Indian Journal of Chemistry
Vol.39A, Jan-March 2000, pp. 262-273

Free energy analysis of enzyme-inhibitor binding: The carboxypeptidase A-inhibitor complexes

Parul Kalra, Achintya Das, Surjit B. Dixit & B Jayaram*
Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi 110 016, India
Email: bjayaram@chemistry.iitd.ernet.in

Received 20 October 1999; accepted 22 November 1999

Developing free energy estimates of biological molecules starting from a molecular description of the solute, solvent and the salt, is currently in the domain of computationally intractable problems. However, structure based drug design efforts involving for instance, designing a suitable inhibitor molecule with desired binding attributes targeted to an active site on an enzyme necessitates free energy estimates. We present here a computationally expedient and rigorous methodology to develop and analyse the thermodynamics of enzyme-inhibitor binding starting from crystal structures. The complexes of carboxypeptidase A with five inhibitors with known structural and binding constant data have been adopted for this study as illustrative cases. The standard free energy of complexation is considered in terms of a thermodynamic cycle of six distinct steps decomposed into a total of 18 well-defined components. The model we employ involves explicit all atom accounts of the energetics of electrostatic interactions, solvent screening effects, van der Waals components and cavitation effects of solvation combined with a Debye-Huckel treatment of salt effects. Estimates of entropy loss due to decreased translational and rotational degrees of freedom in the complex relative to the unbound species based on classical statistical mechanics are included, as well as corresponding changes in the vibrational and configurational entropy. The magnitudes and signs of the various components are estimated using the AMBER parm94 force field, generalized Born theory and solvent accessibility measures. The calculated standard free energies of formation agree with experiment in these systems to within 5-12 kcal/mol. This generates considerable optimism in the potential viability of the methodology for drug design. Fine tuning of the computational protocols, inclusion of structural adaptation effects and a careful examination and minimization of possible errors are some areas for further research. The net binding free energies are a resultant of several competing contributions with 7 out of the 18 terms favouring complexation. A component-wise analysis of the binding free energy for the five carboxypeptidase A-inhibitor complexes studied here indicates that the nonelectrostatic contributions, i.e. the net van der Waals interactions and the differential cavitation effects are favourable to binding. Electrostatic contributions averaged over the five systems turn out to be favourable despite the desolvation expense incurred during binding. Analyses on these lines yield pointers to structural modifications to be attempted to accomplish optimal binding besides presenting a molecular energetic perspective of induced-fit mechanisms.

I. Introduction

Almost a century ago, Emil Fischer[1] proposed that enzymes interact with their substrates via a lock and key fit mechanism. Over the years, this hypothesis received several confirmations but was enlarged to encompass electrostatic complementarity as well as induced fit. However, our understanding of what contributes favourably to enzyme-substrate binding in terms of the intermolecular forces, considered together with medium effects, remains ambiguous. A molecular level appreciation of forces to be regulated for optimal binding and a methodology for expeditious estimation of binding free energies are of immense practical interest in structure based drug design efforts in the pharmaceutical industry, especially in screening a multitude of candidate molecules to identify lead compounds.

The nature of binding and specificity in molecular association is a problem of central interest in chemistry and biology. In chemical systems, binding energies and free energies can be calculated accurately by molecular quantum mechanics in the gas phase[2] and by molecular dynamics and free energy simulations in condensed phases[3-7]. In biological systems, the situation is somewhat more complex, since interactions among charged and conformationally flexible macromolecules in solution must usually be considered. For this class of problems, molecular quantum mechanics is intractable. Free energy simulations while possible in principle remain computationally intensive, and the results may be subject to convergence problems and statistical uncertainties. As a result, linking structure with thermodynamics in macromolecular complexes poses severe computational challenges.

Table 1-The carboxypeptidase A-inhibitor complexes investigated and their PDB codes

	PDB Code	Name of the complexes
1	7CPA	Carboxypeptidase A with O-[[(1R)-[[N-(phenyl methoxycarbonyl)- L-phenylalanyl]amino]isobutyl]hydroxyphosphinyl]-L-3-phenyllactate [ZFV^{p}(O)F]
2	6CPA	Carboxypeptidase A with O-[[(1R)-[[N-(phenyl methoxycarbonyl)- L-alanyl]amino]ethyl]hydroxyphosphinyl]-L-3-phenyllactate[ZAA^{p}(O)F]
3	8CPA	Carboxypeptidase A with O-[[(1R)-[[N-(phenyl methoxycarbonyl)- L-alanyl]amino]methyl]hydroxyphosphinyl]-L-3-phenyllactate [ZAG^{p}(O)F]
4	2CTC	Carboxypeptidase A with L-Phenyllactate [LOF]
5	1CBX	Carboxypeptidase A with L-Benzylsucinate [BZS]

Recent improvements in the description of intermolecular interactions using empirical force fields[8,9] and the development of a new methodology for obtaining estimates of the free energy of solvation simply but accurately using the "generalized Born-solvent accessibility" (GBSA) model[10-15] enable us to determine the various components of binding processes involving non-covalent associations. Using these developments, we describe herein a theoretical "component analysis" of the standard free energy of binding for some enzyme-inhibitor complexes. Enzyme-substrate complexes are a class of systems fundamental to catalytic processes in biology, in which the nature of the exquisite specificity and catalytic efficiency is yet to be fully explained at the molecular level. A proper theoretical account of the thermodynamics of binding is a necessary prerequisite for understanding specificity in general.

The focus of our current investigations is the binding of the enzyme carboxypeptidase A (Fig. 1) with the five inhibitors[16-25] (Figs 2(a-e)) listed in Table 1. Carboxypeptidase A is a digestive enzyme containing 307 amino acid residues in a single polypeptide chain and a zinc ion which is essential for catalysis located in the groove at the surface of the molecule. The zinc ion is coordinated in a tetrahedral fashion to two histidines (HIS 69, HIS 196), one glutamate (GLU 72) and a water molecule. The enzyme hydrolyzes the carboxyl terminal peptide bond in polypeptide chains. Hydrolysis occurs most readily if the carboxyl terminal residue has an aromatic or a bulky aliphatic side chain. The active site of carboxypeptidase A comprises GLU 270, Zn and TYR 248 . Other residues which interact with the substrate are ASN 144, ARG 145, ARG 127, SER 197 and ARG 71 (Figs 1 and 3). The proposed mechanism of catalysis[24] (Fig. 4) is categorized as a "general base catalysis" which is triggered by the activation of a water molecule by GLU 270. The resulting OH^- then attacks the carbonyl of the susceptible peptide bond resulting in the formation of the tetrahedral intermediate which finally collapses after the required proton donation by GLU 270, thus leading to hydrolysis of the substrate.

Carboxypeptidase A has been the focus of intense kinetic, chemical and structural studies for the mechanisms of catalysis and inhibition. Importantly, carboxypeptidase A and the related zinc protease thermolysin have served as paradigms for structure assisted rational drug design. Intensive structural studies have yielded potent inhibitors of angiotensin-converting enzyme, a zinc protease related by the paradigms by convergent evolution, which are useful in hypertension therapy[26-29]. Given the sustained pharmaceutical interest in the regulation of zinc hydrases, it is desirable to explore and compare different inhibitor designs that may lead to the development of novel drug candidates and for this, determination of the free energy of binding becomes a very important factor. The carry over design of various inhibitors can then be extended to other enzymes since all the zinc containing enzymes share a common function of zinc promoted hydration reaction. The free energy analysis of the enzyme-inhibitor complexes which are reported in this paper deal with five inhibitors out of which the first three serve as models for non-covalent transition state analogues. Structural studies have shown that in all these inhibitors (Figs 2(a-e), 3, Table 1), the

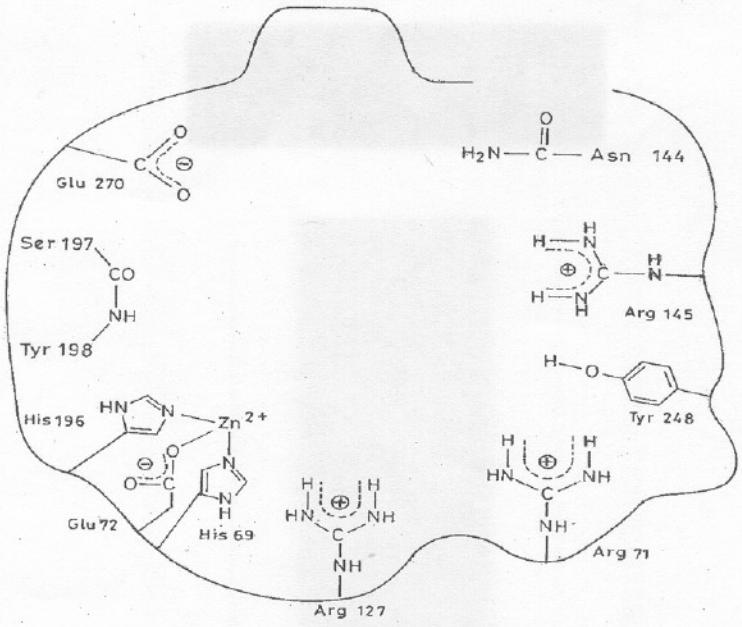

Fig. 1- A schematic representation of the residues in the active site of carboxypeptidase A and other residues reported to interact with the substrate/inhibitor.

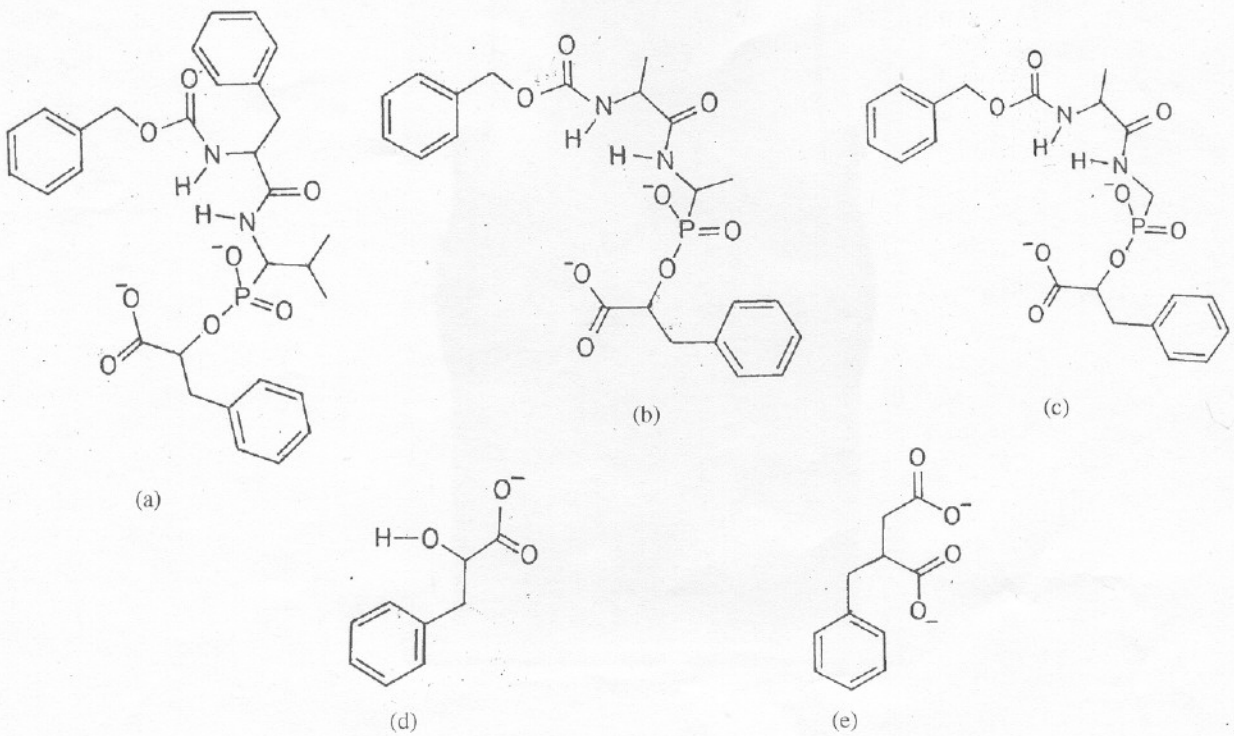

Fig. 2 - Molecular structural formulae of the inhibitors of carboxypeptidase A investigated: (a) ZFV^{p}(O)F; (b) ZAA^{p}(O)F; (c) ZAG^{p}(O)F; (d) LOF; (e) BZS.

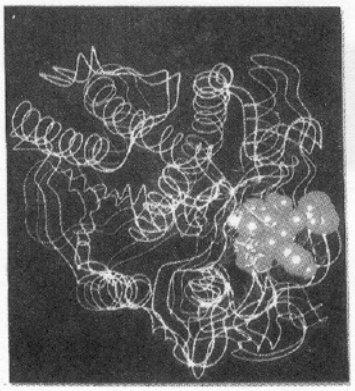

Fig. 3 - An image of the complex of ZFV^{p}(O)F with carboxypeptidase A. Residues that interact with the inhibitor are shown in red and the inhibitor is shown in blue. The metal ion is shown in pink.

terminal carboxylate group interacts with ASN 144, ARG 145 and TYR 248 of the enzyme[16-18]. The phosphonate group of the inhibitors interacts with GLU 270, ARG 127 and Zn in such a way that one oxygen atom is close to zinc while the other oxygen makes a longer bond to zinc. In the case of L-benzylsuccinate[19] (a by-product analogue), the second carboxylate group interacts with zinc in a bidentate fashion. The energetic significance of such contacts advanced as responsible for the enzyme-inhibitor recognition, their contribution to the overall binding process in the presence of solvent and salt and delineation of the forces favourable to binding are thus of general interest in molecular recognition research.

II. Theory and Methodology

This study is carried out in the theoretical framework of the free energy component analysis[30] in which additivity is assumed[31] and the net free energy change is treated as a sum of a near comprehensive set of individual contributions for which best estimates are obtained. The components are enumerated in a way that strategically isolates various contributions to the standard free energy of binding accessible to theoretical calculations via empirical energy functions and simplified models of solvation. With the assumption of additivity and an arbitrary albeit rational selection of terms, component analysis is not theoretically rigorous and one can expect at best only a semi quantitative account[32,33], expectations must be framed accordingly[34]. However, for complex purposes such as protein-inhibitor binding no viable alternative currently exists and simple identification of the important terms, estimates of their relative magnitudes and determination of whether they make favourable or unfavourable contributions to the free energy of complexation provides potentially useful new knowledge.

The thermodynamic cycle for protein-inhibitor binding in solution used in this study is presented in Fig. 5. Here the net binding process is decomposed into six steps. Step I describes the process of converting the uncomplexed protein denoted "P", to the form "P*" in which the protein has adapted its structure to that of the inhibitor bound form. The free energy of this step is

$$\Delta G^{0}_{I} = \Delta G_{1}^{dpt,P} \qquad ...(1)$$

Step II is the corresponding adaptation required of the inhibitor, converting the uncomplexed form "I" to the complexed form "I*" in solution. The free energy is thus

$$\Delta G^{0}_{II} = \Delta G_{2}^{adpt,I} \qquad ...(2)$$

The next two steps (III and IV) involve desolvation of P* and I* from aqueous medium to vacuum. The free energy of each of these steps is written as a sum of four components.

$$\Delta G^{0}_{III} = \Delta G_{3}^{el,P} + \Delta G_{4}^{vdw,P} + \Delta G_{5}^{cav,P} + \Delta G_{6}^{DH,P} \qquad ...(3)$$

$$\Delta G^{0}_{IV} = \Delta G_{7}^{el,I} + \Delta G_{8}^{vdw,I} + \Delta G_{9}^{cav,I} + \Delta G_{10}^{DH,I} \qquad ...(4)$$

with contributions from electrostatic effects of desolvating the macromolecule, the van der Waals interactions with the solvent, elimination of the solvent cavity in which the molecule is accomodated and the change in added salt effects. The transfer from aqueous medium to vacuum in steps III and IV involves the loss of favourable electrostatics and van der Waals interactions with the solvent and a gain from the cavity term, the latter being of course the reverse of the free energy of cavity formation. The free energy of interaction with added salt is also lost on desolvation.

In step V, the protein and the inhibitor associate as a non-covalently bound complex. The thermodynamics of this step can be described as

$$\Delta G^{0}_{V} = \Delta H_{11}^{el,C} + \Delta H_{12}^{vdw,C} - T\Delta S_{13}^{tr+rot} - T\Delta S_{14}^{vib+conf} \qquad ...(5)$$

Complexation involves introducing the electrostatic and van der Waals interactions between the protein and the inhibitor *in vacuo*. A change in external entropy due to loss of translational and rotational degrees of freedom enters this step, which always disfavors complexation. The lost external modes are converted into low

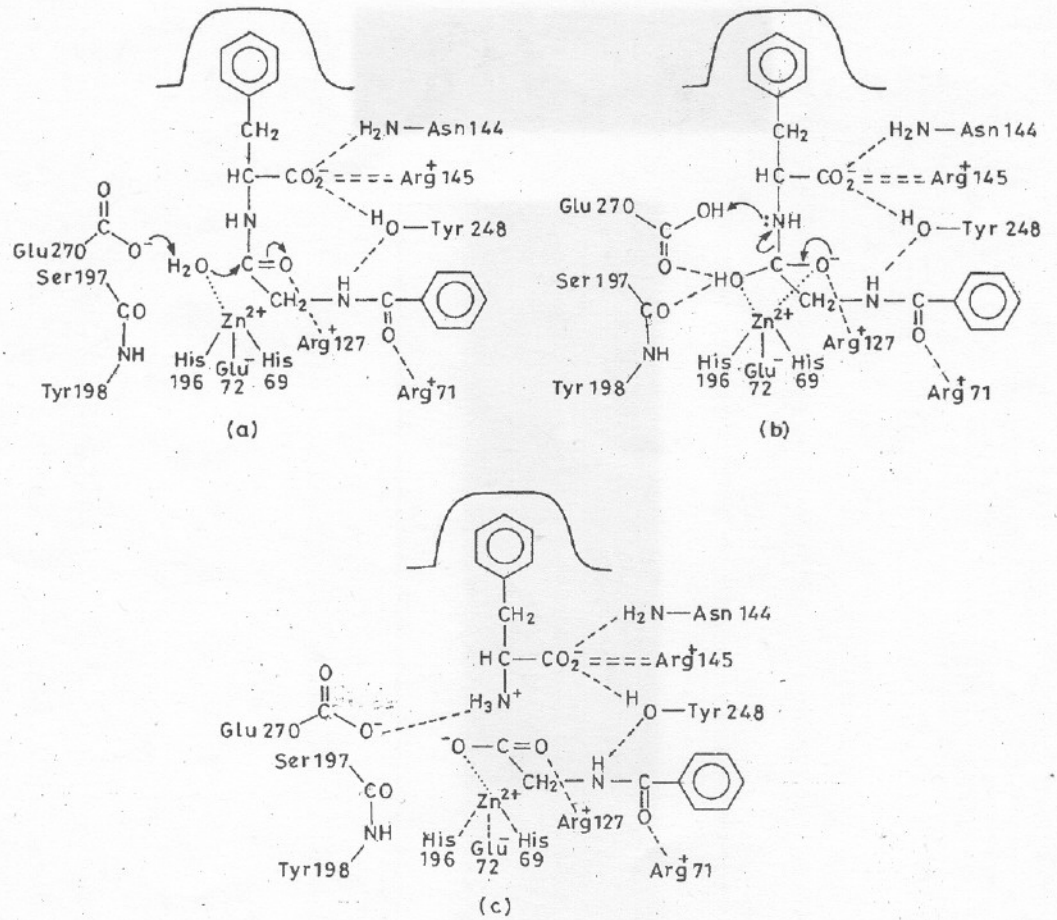

Fig. 4 - Proposed reaction pathway of carboxypeptidase A catalyzed peptide hydrolysis (adapted from Ref. 25).

frequency internal vibrational and configurational degrees of freedom in the complex and are reflected along with motional changes occuring as a consequence of the burial of amino acid side chains on complexation in the corresponding change in vibrational and configurational entropy[35].

In step VI, the complex is transferred from vacuum back to aqueous solution and the free energy change is due to solvation of the complex.

$$\Delta G^0_{VI} = \Delta G_{15}^{el,C} + \Delta G_{16}^{vdw,C} + \Delta G_{17}^{cav,C} + \Delta G_{18}^{DH,C} \quad ...(6)$$

Here again an electrostatic component, a van der Waals component, a cavity formation component and added salt effects are involved. While the cavitation term is unfavourable to complexation, all the other terms are favourable to complexation..

In summary, the binding process in solution as considered here consists of six well-defined thermodynamic steps each of which can be decomposed into physically meaningful thermodynamic components. The total number of individual contributions to the free energy of binding in this model is 18. Following Holtzer[36] and Gilson *et al*[35] no additional entropy of mixing terms are included explicitly.

The theoretical estimates of values for the various contributions proceed as follows. We write the standard free energy change of a given macromolecular structure (chemical potential) in solution G^0 as

$$G^0 = G^0_{int} + g^0_{solv} \quad ...(7)$$

where G^0 is the free energy intrinsic to the molecule or complex and g^0_{solv} is the standard free energy of solvation; the upper and lower case notation for the intrinsic and solvation components is introduced to clearly distinguish these terms. Values of G^0 will be used to obtain free energy differences ΔG^0 between initial and final thermodynamic states defined for complexation in figure 5. The underlying energetics intrinsic to macromolecules and complexes thereof is written in the conventional form of an empirical energy function

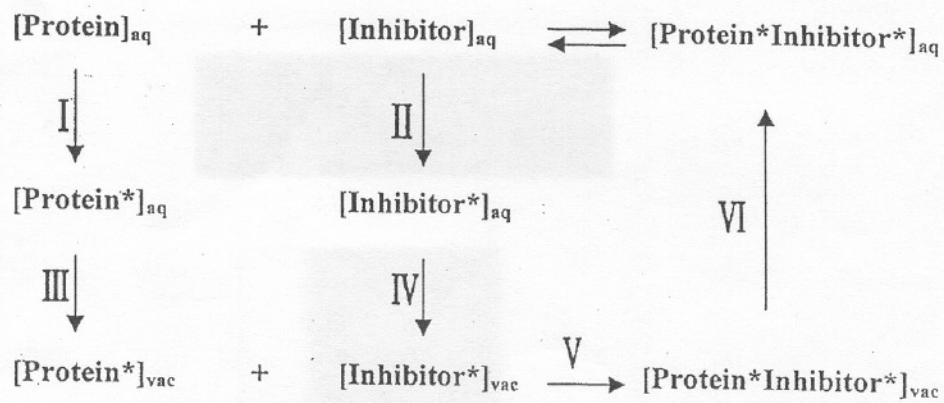

Fig. 5 - The thermodynamic cycle used for a component-wise estimation and analysis of the binding free energies of enzyme-inhibitor complexes.

$$E_{int} = E_{bonds} + E_{angles} + E_{dihedrals} + E_{nb} \quad ...(8)$$

where E_{bonds}, E_{angles}, $E_{dihedrals}$ describe bond stretching, angle bending and dihedral displacements. The nonbonded interaction term E_{nb} is written as a sum of electrostatics (el) and van der Waals (vdw) terms.

$$E_{nb} = E_{el} + E_{vdw} \quad ...(9)$$

Each of these terms is pairwise additive over atoms explicitly considered in this model, viz.

$$E_{el} = \sum_{i<j} q_i q_j / r_i \quad ...(10)$$

where q_i and q_j are any two net atomic charges separated by a distance r_{ij} and

$$E_{vdw} = \sum_{i<j} 4\varepsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^{6}] \quad ...(11)$$

where σ and ε are the Lennard-Jones collision diameter and binding energy from dispersion forces, respectively for each atom pair.

For a thermodynamic step with well-defined initial and final states, the intramolecular energy change is

$$\Delta E_{int} = \Delta E_{bonds} + \Delta E_{angles} + \Delta E_{dihedrals} + \Delta E_{el} + \Delta E_{vdw} \quad ...(12)$$

In the case where a single time averaged crystal structure is used to represent a Boltzmann ensemble of states, we shall assume

$$\Delta H^{o}_{int} \approx \Delta E_{int} \quad ...(13)$$

where ΔH^{o} is the standard enthalpy change for the step.

The external rotational and translational entropies for P*, I* and the complex are required for the free energy of complexation *in vacuo* (Step 5 of Fig. 5). These quantities are calculated from ideal gas partition function Q using classical statistical mechanics[37] viz.

$$S_0 = k \ln Q - (\partial \ln Q/\partial B)_V \quad ...(14)$$

The translational partition function is computed as

$$Q_{trans} = V/h(\beta/2\pi m)^{3/2} \quad ...(15)$$

where V is the volume, $\beta = (kT)^{-1}$ with k the Boltzmann constant, T the temperature and m is the mass. For the rotational partition function

$$Q_{rot} = (\pi^{1/2}/\sigma)\,(1/hc\beta)^{3/2}\,(1/ABC)^{1/2} \quad ...(16)$$

where A, B, C are rotational constants calculated from molecular geometry by standard methods, c is the velocity of light and σ is the symmetry number. The translational and rotational entropies are introduced into the thermocycle at the step of complexation of the protein and the inhibitor in vacuum. Vibrational and configurational contributions to entropy are indistinguishable in this problem and are considered together. Included in this contribution is the increase in vibrational/configurational entropy as a consequence of the new low frequency motions that are interconverted from external degrees of freedom on complex formation and the loss of conformational entropy when an amino acid side chain of the protein is restricted by contacts with the inhibitor on complexation[38].

We write the solvation energy of a structure as

$$g^{0}_{solv} = g^{0}_{el} + g^{0}_{nel} \quad ...(17)$$

Here, the electrostatic contribution to the solvation energy, g^{0}_{GB} is estimated via the generalized Born (GB) model, whose defining equation is

$$g^0_{el} = -166\,(1 - 1/\varepsilon) \sum_{i=1}^{n} \sum_{j=1}^{n} q_i q_j / f_{m2GB} \quad \ldots(18)$$

where f_{m2GB} is an effective atomic size/distance parameter derived from the Born radii α_i and pairwise distance r_{ij}. With suitable values for α_i, the solvation energy of a given molecule in a specified conformation can be computed. The GB solvation energy can be partitioned into contributions from polarization and solvent screening if necessary[13,14].

Added salt effects were incorporated into GB theory via the Debye-Huckel theory, resulting in the expression

$$g^0_{el} = -\,(166 / \varepsilon) \sum_{i=1}^{n} \sum_{j=1}^{n} q_i q_j / f_{m2GBDH} \quad \ldots(19)$$

$$f_{m2GBDH} = (\kappa^{-1} + r_{ij})\,(f_{m2GBDH}/r_{ij}) \quad \text{for } i \neq j \quad \ldots(19a)$$

$$f_{m2GBDH} = (\kappa^{-1} + r_{vdw})\,(\alpha_i/r_{ij}) \text{ for } i = j \quad \ldots(19b)$$

where f_{m2GBDH} is the effective Born radius parameter, including the Debye-Huckel modification. With this addition, the solvent model becomes a combination of GB and Debye-Huckel theory.

The non-electrostatic (nel) contributions to the standard free energy are due to van der Waals interactions between the solute and solvent and the work required to alter the cavitation in water in going from initial to final conditions. The total non-electrostatic free energy is written as a function of the solvent accessible (SA) surface area

$$g^0_{nel} = \gamma_{nel} \Delta A \quad \ldots(20)$$

with an empirical coefficient γ_{nel} defining the proportionality. Still and coworkers found that a value of $\gamma_{nel} = 7.2$ cal/mol/Å^2 gave reasonable results for a large number of cases[10]. The quantity γ_{nel} can be considered as the sum of van der Waals and cavitation terms

$$\gamma_{nel} = \gamma_{vdw} + \gamma_{cav} \quad \ldots(21)$$

with a value of 7.2 cal/mol/Å^2 considered as a resultant of +47cal/mol/Å^2 from the cavity term[39] and –39.8 cal/mol/Å^2 from van der Waals interactions of the solute and the solvent. An independent check on this partitioning comes from noting that the van der Waals contribution is close to the value of 38.75 cal/mol/Å^2 derived from experimental enthalpies of vaporization of hydrocarbons.[40] The surface area A referred to above is however that of all the atoms. The agreement between the GBSA results and the experimental solvation energies for a wide range of molecules is well documented[10-15] and comparable to that obtained with both free energy simulations and finite difference Poisson-Boltzmann calculations while requiring much less computational effort. In a recent study, we have derived α_i parameters consistent with the AMBER parm94 forcefield[8] and experimental solvation energies of small molecules[14]. Further details on the theory and methodology for obtaining thermodynamic indices of macromolecular complexation are provided in ref. 30.

The methodology outlined above leads a way to quantify 16 of the 18 terms in the thermodynamic cycle considered. Quantification of the two missing terms, viz. contributions arising due to structural adaptation of the enzyme and the inhibitor upon complexation await detailed molecular dynamics simulations followed by free energy analyses of the trajectories. Availability of the crystal structures of the (uncomplexed) reactants provides an alternative source.

The atomic coordinates of carboxypeptidase A (isolated from the pancreas of *Bos Taurus)* in complex with inhibitor molecules were obtained from the Brookhaven Protein Data Bank (Table 1). Our calculations are based on all atom models for the enzyme and the inhibitors in which hydrogens are added explicitly to the crystal structure. The protonation state of the ionizable groups was set at that corresponding to *p*H 7 and assumed to be constant. AMBER partial charge assignment for the enzyme atoms proceeds in a straightforward manner. Charges to the inhibitor atoms however, in this first report, are assigned as closely as possible to the existing AMBER atom types. Work on building a library of force field compatible charges for non-standard atom types encountered commonly in inhibitor molecules is in progress. Next, energy minimization of the enzyme-inhibitor complex was performed using the Sander module of the AMBER 4.1 molecular modelling package with the "parm94" force field[8] so as to relieve any unfavourable clashes in the crystal structure. Here 500 steps of minimization restraining heavy atoms (50 steps of steepest descent, SD, followed by 450 steps of conjugate gradient, CG), followed by a further 500 steps (50 SD + 450 CG) of free minimization, were carried out using a sigmoidal distance dependent dielectric function.[41,42] The electrostatic contribution to solvation was calculated via the generalized Born model using the effective radii parameters derived by Jayaram *et al*[14] based on AMBER

Table 2 - Calculated values of the various contributions to the standard free energy of binding (in kcal/mol) for the carboxypeptidase A-Inhibitor complexes

Term	Component	ZFV^{P}(O)F	ZAA^{P}(O)F	ZAG^{P}(O)F	LOF	BZS
	Step I: Structural adaptation of enzyme					
$\Delta G_1^{adpt,P}$	Free energy change for the process P → P*	-	-	-	-	-
	Step II: Structural adaptation of inhibitor					
$\Delta G_2^{adpt,I}$	Free energy change for the process I → I*	-	-	-	-	-
	Step III: Desolvation of enzyme					
$\Delta G_3^{el,P}$	Electrostatic component of P* desolvation	2404.8	2492.7	2318.0	2664.0	2561.9
$\Delta G_4^{vdW,P}$	VdW component of P* desolvation	465.9	473.6	464.9	464.4	472.4
$\Delta G_5^{cav,P}$	Cavity component of P* desolvation	-550.2	-559.2	-549.1	-548.5	-557.9
$\Delta G_6^{DH,P}$	Loss of added salt interactions	22.2	22.4	22.0	23.1	22.1
	Step IV: Desolvation of inhibitor					
$\Delta G_7^{el,I}$	Electrostatic component of I* desolvation	260.1	281.4	259.1	93.6	244.8
$\Delta G_8^{vdW,I}$	VdW component of I* desolvation	34.4	29.6	29.7	14.3	16.3
$\Delta G_9^{cav,I}$	Cavity component of I* desolvation	-40.6	-34.9	-35.1	-16.9	-19.3
$\Delta G_{10}^{DH,I}$	Loss of added salt interactions	1.1	1.1	1.1	0.3	1.0
	Step V: Complex formation in vacuo					
$\Delta H_{11}^{el,C}$	Electrostatic interactions between of P* & I*	-627.1	-682.6	-598.2	-207.3	-611.6
$\Delta H_{12}^{vdW,C}$	VdW interactions between P* & I*	-62.0	-50.2	-50.7	-26.2	-28.8
$-T\Delta S_{13}^{tr+rot}$	Rotational, translational entropy change	25.7	25.4	25.2	22.3	22.9
$-T\Delta S_{14}^{vib+conf}$	Vibrational, configurational entropy change	9.2	10.5	9.8	8.5	7.9
	Step VI: Solvation of complex					
$\Delta G_{15}^{el,C}$	Electrostatic component of complex solvation	-2031.6	-2091.4	-1984.0	-2563.6	-2218.5
$\Delta G_{16}^{vdW,C}$	VdW component of complex solvation	-459.6	-466.1	-457.3	-461.5	-468.3
$\Delta G_{17}^{cav,C}$	Cavity component of complex solvation	542.8	550.4	540.0	545.1	552.4
ΔG_{18}^{DH}	Added salt interactions with complex	-20.6	-20.9	-20.7	-23.1	-20.8
ΔG^{0}_{net}	Net free energy of binding	-25.5	-18.2	-25.3	-11.5	-23.5

charges and sizes. The nonelectrostatic contribution to solvation which involves molecular surface area calculations were performed using the ACCESS program based on the algorithm of Lee and Richards[43]. The added salt concentration employed in the Debye-Huckel term was 0.18*M*. The various free energy components (16 in all) for each complex are then computed and discussed underneath.

III. Results

A complete list of our calculated contributions to the standard free energy of binding for the five Carboxypeptidase A-inhibitor complexes studied is provided in Table 2. Table 3 gives the comparison of the free energies obtained by our studies against the values reported on the basis of the inhibition constants for the various systems. A quantitative comparison of the computed standard binding free energies with experiment necessitates a rigorous consideration of the compatibility of the standard states. We defer this issue to a subsequent study.

Focusing here on the analysis of the free energies, our conventions are defined in such a way that negative values are favourable and positive values unfavourable to binding. Of the 18 components, 7 are found favourable

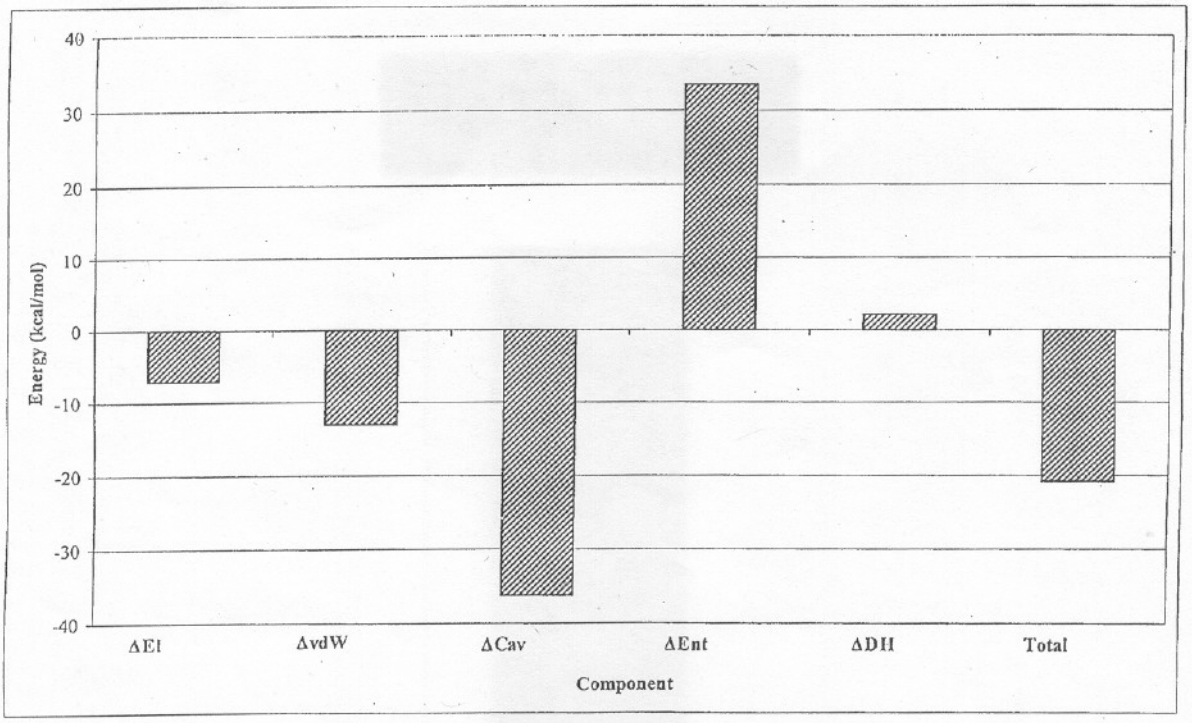

Fig. 6 - A histogram of the calculated contributions to the binding free energy of carboxypeptidase A- inhibitor complexes.

to binding. As is typical of a component analysis based on terms representative of fundamental aspects of the structural chemistry of a process, the net free energy is seen to be a result of large competing terms representing the occurrence of various compensation effects during binding. The compensatory effects appear mainly between (i) the internal and solvation electrostatics, (ii) the direct van der Waals interactions between protein and inhibitor and loss in van der Waals interactions with the solvent and (iii) the cavitation effects and the entropic losses of the enzyme and inhibitor molecules upon complexation.

A more important aspect in a conventional free energy analysis study is the net contribution of electrostatics, shape complementarity, hydrophobic effects, structural adaptation etc. to the binding. The answers to these types of questions can be obtained from a combination of the values associated with the primary terms in Table 2. Specifically, the contribution of structural adaptation to free energy can be written as

$$\Delta G^{adpt} = \Delta G_1^{adpt,P} + \Delta G_2^{adpt,I} \qquad ...(22)$$

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

$$\Delta G^{el} = \Delta G_3^{el,P} + \Delta G_7^{el,I} + \Delta G_{15}^{el,C} + \Delta H_{11}^{el,C} \qquad ...(23)$$

The van der Waals interactions, effectively the net energetics of shape complementarity, are expressed as

$$\Delta G^{vdW} = \Delta G_4^{vdW,P} + \Delta G_8^{vdW,I} + \Delta H_{12}^{vdW,C} + \Delta G_{16}^{vdW,C} \qquad ...(24)$$

The total contribution of cavitation effects to the binding is

$$\Delta G^{cav} = \Delta G_5^{cav,P} + \Delta G_9^{cav,I} + \Delta G_{17}^{cav,C} \qquad ...(25)$$

The entropy change on complexation is described by the combination

$$\Delta G_{trvc} = -T\Delta S_{13}^{tr+rot} - T\Delta S_{14}^{vib+conf} \qquad ...(26)$$

Small ion effects on free energy, due to added salt in the model, can be written as

$$\Delta G_{DH} = \Delta G_6^{DH,P} + \Delta G_{10}^{DH,I} + \Delta G_{18}^{DH,C} \qquad ...(27)$$

The sum of all these terms equals the net standard free energy of binding

$$\Delta G^{o}_{net} = \Delta G_{el} + \Delta G_{vdw} + \Delta G_{cav} + \Delta G_{DH} + \Delta G_{trvc} \qquad ...(28)$$

An analysis of the results of Table 2 based on the net contributions from electrostatics, van der Waals, cavitation, entropy and added salt effects are presented schematically in figure 6 for all the five complexes. Here the differential effects of direct van der Waals interactions between the protein and of cavity formation upon com-

Table 3 - The calculated and experimental free energies of binding (in kcal/mol) of some carboxypeptidase A-inhibitor complexes

	Inhibitor	ΔG^0_{calc}	ΔG^0_{expt}	K_i(pM)
1	ZFV^P(O)F	-25.5	-18.96	0.011
2	ZAA^P(O)F	-18.2	-15.65	3
3	ZAG^P(O)F	-25.3	-12.34	710
4	LOF	-11.5	—	—
5	BZS	-23.5	-8.62	450000

plexation are seen to be favourable to binding. In the protein-inhibitor complex literature, the role of van der Waals forces is somewhat underplayed except for references to lock and key fit or induced fit. Thus, our results introduce a potentially significant new perspective on the binding phenomenon in this class of systems. The change in the size and shape of the solvent cavity on complexation gives rise to water reorganization, a component of which, originating from nonpolar sources, is the hydrophobic effect.

IV. Discussion

Our analysis of the binding free energy for the Carboxypeptidase A-inhibitor complexes indicates that the nonelectrostatic contributions, i.e. van der Waals interactions and differential cavitation effects, are favourable to complexation (Table 4). Electrostatic contribution is favourable for some systems while unfavourable for others. Note that all the inhibitors studied (Fig. 2) carry a net negative charge with four of them carrying a charge of -2. The active site (Fig. 1) is characterized by a preponderance of positive charges. Structurally thus, the electrostatic complementarity is obvious. Energetically however, the bulkier inhibitors (ZFV^P(O)F ZAA^P(O)F & ZAG^P(O)F) appear to gain via van der Waals and cavitation contributions at the expense of electrostatics whereas the smaller inhibitors (LOF, BZS) appear to gain electrostatically. Theories of induced fit and shape / electrostatic complementarities are thus to be understood in their proper energetic context considering the possible compensatory effects. In relation to drug design, we believe that a balance between these competing effects and such others to obtain optimal binding can be best achieved via computer modelling on the above lines with a set of candidate molecules.

The result that the net electrostatics is unfavourable to complexation for some complexes suggests that the electrostatic complementarity of the components in the complex is not dramatic. The fact that electrostatics is net destabilizing to complexation does not imply that electrostatics is unimportant. The net free energy of binding is a fine balance of competing terms and would show a corresponding sensitivity to the magnitude of the electrostatic contribution even if it were destabilizing. Furthermore, in considering the relative binding process of a series of molecular or macromolecular ligands, differential effects of electrostatics may still be critical in the result. Salt effects are unfavourable to complexation in all cases due to the screening of electrostatic interactions between reactants with complemetary charges, as anticipated. The effect of the net entropy change is found to be unfavourable for all the five complexes, which is expected since upon formation of the complex there occurs a motional restriction on the amino acid residues of the protein and the inhibitor.

In contrast to protein-DNA systems[44], a theoretical analysis of the protein-inhibitor complexes is facilitated by the absence of counterions, which in the former class of systems compounds the treatment. Also, the net electrostatics is unfavourable almost always in the protein-DNA case when the counterion effects are considered. The net electrostatics in protein-inhibitor complexes however can be both favourable or unfavourable depending on the inhibitor functional group design.

In this study, the objective was to evolve a methodology and carry out an analysis of the energetics of protein-inhibitor complexation based on crystal structure data in a computationally reliable and expeditious manner. The results obtained here demonstrate the promise of the methodology. They also raise a question as to whether an analysis of this type can be successfully extended to other systems. We are extending this protocol to other protein-ligand systems and only after a judicial analysis of all the results can further conclusions be drawn. In concluding this section, we would like to highlight the issue of errors and uncertainties in the process of developing *ab initio* estimates of binding free energies of complex macromolecular systems based on the constituent phenomenological components. The quantifiable errors in certain estimates plus the qualitative approximations in others lead to a net result that has an uncertainty of the order of the calculated net binding energy, which indicates that neither the magnitude nor the sign of this quantity is secure. Another source of er-

Table 4 -A conventional combination of the computed primary terms contributing to the net free energies of binding (in kcal/mol) of the carboxypeptidase A-inhibitor complexes

	ZFV^{P}(O)F	ZAA^{P}(O)F	ZAG^{P}(O)F	LOF	BZS	Averages
ΔG^0_{el}	6.2	0.1	-5.1	-13.3	-23.4	-7.1
ΔG^0_{vdw}	-21.3	-13.1	-13.4	-9.0	-8.4	-13.0
ΔG^0_{cav}	-48.0	-43.7	-44.2	-20.3	-24.8	-36.2
$\Delta G^0_{entropy}$	34.9	35.9	35.0	30.8	30.8	33.5
ΔG^0_{DH}	2.7	2.6	2.4	0.3	2.3	2.1
ΔG^0_{tot}	-25.5	-18.2	-25.3	-11.5	-23.5	-20.8

ror in this method is the neglect of the deformation energy, which at this stage could not be obtained due to the non-availability of the native enzyme structure. An alternative approach to determine the deformation energy would be to carry out molecular dynamics studies on the uncomplexed and complexed protein / inhibitor and then noting the difference in the energies of the two forms. We also envision improvements to entropy estimates via full scale MD simulations with explicit waters. Other areas currently under our active consideration are the ionization state of amino acid residues on the protein and force field compatible derivation of inhibitor charges. Thus, the results described in this article must be analyzed in the context of the expected uncertainties, but there are numerous new features that can be learnt from these studies such as ideas about relative magnitudes and signs of various contributions, and considerations of both initial and final states in estimating thermodynamic components and above all the rapidity with which such binding free energy calculations can be performed incorporating environmental effects.

V. Conclusions

A computational methodology to estimate and carry out a component analysis of the binding free energies of enzyme-inhibitor complexes is presented, taking into account solvent and salt effects together with a consideration of initial and final states, with five carboxypeptidase A-inhibitor complexes serving as illustrative cases. The net van der Waals and cavitation contributions favour binding in all the systems considered. The sign and magnitude of the electrostatic contributions (including hydrogen bonds) in the overall free energy of formation are noticed to vary as a result of competition with van der Waals interactions (requirements of optimal steric fit). It is hoped that with some more fine-tuning, this methodology would offer a computationally viable tool in structure based drug design research.

Acknowledgement

Funding from the C S I R, India is gratefully acknowledged.

References

1. Lehninger A L, Nelson D L & Cox M M, *Principles of biochemistry* (CBS Publishers and Distributors, New Delhi), 1993, Chap 8.
2. Hehre W J, Radom L, Schleyer P V R &, Pople J, *Ab Initio molecular orbital theory* (Wiley, New York), 1986.
3. Beveridge D L & DiCapua F M, *Ann Rev Biophys biophys Chem*, 18 (1989) 431.
4. Jorgensen W L, *Acc chem. Res*, 22 (1989) 184.
5. McCammon J A, *Curr Opinion Struct Biol*, 1 (1991) 196.
6. Kollman P A, *Chem Rev*, 93 (1993) 2395.
7. van Gunsteren W F & Berendsen H J C, *Angew Chem Int Ed Engl*, 29 (1990) 992.
8. Cornell W D, Cieplak P, Bayly C I, Gould I R, Kenneth M M, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W & Kollman P A, *J Am chem Soc*, 117 (1995) 5179.
9. Pearlman D A, Case D A, Caldwell J W, Cheatham T E, Debolt S, Ferguson D, Seibel G & Kollman P A, *Comput Phys Commun*, 91 (1995) 1.
10. Still W C, Tempczyk A, Hawley R C & Hendricksen T J, *J Am chem Soc*, 112 (1990) 6127.
11. Hawkins G D, Cramer C J & Truhlar D G, *J phys Chem*, 100 (1996) 19824.

12. Hawkins G D, Cramer C J & Truhlar D G, *Chem Phys Lett*, 246 (1995) 246.
13. Jayaram B, Liu Y & Beveridge D L, *J chem Phys*, 109 (1998) 1465.
14. Jayaram B, Sprous D & Beveridge D L, *J phys Chem B*, 102 (1998) 9571.
15. Reddy M R, Erion M D, Agarwal A, Vishwanadhan V N, McDonald Q & Still W C, *J comput Chem*, 19 (1998) 769.
16. Hidong K & Lipscomb W N, *Biochemistry*, 30 (1991) 8171.
17. Hanson J E, Kaplan A P & Bartlett P A, *Biochemistry*, 28 (1989) 6294.
18. Kaplan P & Bartlett P A, *Biochemistry*, 30 (1991) 8165.
19. Mangani S, Carloni P & Orioli P, *J mol Biol*, 223 (1992) 573.
20. Rees D C & Lipscomb W N, *Proc Nat Acad Sci USA*, 78 (1981) 5455.
21. Cappalonga A M, Alexander R S & Christianson D W, *J biol Chem*, 267 (1992) 19192.
22. Christianson D W & Lipscomb W N, *Proc Nat Acad Sci USA*, 83 (1986) 7568.
23. Rees D C & Lipscomb W N, *Proc Nat Acad Sci USA*, 80 (1983) 7151.
24. Hidong K & Lipscomb W N, *Biochemistry*, 29 (1990) 5546.
25. Christianson D W & Lipscomb W N, *Acc chem Res*, 22 (1989) 62.
26. Cushman D W, Cheung H S, Sabo E F & Ondetti M A, *Biochemistry*, 16 (1977) 5484.
27. Maycock A L, Desousa D M, Payne L G, ten Broeke J, Wu M T & Patchett A A, *Biochem Biophys Res Commun*, 102 (1981) 963.
28. Ondetti M A, Rubin B & Cushman D W, *Science*, 196 (1977) 441.
29. Patchett A A, Harris E, Tristham E W, Wyvratt M J, Wu M T, Taub D, Peterson E R, Ikeler T J, ten Broeke J, Payne L G, Ondeyka D L, Thorsett E D, Greenlee W J, Lohr N S, Hoffsommer R D, Joshua F M & Hirschmann R, *Nature*, 288 (1980) 280.
30. Jayaram B, McConnell K J, Dixit S B & Beveridge D L, *J comput Physics*, 151 (1999) 333.
31. Dill K A, *J biol Chem*, 272 (1997) 70.
32. Janin J, *Proteins*, 21 (1995) 30.
33. Janin J, *Prog Biophys mol Biol*, 64 (1995) 145.
34. Novotny J, Bruccoleri R E, Davis M & Sharp K A, *J mol Biol*, 268 (1997) 401.
35. Gilson M K, Given J A, Bush B L & McCammon J A, *Biophys J*, 72 (1997) 1047.
36. Holtzer, *Biopolymers*, 35 (1994) 595.
37. Atkins P W, *Physical chemistry* (Freeman, New York), 1990.
38. Finkelstein A V & Janin J, *Protein Engg*, 3 (1989) 1.
39. Sharp K A, Nicholls A, Fine R F & Honig B, *Science*, 252 (1991) 106.
40. Ohtaki H & Fukushima N, *J soln Chem*, 21 (1992) 23.
41. Arora N & Jayaram B, *J phys Chem*, 102 (1998) 6139.
42. Arora N &Jayaram B, *J comput Chem*, 18 (1997) 1245.
43. Lee K & Richards F M, *J mol Biol*, 55 (1971) 379.
44. Jayaram B, McConnell K J, Dixit S B & Beveridge D L, communicated, 1999.