔG°_{comp}), enthalpy (ΔH°_{comp}), and entropy (ΔS°_{comp}) can be calculated for every docked complex from the differences in ASA ($ASA_{\text{docked complexes}} - ASA_{\text{X-ray complex}}$). Similarly, one can obtain ΔG°_{MLR}, ΔH°_{MLR}, and ΔS°_{MLR} between the different NMR MLR and X-ray MLR from differences in ASA ($ASA_{\text{NMR MLR}} - ASA_{\text{X-ray MLR}}$). $\Delta G^\circ_{d,corr}$ ($\Delta H^\circ_{d,corr}$ or $\Delta S^\circ_{d,corr}$) is obtained by subtracting ΔG_{MLR} (ΔH°_{MLR} or ΔS°_{MLR}) from the ΔG°_d (ΔH°_d or ΔS°_d) of the different docked complexes. $\Delta G^\circ_{PP\text{-}1c} = 0$.

dissociation. It has to be noted that other factors that could contribute to the enthalpy of dissociation (and the entropy) like putative proton transfer(s) (Gomez & Freire, 1995) are not treated explicitly here. No change in pK_a of ionizable groups in the catalytic domain or for MLR has been reported. Moreover, no experimental enthalpy (or entropy) of dissociation is available so far to allow us to compare the enthalpy (and entropy) of dissociation calculated with the parameterization used here. In addition, no experimental and conventional K_d is available because of the very high affinity of PP-1c for MLR (Takai et al., 1995). On the other hand, it is interesting to note that the K_d of 4.0×10^{-11} calculated (Table 1) is excellent agreement with the K_i of 3.9×10^{-11} reported by Takai et al. (1995). Since, formally, a K_d and a K_i are different quantities, the agreement should be taken as a matching of order of magnitude between the experiment and the calculation. But more importantly, this agreement suggests that the present parameterization satisfactorily describe the energetics of the dissociation of MLR and PP-1c and also suggests that all the assumptions made above appear to be justified.

Assessment of the docked complexes obtained from an ensemble of NMR solution structure of MLR

As mentioned previously, the development of methods to obtain information about the "bound structure" or binding competent configurations from the structure of the free ligand in solution is very important for our understanding of association (dissociation) reactions and for the rational design of ligands of pharmaceutical interest. We explore in this section the potential use of the structure-based approach to address these issues by analyzing the dissociation thermodynamic parameters obtained for an ensemble of complexes generated from the docking of 47 solution NMR structures of MLR (NMR MLR) onto the crystal structure of PP-1c (X-ray PP-1c) as described in Materials and methods.

Before presenting our results, we want to stress a few points about the calculation of the energetic parameters of the ensemble of docked complexes. As is evident in the bottom equilibrium of Figure 3, the structures of the different docked complexes as well as the structures of the NMR MLR are all going to be different to that of the X-ray complex and MLR X-ray, respectively. Since the energetic calculations are based in differences in structure, it is to be expected that differences in energetic parameters (e.g., G, H, and S) are going to exist between the different MLR NMR and the MLR X-ray (ΔG°_{mlr}) and the different docked complexes and the X-ray complex (ΔG°_{comp}, see Fig. 3).

We present in Figure 5A, the calculated ΔG°_d of the docked complexes as a function of the positional RMSD. Positional RMSD is defined as the RMS difference between the location of the heavy atoms of the MLR in a docked complex as a function of their location in the X-ray complex when the PP-1c molecules are superimposed. Positional RMSD takes into account the structural and location differences of MLR between the complexes. One can clearly see that there is a general trend for the ΔG°_d to be larger as the RMSD becomes smaller. This relationship indicates that there is a clear tendency for the calculated ΔG°_d to be closer to the one calculated from the X-ray complex as the structure of the docked

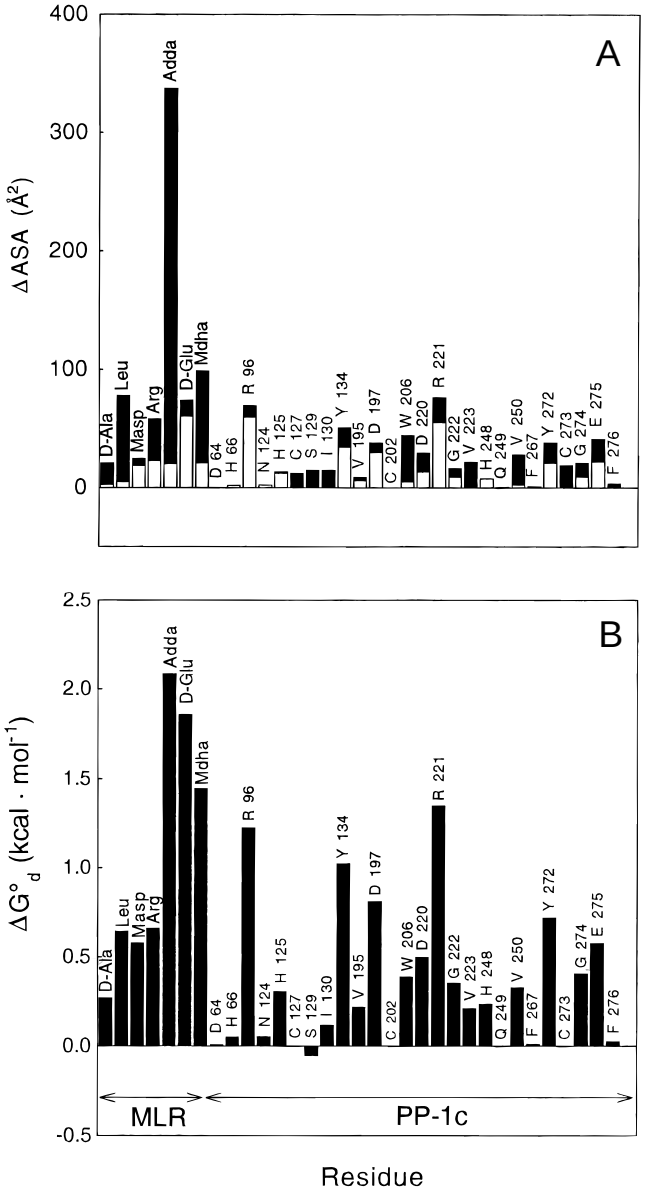

Fig. 4. Results for the structure-based thermodynamic calculation from the X-ray complex of MLR and PP-1c. **A:** Changes in ASA per residue (nonpolar: black and polar: white) upon dissociation of the X-ray complex. **B:** Corresponding ΔG°_{d}. MLR exposes 542 and 150 Å^2 of nonpolar and polar ASA, respectively, while PP-1c exposes 282 Å^2 of nonpolar and 296 Å^2 polar ASA upon dissociation.

complexes become closer to it. There is, however, a significant scatter in the RMSD, i.e., several complexes with higher RMSD have higher affinities than some with a lower RMSD. This is somewhat similar to the case of false positives encountered in docking calculations with potential energy force fields (Cummings et al., 1995). On the other hand, this could be an indication that there are, in fact, different possible configurations for the complexes that lead to similar decreases in free energy in solution.

Formally, the docked complexes that give larger ΔG°_{d} than the X-ray complex must have lower free energy (more stable than the X-ray complex) and/or MLR NMR that have higher free energy (less stable) than the X-ray MLR. Since we make the assumption that the structure of PP-1c does not change upon dissociation, its free energy is not changing either in the calculations. To explore the origins of the higher ΔG°_{d} values than the X-ray complex, we have calculated the ΔG°_{comp} for all the docked complexes and the ΔG°_{mlr} (see Fig. 3) for all the NMR MLR. Values for ΔG°_{comp} and ΔG°_{mlr} were calculated from the differences in ASA between the docked complexes and the X-ray complex and the NMR MLR and the X-ray MLR. A positive value for both ΔG°_{comp} and ΔG°_{mlr} indicates that the particular docked complex and NMR MLR are less stable that the X-ray complex and X-ray MLR, respectively.

We present in Figure 5B the plot of the ΔG°_{d} as a function of ΔG°_{comp}. As one can observe, a clear correlation between ΔG°_{d} and ΔG°_{comp} exists i.e., the more stable the complex the more unfavorable the ΔG°_{d}. On the other hand, none of the complexes are more stable than the crystal structure (no negative ΔG°_{comp}). Therefore, it is clear that the ΔG°_{d}s greater than that of the X-ray complex have to originate from MLR solution structures having positive ΔG°_{mlr}. The ΔG°_{mlr} were all found to be positive as expected from the argument above. This could indicate that the X-ray MLR is a more stable conformation than all of the NMR MLR. Although this might be true, we notice that relatively small changes in ASA (e.g., $\Delta ASA_{pol} = 50$ Å^2 and $\Delta ASA_{np} = 50$ Å^2) lead to high ΔG°_{mlr} values (e.g., 2 kcal·mol^{-1}). These changes originate from small conformational changes (fluctuations) in the backbone and in side-chain dihedral angles that are unlikely to lead to such important changes in free energy as the one calculated. These fluctuations should, to a first approximation, be nearly isoenergetic and not likely to affect significantly the relative population of the different members of the ensemble of solution structures. Unstable conformations of the toxin will increase the ΔG°_{d} but, on the other hand, will not reflect the most probable conformation of the toxin nor a realistic representation of the dissociation reaction. We assume here that the ensemble of NMR MLR structures are equally populated and consider that the ΔG°_{mlr} corresponds more or less to noise inherent to the present method to calculate differences in free energy between structurally fluctuating small peptides. We have, therefore, subtracted the corresponding ΔG°_{mlr} (and similarly differences in enthalpy ΔH°_{mlr} and entropy ΔS°_{mlr}) from the ΔG°_{d} (ΔH°_{d} and ΔS°_{d}) of the different docked complexes to yield a corrected ΔG°_{d}: $\Delta G^{\circ}_{d,corr}$ (Fig. 5). We present on Figure 5C the plot of the corrected $\Delta G^{\circ}_{d,corr}$ as a function of ΔG°_{comp}. We can see that there is a nearly perfect correlation. This is explained from the fact that the relative $\Delta G^{\circ}_{d,corr}$ depends only on the relative ΔG°_{comp} or in other words the more stable the complex the larger the ΔG°_{d}. Finally, we present on Figure 5D the plot of $\Delta G^{\circ}_{d,corr}$ as a function of positional RMSD depicting that none of the $\Delta G^{\circ}_{d,corr}$ are larger than the ΔG°_{d} of the X-ray complex but still highlighting the scatter in RMSD. This is particularly evident for the cluster (open circles) with $\Delta G^{\circ}_{d,corr}$ closest to the ΔG°_{d} of the X-ray complex.

Enthalpy–entropy compensation phenomenon evidenced from structure-based thermodynamics

Figure 6 displays the structures of the six complexes highlighted in Figure 5D as open circles including the average structure (purple). The X-ray MLR is also displayed (yellow). One can notice that the backbones of all the MLR molecules are close to each other and that the high RMSD values come from differences in the conformations of the long Adda and Arg side chains. In this regard, one can see that the NMR MLRs have Adda side chains that are more exposed and Arg side chains that lie closer to the enzyme poten-

Table 1. *Thermodynamic parameters calculated for the dissociation of the PP-1C:MLR X-ray complex and the individual residues of MLR*[a]

	ΔASA_{np} (Å^2)	ΔASA_{pol} (Å^2)	ΔC_p (kcal·mol^{-1}·K^{-1})	ΔH°_d (kcal·mol^{-1})	$-T\cdot\Delta S^\circ_d$[b] (kcal·mol^{-1})	$-T\cdot\Delta S^\circ_{sol}$ (kcal·mol^{-1})	$-T\cdot\Delta S_{conf}$ (kcal·mol^{-1})	ΔG°_d (kcal·mol^{-1})	K_d
X-ray complex	824	446	0.255	−1.87	16.0	21.9	−5.89	14.1	4.05×10^{-11}
D-Ala	18	3	0.0074	−0.32	0.60	0.60	0.000	0.27	—
Leu	72	5	0.0313	−1.54	2.20	2.45	−0.257	0.65	—
Masp	6	18	−0.0020	0.60	−0.02	0.05	−0.065	0.58	—
Arg	36	23	0.0102	0.05	0.61	1.02	−0.413	0.66	—
Adda	317	20	0.1376	−6.87	8.95	10.71	−1.761	2.09	—
D-Glu	14	60	−0.0095	2.11	−0.26	−0.08	−0.179	1.86	—
Mdha	78	21	0.0297	−1.05	2.50	2.50	0.000	1.45	—

[a]Temperature = 25 °C.
[b]The $-T\cdot\Delta S^\circ_d$ for the X-ray complex contains contribution ΔS_{rt} but not the $-T\cdot\Delta S^\circ_d$ of individual residues of MLR.

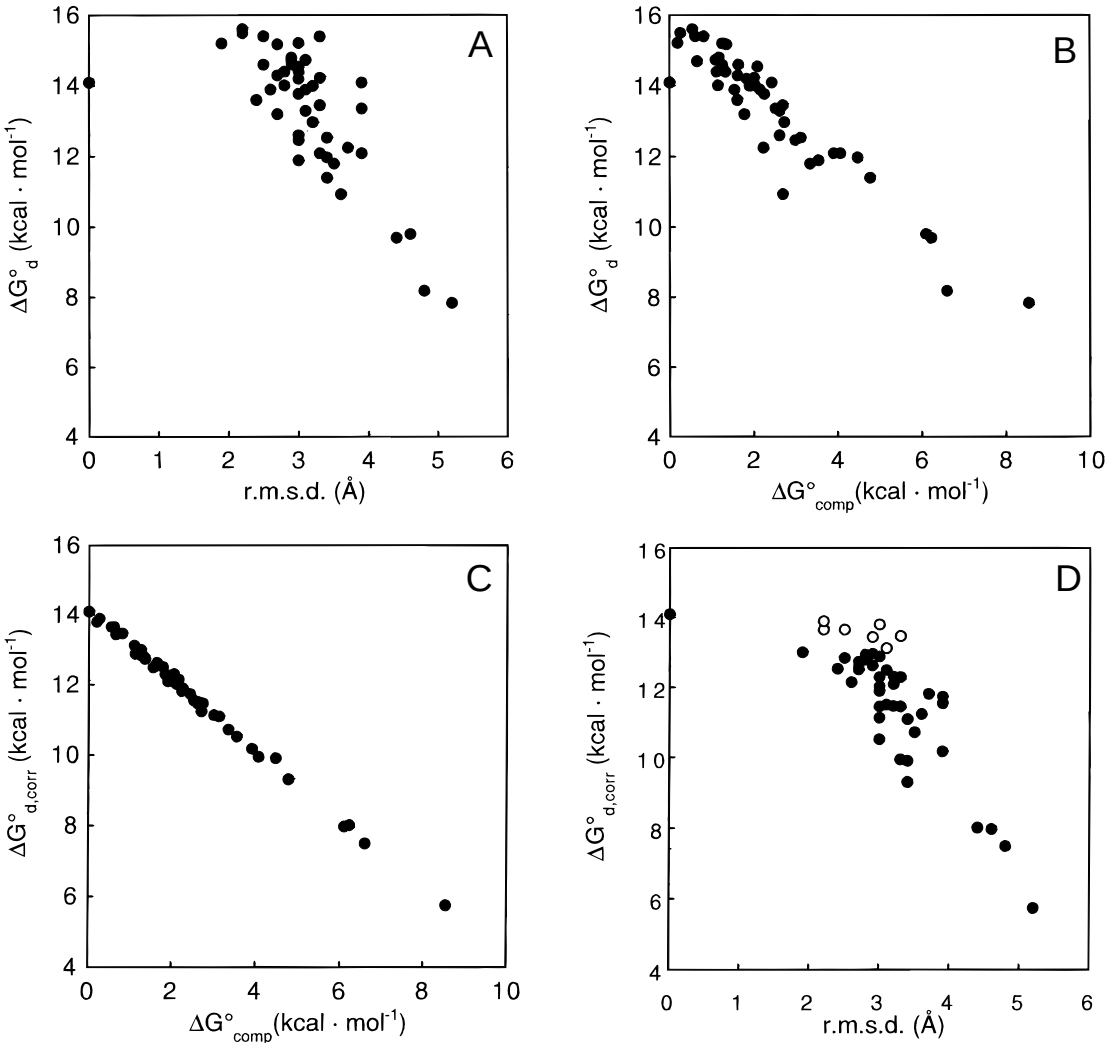

Fig. 5. Results for the structure-based thermodynamic calculations from the docked complexes of MLR and PP-1c. **A:** The calculated ΔG°_d of the docked complexes as a function of the positional RMSD. Positional RMSD is defined in the text. **B:** A plot of the ΔG°_d as a function of ΔG°_{comp} of the complexes as described in Figure 4. **C:** The plot of the $\Delta G^\circ_{d,corr}$ as a function of ΔG°_{comp}. **D:** The plot of the corrected $\Delta G^\circ_{d,corr}$ as a function of positional RMSD. See text for further details.

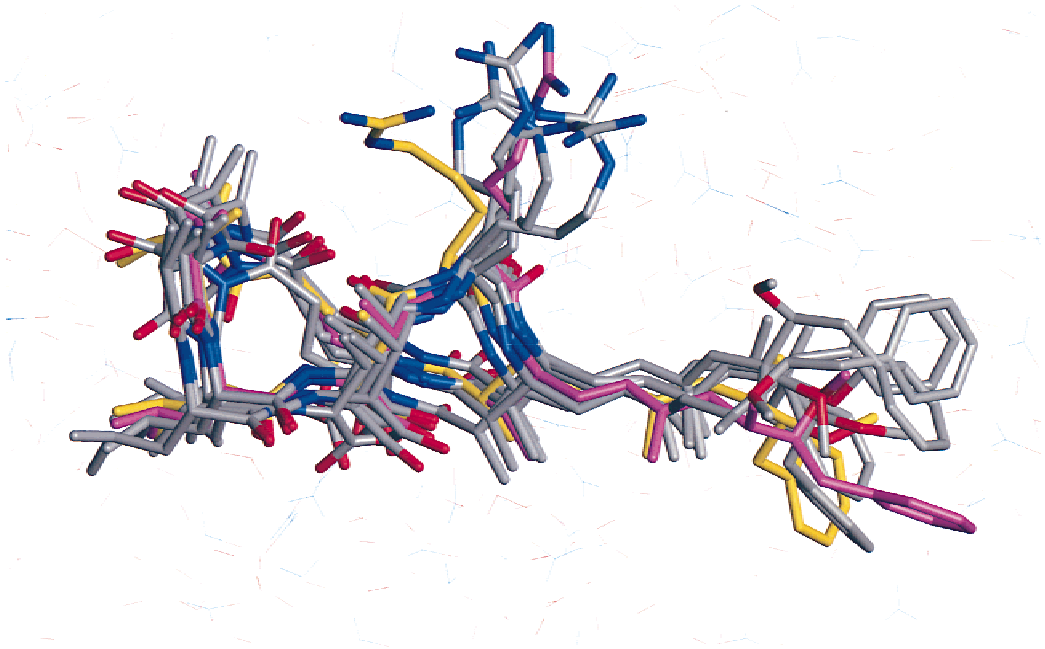

Fig. 6. Multiple complexes with similar free energies. The structure of the seven complexes representing the highest $\Delta G_{d,corr}$ highlighted in Figure 5D including the bound crystal structure of MLR (yellow) and the geometric average of the 46 NMR MLRs (purple). The X-ray PP-1c molecules are superimposed onto one another.

tially forming H-bonds and/or salt bridges with either Glu275 and/or Asp220 (Fig. 2).

It is clear the present parameterization is able to recognize docked complexes that are close to the X-ray complex on the basis of the $\Delta G^{\circ}_{d,corr}$ calculated here but it still shows some significant structural differences. This could indicate that this approach does not have a high degree of discrimination. On the other hand, it is possible that the complexes with $\Delta G^{\circ}_{d,corr}$ similar to the X-ray complex could be populated in solution and that the differences in structure displayed in Figure 6 are nearly isoenergetic.

To investigate why complexes with relatively high and scattered RMSD can give rise to similar ΔG°_{d} (similar free energies), we have analyzed the normalized $\Delta H^{\circ}_{d,corr}$ and $\Delta S^{\circ}_{d,corr}$ (corrected for the variations in the enthalpy and entropy of the different NMR MLR) and the structural features of the complexes of the high $\Delta G^{\circ}_{d,corr}$ cluster. Figure 7A shows a plot of the $\Delta H^{\circ}_{d,corr}$ as a function of the $\Delta S^{\circ}_{d,corr}$ and Figure 7B exhibits a plot of ΔH°_{comp} as a function of the ΔS°_{comp} calculated for the subset of seven complexes. The larger the RMSD of the complexes, the more they lie on the left (Fig. 7A) and on the right (Fig. 7B) of the plots. One can clearly observe a linear relationship between the two quantities in both instances. This indicates the existence of an enthalpy–entropy compensation phenomenon in the calculations. Experimental enthalpy–entropy compensation effects are ubiquitous and have been reported for the binding of series of ligands to their specific enzymes or for binding studies carried on at different pHs or ionic strengths (Lumry & Rajender, 1970; Lumry, 1995). Experimental enthalpy–entropy compensations have also been reported for the stability of protein mutants (Hawkes et al., 1984; Shortle et al., 1988). It is thought that water plays a key role in the mechanism of the compensation effect with typical compensation temperatures or T_c (slope of the ΔH°_{d} vs. ΔS°_{d} curves) between 270 and 320 K (Lumry & Rajender, 1970; Lumry, 1995). Eftink et al. (1983) have presented a thermodynamic model for the enthalpy–entropy compensation in ligand-enzyme systems. In their model the compensation can be modeled if ligands can bind different microstates of the enzyme with different affinities that can vary under the different experimental conditions. Such a model is also applicable to different microstates of the ligand. The compensation effect noted here is somewhat different. In fact, we observe that different "microstates" of a ligand can have the same affinity for one conformation (or microstate) of an enzyme and that different "microstates" of a complex can have the same stability under one "experimental" condition. The fact that the calculations give rise to the compensation should be informative as to the mechanism of the experimental compensations observed as well as the relevance of our calculations in the understanding of the dissociation of protein complexes in solution.

Since the main structural differences are located at the Arg and Adda side chains (Fig. 6), we will focus on the implications of these side chains in a putative compensation mechanism. For the dissociation of docked complexes with the Adda side chain more solvent exposed and the Arg side chain lying closer to PP-1c (forming salt-brides and/or H-bonds, see Fig. 2) compared to the X-ray complex, the relative ratio $\Delta ASA_{pol}/\Delta ASA_{np}$ will increase. This leads to a more unfavorable $\Delta H^{\circ}_{d}(25)$. In other words, the enthalpy of such a docked complex (PP-1c:MLR and the solvent) will decrease compared to the X-ray complex. Indeed, the proportion of polar and more enthalpically favorable (Makhatadze & Privalov, 1995) interactions per Å^2 of buried surface in the complex will increase. Therefore, more positive (unfavorable) ΔH°_{d} values for the docked complexes are obtained. On the other hand, we observe that the reduction in entropy of the solvent upon dissociation becomes less unfavorable for such complexes, i.e., $\Delta S^{\circ}_{solv}(25)$ is less negative than for the X-ray complex. In other words, the entropy of the complexes (PP-1c, MLR, and the solvent) will decrease. Variations in ΔS_{conf} for the $\Delta G^{\circ}_{d,corr}$ and ΔG°_{comp}

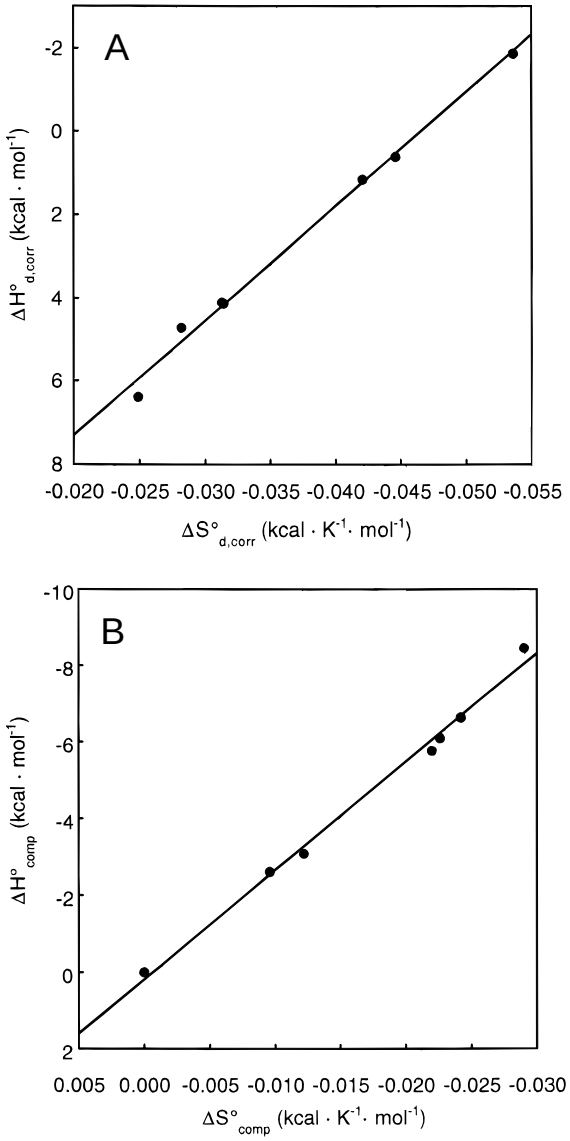

Fig. 7. Enthalpy–entropy compensation. **A:** Plot of the $\Delta H_{d,corr}$ as a function of the $\Delta S_{d,corr}$ for the seven complexes represented in Figure 6. The slope of the line is 280 K. **B**: Plot of the ΔH_{comp} as a function of the ΔS_{comp} for the seven complexes presented in Figure 6. The slope of the line is 281 K.

were found to be about one order of magnitude lower than that of the ΔS_{solv}. Therefore, in the system studied here (PP-1c, MLR, and the solvent), the response to a decrease in enthalpy (more favorable PP-1c:MLR interactions) is a decrease in entropy of water. The variations in enthalpy ($\Delta H^{\circ}_{d}(25)$ or $\Delta H^{\circ}_{comp}(25)$) are almost exactly compensated by the variations in entropies ($-T\cdot\Delta S^{\circ}_{d}(25)$ or $-T\cdot\Delta S^{\circ}_{comp}(25)$) giving rise to the linear relationships observed (Fig. 7) and leaving $\Delta G^{\circ}_{d,corr}$ or ΔG°_{comp} unchanged. It is interesting to note that the slopes of Figure 7A and 7B are equal to ~280 K, which is of the same order of magnitude of the values for T_c values reported in the literature and which were attributed to the implication of water (Lumry & Rajender, 1970; Lumry, 1995). The coincidence in T_c values could be an indication that the compensation mechanism described above is realistic as far as the role of water and that the parameterization used here is faithful enough to simulate entropy–enthalpy compensations that occur in the PP-1c:MLR complex in solution (and complex dissociation). The mechanism described could also apply for the compensation effects observed under different experimental conditions referred to above. The existence of structural and energy fluctuations in proteins are both documented on theoretical (Cooper, 1976) and experimental backgrounds (Frauenfelder et al., 1988). These fluctuations remind us that the structure and thermodynamic parameters (such as the enthalpy of dissociation of a complex) observed for proteins and protein complexes are weighted mean values or ensemble averages.

Cooperativity of dissociation and multiple conformations (microstates) for the PP-1c:MLR complex

According to Boltzmann's statistics, the docked complexes that give rise to the largest ΔG°_{d} (or that are the most stable) should correspond to the most probable conformations within the limited ensemble of complexes generated here. We have calculated the relative population (probabilities) of the complexes according to the following equation:

$$P_i = \frac{\exp(-\Delta G_{comp,i}/RT)}{\sum_i \exp(-\Delta G_{comp,i}/RT)}, \quad (1)$$

where $\Delta G^{\circ}_{comp,i}$ is the relative free energy of the complex compared to the crystal structure, R is the gas constant, and T is the absolute temperature.

Figure 8A shows the P_i of every complex as a function of the positional RMSD of all heavy atoms of the MLR. As can be seen, there is an increase in P_i as the RMSD reaches around 3 Å. But one can also notice the scatter in RMSD. As discussed before, this is a manifestation of the enthalpy–entropy compensation. We show in Figure 8B the same P_I as a function of the positional RMSD of the backbone of the toxin. Here, the increase is sharper indicating a cooperative role for the cyclic portion of the MLR molecule. This is quite interesting because the cyclic portion (Masp, D-Glu) is involved in the molecular recognition. Since the D-Glu and Masp side chains have only one rotable bond (Fig. 1A), the position of the backbone controls the position and the burial of the two carboxylic groups at the interface of the complex. Therefore, Figure 8B demonstrates that if the residues that are involved in making the specific interactions are not properly buried by not being in the right position, the probability of the particular complex will be low. Moreover, Figure 8A indicates that the structure of hydrophobic side chains (e.g., Adda) is more likely to fluctuate or to have different conformations in the significantly populated complexes. Therefore, Figure 6 could represent a dynamic rendition of the complex on a short time scale as this type of motion occurs between states of similar free energy. It has to be noted that this rendition would be partial as all the different possible configurations are not present in the limited ensemble of complexes used here.

Are the motions in the complex and their amplitude as seen on Figure 6 in contradiction with the more "static" structure of the X-ray complex? In the light of the work of Cooper (1976) and Frauenfelder et al. (1988), we think not. Moreover, issues such as the temperature at which the diffraction data has been recorded (100 K; Goldberg et al., 1995) certainly have to be considered. On

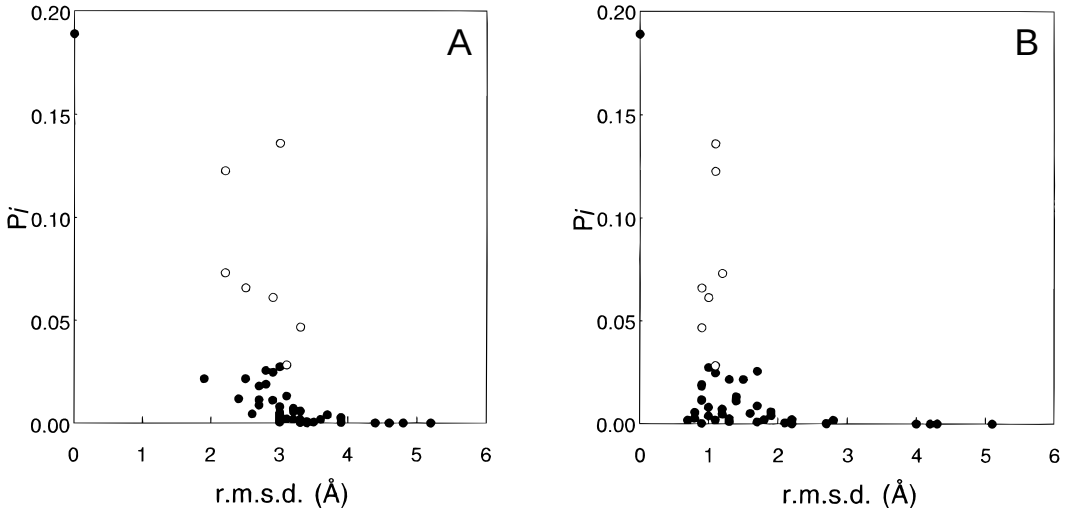

Fig. 8. The position of the backbone of MLR dictates the cooperativity of dissociation. **A:** The P_i of every complex as a function of the positional RMSD of all heavy atoms of the MLR. P_i is defined in the text. **B:** The same probabilities as in **A**, except this time as a function of the positional RMSD of the backbone of the toxin. The open circles refer to the complexes on Figure 6.

the other hand, to confirm our results suggesting that the complex between PP-1c and MLR shows significant structural fluctuations, NMR relaxation studies should be done on MLR bound to PP-1c. Experimental efforts toward such measurements are being pursued in our laboratory but are complicated by the high molecular weight of the complex and solubility issues. Experimental thermodynamic measurements could also shed light on the existence and the extent of the structural fluctuations. Indeed, if the experimental ΔH_d° for the PP-1c:MLR complex (corrected for proton transfer effects) could be obtained and observed to be more unfavorable than the one calculated from the X-ray complex, this would indicate the population of different microstates of the complexes with different enthalpies. It is worth recalling that the ΔH_d° measured would correspond to a weighted mean value or ensemble average ($\langle \Delta H_d^\circ \rangle$):

$$\langle \Delta H_{d,corr}^\circ \rangle = \sum_i P_i \cdot \Delta H_{d,corr,i}^\circ \qquad (2)$$

where $\Delta H_{d,corr,i}^\circ$ is the enthalpy of dissociation for each complex or microstates of an ensemble. For example, the $\langle \Delta H_{d,corr}^\circ \rangle$ for the limited ensemble of complexes generated here is 2.05 kcal·mol^{-1} compared to -1.87 kcal·mol^{-1} calculated from the X-ray complex.

Implications for molecular recognition

It is clear from our results that different conformations (microstates) than the X-ray MLR can bind to one microstate of PP-1c (X-ray PP-1c) with similar affinities. The major structural differences between these microstates are located at side chains (e.g., Adda and Arg) that are not involved in forming specific interactions as evidenced from the X-ray complex (Fig. 2). On the other hand, the side chains that are involved in molecular recognition (Masp and D-Glu) occur in the best-defined region of MLR in solution (Fig. 1). The fact that the portion or domain responsible for molecular recognition is already folded is an advantage for effective rates of binding (k_{on}). Indeed, all the microstates are potentially able to recognize PP-1c. In addition, the fact that the Adda and Arg side chains can exist many different conformations without affecting the ΔG_d° should also help for elevated k_{on} rates. There is a limit, however, to which the Adda side chain can vary its conformation without changing ΔG_d° significantly. Indeed, a complex generated by manually changing the conformation of the Adda side chain in the X-ray complex so that it is maximally exposed had ΔG_d° decrease by 2 kcal·mol^{-1} or a 100-fold decreased K_d (data not shown).

Our results might also have general implications for the study of free ligands by solution state NMR. Indeed, our study indicates that all the members of the ensemble of structures usually calculated should be considered as being potentially able to bind a receptor and not only the geometric average. Although the geometric average structure of MLR here was found to be in the cluster of seven complexes of high ΔG_d°, it could well not have been. In fact, for the calculation of a geometric average, all the members of the ensemble are given the same weight and therefore can lead to a biased structure by conformers that may not be populated in solution although they satisfy the experimental NMR restraints. The best defined regions, as evidenced from the ensemble, could potentially be involved in molecular recognition. Moreover and as shown here, the worst defined regions of a ligand can contribute significantly to the affinity of binding and as such cannot be ignored for their potential importance in the binding process.

Conclusion

Using a structure-based approach, we have calculated form the X-ray complex of PP-1c and MLR a free energy of dissociation that is in close agreement with the reported K_i. We also note that the residues with the largest contribution to the overall dissociation free energy to be Arg96, Arg221 on PP-1c and D-Glu and Adda on MLR. This is in accordance with experimental data that shows that they are indeed critical for complex formation. We notice that the reported high affinity of MLR is due to its cyclic nature, i.e., the dissociation reaction is not linked to a conformational change in the backbone and therefore leads to a minimal conformational entropy gain compared to a fictitious linear version of the toxin. The analysis of 47 complexes, obtained from docking a family of

NMR solution structures of MLR onto the crystal structure of the PP-1c, predicts structural fluctuations for the bound form of MLR especially for the long Adda and Arg side chains. This analysis also suggests that the cyclic part of the MLR is more important that the long hydrophobic Adda side chain in the cooperativity of dissociation (binding). We observe an enthalpy–entropy compensation phenomenon for which we describe a putative mechanism that could be applicable for experimentally observed compensation effects and to understand the mechanisms of structural fluctuations and molecular dynamics in solution.

Materials and methods

Docking procedure

Metropolis Monte Carlo docking of the 46 calculated solution structures and the average minimized solution structure of MLR to the crystal structure of PP-1c was accomplished using the Monte Carlo macro in Insight II version 2.3. This technique proved useful in docking other toxins as well (Bagu et al., 1997). Based on previous work (Bagu et al., 1997), the starting positions of all calculated MLR solution structures were determined by superimposing their backbone atoms onto the backbone of bound MLR (Goldberg et al., 1995), which was then removed. Docking calculations were performed using 2,000 trials at a temperature of 50 K (optimization from Bagu et al., 1997). The force field used was CVFF (Dauber-Ogusthorpe et al., 1988).

This docking technique involves the minimization of the Van der Waals and Coulomb potential energies between two rigid bodies by altering their relative positions (in this case the MLR and PP-1c structures). The new state is rejected or accepted based on the new potential energies. Normally accepted states have reduced energies; however, the docking procedure allows higher energy states to be accepted on occasion in order that the docked structures are not trapped in a local energy minimum. However, the jump to higher energy states that are accepted become reduced as the number of docking iterations increases. By about 2,000 trials accepted states at higher energy than the previous accepted state have only minor increases in energy.

Structure-based free energy calculations

We describe in the following paragraphs the approach that we used to calculate the free energy of dissociation from the parameterization of the heat capacity, enthalpy, and entropy developed by Murphy and Freire (Murphy & Freire, 1992; Baker & Murphy, 1998; Luque & Freire, 1998). This parameterization has been verified extensively against experimental heat capacities, enthalpies, and entropies of unfolding and/or binding (Murphy & Freire, 1992; Baker & Murphy, 1998; Luque & Freire, 1998, and references therein).

The free energy of dissociation of a complex, $\Delta G_d^\circ(T)$, is classically given by the following equations:

$$\Delta G_d^\circ(T) = -RT\cdot\ln K_d, \tag{3}$$

$$\Delta G_d^\circ(T) = \Delta H_d^\circ(T) - T\cdot\Delta S_d^\circ(T), \tag{4}$$

where $\Delta H_d^\circ(T)$ is the temperature dependent standard enthalpy of dissociation and $\Delta S_d^\circ(T)$ is the temperature dependent standard entropy of dissociation.

The calculation of the $\Delta H_d^\circ(T)$

The following description addresses the structural parameterization of protein unfolding enthalpy calculation. Since protein unfolding and protein dissociation are governed by the same molecular forces, the parameterization developed for protein unfolding is assumed to apply to protein dissociation (Gomez & Freire, 1995) throughout the text.

$\Delta H_d^\circ(T)$ is given by

$$\Delta H_d^\circ(T) = \Delta H_d^\circ(T^\circ) + \Delta C_p\cdot(T - T^\circ), \tag{5}$$

where $\Delta H_d^\circ(T^\circ)$ is a standard reference enthalpy of dissociation at some reference temperature (T°). $\Delta C_p(T)$ is the temperature independent heat capacity change upon dissociation of the complex. It is a good approximation to consider ΔC_p temperature independent from 0 to 85 °C (Gomez et al., 1995).

The experimental ΔC_p of unfolding of a series proteins for which the both the thermodynamics of unfolding and the crystal structure were well characterized can be reliably calculated by a linear combination of the change in solvent ASA of polar ($\Delta\mathrm{ASA_{pol}}$) and nonpolar ($\Delta\mathrm{ASA_{np}}$) atoms through the following empirical relationship (Murphy & Freire, 1992):

$$\Delta C_p = 0.45\cdot\Delta\mathrm{ASA_{np}} - 0.26\cdot\Delta\mathrm{ASA_{pol}}, \tag{6}$$

where the parameters have units of $\mathrm{cal\cdot K^{-1}\cdot mol^{-1}\cdot Å^{-2}}$. ΔC_p of protein unfolding has been shown to come mainly from the hydration of atoms that become exposed to the solvent upon unfolding and as can be seen in the preceding relationship the hydration polar atoms and nonpolar atoms have opposite contributions (Murphy & Freire, 1992; Gomez et al., 1995).

Similarly, it has been shown that the experimental unfolding enthalpy $\Delta H_d^\circ(T)$ at 60 °C of the same series of proteins could be calculated, within 6% error, with the following empirical rule (Xie & Freire, 1994; Hilser et al., 1996):

$$\Delta H_d^\circ(60) = 31.4\cdot\Delta\mathrm{ASA_{pol}} - 8.44\cdot\Delta\mathrm{ASA_{np}}, \tag{7}$$

where the parameters have units of $\mathrm{cal\cdot mol^{-1}\cdot Å^{-2}}$. It is implicit that this empirical function accounts for the change in enthalpy resulting for breaking noncovalent bonds (H-bonds, salt bridges, and van der Waals interactions, etc.) and solvating these atoms (Hilser et al., 1996). Other contributions such as proton transfer have to be accounted for if they are coupled to the unfolding or dissociation process (Gomez & Freire, 1995; Baker & Murphy, 1997).

Therefore, once the changes in ASA for a dissociation are calculated, a corresponding $\Delta H_d^\circ(T)$ can be calculated by combining Equations 6 and 7:

$$\Delta H_d^\circ(T) = \Delta H_d^\circ(60) + \Delta C_p\cdot(T - 60), \tag{8}$$

The calculation of the $\Delta S_d^\circ(T)$

In the absence of proton transfer, the standard entropy of dissociation ΔS_d° can be assumed to correspond to the sum of the changes in solvation entropy (ΔS_{sol}(T)), conformational entropy (ΔS_{conf}), and overall rotational/translational entropy (ΔS_{rt}) to account for the appearance of an additional kinetic unit upon dissociation:

$$\Delta S_d^\circ(T) = \Delta S_{sol}(T) + \Delta S_{conf} + \Delta S_{rt}. \tag{9}$$

Of the three contributions, only ΔS_{sol} is assumed temperature dependent and can be broken into contributions arising from the change in solvation entropy resulting from the solvation of polar and nonpolar atoms that become exposed upon dissociation. It has been shown that the solvation entropy of nonpolar atoms is zero at 112 °C (Baldwin, 1986) and that the temperature at which solvation entropy of polar atoms equals zero is close to 62 °C (D'Aquino et al., 1996). Therefore, it has been proposed that $\Delta S_{sol}(T)$ (D'Aquino et al., 1996) can be parameterized by the following relationship:

$$\Delta S_{sol}(T) = 0.45 \cdot \Delta \mathrm{ASA}_{\mathrm{np}} \cdot \ln(T/384.15) - 0.26 \cdot \Delta \mathrm{ASA}_{\mathrm{pol}} \cdot \ln(T/335.15) \quad (10)$$

where the coefficient 0.45 and −0.26 are the ones described in Equation 6.

Murphy et al. (1994) proposed the following scheme to account for the change in conformational entropy (ΔS_{conf}) for protein dissociation:

$$\Delta S_{conf} = \Delta S_{bu \to ex} + \Delta S_{ex \to u} + \Delta S_{bb} \quad (11)$$

where $\Delta S_{bu \to ex}$ is the gain in conformational entropy of a side chain when it becomes exposed after disruption of tertiary or quaternary interactions, $\Delta S_{ex \to u}$ is the change in conformational entropy of the side chain when the secondary structure unfolds, and ΔS_{bb} is the gain in conformational entropy from the backbone upon unfolding. $\Delta S_{bu \to ex}$, $\Delta S_{ex \to u}$, and ΔS_{bb} values for all the amino acids have been estimated from a statistical mechanical analysis (Lee et al., 1994; D'Aquino et al., 1996). In our calculations, we use the values reported by (D'Aquino et al., 1996).

In the present case, it is assumed that no conformational change in the backbones of the PP-1c and MLR occur upon dissociation. This is supported by the fact that MLR is a cyclic peptide and our earlier findings that the conformation of the cyclic backbone of the solution structure of MLR is identical to that of the bound state of MLR in the crystal structure of the PP-1c:MLR complex (Bagu et al., 1995, 1997; Goldberg et al., 1995). The crystal structure of the free PP-1c was shown to be almost identical to that of the complexed form (Egloff et al., 1995; Goldberg et al., 1995) indicating that there should not be any major conformational change occurring in the free form of the enzyme. Therefore, the change in conformational entropy of dissociation of PP-1c is assumed to originate solely from the gain in conformational entropy of the side chains that become exposed upon dissociation ($\Delta S_{bu \to ex}$) and is scaled as the fraction of the total ASA of the side chain that is gained and computed according to the following equation:

$$\Delta S_{conf} = \sum_i \frac{\Delta \mathrm{ASA}_i}{\mathrm{ASA}_i} \cdot \Delta S_{bu \to ex} \quad (12)$$

where $\Delta \mathrm{ASA}_i$ is the change in ASA of the side chain of residue i and ASA_i is the ASA of the corresponding side chain in a fully exposed state. Here, we used the ASA values reported by Miller et al. (1987). Special care had to be taken for the conformational entropy and the ASA of the nonnatural side chains of the toxins (see Fig. 1). For the Adda side chain, we used the empirical equation proposed by Bardi et al. (1997) for nonnatural peptidyl side chains that relates the number of rotable bonds (1.76 $\mathrm{cal \cdot K^{-1} \cdot mol^{-1}}$ per rotable bond) and the number of atoms (0.414 $\mathrm{cal \cdot K^{-1} \cdot mol^{-1}}$ per atom) to correct for excluded volume effects. A value of 7.8 $\mathrm{cal \cdot K^{-1} \cdot mol^{-1}}$ is calculated for the Adda side chain. The ASA of Adda side chain was calculated in the free toxins and amounts to 387 Å^2. The other side chains that had to be ascribed with a ΔS_{conf} where the D-Glu and Masp residues (see Fig. 1). As can be seen, these side chains consists in a single carboxylate that can rotate around their Cα(sp3)-COO(sp2) bond. As discussed by Pickett and Sternberg (1993), the $\mathrm{COO^-}$ has a symmetry number of 2 leading to distinguishable rotamers on only 180° and approximating that the conformational entropy in the buried state is $R \cdot \ln 2$ rather than 0 ($R \cdot \ln 1$). We assume the change in the conformational entropy of the D-Glu and Masp side chains to be equal to 0.8 $\mathrm{cal \cdot K^{-1} \cdot mol^{-1}}$ ($-R \cdot [\ln 2 - \ln 3]$) by supposing that three distinguishable rotamers can be adopted by the side chain when free to rotate in the free form of the toxins. We also calculated an ASA of 90 Å^2 for the side chains.

The gain in translational and rotational entropy (ΔS_{rt}) seems to be well accounted for by to the cratic entropy (Murphy et al., 1994; Gomez & Freire, 1995). The cratic entropy is equal to $R \ln(1/55)$ where the ratio is the mole fraction of the additional particle appearing (mixing ideally) upon dissociation at a fictitious 1 M standard state in water (Kauzmann, 1959). The cratic entropy amounts to 8 $\mathrm{cal \cdot K^{-1} \cdot mol^{-1}}$ or 2.4 $\mathrm{kcal \cdot mol^{-1}}$ at 25 °C. On the other hand, the use of this value and its physical basis is a matter controversy (Holtzer, 1995; Gilson et al., 1997). However, recent experimental evidence (Tamura & Privalov, 1997; Yu et al., 1998) and theoretical arguments (Amzel, 1996) indicate that the loss in rotational and translational entropy is numerically close to the cratic entropy. Therefore, we are also using a value of 8 $\mathrm{cal \cdot K^{-1} \cdot mol^{-1}}$ to account for the ΔS_{rt}.

The STC program suite

To perform the free energy calculations from the structure of the different complexes, we developed a suite of programs called STC (structure-based thermodynamics calculation). In essence, STC consists of two modules. The first module, *CALCASA*, calculates the change in ASA for the dissociation process from the coordinate files in the Brookhaven Protein Data Bank (PDB) format using the algorithm ANAREA (Richmond, 1984) as implemented in the program VADAR (Wishart et al., 1994). The output files consists in the tabulated ASA of every atom of the complex and of both the free forms of the enzyme and the ligand as well the difference in ASA for each atom. The total changes in nonpolar (all carbon atoms and sulfur atoms) and polar (all oxygen and nitrogen atoms) are summed up. In addition, the atomic ΔASA are regrouped per residue (and per side chain) for the calculation of ΔS_{conf} as described above in the next module.

The module *THERMO* calculates the energetics from the ΔASA. From the total changes in $\Delta \mathrm{ASA}_{\mathrm{np}}$ and $\Delta \mathrm{ASA}_{\mathrm{pol}}$, the contribution of nonpolar and polar atoms to ΔC_p and $\Delta H_d^\circ(60)$ are calculated. Then according to Equation 8, the ΔH_d° at the desired T is calculated. In the present study, all the calculations are done at 25 °C. Similarly, using the proper reference temperatures, the ΔS_{sol} is extrapolated at 25 °C using Equation 10. From the ΔASA of the different side chains involved in the dissociation, the conformational entropy gained for the ligand and the enzyme is calculated with Equation 12. The total entropy change is then taken to be the sum of all the entropic contributions listed in Equation 9. A $\Delta G_d^\circ(25)$ is calculated using Equation 3. The program STC can be downloaded from the following site: http://www.pence.ca/ftp.

Acknowledgments

This work was supported by the Protein Engineering Network of Centres of Excellence of Canada. Pierre Lavigne acknowledges the Medical Research Council of Canada for a postdoctoral fellowship. John R. Bagu acknowledges the Alberta Heritage Foundation for Medical Research for a postgraduate scholarship.

References

Abagyan R, Totrov M. 1994. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptide and proteins. *J Mol Biol 235*:983–1002.

Amzel LM. 1996. Loss of translational entropy in binding, folding and catalysis. *Proteins Struct Funct Genet 29*:1–6.

Bagu JR, Sönnichsen FD, Williams D, Andersen RJ, Sykes BD, Holmes CFB. 1995. Comparison of the solution structures of microcystin-LR and motuporin. *Nature Struct Biol 2*:114–116.

Bagu JR, Sykes BD, Craig MM, Holmes CFB. 1997. A molecular basis for different interactions of marine toxinse with protein phosphatase-1. *J Biol Chem 272*:5087–5097.

Baker BM, Murphy KP. 1997. Dissecting the energetics of a protein-protein interaction: The binding of ovomucoid third domain to elastase. *J Mol Biol 268*:557–569.

Baker BM, Murphy KP. 1998. Prediction of binding energetics from structure using empirical parameterization. *Methods Enzymol 295*:295–315.

Baldwin RL. 1986. Temperature dependence of the hydrophobic interaction in protein folding. *Proc Natl Acad Sci USA 83*:8069–8072.

Bardi J, Luque I, Freire E. 1997. Structure-based thermodynamic analysis of HIV-1 protease inhibitor. *Biochemistry 36*:6588–6596.

Barford D. 1995. Protein phosphatases. *Curr Opin Struct Biol 5*:728–734.

Brady GP, Sharp KA. 1997. Entropy in protein folding and in protein-protein interactions. *Curr Opin Struct Biol 7*:215–221.

Cohen P. 1989. The structure and regulation of protein phosphatases. *Annu Rev Biochem 58*:453–508.

Cooper A. 1976. Thermodynamic fluctuations in protein molecules. *Proc Natl Acad Sci USA 73*:2740–2741.

Craig M, McCready RL, Luu HA, Smillie MA, Dubord P, Holmes CFB. 1993. Identification and characterization of hydrophobic microcystins in canadian freshwater cyanobacteria. *Toxicon 31*:1541–1549.

Cummings MD, Hart TN, Read RJ. 1995. Atomic solvation parameters in the analysis of protein-protein docking results. *Protein Sci 4*:2087–2099.

D'Aquino JA, Gomez J, Hilser VJ, Lee KH, Amzel LM, Freire E. 1996. The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct Funct Genet 25*:143–156.

Dauber-Ogusthorpe P, Roberts VA, Ogusthorpe DJ, Wolff J, Genest M, Hagler AT. 1988. Structure and energetics of ligand binding to proteins: *E. coli* dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins Struct Funct Genet 4*:31–47.

Eftink MR, Anusiem AC, Biltonen RL. 1983. Enthalpy-entropy compensation and heart capacity changes for protein-ligand interactions: General thermodynamic models and data for the binding of nucleotides to ribonuclease A. *Biochemistry 22*:3884–3896.

Egloff MP, Cohen PTW, Reinemer P, Barford D. 1995. Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *J Mol Biol 254*:942–959.

Frauenfelder H, Parak F, Young RD. 1988. Conformational substates in proteins. *Ann Rev Biophys Chem 17*:451–479.

Gilson MK, Given AG, Bush BL, McCammon JA. 1997. The statistical thermodynamic basis for computation of binding affinities: A critical review. *Biophys J 72*:1047–1069.

Goldberg J, Huang H-B, Kwon Y-B, Greengard P, Nairn AC, Kuriyan J. 1995. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature 376*:745–753.

Gomez J, Freire E. 1995. Thermodynamic mapping of the inhibitor site of the aspartic protease endopepstatine. *J Mol Biol 252*:337–350.

Gomez J, Hilser VJ, Xie D, Freire E. 1995. The heat capacity of proteins. *Proteins Struct Funct Genet 22*:404–412.

Hawkes R, Grutter MG, Schellman J. 1984. Thermodynamic stability and point mutations of bacteriophage T4 lysozyme. *J Mol Biol 175*:195–212.

Hilser VJ, Gomez J, Freire E. 1996. The enthalpy change in protein folding and binding: Refinement of parameters for structure-based calculations. *Proteins Struct Funct Genet 26*:123–133.

Holtzer A. 1995. The "cratic correction" and related fallacies. *Biopolymers 35*:595–602.

Huang H-B, Horiuchi A, Goldberg J, Greengard P, Nairn AC. 1997. Site-directed mutagenesis of amino acid residues of protein phosphatase 1 involved in catalysis and inhibitor binding. *Proc Natl Acad Sci USA 94*:3530–3535.

Jackson RM, Sternberg MJE. 1995. A continuum model for protein-protein interactions: Application to the docking problem. *J Mol Biol 250*:258–275.

Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv Protein Chem 14*:1–63.

Lee KH, Xie D, Freire E, Amzel LM. 1994. Estimation of changes in side-chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins Struct Funct Genet 20*:68–84.

Lumry R. 1995. The new paradigm for protein research. In: Gregory RB, ed. *Protein-solvent interactions*. New York: Marcel Dekker, Inc. pp 1–136.

Lumry R, Rajender S. 1970. Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous property of water. *Biopolymers 9*:1125–1227.

Luque I, Freire E. 1998. Structure-based prediction of binding affinities and molecular design of peptide ligands. *Methods Enzymol 295*:295–315.

MacKintosh RW, Dalby KN, Campbell DG, Cohen PTW, Cohen P, MacKintosh C. 1995. The cyanobacterial toxin microcystin binds covalently to cysteine-273 on protein phosphatase 1. *FEBS Lett 371*:236–240.

Makhatadze GI, Privalov P. 1995. Energetics of protein structure. *Adv Protein Chem 47*:307–417.

Miller S, Janin J, Lesk AM, Chothia C. 1987. Interior and surface of proteins. *J Mol Biol 196*:641–656.

Moult J. 1997. Comparison of database potentials molecular mechanics force fields. *Curr Opin Struct Biol 7*:194–199.

Murphy KP, Freire E. 1992. Thermodynamics of structural stability and cooperative folding behaviour in proteins. *Adv Protein Chem 43*:313–361.

Murphy KP, Xie D, Thompson KS, Amzel LM, Freire E. 1994. Entropy in biological processes: Estimation of the translational entropy loss. *Proteins Struct Funct Genet 18*:62–67.

Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct Funct Genet 11*:281–296.

Nishiwaki-Matsushima R, Nishiwaki S, Ohta T, Yoshizawa S, Suganuma M, Harada K-I, Watanabe MF, Fujiki H. 1991. Structure-function relationships of microcystin, liver tumor promoters, in interaction with protein phosphatase. *Jpn J Cancer Res 82*:993–996.

Novotny J, Bruccoleri RE, Davis M, Sharp KA. 1997. Empirical free energy calculations: A blind test and further improvement of the method. *J Mol Biol 268*:401–411.

Ohta T, Nishiwaki R, Yatsunami J, Komori A, Suganuma M, Fujiki H. 1992. Hyperphosphorylation of cytokeratins 8 and 18 by microcystin-LR, a new liver tumor promoter, in primary cultured rat hepatocytes. *Carcinogenesis 13*:2443–2447.

Pickett S, Sternberg MJE. 1993. Empirical scale of side-chain conformational entropy in protein folding. *J Mol Biol 231*:825–839.

Richmond TJ. 1984. Solvent accessible surface area and excluded volumes in proteins. *J Mol Biol 82*:63–89.

Runnegar M, Berndt N, Kong SM, Lee EYC, Zhang L. 1995. In vivo and in vitro binding of microcystin to protein phosphatase 1 and 2A. *Biochem Biophys Res Commun 216*:162–169.

Shenolikar S. 1994. Protein serine/threonine phosphatases-new avenues for cell regulation. *Annu Rev Cell Biol 10*:55–86.

Shortle D, Meeker AK, Freire E. 1988. Stability mutants of staphylococcal nuclease: Large compensation enthalpy-entropy changes for the reversible denaturation reaction. *Biochemistry 27*:4761–4768.

Stotts RR, Namikoshi M, Haschek WM, Rinehart KL, Carmichael WW, Dahlem AM, Beasley VR. 1993. Structural modifications imparting toxicity in microcystins from microcystis spp. *Toxicon 41*:783–789.

Takai A, Sasaki K, Nagai H, Mieskes G, Isobe M, Isono K, Yasumoto T. 1995. Inhibition of specific binding of okadaic acid to protein phosphatase-2A by microcystin-LR, calyculin-A and tautomycin: Method of analysis of interactions of tight-binding ligands with target protein. *Biochem J 306*:657–665.

Tamura A, Privalov PL. 1997. The entropy of protein association. *J Mol Biol 273*:1048–1060.

Vajda S, Sippl M, Novotny J. 1997. Empirical potentials and functions for protein folding and binding. *Curr Opin Struct Biol 7*:222–228.

Weng Z, Vajda D, DeLesi C. 1996. Prediction of complexes using empirical free energy functions. *Protein Sci 5*:614–626.

Wishart DS, Willard L, Richards FM, Sykes BD. 1994. VADAR: A comprehensive program for protein structure evaluation. Version 1.2, University of Alberta, Edmonton, Alberta, Canada.

Xie D, Freire E. 1994. Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. *Proteins Struct Funct Genet 19*:291–301.

Yu Y, Lavigne P, Kay CM, Hodges RS, Privalov PL. 1998. Contribution of translational and rotational entropy to the unfolding of a dimeric coiled-coil. *J Phys Chem B 103*:2270–2278.