We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design

Agnieszka K. Bronowska
Heidelberg Institute for Theoretical Studies Heidelberg,
Germany

1. Introduction

Biologically relevant macromolecules, such as proteins, do not operate as static, isolated entities. On the contrary, they are involved in numerous interactions with other species, such as proteins, nucleic acid, membranes, small molecule ligands, and also, critically, solvent molecules. These interactions often display a remarkable degree of specificity and high affinity. Fundamentally, the biological processes rely on molecular organisation and recognition events. Binding between two interacting partners has both enthalpic (ΔH) and entropic ($-T\Delta S$) components, which means the recognition event is associated with changes of both the structure and dynamics of each counterpart. Like any other spontaneous process, binding occurs only when it is associated with a negative Gibbs' free energy of binding (ΔG), which may have differing thermodynamic signatures, varying from enthalpy- to entropy-driven. Thus, the understanding of the forces driving the recognition and interaction require a detailed description of the binding thermodynamics, and a correlation of the thermodynamic parameters with the structures of interacting partners. Such an understanding of the nature of the recognition phenomena is of a great importance for medicinal chemistry and material research, since it enables truly rational structure-based molecular design.

This chapter is organised in the following way. The first part of it introduces general principles which govern macromolecular associations under equilibrium conditions: the free energy of binding and its enthalpic and entropic components, the contributions from both interacting partners, interaction energy of the association, and specific types of interactions – such as hydrogen bonding or van der Waals interactions, ligand and protein flexibility, and ultimately solvent effects (e.g. solute-solvent interactions, solvent reorganisation). The second part is dedicated to methods applied to assess particular contributions, experimental as well as computational. Specifically, there will be a focus on isothermal titrational calorimetry (ITC), solution nuclear magnetic resonance (NMR), and a discussion of computational approaches to the estimation of enthalpic and entropic contributions to the binding free energy. I will discuss the applicability of these methods, the approximations behind them, and their limitations. In the third part of this chapter, I will provide the reader several examples of ligand-protein interactions and focus on the forces driving the associations, which can be very different from case to case. Finally, I will address several practical aspects of assessing the thermodynamic parameters in molecular design, the

bottlenecks of methods employed in such process, and the directions for future development.

The information content provided by thermodynamic parameters is vast. It plays a prominent role in the elucidation of the molecular mechanism of the binding phenomenon, and – through the link to structural data – enables the establishment of the structure-activity relationships, which may eventually lead to rational design. However, the deconvolution of the thermodynamic data and particular contributions is not a straightforward process; in particular, assessing the entropic contributions is often very challenging.

Two groups of computational methods, which are particularly useful in assessment of the thermodynamics of molecular recognition events, will be discussed. One of them are methods based on molecular dynamics (MD) simulations, provide detailed insights into the nature of ligand-protein interactions by representing the interacting species as a conformational ensemble that follows the laws of statistical thermodynamics. As such, these are very valuable tools in the assessment of the dynamics of such complexes on short (typically, picosecond to tens of nanosecond, occasionally microsecond) time scales. I will give an overview of free energy perturbation (FEP) methods, thermodynamic integration (TI), and enhanced sampling techniques. The second group of computational methods relies on very accurate determinations of energies of the macromolecular systems studied, employing calculations based on approximate solutions of the Schrödinger equation. The spectrum of these quantum chemical (QM) methods applied to study ligand-protein interactions is vast, containing high-level ab initio calculations: from Hartree-Fock, through perturbational calculations, to coupled-clusters methods; DFT and methods based on it (including "frozen" DFT and SCC-DFTTB tight binding approaches); to semi-empirical Hamiltonians (such as AM1, PM3, PM6, just to mention the most popular ones) (Piela, 2007, Stewart, 2009). Computational schemes based the hybrid quantum mechanical –molecular mechanical (QM/MM) regimes will also be introduced. Due to the strong dependence of the molecular dynamics simulations on the applied force field, and due to the dependence of both MD simulations and QM calculations on the correct structure of the complex, validation of results obtained by these methodologies against experimental data is crucial.

Isothermal titration calorimetry (ITC) is one of the techniques commonly used in such validations. This technique allows for the direct measurement of all components of the Gibbs' equation simultaneously, at a given temperature, thus obtaining information on all the components of free binding energy during a single experiment. Yet since these are *de facto* global parameters, the decomposition of the factors driving the association, and investigation of the origin of force that drives the binding is usually of limited value. Nonetheless, the ITC remains the primary tool for description of the thermodynamics of ligand-protein binding (Perozzo *et al.*, 2004). In this chapter, I will give a brief overview of ITC and its applicability in the description of recognition events and to molecular design.

Another experimental technique, which has proven very useful in the experimental validation of computational results, is NMR relaxation. These measurements are extremely valuable, as they specifically investigate protein dynamics on the same time scales as MD simulations. As such, the results obtained can be directly compared with simulation outputs. In addition, the Lipari-Szabo model-free formalism (Lipari and Szabo, 1982) is relatively free of assumptions regarding the physical model describing the molecular motions. The only requirement is the internal dynamics being uncorrelated with the global tumbling of the system under investigation. The results of the Lipari–Szabo analysis, in the form of generalised order parameters (S_{LS}^2), can be readily interpreted in terms of the

conformational entropy associated with the measured motions (Yang and Kay, 1996). It has been shown that for a wide range of motion models, the functional dependence of the conformational entropy on the order parameter is similar, suggesting that changes in order parameters can be related to changes in entropy in a model-independent manner. I will introduce the application of this model-free formalism to MD simulation, for the study of dynamical behaviour of ligand-protein complexes and the estimation of changes in the conformational entropy upon ligand-protein association. The MD simulations, performed on several proteins in complexes with their cognate ligands, indicate that the molecular ensembles provide a picture of the protein backbone dynamics that show a remarkably high degree of consistency with NMR relaxation data, regardless of the protein's size and structure (Schowalter and Brüschweiler, 2007).

In this chapter I will also address the enthalpy-entropy compensation phenomenon and the challenges it imposes on molecular design. The generality of this phenomenon have been a subject of debate for many years. Although this compensation is not a thermodynamic requirement as such (Ford, 2005, Sharp, 2001), it has been very frequently observed in protein-ligand interactions (Whitesides and Krishnamurthy, 2005). Briefly, stronger and more directed interactions are less entropically favourable, since the tight binding constricts molecular motions. The detailed mechanism of enthalpy-entropy compensation is, nonetheless, highly system-dependent, and this compensation does not obey a single functional form. An example of enthalpy-entropy compensation and its consequences to the design process will be provided.

A discussion of the thermodynamics of protein-ligand interactions would not be complete without commenting on dynamic allostery and cooperativity. The mechanism of allostery plays a prominent role in control of protein biological activity, and it is becoming accepted that protein conformational dynamics play an important role in allosteric function. Changes of protein flexibility upon ligand binding affect the entropic cost of binding at distant protein regions. Counter-intuitively, proteins can increase their conformational entropy upon ligand binding, thus reducing the entropic cost of the binding event (MacRaild *et al.*, 2007). I will discuss these phenomena, illustrating them through several examples of biologically-relevant protein-ligand interactions.

The overall aim of this chapter is to introduce the forces driving binding events, and to make the reader familiar with some general rules governing molecular recognition processes and equally to raise awareness of the limitations of these rules. Combining the structural information with equilibrium thermodynamic data does not yield an understanding of the binding energetics under non-equilibrium conditions, and global parameters, obtained during ITC experiments, do not enable us to assess the individual contributions to the binding free energy. Certain contributions, such as entropy, may behave in a strongly non-additive and highly correlated manner (Dill, 1997). This chapter will discuss the boundaries of rational molecular design guided by thermodynamic data.

2. Principles

2.1 Enthalpic and entropic components of free binding energy

A non-covalent association of two macromolecules is governed by general thermodynamics. Similarly to any other binding event (or – in a broader context – to any spontaneous process), it occurs only when it is coupled with a negative Gibbs' binding free energy (1), which is the sum of an enthalpic, and an entropic, terms:

$$\Delta G = \Delta H - T\Delta S \tag{1}$$

where ΔG is free binding energy, ΔH is enthalpy, ΔS entropy, and T is the temperature.

The enthalpic contribution to the free energy reflects the specificity and strength of the interactions between both partners. These include ionic, halogen, and hydrogen bonds, electrostatic (Coulomb) and van der Waals interactions, and polarisation of the interacting groups, among others. The simplest description of entropic contribution is that it is a measure of dynamics of the overall system. Changes in the binding entropy reflect loss of motion caused by changes in translational and rotational degrees of freedom of the interacting partners. On the other hand, changes in conformational entropy may be favourable and in some cases these may reduce the entropic cost of binding (MacRaild *et al.*, 2007). Solvation effects, such as solvent re-organisation, or the release of tightly bound water upon ligand binding can contribute significantly to the entropic term of the binding free energy.

The Gibbs equation can be also written as in equation (2):

$$\Delta G = -RT \ln K_d \tag{2}$$

where R is a gas constant, T is the temperature, and K_d is binding constant. This formulation emphasises the relationship between Gibbs energy and binding affinity. The ligand-protein association process can be represented in the form of a Born-Haber cycle. A typical cycle is showed in Figure 1. The 'intrinsic' free energy of binding between ligand L and protein P is represented by ΔG_i, whereas the experimentally observable free energy of binding is represented by ΔG_{obs}.

Fig. 1. An example of Born-Haber cycle for ligand-protein (LP) association. It relates the experimentally observed free energy of binding (ΔG_{obs}) with 'intrinsic' free energy of binding (ΔG_i) between ligand (L) and protein (P) and with solvation free energies of free interactors (ΔG_{sf}) and the resulting complex (ΔG_{sb}). X, Y, Z, and B refer to the number of water molecules involved in solvation of the unbound ligand (X), unbound protein (Y), ligand-protein complex (Z), and to the bulk solvent (B).

Two additional processes can be defined: the free energy of solvation of the free (unbound) interacting partners (ΔG_{sf}), and the free energy of solvation of the ligand-protein complex (ΔG_{sb}). Since the free energy is a state function, it is independent of the path leading from from one state of the system to another. Hence, the observable free energy of binding can be written as in equation (3):

$$\Delta G_{obs} = \Delta G_i + \Delta G_{sb} - \Delta G_{sf} \qquad (3)$$

The equation above shows how the observable free energy of binding can be decomposed into the 'intrinsic' term, and the solvation contributions from the ligand-protein complex and unbound interactors. Similar decomposition can be done for the enthalpic and entropic terms separately, as these terms are also state functions.

Since the enthalpic and entropic contributions to the binding free energy depend on many system-specific properties (such as protonation states, binding of metal cations, changes in conformational entropy from one ligand to another in a way which is very difficult to predict, etc), the conclusion is that optimising the overall free energy remains the most viable approach to rational (structure-based) molecular design. Attempting to get an insight into individual components of the free energy requires re-thinking the whole concept of ligand-protein binding. This means regarding ligand-protein complexes as specifically interacting yet flexible ensembles of structures rather than rigid entities, and the role of solvation effects. The significant contribution of specific interactions and flexibility to the 'intrinsic' component of binding free energy, and solvation effects will be discussed next in this chapter.

2.2 Specific interactions
2.2.1 Electrostatic interactions
Electrostatic interactions, involved in ligand-protein binding events, can be roughly classified into three types; charge-charge, charge-dipole, and dipole-dipole. Typical charge-charge interactions are those between oppositely charged atoms, ligand functional groups, or protein side chains, such as positively charged (amine or imine groups, lysine, arginine, histidine) and negatively charged (carboxyl group, phosphate groups, glutamate side chain). An important contribution to the enthalpy change associated with a binding event arises from charge-dipole interactions, which are the interactions between ionised amino acid side chains and the dipole of the ligand moiety or water molecule. The dipole moments of the polar side chains of amino acid also affect their interaction with ligands.

2.2.2 Van der Waals interactions
Van der Waals interactions are very important for the structure and interactions of biological molecules. There are both attractive and repulsive van der Waals interactions that control binding events. Attractive van der Waals interactions involve two induced dipoles that arise from fluctuations in the charge densities that occur between adjacent uncharged atoms, which are not covalently bound. Repulsive van der Waals interactions occur when the distance between two involved atoms becomes very small, but no dipoles are induced. In the latter case, the repulsion is a result of the electron-electron repulsion that occurs in two partly-overlapping electron clouds.

Van der Waals interactions are very weak (0.1- 4 kJ/mol) compared to covalent bonds or electrostatic interactions. Yet the large number of these interactions that occur upon molecular recognition events makes their contribution to the total free energy significant.

Van der Waals interactions are usually treated as a simple sum of pairwise interatomic interactions (Wang *et al.*, 2004). Multi-atom VdW interactions are, in most cases, neglected. This follows the Axilrod-Teller theory, which predicts a dramatic (i.e. much stronger than for pairwise interactions) decrease of three-atom interactions with distance (Axilrod and Teller, 1943). Indeed, detailed calculations of single-atom liquids (Sadus, 1998) and solids

(Donchev, 2006) indicate that multi-body effects amount to only 5% of the total energy (Finkelstein, 2007). However, Finkelstein (2010) shows that those largely ignored multi-atom Van der Waals interactions may lead to significant changes in free energy in the presence of covalent bonds. Those changes can be comparable to those caused by the substitutions of one atom by another one in conventional pairwise Van der Waals interactions. Thus, the currently used force fields (applied in MD simulations) need to be revised.

2.2.3 Hydrogen bonds

Hydrogen bonds are non-covalent, attractive interactions between a hydrogen covalently bonded to some electronegative group ("donor"), and another electronegative atom, such as oxygen or nitrogen ("acceptor"). The hydrogen bond can be described as an electrostatic dipole-dipole interaction. However, it also has some features of covalent bonding: it is specific, directional, it produces interatomic distances shorter than sum of van der Waals radii, and usually it involves a limited number of interaction partners, which can be interpreted as a type of valence.

Proteins contain ample hydrogen bond donors and acceptors both in their backbone and in the side chains. The environment (aqueous solvent, protein-protein network, lipid bilayers) in which proteins of interest are immersed also contains numerous proton donors and acceptors – be it water molecule, interacting proteins, lipid headgroups, or DNA/RNA. Hydrogen bonding, therefore, occurs not only between ligand and protein and within the protein itself, but also within the surrounding medium.

Like all non-covalent interactions, hydrogen bonds are fairly weak: in biological conditions, the strength of hydrogen bonds varies between 5-30 kJ/mol (outside of biological systems, the strength of hydrogen bonds may vary from 2 kJ/mol to even 155 kJ/mol for HF2-) (Emsley, 1980), which is weaker than ionic or covalent bonds. However, because of their relative weakness, they can be formed and broken rapidly during binding event, conformational changes, or protein folding. Thus, hydrogen bonds in biological systems may be switched on or off with energies that are within the range of thermal fluctuations. This is one of the prime factors that facilitates macromolecular association events, and biological activity. Another key factor is related to the strict geometric rules, followed by hydrogen bonds in biological systems. Namely, their orientations, lengths, and angular preferences, which make hydrogen bonding very specific. Due to these properties, the role of hydrogen bonds in governing specific interactions in biological recognition processes is absolutely crucial. Hydrogen bonds, both intra-and inter-molecular, are partly responsible for the secondary, tertiary, and quaternary structures of proteins, nucleic acids, and also some synthetic polymers. They play a pivotal role in molecular recognition events, and they tune the properties of the macromolecular system (e.g. mechanical strength, binding specificity). These geometric rules were among the first to be extracted from crystal structure databases (Bissantz et al., 2010). While the preferred geometries of hydrogen bonds are easily defined, their contributions to binding free energy are system-specific (Davis and Teague, 1999, Williams and Ladbury, 2003). Hydrogen bonds always convey specificity to a recognition process but do not always add much binding free energy (Bissantz et al., 2010).

Hydrogen bonds can vary quite considerably in their strength. Often, a stronger hydrogen bond implies higher penalty of desolvation, so the net free energy gain of a stronger hydrogen bond might be seriously compromised. However, such a picture is not always the case. Hydrogen bond strength, in the context of the free energy changes, should be carefully

examined, as it is likely to vary considerably from one ligand-protein system to another one (Barratt *et al.*, 2005, 2006).

Regarding weak hydrogen bonds, the most prominent donor is the CH group. These interactions, despite of their weakness, play an important role in stabilising appropriate conformations of ligand-protein complexes, for instance among the complexes between protein kinases and their inhibitors (Bissantz *et al.*, 2010). Protonated histidines can also act as strong CH donors (Chakrabarti and Bhattacharyya, 2007). Weak hydrogen bonds, their nature, and their role in ligand-protein interactions have been extensively reviewed by Panigrahi and Desiraju (2007).

2.2.4 Halogen bonds and multipolar interactions

The concept of halogen bonds is similar to hydrogen bonds: both types of interactions involve relationships between an electron donor and electron acceptor. In hydrogen bonding, a hydrogen atom acts as the electron acceptor and forms a non-covalent bond by accepting electron density from an electronegative atom ("donor"). In halogen bonding, a halogen atom is the donor.

Despite of their prevalence in complexes between proteins and small organic inhibitors (many of them contain halogen atoms due to solubility and bioavailability) and their importance for medicinal chemistry, the significance of halogen bonds in biological context has been overlooked for a long time (Zhou *et al.*, 2010). For a number of years, halogen atoms were regarded as hydrophobic appendages, convenient – from the molecular design point of view - to fill apolar protein cavities. The nature of halogen interactions (such as directionality, sigma-holes) was not studied in detail and not regarded as very important. Indeed, halogen bonds are, in general, fairly weak interactions. On the other hand, in some cases they can compete with hydrogen bonds, thus should be considered in more details, given the importance of hydrogen bonds in ligand-protein interactions and given that many of synthesised small organic compounds contain halogen bonds in their structure (Bissantz *et al.*, 2010, Zhou *et al.*, 2010).

Halogens involved in halogen bonds are chlorine, bromine, iodine, and fluorine (not very often). All four halogens are capable of acting as donors (as proven by computational and experimental data) and follow the general trend: F < Cl < Br < I, with iodine normally forming the strongest bonds, as the strength increases with the size of the halogen atom. From the chemical point of view, the halogens, with the exception of fluorine, have unique electronic properties when bound to aryl or electron withdrawing alkyl groups. They show an anisotropy of electron density distribution with a positive area (so-called σ-hole) of electrostatic potential opposite the carbon-halogen bond (Clark *et al.*, 2007). The molecular origin of the σ-hole can be explained quantum chemically and the detailed description is provided in the work by Clark and coworkers (2007). Briefly, a patch of negative charge is formed around the central region of the bond between carbon and halogen atom, leaving the outermost region positive (hence the "hole").

Available experimental data show the strong influence of halogen bonds on binding affinity. Replacement of hydrogen by halogen atom is often used by medicinal chemists in order to increase the affinity. Indeed, in a series of adenosine kinase inhibitors, a 200-fold affinity gain from hydrogen to iodine has been observed (Iltzsch *et al.*, 1995). Another spectacular, 300-fold affinity difference upon iodine substitution was observed in a series of HIV reverse-transcriptase inhibitors (Benjahad *et al.*, 2003). Unsurprisingly, substitution of hydrogen by iodine typically leads to the largest affinity gain, since the strength of the halogen bond increases with the size of halogen atom.

Halogen atoms can interact with the oxygen and with the carbon atoms of C=O groups, as well. The former attributes to the halogen bond formation, the latter is a hallmark of so-called orthogonal multipolar interactions. These interactions are formed by two dipolar functional groups, which are in a close distance from each other. Only recently it received attention in the field of medicinal chemistry and ligand-protein interactions (Paulini *et al.*, 2005), even though it has been described for a long time. This interaction is known to contribute to ligand-protein stabilisation (Fischer *et al.*, 2008), and it is particularly important in the context of halogen bonds (Bissantz *et al.*, 2010 and references therein). It is worth bearing in mind that in an orthogonal (perpendicular) orientation of two dipoles, the actual dipole contribution to interaction energy is zero. Thus, higher order electrostatic and dispersion terms must be responsible for this type of interaction. The disappearance of the dipole term may turn a repulsive electrostatic interaction into an attractive one. Because of its high electron density and low polarisability, fluorine's preference for dipolar interactions is more pronounced than for the other halogens (Bissantz *et al.*, 2010). Chlorine and other heavy halogens also form multipolar interactions with carbonyl groups, but they show a tendency for the C-X bond to be parallel rather than orthogonal to the amide plane, a consequence of the σ -hole (Bissantz *et al.*, 2010).

2.2.5 Hydrophobic interactions
The interactions between ligands and the hydrophobic side chains of proteins contribute significantly to the binding free energy. The hydrophobic residues mutually repel water and other polar groups and results in a net attraction of the non-polar groups of ligand. In addition, apolar and aromatic rings of tryptophan, phenylalanine, and tyrosine participate in "stacking" interactions with aromatic moieties of ligand. Many studies have demonstrated that the hydrophobic interactions, quantified by the amount of hydrophobic surface buried upon ligand binding, is the structural parameter correlating best with binding free energy (Bissantz *et al.*, 2010, Perozzo *et al.*, 2004). It holds well for very diverse sets of ligands (Boehm and Klebe, 1996) as well as for protein-protein interactions (Vallone *et al.*, 1998). It should be emphasised, though, that a considerable part of the affinity gain caused by hydrophobic interactions in hydrophobic binding pockets comes from sub-optimal solvation of the pocket in the unbound (apo) state.

Aromatic interactions, hydrophobic effect, and other solvent effects will be discussed further in the following parts of this chapter.

2.2.6 Interactions mediated by aromatic rings
Aromatic rings deserve special attention in the context of ligand-protein interactions. Interactions between ligands and protein aromatic side chains (Phe, Trp,and Tyr) are widespread in ligand-protein complexes (Bissantz *et al.*, 2010). The unique steric and electronic properties of these side chains, which give rise to large polarizabilities and quadrupole moments, result in preferred geometries upon interactions.

For interactions between two aromatic systems, two geometries are predominant: one, where two rings are parallel to each other, and the perpendicular, edge-to-face arrangement. High-accuracy *ab initio* CCSD(T) quantum chemical calculations of the dimerisation energy of benzene predict these two geometries to be isoenergetic (Hobza *et al.*, 1996), which agrees with experimental results qualitatively and quantitatively (Grover *et al.*, 1987, Krause *et al.*, 1991).

An introduction of heteroatoms into aromatic ring affects the ratio of both geometries. The preference to perpendicular interactions increases when the acidity of the interacting "side" atoms increases; this happens upon introduction of a strongly electron-withdrawing substituent in either ortho- or para-position. This was demonstrated by high-accuracy quantum chemical calculations by Sinnokrot and Sherrill (2004): The interaction between benzene as a donor and fluorobenzene as the acceptor, while both compounds were perpendicular to each other, was ~0.3 kcal/mol weaker than that of the benzene dimer. With reverse of roles (fluorobenzene as the donor), the interaction became ~0.6 kcal/mol stronger as compared to the benzene dimer.

For perpendicularly-oriented aromatic-aromatic interactions, studies on several model systems showed that aliphatic-aromatic interactions in the same orientation provide a favourable contribution to the free energy of the same magnitude as aromatic-aromatic interactions (Turk and Smithrud, 2001). For aliphatic-aromatic interactions, interactions energy becomes more favourable when acidity of the interacting CH unit of aliphatic counterpart increases. Study conducted by Tsuzuki *et al.* (2000) showed that ethane (sp3 hybridisation of carbon atom, less acidic) is a worse binder of benzene than acetylene (sp hybridisation of carbon atom, more acidic), and the difference in dissociation energies between acetylene-benzene and ethane-benzene complexes is around 1 kcal/mol. In ligand-protein complexes, this type of interaction can be found in interactions between aromatic side chains and methyl groups. The strength of such interactions depends on the group to which the interacting methyl group is bound: the more electronegative the group, the more the preference towards perpendicular geometry of interacting methyl-aromatic side chain is pronounced (Bissantz *et al.*, 2010). $\Pi - \pi$ interactions are also displayed by amide bonds of protein backbone (namely, their pi faces) and ion pairs - interactions between acidic (Asp, Glu) and basic (Lys, Arg) side chains.

Aromatic interactions are not limited to $\pi - \pi$ interactions. Recently, the nature of favourable interactions between heavier halogens and aromatic rings has been studied, in particular in the context of halogen bonds. C-H - halogen interactions can be regarded as "very weak hydrogen bonds" (Desiraju, 2002).

2.3 Solvent effects, structural waters, and the bulk water

Any binding event displaces water molecules from the interaction interface or from the binding pocket, while simultaneously desolvating the ligand (or a part of it). Although most of those waters are disordered and loosely associated with protein structure, such displacement affects the whole solvation shell around the ligand-protein complex (Poornima and Dean, 1995b).

While the vast majority of those water molecules are mobile and easily displaceable, some are tightly bound to the protein structure. Tightly bound water molecules are often conserved across multiple crystal structures of ligand-protein complexes (Poornima and Dean, 1995c). Often, those water molecules play an important role in tuning the biological activity of the protein, as in the case of many enzymes (Langhorst *et al.*, 1999, Nagendra e*t al.*, 1998, Poornima and Dean, 1995a). Those water molecules may be regarded as part of the protein structure. Ligand-protein interactions are often mediated by water molecules buried in the binding site and forming multiple hydrogen bonds with both binding partners (Poornima and Dean, 1995a-c). In other cases, those bound water molecules are released to

the bulk upon ligand binding. Such displacement may affect the thermodynamic signature of the binding event in a dramatic way. It is generally assumed that the release of a water molecule from a rigid environment should be entropically favorable. The upper limit of the entropy gained for transferring a water molecule from a protein to bulk solvent was estimated to be 2 kcal/mol at room temperature (Dunitz, 1994). This gain would be compensated by loss of enthalpy, so the total contribution to the free energy (as a sum of its enthalpic and entropic terms) of a single water molecule released from the protein to the bulk is difficult to guess. Moreover, in order to reach this 2 kcal/mol limit the water molecule would have to be fixed very rigidly while bound. This is often not the case, and it has been observed in numerous occasions that even very tightly bound, "structural" waters may retain a significant amount of residual mobility (Denisov et al., 1997, Fischer and Verma, 1999, Matthews and Liu, 2009, Smith et al., 2004).

"Structural" water molecules affect their surrounding not only via direct interactions (such as hydrogen-bonding network), but also by influencing the dynamical behaviour of their environment. Numerous cases have been reported when binding of the structural water affected protein flexibility (Fischer and Verma, 1999, Smith et al., 2004). The direction of such influence cannot be predicted by simple rules, as it is heavily dependent on the details of the binding site – some protein become more dynamic upon water binding (Fischer and Verma, 1999), while other ones become more rigid (Mao et al., 2000). Yet ignoring those water effects is likely to lead to substantial errors in the free energy predictions. The importance of the contributions of "structural" water molecules to binding events and its implications for drug design have been emphasised in a study by Michel et al (Michel et al., 2009).

The traditional, enthalpy-dominated view of ligand-protein association largely neglects solvation effects, which strongly affect the thermodynamic profile of a binding event. Recently it became clear that studying the hydration state of a protein binding pocket in the apo (unbound) state should be a routine procedure in rational drug design, as the role of solvation in tuning binding affinity is critical. Solvation costs are a plausible reason why some ligands, despite fitting into a binding site, fail during experimental tests as inhibitors. Young and coworkers showed that an optimised inhibitor of factor Xa turns virtually inactive when the isopropyl group interacting in the S4 pocket of factor Xa is substituted by hydrogen: The compound (PDB code 2J4I) is characterised by Ki of 1 nM. Replacing the isopropyl group by hydrogen reduces its affinity to 39 μM. Substitution of this group by hydrogen, apart from reducing the number of favourable hydrophobic interactions, leads to unfavourable solvation of the binding pocket (Young et al., 2007, and references therein).

Desolvation of the ligand itself may sometimes control the binding free energy. For highly hydrophilic ligands, the desolvation costs may be very high and make unfavourable contributions to the binding (Daranas et al. 2004, MacRaild et al., 2007, Syme et al., 2010). The calorimetric study of β-galactose derivatives binding to arabinose binding protein (ABP) showed dramatic differences in binding free energy between several deoxy derivatives (Daranas et al., 2004). The most likely reason of 4-deoxygalactose failing to bind to ABP is the unfavourable desolvation cost (Bronowska and Homans, unpublished data).

Spectroscopic evidence shows that (1) water molecules in the first solvation shell (surrounding the hydrophobic solute) are more flexible that it was originally thought (Finney and Soper, 1994) and (2) hydrogen bonds at hydrophobic surfaces are weaker than it was assumed (Scatena et al., 2001). In addition, the properties of the water molecules from first two solvation shells are very different from these of bulk water, as emerged from terahertz spectroscopy results (Ebbinghaus et al., 2007, Heugen et al., 2006).

2.4 Classical and non-classical hydrophobic effect

The concept of the classical hydrophobic effect relies on a hydrophobic solute disrupting the structure of bulk water. This decreases entropy due to ordering of water molecules around the hydrophobic entity. Such unfavourable effects can be minimised if solute molecules aggregate. Upon aggregation, water molecules form one larger "cage" surrounding the hydrophobic aggregate, and the surface area of such aggregate is smaller than the sum of surface areas of individual (non-aggregated) solutes. This makes the entropic contribution less unfavourable and hence makes the free energy more favourable (Homans, 2007).

If this mechanism was the sole driving force for a protein-ligand interaction, all binding events involving hydrophobic ligands would be entropy-driven. This is not the case. Several years ago, in the group of Steve Homans (University of Leeds), we studied the thermodynamics signature of ligand binding by the mouse major urinary protein (MUP). This protein is characterised by a strongly hydrophobic binding pocket and it binds a handful of very different hydrophobic ligands – long-chain alcohols and pyrazine derivatives, among others. Surprisingly, the ITC data showed that the binding was enthalpy-driven (Barratt *et al.*, 2005). This was combined with a negative change in heat capacity upon binding - a hallmark of the hydrophobic effect.

In order to elucidate the molecular origin of this unusual binding signature, we employed computational methods, such as molecular dynamics (MD) simulations. I will discuss the results in more details later in this chapter. The data showed that the key to this favorable enthalpy of binding of ligands to MUP seems to be the sub-optimal solvation of the binding pocket in apo (unbound) state: only a few water molecules remained there prior to ligand binding. The favourable enthalpic component was, thus, largely determined by ligand desolvation, with only a minor contribution from desolvation of the protein. Such complexation thermodynamics driven by enthalpic components have been referred to as the "non-classical hydrophobic effect".

2.5 Enthalpy-entropy compensation, binding cooperativity, and protein flexibility

The enthalpic and entropic contributions are related. An increase in enthalpy by tighter binding may directly affect the entropy by the restriction of mobility of the interacting molecules (Dunitz, 1995). This phenomenon, referred to as enthalpy-entropy compensation, is widely observed, although its relevance is disputed (Ford, 2005). Such compensation, although frequently observed, is not a requirement: if it was, meaning that changes in ΔH were always compensated by opposing changes in $T\Delta S$, optimisation of binding affinities would not be possible, which is clearly not the case.

In connection to the enthalpy-entropy compensation, ligand-protein interactions can be cooperative, which means the binding energy associated with them is different than the sum of the individual contributions to the binding free energies. Cooperativity provides a medium to transfer information, enhance or attenuate a response to changes in local concentration and regulate the overall signalling/reaction pathway. Its effects are either positive (synergistic) or negative (interfering), depending on whether the binding of the first ligand increases or decreases the affinity for subsequent ligands. Noncooperative (additive) binding does not affect the affinity for remaining ligands and the subsequent binding sites can be regarded as independent.

Cooperativity is often linked to pronounced conformational changes in the structure of the protein. It can be, in some cases, caused by structural tightening through the presence of additional interactions; inter-atomic distances become shorter and interaction becomes

enthalpically more favorable. Evidence for such a mechanism has been reported for many ligand-protein complexes; biotin-streptavidin being one of the most extensively studied (Williams *et al.*, 2003). In other cases, cooperativity can occur in the absence of any conformational changes of the protein, and be driven solely by changes in protein dynamics (Homans, 2005, Wand, 2001). Catabolite-activated protein (CAP) is a very good example of such dynamic allostery. CAP is a transcriptional activator that exists as a homodimer in solution, with each subunit comprising a ligand-binding domain at the N-terminal domain and a DNA-binding domain at the C-terminal domain (Harman, 2001). Two cyclic AMP (cAMP) molecules bind to CAP dimer, and this binding increases affinity of CAP for DNA (Harman, 2001). Binding of each cAMP molecule shows negative cooperativity, i.e. binding of the first cAMP molecule decreases affinity of binding of the second cAMP molecule to CAP. This is accompanied by absence of long-range structural changes. Thermodynamic analysis, performed by a combination of ITC and solution NMR, confirmed that the observed negative cooperativity was entirely driven by changes in protein entropy (Popovych *et al.*, 2009). Thus, it is more appropriate to describe the phenomenon of cooperativity in terms of thermodynamics rather than merely conformational changes (if any such changes can be observed), since it is fundamentally thermodynamic in its nature.

Examples above illustrate the importance of protein dynamics in binding events. Proteins tend to compensate the unfavourable entropic contribution to ligand binding by increasing their dynamics in regions distant from the ligand binding site (Evans and Bronowska, 2010, MacRaild *et al.*, 2007) Flexible binding sites may require more flexible ligand moieties than 'stiffer' ones. The traditional focus on the enthalpic term (direct and specific interactions) and dominance of the 'induced fit' model has led to an overly enthalpic view of the world that neglects protein flexibility. Such view of the ligand-protein binding events, although very intuitive, is flawed by neglect of entropic contributions and – as a consequence – an impairment to correct predictions of free binding energy. Although it is true that tighter interactions make binding more favourable, the thermodynamic signature of a "good" binder does not need to be dominated by an enthalpic term.

3. Methods

3.1 Experimental methods

Many experimental techniques have been developed to study various aspects of ligand-protein thermodynamics. X-ray crystallography provides very valuable information about the enthalpic contribution (hydrogen and halogen bonds, electrostatic interactions, etc). Although it focuses on static structures of ligand-receptor complexes, it also yields some information on entropic contribution. B-factors (temperature factors), obtainable for heavy (non-hydrogen) atoms of the complex under investigation, are sensitive to the mean square displacements of atoms because of thermal motions, therefore they reflect on ligand-protein dynamics. However, B-factors do not distinguish time scales of the motions and their interpretation is not straightforward. X-ray (Makowski *et al.*, 2011) and neutron scattering (Frauenfelder and Mezei, 2010) also reflect on ligand-protein dynamics. The former one focuses on global changes in protein size and shape in a time-resolved manner, while the latter reports on motion amplitudes and time scales for positions of hydrogen atoms. Another technique useful in understanding protein dynamics both in unbound (apo) and bound (holo) forms is fluorescence spectroscopy (Weiss, 2000). Single molecule techniques

also offer an opportunity to measure contributions to binding events from interacting partners individually. Hydrogen-deuterium exchange mass spectrometry (HX-MS) and related methods, have been very successful in studying protein dynamics in large supramolecular complexes (Wales and Engen, 2006). The motion of the entire complex and individual contributors, and the dynamics of the binding events can be investigated by time-resolved HX-MS (Graf *et al.*, 2009). Another technique frequently used to study binding events is surface plasmon resonance (SPR), which allows for straightforward determination of equilibrium binding constants (Alves *et al.*, 2005). Terahertz spectroscopy is a relatively new technique, used primarily to probe solvation of macromolecules and their complexes (Ebbinghaus *et al.*, 2007). It is very sensitive to changes of the collective water network dynamics at the at the macromolecule-water interface. Terahertz absorption spectroscopy can also be used to probe collective modes in ligand-protein complexes (Xu et al., 2006).

There are two groups of methods that deserve special attention in the context of thermodynamics of binding events and will be discussed more in details in the following part of this chapter. One of these is NMR spectroscopy, especially powerful for the study of ligand-protein dynamics, hence the entropic contribution to the binding free energy (Meyer and Peters, 2003). The other group contains calorimetric techniques, which are very important for the study of biological systems, their stability, and the thermodynamics of macromolecular interactions. Currently, two most popular techniques applied to investigate biological systems are differential scanning calorimetry (DSC) and isothermal titration calorimetry (ITC). The former quantifies the heat capacity and enthalpy of thermal denaturation, the latter measures the heat exchanged during macromolecular association. While DSC provides the way to estimate the stability of the system (protein, nucleic acid, ligand-protein complex, etc), ITC is an excellent tool to study the thermodynamics of binding events (Perozzo *et al.*, 2004). Since this chapter is dedicated to the thermodynamics of macromolecular associations, in the course of this chapter I will focus mainly on ITC and its applications to study biological systems.

3.1.1 Isothermal titration calorimetry (ITC)

ITC measures the heat evolved during macromolecular association events. In an ITC experiment, one binding partner (ligand) is titrated into a solution containing another binding partner (protein), and the extent of binding is determined by direct measurement of heat exchange (whether heat is being generated or absorbed upon the binding). ITC is the only experimental technique where the binding constant (K_d), Gibbs free energy of binding (ΔG), enthalpy (ΔH) and entropy (ΔS) can be determined in a single experiment (Perozzo *et al.*, 2004). ITC experiments performed at different temperatures are used to estimate the heat capacity change (ΔC_p) of the binding event (Perozzo *et al.*, 2004).

During last few decades, ITC has attracted interest of broader scientific community, as a powerful technique when applied in life sciences. Several practical designs emerged, but the greatest advances have happened during last 10 years. Development of sensitive, stable, and – last but not least - affordable calorimeters made calorimetry a very popular analytical procedure and ITC became the gold standard in estimations of macromolecular interactions. Given the ability of ITC to obtain a full thermodynamic description of the system studied, the technique has found widespread applicability in the study of biological systems. Apart from its versatility and simple experimental setup, ITC also has advantages over some other

techniques: the experiments can be performed in a physiologically relevant buffer, no surface effects have to be taken into account, and the interacting species do not require immobilisation or labelling.

ITC is also used for determination of binding affinity-independent reaction stoichiometry. The reaction stoichiometry is estimated from the titration equivalence point. Provided this, ITC is increasingly used in the analysis of macromolecular complexes involving multiple binding events (e.g. protein aggregation or the formation of multi-protein complexes). Systems that involve multiple binding events that occur at two or more interacting sites often demonstrate cooperativity, which is an important mechanism of regulation in biological systems (Brown, 2009).

Using ITC it is also possible to study protonation effects, in cases when protein-ligand binding is coupled to changes in the protonation state of the system. If the formation of the complex changes the protonation state of ligand as well as that of the protein (whether free or bound), proton transfer with the solvent occurs. As a result, the signal measured by ITC will contain the heat effect of protonation/ deprotonation, contributing to the overall heat of binding. Repeating the experiment at the same pH in buffers with different ionisation enthalpies but otherwise under the same conditions allows for the determination of the number of protons released/ accepted by buffer solution. From this, the intrinsic binding enthalpy corrected by protonation heats, can be established (4).

$$\Delta H_{obs} = \Delta H_{bind} + nH + \Delta H_{ion} \tag{4}$$

ITC can also provide information about solvation effects. If ΔH is determined at a range of temperatures, the change in the constant pressure heat capacity (ΔC_p) for an interaction is given by the slope of the linear regression analysis of ΔH_{obs} plotted vs. temperature. There is a strong correlation between ΔC_p and the amount of desolvated (buried) surface area of a macromolecular complex. Thus, for the ligand-protein binding events, ΔC_p is most often negative, when the complex is regarded as a reference state. Through this correlation, changes in ΔC_p are measure of solvation state of the macromolecule and involvement of solvent effects in binding event (Perozzo et al., 2004).

3.1.1.1 Experimental setup

In a typical ITC experiment, a solution of ligand is injected (titrated) into a solution of the protein, in small volumes, over the time. During that time, the changes in heat resulting from the interaction are monitored (Figure 2, upper panel). Each peak represents a heat change associated with the injection of a ligand sample into the protein solution inside the ITC reaction cell. Concentrations of both ligand and protein in their respective solutions are known. As the ligand-protein system reaches saturation, the heat changes diminish until only heats of dilution are observed. A binding curve is then obtained from a plot of the heats from each titration against the ratio of ligand and protein inside the ITC cell (Figure 2, lower panel). The binding curve is analysed with the appropriate binding model to determine the thermodynamic parameters.

ITC is a straightforward technique to accurately measure binding events with affinity range from mM up to high-nM. Problems occur when the ligand binds very tightly, in a single-digit nM and below. This is due to the titration curve becoming too steep to fit accurately. In such cases, the displacement experiments are commonly used. Such experimental setup consists of binding a low-affinity binding ligand first and then displacing it during titration

with a stronger binder of interest. However, this method requires precise knowledge of binding constants of those weak binders. The experimental setup of the displacement assay is often challenging, as there are several factors increasing the error of measurements.
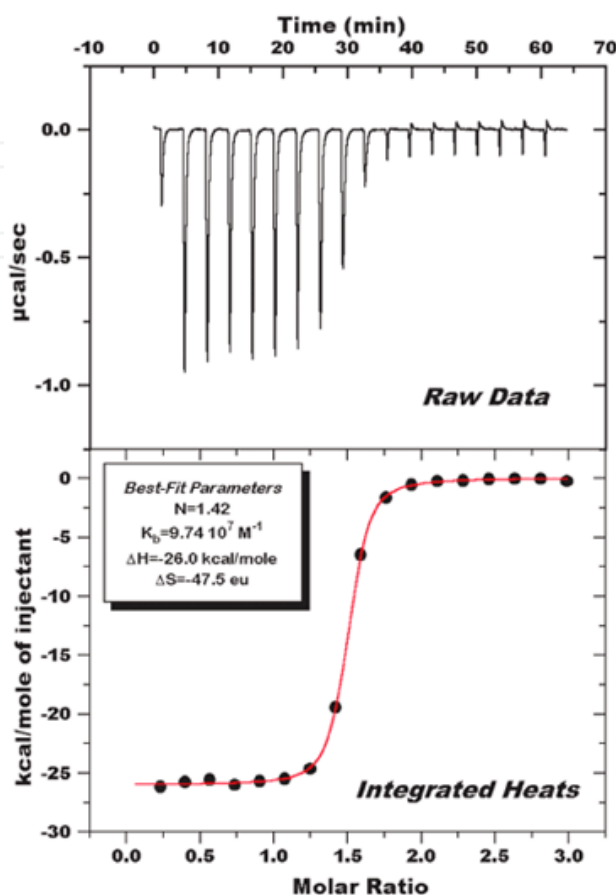

Fig. 2. An example of ITC data. Raw data, representing observed changes in heat resulting from interactions are shown in the upper panel. The resulting binding curve is displayed in the lower panel (from MicroCal materials http://www.microcal.com/technology/itc.asp).

It is worth remembering that ITC experiment not only measures the heat absorbed or released during binding reactions, but it also detects the total heat effect in the calorimetric cell upon titration of ligand. Thus, the experimental data contain contributions arising from non-specific effects, such as dilution of ligand and protein, mixing two solutions of slightly different compositions, temperature differences between the ITC cell and the titrating syringe, and so forth. In order to determine these contributions the control experiments need to be performed in order to extract the heat of ligand-protein complex formation.

3.1.1.2 Thermodynamic content of ITC data

The ΔG determines the stability of any ligand-protein complex of interest, which makes it very useful for studies and predictions of structure–activity relationships. The conventional analysis of ITC data involves fitting an appropriate model (*i.e.* single- or two-site binding model) to the data, and obtaining the binding constant. Quite often, though, more sophisticated models (such as multiple interacting-site models) must be applied, if the behaviour of the system is more complex.

As mentioned earlier, observed overall ΔG can be very similar regardless of the driving force, which can be very different from one case to another. ΔG can be the same for an interaction with positive ΔS and ΔH (entropy-driven, binding signature dominated by the classical hydrophobic effect), an interaction with negative ΔS and ΔH (enthalpy-driven binding signature), or all sort of combinations of negative ΔH and positive ΔS. As described in the previous section, ligand-protein complexes tend to compensate for enthalpic and entropic contributions, making changes in ΔG less sensitive to the molecular details of the interactions. Therefore, dissection of ΔG into enthalpic and entropic contributions is of a fundamental importance for understanding of the binding energetics.

3.1.1.2.1 Enthalpic contributions

The change in the enthalpy represents the changes in energy associated with specific, non-covalent interaction. However, such an interpretation is too simplistic to describe experimental ΔH values, and the physical meaning of observed ΔH seems to be more complex. The measured changes in enthalpy are the result of the formation and breaking of many individual bonds; it reflects the loss of protein–solvent hydrogen bonds and van der Waals interactions, the loss of ligand-solvent interactions, the formation of ligand-protein bonds, salt bridges and van der Waals contacts, the re-organisation of the intra-molecular hydrogen-bonding network of the protein, solvent reorganisation near the protein surface, conformational changes at the binding site due to the binding event, and many more. These individual components may produce either favourable or unfavourable contributions, depending on the system.

The treatment of each component individually is very challenging since the global heat effect of a particular interaction is a balance between the enthalpy of the ligand binding to the protein and to the solvent. Several approaches have been employed to investigate the energetics of individual bonds, including alanine scanning mutagenesis (Perozzo *et al.*, 2004 and references therein), and removal of particular hydrogen bonds at the binding site (Connelly *et al.*, 1994). However, these approaches suffer from the major bottleneck, resulting from the fact that a direct relation between the change in enthalpy and the removal of the corresponding specific interactions cannot be made *a priori*.

A large part of the observed ΔH is due to a bulk hydration effect, as emerged from ITC studies carried out in water and deuterium (Connelly *et al.*, 1993). Frequently, water molecules are located at complex interfaces, improving the complementarity of the surfaces and extending hydrogen-bonding networks. This should contribute favourably to the enthalpy, but it may be offset by an unfavourable entropic contribution (Perozzo *et al.*, 2004). The role of interfacial water was studied by lowering water activity by adding osmolytes such as glycerol to the solution. It was found that complexes with a low degree of surface complementarity and no change in hydration are tolerant to osmotic pressure (Perozzo *et al.*, 2004, and references therein).

3.1.1.2.2 Entropic contributions

ΔS may be calculated directly from ΔG and ΔH, according to the Gibbs' equation. Its physical representation is not straightforward. It is often related to the dynamics and flexibility of the system (Diehl *et al.*, 2010, Homans, 2007), sometimes dubbed as a 'measure of the system's disorder' (which is incorrect). It has been proposed that the ΔS associated

with ligand-protein binding can, at a given temperature, be expressed as the sum of several contributing effects. The main one is related to solvent effects. The burial of water-accessible surface area upon binding event should result in release of confined or interfacial water molecules to the bulk. This should contribute favourably to the total entropy of interaction. A positive entropy change is usually a strong indication that water molecules have been released from the complex surface (Jelesarov and Bosshard, 1999). On the other hand, interfacial water remaining upon binding can also contribute positively to the total entropy of the interaction (Fischer and Verma, 1999).

Another important entropic contribution is related to the reduction of conformational (rotational and vibrational) degrees of freedom of protein side-chains. In addition to these, the ligand loses translational degrees of freedom upon binding. All these contribute unfavourably to the overall entropy of interaction. However, in some cases the protein increases the number of conformational degrees of freedom upon ligand binding, as observed by NMR and deduced from MD simulations (MacRaild *et al.*, 2007, Stoeckmann *et al.*, 2008). This is likely to happen in order to partly offset the unfavourable entropic contribution from ligand binding and thus to reduce the overall thermodynamic cost of that process.

3.1.1.2.3 Enthalpy-entropy compensation

As mentioned in the previous section of this chapter, this phenomenon is described by the linear relationship between the change in enthalpy and the change in entropy. This means that favourable changes in binding enthalpy are compensated by opposite changes in binding entropy and vice versa, resulting in very small changes in overall free binding energy. Enthalpy–entropy compensation is an illustration of the 'motion opposes binding' rule, and it is believed to be a consequence of altering the weak inter-molecular interactions as well as being related to solvent effects. Since both ΔH and ΔS are connected to ΔC_p, the correlation between enthalpy/entropy and heat capacity changes is clear.

Enthalpy-entropy compensation is a difficult problem to address in the context of rational molecular design. In such framework, the goal is to maximise the binding affinity of a complex of the designed compound and the protein target. The optimisation strategy requires simultaneous minimisation of both enthalpic and entropic penalties. However, reducing one of them usually means increasing the other.

3.1.2 Nuclear Magnetic Resonance (NMR) spectroscopy

Thermodynamics of biologically-relevant macromolecules and their complexes can be characterised by measurements using NMR spectroscopy. The basis of NMR spectroscopy is the non-zero nuclear magnetic moment of many elements, such as ^{1}H, ^{13}C, ^{15}N, or ^{19}F. When put into an external static magnetic field (B), the different nuclear spin states of these elements become quantised with energies proportional to their projections onto vector B. The energy differences are also proportional to the field strength and dependent on the chemical environment of the element, which makes NMR an ideal technique to study 3D structural and dynamical properties of the systems.

A variety of NMR methods have been introduced to study ligand-protein interactions. These methods include one-, two- and three-dimensional NMR experiments. Many studies, to date, proved the power of stable-isotope labelling and isotope-edited NMR in the investigation of ligand-protein interactions. Recent development of techniques allowed for

the study of ligand-induced conformational changes, investigating positions and dynamic behaviour of bound water molecules, and for quantification of conformational entropy. The steady-state heteronuclear Overhausser effects (NOEs) are very useful for structural analysis of three-dimensional structures of macromolecules in solution (Boehr *et al.*, 2006, Meyer and Peters, 2003). It is important to note that the NOE occurs through space, not through chemical bonds, which makes it applicable to characterise non-covalent binding events. When ligand binds the NOEs change dramatically, and transferred NOEs (trNOEs), relying on different tumbling times of free and bound interactors, can be observed.

Another NMR technique commonly used for identification of the ligand binding is chemical shift mapping (Meyer and Peters, 2003). Briefly, chemical shifts describe the dependence of nuclear magnetic energy levels on the electronic environment in the given macromolecule. Electron density, electronegativity, and aromaticity are among the factors affecting chemical shifts. Not surprisingly, binding event changes the chemical shifts of both interacting partners, particularly in the area of the association (e.g. protein binding pocket, protein-peptide interaction interface). Thus, changes in chemical shifts can be used to identify binding events and to describe the location of the binding.

Ligand-protein thermodynamics can be investigated using NMR relaxation analysis, which provides an insight into protein dynamics in the presence and the absence of ligand. These results can be integrated with thermodynamic data obtained from isothermal titration calorimetry (ITC) experiments and computational results (e.g. MD simulations). For proteins, the relaxation rates of backbone (^{15}N) and side chains (^2H and ^{13}C), can be obtained. The time scales available to NMR ranges over 17 orders of magnitude, reflecting protein motions on timescales from picoseconds to milliseconds (Boehr *et al.*, 2006). This covers all the relevant motions of proteins and their complexes.

Backbone and side chain (methyl groups) NMR relaxation measurements revealed the role of protein dynamics in ligand binding and protein stability (Boehr *et al.*, 2006). Development of molecular biology techniques for incorporation of stable, ^{13}C and ^{15}N isotopes into expressed proteins allowed for design and application of modern multidimensional heteronuclear NMR techniques. As a consequence, the maximum size of the macromolecule studied using these techniques rose from about 10 kDa (when ^1H homonuclear NMR is used) to 50 kDa and beyond (using ^{13}C and ^{15}N heteronuclear NMR with fractional ^2H enrichment). Application of modern TROSY (transverse relaxation optimized spectroscopy) techniques further expanded the size limitations of NMR, reaching up to the 900 kDa (Fernandez and Wider, 2003).

While NMR methodologies are being developed to study ligand-protein complexes in solid state, special techniques have been developed specifically to study protein stability and folding (Baldus, 2006), or in-cell NMR (Burz *et al.*, 2006), providing complementary information to fluorescence studies in biological settings. In this chapter I will briefly discuss only application of relaxation analysis in solution for the study of ligand-protein thermodynamics, specifically intrinsic entropic contributions.

3.1.2.1 Slow and fast dynamics: from dynamics to entropy

Conformational changes that may be associated with ligand binding events generally occur on 'slow' (microsecond to millisecond) time scales and thus report on slower motions than protein backbone and side chain fluctuations (pico-to-nanoseconds). There is no straightforward relationship between 'slow' and 'fast' motions. Experiments on several ligand-enzyme systems have shown that binding events, which decrease the 'fast' motions, may increase, decrease, or not affect the 'slow' motions (Boehr *et al.*, 2006). This obviously has an effect on the overall entropy contribution, but this has not been fully explored.

NMR relaxation techniques have been used to study multiple time scale dynamics of ligand-protein complexes. Their results show that even though large conformational changes occur on 'slow' time scale, 'fast' (pico-to-nanosecond) protein motion plays important roles in all aspects of binding event. These are typically probed by measuring three relaxation rates: the longitudinal relaxation rate (R1), the transverse relaxation rate (R2), and the NOE. These relaxation rates are directly related to the spectral density function, $J(\omega)$. This function is proportional to the amplitude of the fluctuating magnetic field at the frequency ω. Such fluctuating magnetic fields are caused by molecular motion in an external magnetic field, which is closely coupled to nuclear spin relaxation (Boehr *et al.*, 2006, and references therein).

In early studies of ps-ns time scale protein dynamics, various models for protein internal motion were used to generate different spectral density functions that were then compared to the experimental data. Subsequently, Lipari and Szabo (Lipari and Szabo, 1996) generated a spectral density function (5) that is independent of any specific physical model of motion, which is shown in equation 5 and is referred to as model free formalism.

$$J(\omega) = \frac{S^2 \tau_m}{1 + \omega^2 m \tau^2} + \frac{\left(1 - S^2\right)\tau}{1 + \omega^2 \tau^2} \tag{5}$$

For isotropic tumbling (ligand-protein complex tumbles in the water solution), where τ_m is the correlation time for the overall rotational diffusion of the macromolecule, S^2 is the order parameter, and $\frac{1}{\tau} = \frac{1}{\tau_m} + \frac{1}{\tau_e}$, where τ_e is the time scale (ns) for the bond vector internal motions. An order parameter of 1 indicates complete restriction of internal motion, and $S^2 = 0$ indicates unrestricted isotropic internal motion. It should be emphasised that S^2 parameters have a straightforward physical interpretation. The simplest model relates S^2 to 'diffusion in a cone' with semi-angle θ, and is shown in Figure 3.

$$S^2 = \frac{\cos^2 \theta (1 + \cos \theta)^2}{4}$$

Fig. 3. Physical interpretation of S^2 order parameters. S^2 can be interpreted as a measure of a free rotation of a bond vector (here – N-H) in a cone. The semi-angle θ is displayed. Smaller S^2 parameters correspond to more flexible bond vectors. $S^2 = 0$ means unrestricted rotation of the bond vector.

There were attempts to relate order parameters to structural characteristics of proteins and ligand-protein complexes. It was observed that amino acids with smaller side chains tend to show – intuitively - greater backbone flexibility than those with bulkier side chains (Goodman *et al.*, 2000). However, the variation of backbone amide S^2 parameters is larger than the differences between the averages for different amino acid types. Backbone amide order parameters are also only weakly affected by secondary structure elements, with loops having only slightly smaller average S^2 N-H values than helices or beta-turns (Kay et al., 1989). Backbone S^2 N-H values can be predicted from structures using a simple model that takes account of local contacts to the N-H and C=O atoms of each peptide group (Zhang and Brüschweiler, 2002).

A more sophisticated model for predicting dynamics from structure has recently been reported (tCONCOORD) (Seeliger *et al.*, 2007). tCONCOORD allows for a fast and efficient sampling of protein's conformational degrees of freedom based on geometrical restraints. Weak correlation between side chain order parameters and contact distance between the methyl carbon and neighboring atoms, with solvent exposure (Ming and Brüschweiler, 2004), and amino acid sequence conservation patterns (Mittermaier *et al.*, 2003) have been reported in literature. These results demonstrate that protein dynamics are strongly affected by the unique architecture of the protein as well as the environment. Thus, it cannot be readily predicted by the bioinformatic techniques, based on the primary/secondary sequence analysis. Developing a fast and reliable method of assessment of protein dynamics is, nevertheless, crucial for predictions of ligand-protein interactions - as it will be shown in the course of this chapter, dynamics affects all stages of molecular recognition events.

Order parameters can be related to entropy through the relationship developed by Yang and Kay (1996). This formalism quantifies the conformational entropy associated with observable protein motions by means of a specific motion model. For a wide range of motion models, the functional dependence of entropy on the S^2 parameter was demonstrated to be similar (Yang and Kay, 1996). This suggests that changes in S^2 can be related to the conformational entropy change in a model-independent manner. This approach has many advantages: it is straightforward, relatively free of assumptions (the requirement is that the internal motions are uncorrelated with the global tumbling of the macromolecule), and applicable to both NMR experiments and theoretical approaches (MD simulations). Moreover, since S^2 parameters are measured per bond vector, this approach enables site-specific reporting of any loses, gains, and redistributions of conformational entropy through different dynamic states of the ligand-protein complex.

However, the model-free formalism can give only a qualitative view of micro-to-millisecond time scale motions. Failure to correctly account for anisotropic molecular tumbling and the assumption that all motions are un-correlated seriously compromises the usefulness of this approach for studying dynamics associated with large conformational changes or concerted motions. Because of the time scales, alternative approaches must be implemented to study motions occurring at a millisecond time scale (e.g. R2 relaxation dispersion).

3.1.3 Combination of ITC and NMR

As described, ITC obtains free energy as the global parameter, thus, effects like ligand-induced conformational changes, domain-swapping, or protein oligomerisation, which contribute to the overall ΔG, will not be resolved. In order to assess the role those factors

play in a binding event, a combination of ITC and other techniques (such as NMR) need to be used. A combination of ITC and NMR proves useful in studying cooperativity phenomena. Heteronuclear NMR spectroscopy is one of the few experimental techniques capable of measuring the occupancies of individual binding sites on proteins and therefore determining microscopic binding affinities. Coupling this site-specific data (e.g. chemical shift mapping and/or relaxation analysis data) with the macroscopic binding data from ITC allows a complete description of the binding properties of the system. A method of determining cooperativity using heteronuclear solution NMR spectroscopy has been described using an isotope-enriched two-dimensional heteronuclear single-quantum coherence experiment (2D HSQC) (Tochtrop *et al.*, 2002). The ligands are isotopically labelled (usually ^1H, ^{15}N, or ^{13}C), while the receptor remains unlabelled. Spectra are acquired at different molar ratios and the peak volumes are integrated. Isotherms are generated by plotting the peak volume integration against molar ratio. The data is then fitted to site-specific binding models to obtain the thermodynamic parameters (Brown, 2009).

3.2 Computational approaches

Computational approaches to ligand-protein interaction studies have great potential and the development of various methods, briefly described in this chapter, have been truly outstanding. However, every method – computational, experimental alike - has its limitations and computational methods should not be used in a 'black box' manner; one should beware of the 'Garbage In Garbage Out' phenomenon. Yet it is evident that theoretical approaches have finally come to the stage that makes rational molecular design truly rational.

During a binding event, the ligand may bind in multiple orientations. The conformation of either of the interacting partners can change significantly upon association. The network of intramolecular interactions (e.g. hydrogen bonds, salt bridges) can dramatically change (breaking and/or creating new contacts), and new intermolecular interactions occur. Water molecules and ions can be expelled upon binding, or – on the contrary – bind more tightly. Finally, conformational or solvation entropic contributions may play significant role, affecting the free energy in a way which is difficult to predict.

Growing amount of calorimetric data available allows the investigation of the thermodynamic profiles for many ligand-protein complexes in detail. When structural data (crystal, NMR) are available as well – and often it is the case - it is very appealing to speculate about the link between the structure of the complex and the thermodynamics of the binding event. However, such speculations are challenging. It is important to bear in mind that both enthalpic and entropic contributions to the free energy terms obtained from ITC experiments are global parameters, containing a mixture of different contributions, which can have either equal or opposing signs and different magnitudes. This may lead to various thermodynamic signatures of a binding event. Moreover, 'structural' interpretation of intrinsic entropic contributions is notoriously difficult. Hence, the experimental thermodynamic data cannot be easily interpreted on the basis of structural information alone. Last but not least, the contribution from the solvation effects is difficult to get insight into, and although direct experimental estimations of solvation free energy have been attempted, these always require additional assumptions (Homans, 2007, Shimokhina *et al.*, 2006).

No doubt, a great advantage of theoretical approaches lies with gaining an insight about each of those contributions and their de-convolution. Binding events (ligand-protein

binding pose and the strength of their interactions) can be predicted by molecular docking, albeit intrinsic entropic contributions and solvation effects are usually ignored. Dynamic behaviour of proteins and ligands can be studied using extensive molecular dynamics (MD) simulation, which, combined with experimental NMR and ITC data, provide extremely valuable information on configurational entropy changes upon binding event, and hence about the intrinsic entropic contribution to the free energy. The global free energy changes can be studied by free energy perturbation (FEP) calculations, or related methods, such as thermodynamic integration (TI). Molecular docking methods allow for a quick assessment of enthalpic contributions, while solvent effects can be studied either by quantum chemical (QM) calculations (e.g. COSMO model), hybrid QM/MM schemes, or FEP-related methods. Theoretical approaches allow also for investigations of transient phenomena, e.g. short-living alternative conformers from an ensemble that contribute to the binding event but which cannot be readily observed. In a situation – which is not uncommon – when an experimental structure of the protein target or a part of it is missing (such as in cases of most G-protein-coupled receptors), computational approaches allow the generation of such structures (e.g. by homology modelling, threading, or *ab initio* predictions) and its use for predictions which can be validated experimentally, despite of the absence of protein structural data. Therefore, usage of theoretical methods is indispensable – not only for the interpretation of the existing experimental data, but also to direct and design new experiments.

Because of space limitations, only two theoretical methods, which are the most relevant for thermodynamics of molecular binding events, will be briefly discussed: MD-related methods (which includes MD simulations, FEP-like approaches, methods which use MD algorithms with enhanced sampling, and hybrid QM/MM schemes), and quantum chemical (QM) calculations. This division is not strict and many of these methods overlap, e.g. QM/MM methods use both MD simulations and QM calculations, and FEP-like methods have many flavours, including hybrid QM/MM-FEP.

3.2.1 Molecular Dynamics (MD) simulations

Molecular dynamics (MD) simulation consists of the numerical solution of the Newton's equations of motion of a system (e.g. protein, or a ligand-protein complex in water environment). The potential energy of the particle system is described by a function called force field (6). The potential energy of the system (U) is described as a sum of energy terms for covalent bonds, angles, dihedral angles, a van der Waals non-bonded term, and a non-bonded electrostatic term (Cornell *et al.*, 1995). Since the kinetic energy is also taken into account, the system is able to move across the energy barriers on the potential energy surface, which implies substantial changes (e.g. conformational) during the simulation.

$$
U = \sum_{\text{bonds}} k_r \left(r - r_0 \right)^2 + \sum_{\text{angles}} k_\theta \left(\theta - \theta_0 \right)^2 + \sum_{\text{dihedrals}} k\phi \left[1 + \cos\left(n\phi + \phi_0 \right) \right]
$$
$$
+ \sum_{\text{atom } i} \sum_{j \neq i} 4\varepsilon_{i,j} \left[\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^6 \right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\varepsilon_0 r_{i,j}}
\tag{6}
$$

The principles of MD simulations, algorithms used, and different types of force fields applied (all-atom, united atom, coarse-grain, etc) have been described in many publications

(Klepeis *et al.*, 2009 and references therein). MD methods rely on quality of the force field (parameters, inclusion of non-additive effects, etc), description of solvent effects, adequate sampling, and quality of initial structures used for the simulations. The quality of the results relies also on the duration of the simulation. There are limits on the time scales at which the system of interest can be considered. Simulation runs are fairly short: typically nanoseconds to microseconds, rarely extending to miliseconds, if super-fast computers are employed. Since biological processes (ligand-protein binding, large conformational changes, etc) typically occur ar micro-to-milisecond scales, one needs to assess whether or not a simulation has reached equilibrium before the averages calculated can be trusted. Furthermore, the averages obtained need to be subjected to a statistical analysis, to make an estimate of the errors.

MD methods have been widely employed to study ligand-protein binding phenomena, conformational changes, solvent effects, and to assess individual contributions to the binding free energy. These methods are particularly useful in assessing the conformational entropic contribution to the free energy. Information about ps-to-ns time scale molecular motions can be readily obtained from the MD simulation trajectory and analysed either through diagonalisation of the covariance matrix of displacements of atomic Cartesian coordinates - quasi-harmonic analysis, Schlitter's approach (7,8), analysed through principal component analysis (PCA), or quantified NMR-like via generalised order parameters. Entropy changes can be estimated from the MD trajectory through Yang and Kay's relationship (9). The order parameter analysis has the advantage of being able to calculate order parameters by-vector, thus providing site-specific information on flexibility and hence intrinsic entropic contribution. Computed parameters can be also directly compared to the experimental results of NMR relaxation analysis (Best and Vendruscolo, 2004). In last few years several studies proved the success of this methodology in estimating of entropic contributions to the binding thermodynamics.

$$S < S' = \frac{1}{2} k_B \ln \det \left[\mathbf{1} + \frac{k_B T e^2}{\hbar^2} \mathbf{M} \sigma \right] \qquad (7)$$

$$\sigma_{ij} = \left\langle \left(x_i - \langle x_i \rangle \right) \left(x_j - \langle x_j \rangle \right) \right\rangle \qquad (8)$$

$$S^2 = \frac{3}{2} \left[\langle x^2 \rangle^2 + \langle y^2 \rangle^2 + \langle z^2 \rangle^2 + 2\langle xy \rangle^2 + 2\langle yz \rangle^2 + 2\langle yz \rangle^2 \right] - \frac{1}{2} \qquad (9)$$

MD-based approaches used for binding thermodynamics calculations include free energy perturbation (FEP) (Foloppe and Hubbard, 2006), thermodynamic integration (TI) (Straatsma and Berendsen, 1988), lambda-dynamics simulations (Knight and Brooks, 2009), Molecular Mechanics-Poisson-Boltzmann Surface Area (MM-PBSA) (Gilson and Zhou, 2007), Linear Interaction Energy (LIE) (Gilson and Zhou, 2007), and hybrid quantum chemical/molecular mechanical (QM/MM) (Senn and Thiel, 2009) methods.

Free energy perturbation (FEP) is used to calculate free energy differences between two states from MD simulations. These two states can represent, for instance, unbound (apo) protein and a ligand-protein complex (holo), or two ligand-protein complexes with different ligands. In the framework of FEP, the difference in the free energy difference for two states is obtained from the Zwanzig equation (10).

$$\Delta G\left(A \rightarrow B\right)= G_B - G_A = -k_B T \ln \left\langle \exp\left(-\frac{E_B - E_A}{k_B T} \right) \right\rangle_A \qquad (10)$$

where A and B represent two states (e.g. apo and holo protein), G is the difference between free energies of both states, k_B is the Boltzmann's constant, T is the temperature, and the triangular brackets denote an average over a simulation run for state A.

FEP calculations converge properly only when the difference between these two states is small enough; therefore it is usually necessary to divide a perturbation into a series of smaller 'steps', which are calculated independently.

Thermodynamic integration (TI) is a related method to calculate free energy differences. Since the free energy can be expressed by its relation to the canonical partition function, the free energy difference in two different states can be used to calculate the difference of potential energy. TI calculations are usually executed by designing a thermodynamic cycle (Figure 1), and integrating along the relevant path. The path can be either a real chemical process or an artificial change (e.g. substitution of a methyl group by hydrogen atom).

The MD methods, despite their numerous successes, suffer of two major bottlenecks. One is the results are critically dependent on the force field used, therefore requires caution when use of appropriate force field and parameters. Many modern force fields are parametrised on experimental NMR data, some are able to include – to some extent – non-additive effects (electronic polarisation). Application of QM/MM schemes allow the inclusion of quantum effects to some extent. Another bottleneck is related to the adequacy of sampling. It is known that – due to relatively short time scales investigated – only some subsets of potential conformational changes can be observed, and often the system gets 'stuck' in a minimum, which does not have to be the global one. This makes the results heavily biased towards the starting structure and is very likely to underestimate the degree of molecular motions observed in the system. Prolonging the simulation time helps to solve the sampling problem only to some extent, and significantly increases the computational cost of MD simulations. Thus, in order to overcome the sampling issue, various enhanced sampling techniques have been employed. One of such methods, frequently used, is replica exchange MD (REMD), which attempts to overcome the problem of multiple-minima by exchanging temperatures of several replicas of the system. These replicas are non-interacting with each other and they run at different temperatures. REMD is also called "parallel tempering" (Earl and Deem, 2005). Another approach used to improve sampling is to construct the bias potential and add it to the potential energy function of the system (force field). This group of methods, referred to as umbrella sampling methods, consist of metadynamics (Laio and Gervasio, 2008), conformational flooding, and accelerated dynamics (Lange et al., 2006). The core feature of metadynamics is the construction of so-called reference potential, which is one that is the most similar to the actual potential. That is, repulsive markers are placed in a coarse time line in a space that is spanned by a small number of relevant collective variables. These markers are then placed on top of the underlying free energy landscape in order to push the system to rapidly accumulate in the initial basin by discouraging it from revisiting points in configurational space. In this way, the system is allowed to escape the lowest transition state as soon as the growing biasing potential and the underlying free energy well exactly counterbalance each other, effectively allowing the simulation to escape free energy minima.

3.2.2 Quantum mechanical (QM) calculations

Ligand-protein interactions can be driven by quantum effects. These include charge transfer, halogen bonds, or polarisation. Stabilisation energy related to charge transfer can be several kcal/mol and force field-based schemes cannot describe this stabilisation correctly.

'Conventional' QM calculations, using HF, DFT, or semi-empirical methods provide a way to obtain the ground state energy of a ligand-protein system or a part of it. Most programs based on these are capable of studying molecular properties such as atomic charges, multipole moments, vibrational frequencies, and spectroscopic constants. In addition, there are methods allowing the study of excited-state processes, such as time-dependent DFT or restricted open-shell Kohn-Sham (ROKS) (Li and Liu, 2010).

The application area of QM methods is vast. QM calculations are used for charge derivation for molecular dynamics simulations, for description of direct interactions (hydrogen bonds, halogen bonds, aromatic stacking), for calculations of pKa, protonation, redox states, and for studying solvation effects, such as computing free solvation energies.

Derivation of accurate charges for a system being studied is an important step in preparation for MD simulation. Failure in charge representation will inevitably lead to incorrect results. Derivation of charges is done using QM calculations, usually in several steps, involving optimisation, electrostatic potential generation, and fitting charges into atoms. RESP methodology, based on charges derived from *ab initio* HF/6-31G* level of theory has been for many years a standard in deriving charges for MD simulations (Bayly *et al.*, 1993, Cieplak *et al.*, 1995, Cornell *et al.*, 1993).

Charge distribution is also required for the calculation of the solvation properties using conductor-like screening model (COSMO) (Klamt and Schüürmann, 1993). COSMO, just like any other continuum solvent approach, approximates the solvent by a dielectric continuum, surrounding the solute molecules outside of a molecular cavity. In COSMO, the polarisation charges of the continuum, caused by the polarity of the solute molecule, is derived from a scaled-conductor approximation (hence the name). In this way, the charge distribution of the molecule, which can be obtained from the QM calculations, and the energy of the interaction between the solvent and the solute molecule can be determined.

QM calculations are also used for force field development, such as adding new parameters and incorporating non-additive effects. Several studies indicate that non-additive effects (.e.g. electronic polarisation) significantly affect binding affinities of many ligands (Ji and Zhang, 2008, 2009 and references therein). Electrostatic interactions are critically dependent on charge distribution around both interacting species, and this distribution is heavily dependent on the conformation (geometry) of the complex. Description of hydrogen bonding is also affected by electronic polarisation – some hydrogen bonds, which are found broken during MD simulation using 'conventional' force fields are found to be stable, when non-additive force field is used (Ji and Zhang, 2009). Corrections for polarisation can be added to MD force fields in order to derive protein charges more accurately and provide a better description of electrostatic interactions. Protein polarisation is important for stabilisation of the native structures of proteins. MD simulations indicate that inclusion of polarisation effects not only improves the description of protein native structures, but also distinguishes native from decoy dynamically: the former are more stable than the latter under the polarised force fields. These observations provide strong evidence that inclusion of polarisation effects in calculations of ligand-protein interactions is likely to greatly improve accuracy of such calculations.

QM methods are also used in molecular docking. The application of QM methods to molecular docking was pioneered by Raha and Merz (2004), who developed a semi-empirical QM-based scoring function and studied ion-mediated ligand binding processes. Their conclusion was that quantum chemical description is required for metal-containing systems, mainly because of poorly-defined atom types of metal atoms in most of the force field parameters, which cannot describe the interactions between a small molecule ligand and a metal ion in the active site of the protein.

Applicability of QM methods to study ligand-protein system has been discussed in literature (Raha *et al.*, 2007, Stewart, 2007). As well as these successes, many examples of the failure of QM approaches have been demonstrated. However, it should be kept in mind that most of these studies were based on either DFT, or semi-empirical Hamiltonians, which do not describe van der Waals interactions and hydrogen bonding terms of ligand-protein interactions correctly. This is, indeed, a serious limitation of "fast", hence more popular QM methods. A straightforward way to solve this problem is to add additional correction terms to the QM energy. It has been demonstrated that the addition of the dispersion energy and corrections for hydrogen bonds improved the performance of semi-empirical QM methods dramatically (Rezac *et al.*, 2009). The recently developed PM6-DH2 method (Fanfrlik *et al.*, 2010) yields, to date, the most accurate results for non-covalent interactions among the semi-empirical QM methods. For small non-covalent complexes, the results obtained were comparable to the high-level wave-function theory-based calculations within chemical accuracy (1 kcal/mol) (Rezac *et al.*, 2009).

Another major bottleneck of QM methods applied to studying ligand-protein interactions thermodynamics is the size of the system. DFT can handle up to 150 atoms, highly-accurate methods such as coupled-clusters can handle a few tens of atoms and require very fast computers and long computing times. This limitation of the size of the systems that can be studied seriously compromises its usage in the study of ligand-protein thermodynamics. For instance, the usage of the linear scaling algorithm MOZYME (Stewart, 2009) based on the localised orbitals allows size increases to systems as large as 18 000 atoms and above, which allows calculation of very large ligand-protein complexes. Due to these developments QM methods have become, therefore, very useful for fast and highly accurate predictions of ligand-protein interactions energetics.

An alternative approach is to use so-called divide-and-conquer (DIVCON) algorithm (Dixon and Merz, 1996, 1997). The principle is to divide a large system into many smaller subsystems, separately determine the electron density of each of these subsystems, and then to add the corresponding contributions from each subsystem in order to obtain the total electron density and the energy.

4. Examples

In the previous sections of this chapter, I briefly introduced the forces governing macromolecular associations and characterised methods commonly used to assess these contributions. Here, I will illustrate on several examples of 'real' ligand-protein systems and the way how their binding thermodynamics is studied.

4.1 Hydrophobic versus hydrophilic binding pocket: MUP and HBP

Both histamine-binding protein (HBP) and mouse major urinary protein (MUP) are members of lipocalin family of proteins, so their overall structures are similar (Figure 4). The

binding pocket of HBP contains a number of polar and charged residues, hence it is an example of a 'hydrophilic' binder. In contrast to HBP, the binding pocket of MUP is very 'hydrophobic'. Surprisingly, both HBP and MUP are characterised by similar overall entropy of ligand binding.

In our recent study (Syme *et al.*, 2010) we compared the driving forces for binding between these two proteins in terms of entropic contributions from ligand, protein, and solvent. We performed an extensive study combining x-ray crystallography, NMR spectroscopy, ITC, MD simulations, and QM calculations.

4.1.1 Structures of HBP and MUP

The structure of MUP was solved both by x-ray crystallography and solution NMR (Barratt *et al.*, 2005, Kuser *et al.*, 2001, Timm *et al.*, 2001). Several ligand-MUP complexes were studied, including long-chain alcohols, pyrazine derivatives, and pheromones as ligands. Regardless of the chemical nature of ligand, the protein structures are very similar to each other: the desolvated ligand, which occupies the central, hydrophobic binding pocket, causes very few conformational changes (Figure 4, left panel).

The crystal structure of HBP complexed with histamine revealed two binding sites for the ligand: one of them possessing considerably higher affinity than the other (Figure 4, right panel) (Syme *et al.*, 2010). Therefore, in order to simplify the thermodynamic analysis of ligand binding, a mutant of HBP was designed. In this mutant, denoted as HBP-D24R, negatively charged aspartic acid D24 inside the "low" affinity site was replaced by larger and positively charged arginine. This abolished binding of ligand to the "low" affinity site.
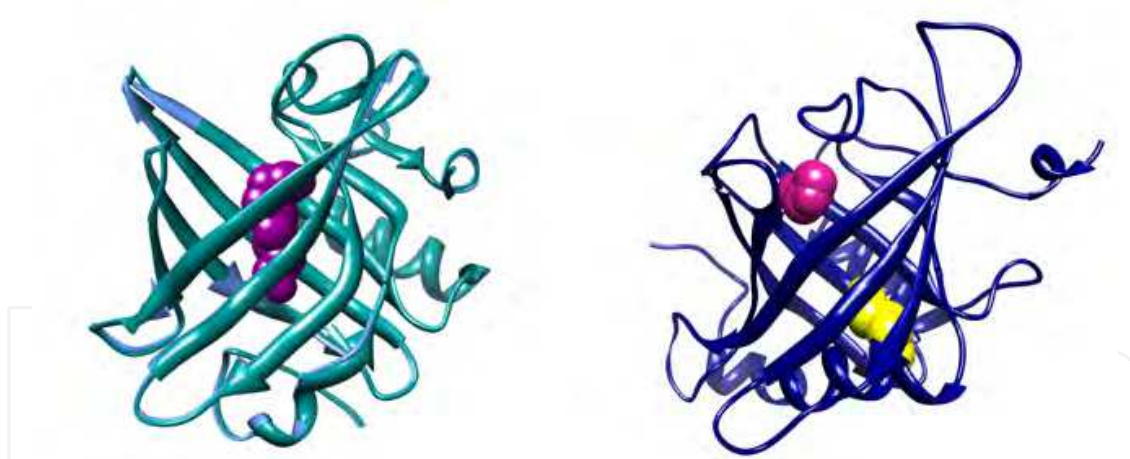
Fig. 4. Crystal structures of MUP (left panel) and HBP (right panel). Both MUP and HBP are displayed with their ligands bound: octanediol (purple), and histamine (pink). Ligands are represented as VDW spheres. For MUP, superimposed structures of apo (blue) and holo (dark cyan) protein are showed, in order to display a lack of major conformational changes associated with ligand binding. For HBP, the second (low affinity) binding site is showed and coloured yellow.

4.1.2 Calorimetric studies of MUP

Given that the binding pocket of MUP is very hydrophobic, an entropy-driven binding signature might have been expected for ligand-MUP interactions. Surprisingly, global

thermodynamics data obtained for pyrazine ligands (Barratt *et al.* 2004) and alcohols (Barratt *et al.*, 2006) showed that binding is driven by favourable enthalpic contributions, rather than the classical hydrophobic effect. The only hydrogen bond that could be formed between a ligand and the protein binding site involved the hydroxyl group of tyrosine Y120. Barratt *et al.* (2004) reported that ITC measurements on the binding of isobutyl-methoxypyrazine (IBMP) to the Y120F (phenylalanine side chain lacks hydroxyl group) mutant showed slightly reduced enthalpy of binding compared to wild-type MUP, but the binding was nonetheless enthalpy-driven.

Binding of long-chain alcohols, such as n-octanol, n-nonanol, and 1,8 octan-diol was characterised by similar thermodynamic signature. Contrary to expectations, binding was enthalpy-driven (Barratt *et al.*, 2006). Each complex was characterised by a bridging water molecule between the hydroxyl group of Y120 and the hydroxyl group of ligand. The thermodynamic penalty to binding derived from the unfavourable desolvation of 1,8 octan-diol (+21 kJ/mol with respect to n-octanol, which came from an additional hydroxyl group facing a hydrophobic pocket) was partially offset by a favourable intrinsic contribution.

Ligand		Temp (K)	$\Delta H°$ (kJ/mol)	$T\Delta S°$ (kJ/mol)	$\Delta G°$ (kJ/mol)	K_d (nM)
HBP(D24R)						
		278	-42.7[a] ± 0.9[b]	6.5 ± 1	-49.2 ± 0.6	2.4 ± 0.6
		288	-48.3 ± 3.1	0.4 ± 4.5	-48.8 ± 1.1	2.9 ± 1.3
		298	-58.3 ± 1.2	-9.3 ± 2.5	-49.1 ± 1.4	2.5 ± 1.4
HBP(WT)						
	HA	298	-70.9 ± 7.1	-24.4 ± 7.3	-46.5 ± 1.3	7.0 ± 3.6
	LA	298	-53.2 ± 3.3	-12.0 ± 3.6	-41.2 ± 1.5	59.6 ± 34.8
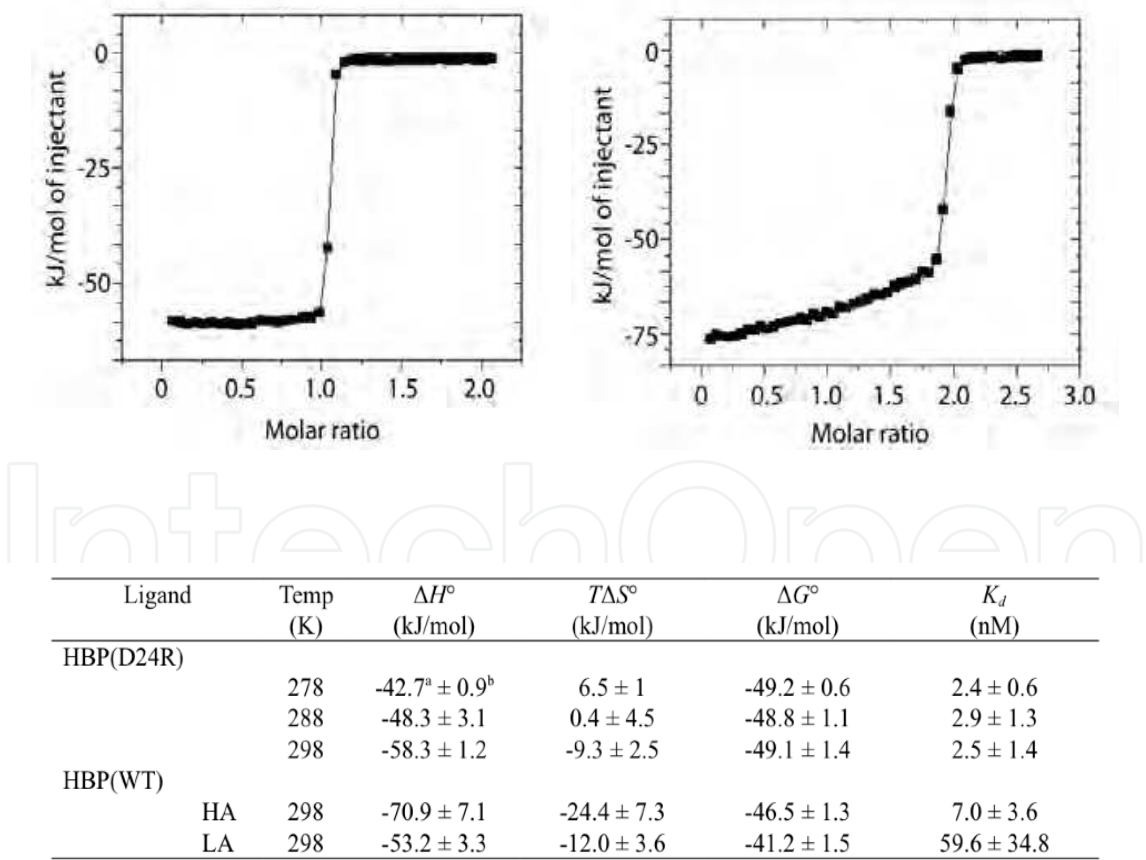
Fig. 5. ITC data for obtained for HBP. Binding curves for HBPD24R mutant and wild-type BP are displayed in left and right panel, respectively. Bottom panel shows thermodynamic parameters for mutated and wild-type HBP, obtained from ITC measurements.

4.1.3 Calorimetric studies of HBP

Typical ITC isotherms for the binding of histamine to the HBP-D24R mutant (top panel) and wild-type HBP (middle panel), and the resulting thermodynamic parameters (bottom panel) are shown in Figure 5. Consistent with structural data, histamine bound to wild-type HBP with 2:1 stoichiometry (both binding sites occupied), while to mutant with 1:1 stoichiometry. In both cases, histamine bound with high affinity (Kd in nanomolar range), but the binding enthalpies and entropies were somewhat different, even though binding was largely enthalpy-driven in both cases. These differences in thermodynamic details are a manifestation of the enthalpy-entropy compensation, introduced in the previous sections of this chapter: wild-type HBP binds histamine with similar affinity than D24R mutant, but its enthalpic contribution is more favourable than that of D24R, at the expense of the entropic contribution, which is less favourable in wild-type HBP than in D24R mutant.

A major concern regarding the data shown in Figure 5 was the possibility of proton exchange (release or binding) during the binding event. In such case, the observed enthalpy change ΔH_b would contain contributions from the ionisation of the buffer (ΔH_{ion}), as explained previously. To assess this effect, titrations between histamine and HBP-D24R mutant were performed in two different buffers, characterised by very different values of the ionisation enthalpy (ΔH_{ion} 47.45 kJ/mol and 3.6 kJ/mol, respectively). Thus, any contribution from proton exchange should be easily detectable on the basis of the differences in ΔH_b for the histamine-HBP interaction. Obtained enthalpies (-58 and -61 kJ/mol, respectively) showed that there were no significant protonation effects associated with histamine-HBPD24R binding.

4.1.4 NMR relaxation measurements

In order to gain deeper insight into the entropic contribution to binding of ligands to HBP and MUP, ^{15}N NMR relaxation measurements were employed to probe per-residue conformational entropies for backbone amides for the free (apo) protein and for the ligand-protein complexes. Backbone 15N longitudinal and transverse relaxation rates ($R1 = 1/T1$ and $R2 = 1/T2$, respectively) were determined for the free protein and the complexes with ligands. Amide ^{15}N and [1]H-^{15}N resonance assignments in the apo-HBP, apo-MUP, and the ligand-protein complexes were determined by use of conventional three-dimensional triple-resonance experiments.
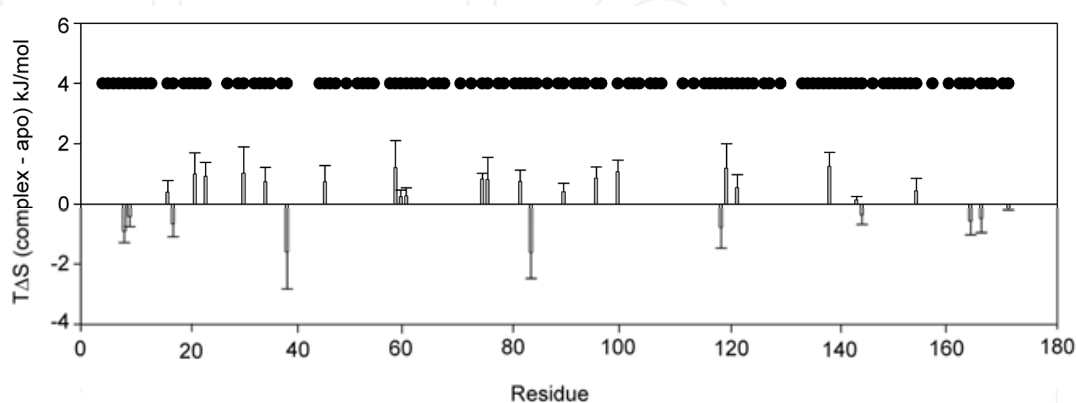

Fig. 6. Backbone amide entropy changes for HBP-D24R, quantified as differences between entropy of holo and apo protein for each protein residue. The plot shows histamine binding – induced changes in entropy assessed by 15N relaxation measurements.

For both HBP and MUP, both positive and negative changes in local backbone entropy were observed. Entropy changes were not restricted to the binding pocket but were dispersed over the protein (Figure 6). For HBP as well as MUP, $T\Delta S$ summed over backbone amides was close to the error value, i.e., an overall change in backbone entropy that was not statistically different from zero (Syme *et al.*, 2010).

4.1.5 Molecular dynamics (MD) simulations

MD simulations were used to examine the solvation of the binding pocket of MUP and HBP. The MUP binding pocket was found virtually devoid of water, even without any ligand bound (Barratt *et al.*, 2006). Even if water molecules were artificially "forced in" at the beginning of the MD simulation, they did not remain inside the pocket. This observation contributed to the explanation of the unexpected thermodynamic signature of the ligand binding to MUP, measured by ITC.

An average of five to six water molecules is observed in the binding pocket of HBP over the simulation time in the absence of ligand. Analysis of the diffusion and rotational correlation functions of these solvent molecules suggested that their dynamic behaviour was very similar to those of bulk water (Syme *et al.*, 2010), which suggests that the return of these water molecules to bulk solution on ligand binding would not offer any significant contribution to the binding entropy.

For HBP, side chain entropies were computed from the MD simulation. Moreover, as a test of the robustness of these simulations, backbone amide entropy changes were calculated from the MD simulation and compared to NMR data. The comparison was very favourable (16.4 ± 1.0 kJ/mol versus 12.4 ± 9.8 kJ/mol), which raised confidence in the applied methodology. Contribution from the protein side chains was estimated at +17.4± 1.8 kJ/mol (Syme *et al.*, 2010). Taken together, these data strongly indicate an overall increase in entropy on ligand binding. This observation, although counter-intuitive, it is not without precedent in the literature (MacRaild *et al.*, 2007, Stoeckmann *et al.*, 2008).

MD simulations were also used to estimate the entropic contributions from the ligand. A contribution from the loss in vibrational degrees of freedom of histamine on binding to HBP was estimated using the Schlitter's method (Schlitter, 1993), leading to an unfavorable contribution of $\approx -22 \pm 2.4$ kJ/mol.

In addition to this contribution, there was an assumption that internal degrees of freedom of the ligand are heavily constrained upon binding. The unfavorable contribution from the three relevant internal degrees of freedom of histamine amounted to $\approx$ -12 kJ/mol (Lundquist *et al.*, 2000). In addition, the entropic contribution from the loss of translational and rotational degrees of freedom of the ligand depends on the logarithm of the molecular mass, and on the basis of earlier work this represents an unfavorable contribution that can be estimated as $\approx$ -25 kJ/mol (Turnbull *et al.*, 2004).

4.1.6 Solvation thermodynamics estimation

For MUP ligands, it was possible to measure their solvation free energies directly, using the water/vapor partitioning experiments (Shimokhina *et al.*, 2006). For histamine, experimental measurements could not be done due to non-volatility of histamine, and hence the ligand free solvation energies were calculated quantum-chemically, using the COSMO model (Klamt and Schüürmann, 1993). Prior to running calculations, the ligand was optimised using *ab initio* QM calculations (RFH/6-31G* basis set). The optimised structure was

subjected to COSMO calculations at three different temperature settings (270, 300, and 330 K) in order to extract the enthalpic and entropic contributions to the free energy of solvation, using the finite-difference approach. The approximation used therein was based on the assumption that the heat capacity is constant over a certain range of temperatures near the target temperature, T. In the case of solute molecules solvated in water, this approximation usually holds near room temperature for temperature ranges (denoted as ΔT) as wide as 50 K. Using the finite-difference approximation, the entropy can be approximated at the target temperature as in equation (11):

$$\Delta S(T) = -\left(\frac{\Delta G(T + \Delta T) - \Delta G(T - \Delta T)}{2\Delta T} \right) \tag{11}$$

where $\Delta S(T)$ denotes entropy at the target temperature, $\Delta G(T)$ is the free energy of solvation energy, and ΔT is the temperature difference. For the calculations presented here ΔT was 30 K.

Obviously, it was very difficult to assess the accuracy of such calculations, but the experimental solvation thermodynamics of two related "fragments" of histamine have been reported (Cabani et al., 1981). Comparison (Table 1) shows that the solvation free energies are reproduced very well, and the solvation entropies and enthalpies reasonably well (for n-propylamine) compared with experiment, which lends some confidence in the computed values for histamine.

Ligand	ΔG^0 (kJ/mol)	ΔH^0 (kJ/mol)	$T\Delta S^0$ (kJ/mol)
Histamine	-70.1	-137.2	-67.1
n-propylamine	-16.5	-65.5	-49
n-propylamine (exp)	-18.4	-55.8	-37.4
2-methyl imidazole	-43	-96.3	-53.3
2-methyl imidazole (exp)	-42.9	n.a.	n.a.

Table 1. Solvation parameters: free energy, enthalpy, and entropy at temperature 300 K for histamine and related ligands. All values are given in kJ/mol. Experimental values (exp) are taken from Cabani et al. (1981).

4.1.7 Driving forces for ligand binding by MUP and HBP

Ligand binding to both HBP and MUP was strongly enthalpy-driven. The overall entropy of binding was the same within error for both MUP and HBP, yet the contributions from protein, ligand, and solvent are very different.

In the case of HBP, the dominant entropic contribution to binding arises from ligand desolvation, with a significant contribution from protein degrees of freedom. This contribution is favourable. However, the overall entropic contribution to binding is unfavourable, which indicates that the entropic contribution from desolvation of the protein binding pocket is strongly unfavourable.

In MUP, a favourable contribution to binding entropy is also derived from ligand desolvation, but is unable to overcome the unfavourable contribution from "freezing"

ligand degrees of freedom upon binding. The favourable entropic contribution from desolvation of the protein binding pocket that one would predict in a "classical" hydrophobic interaction is absent in MUP, since the occluded binding pocket is substantially desolvated prior to binding. This phenomenon has subsequently been observed in other proteins, such as streptavidin and HIV-protease receptors (Young et al., 2007).

As suggested by MD simulations, there are around 6 water molecules (on average) occupying the binding site of HBP prior to ligand binding. As already stated, no significant change in entropic contribution should arise from displacing these waters upon ligand binding, since their dynamic behaviour inside the pocket is similar to the behaviour of the bulk water. However, four water molecules are sequestered in the binding pocket in the histamine-HBP complex. These waters are significantly more ordered than the bulk water (Syme et al., 2010), which contributes negatively to the entropic term of binding free energy. This unfavourable entropic contribution can be estimated as about −30 to −40 kJ/mol, which is qualitatively consistent with the observed sign of ΔC_P on binding.

In conclusion, in the case of HBP, we found favourable entropic contributions to binding from desolvation of the ligand. However, the overall entropy of binding was unfavourable due to a dominant unfavourable contribution arising from the loss of ligand degrees of freedom, together with the sequestration of solvent water molecules into the binding pocket in the complex. This can be contrasted with MUP, where desolvation of the protein binding pocket made a minor contribution to the overall entropy of binding given that the pocket is substantially desolvated prior to binding.

4.2 An integrated picture of galactose binding to Arabinose Binding Protein (ABP)

The arabinose binding protein is found in the periplasm of Gram-negative bacteria. It belongs to the family of proteins, which extert their biological function as components of osmotic shock-sensitive transport systems for sugars and amino acids (Vyas et al., 1991). Besides its biological context, ABP is a very well-defined model system for structure-activity relationships in the hydrophilic ligand binding systems. As demonstrated by ITC, ABP interacts with its natural ligands, namely L-arabinose and D-galactose, and their deoxy derivatives (Daranas et al., 2004). The interactions are enthalpy-driven. The galactose-ABP interactions served as a model for interactions between hydrophilic ligands and hydrophilic binding pockets. Such choice of the model system was made mainly because of the large unfavourable entropic contribution to binding, whose origin was difficult to understand in the framework of the current ligand-protein interaction paradigms. Such a large entropy change upon ligand binding is frequently observed in proteins interacting with carbohydrate ligands, and it contributes to the fact that these interactions are notoriously challenging for predictions and design. In order to address those issues, we employed a combination of solution NMR and MD simulations.

4.2.1 NMR measurements of ABP

An essential prerequisite to NMR studies of protein dynamics was the establishment of the ^{1}H, ^{15}N and ^{13}C resonances that contributed to the investigated spectra. These assignments have been obtained for the complex of ABP with the ligand D-galactose using conventional triple-resonance assignment strategies (Daranas et al, 2004). Assignments for the unbound (apo) ABP were determined from these results, using an approach that combined conventional triple-resonance assignment strategies and [1]H-^{15}N heteronuclear single

quantum coherence (HSQC) titrations of ABP with 1-deoxy-galactose, which is a fast-exchanging ligand (MacRaild *et al.*, 2007).

From these assignments, a comparison was made of the chemical shifts of the backbone amide resonances of ABP in the unbound state and in the complex (Figure 7). Large chemical shift changes were observed in the binding site area and in the region linking the two domains of ABP. This suggested that ligand binding might be associated with a substantial conformational change in ABP (Figure 8). Such conformational change (domain reorientation) upon binding is observed in other members of the periplasmic-binding protein family, and have been proposed for ABP on the basis of the results of small-angle X-ray scattering and theoretical studies (Mao *et al.*, 1982, Newcomer *et al.*, 1981). Smaller changes in chemical shift were also observed at sites distal to the binding site, which suggests that small conformational changes, resulting from protein dynamic behaviour, occur upon the binding event.
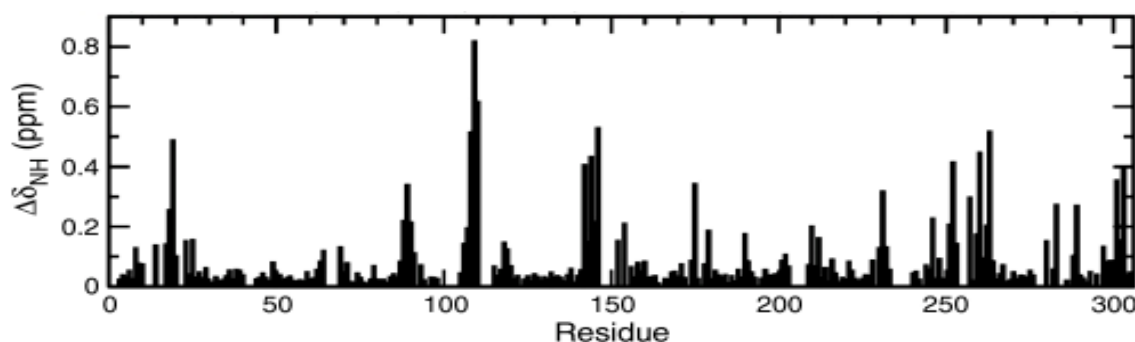
Fig. 7. Chemical shifts for galactose binding to ABP. Changes in backbone amide chemical shifts are plotted against protein residue number.
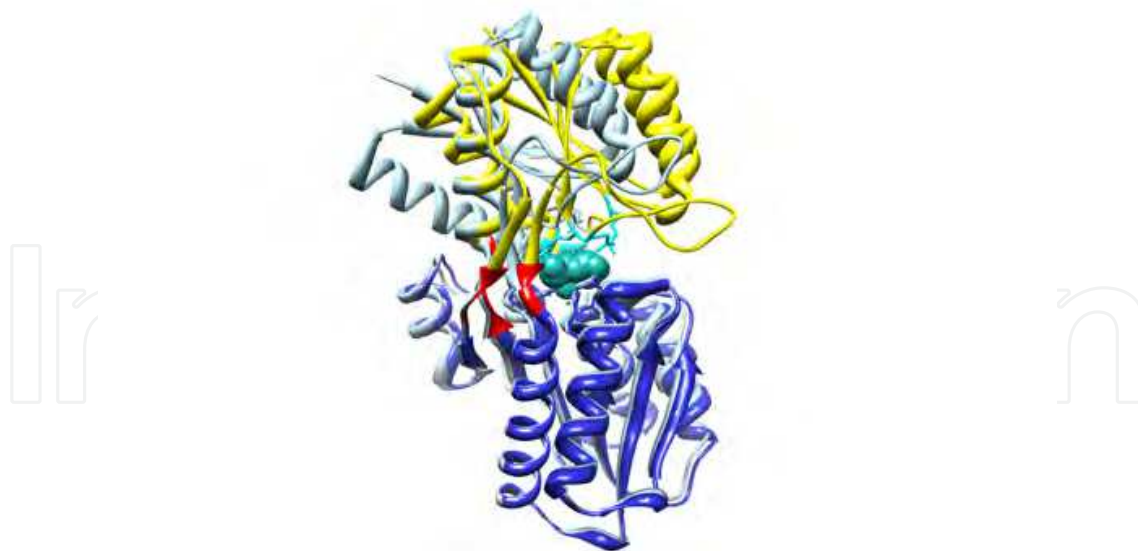
Fig. 8. Conformational changes between apo-ABP (light blue) and galactose-ABP complex. The apo protein is much more 'open' than the complex. Large conformational changes are observed in the hinge region (coloured red), and the reorientation of N-domain (yellow) and C-domain (dark blue) towards each other is quite pronounced. Bound galactose is coloured cyan and displayed as VDW spheres. Side chains of several residues involved in direct interactions with ABP are showed and coloured dark cyan.

Full sets of ^{15}N relaxation measurements were made for ABP in its apo form and in complex with D-galactose, L-arabinose and D-fucose. The analysis of these measurements, assessed by means of Lipari–Szabo model-free approach (1982), allowed for extraction of the information on the extent of 'fast' (ps-ns time scale) motion of the protein backbone and 15N-containing side chains. Differences in Lipari-Szabo generalised order parameters (S^2) between apo- and holo-ABP and were interpreted in terms of changes in dynamics accompanying the binding event. Thus, observed dynamic changes could be related to binding thermodynamics by means of the relationship between changes in order parameters derived from NMR relaxation and changes in conformational entropy. Because S^2 parameters are specific to individual bond vectors, the described approach offered an unprecedented degree of structural resolution in thermodynamic analysis of protein function.

Surprisingly, generalised order parameters for apo ABP were, in general, larger than for the ABP–galactose complex. This suggests that 'fast' (pico- to nanosecond time scale) motions are more extensive in the ligand-protein complex than in the unbound protein.

4.2.2 Molecular dynamics (MD) study

To confirm and further explore NMR relaxation result, we performed MD simulations of ABP in complex with galactose and its unbound state. In order to make a direct comparison between the observed order parameters and the simulations, backbone amide order parameters were calculated from the MD trajectory (MacRaild et al., 2007). A very good agreement between measured and calculated order parameters was observed, albeit with a small tendency for the simulation to underestimate the experimental values. The calculated changes in order parameter upon ligand binding reproduced the changes measured by NMR excellently, showing an approximately uniform decrease in order parameter (and, hence, increase of dynamics) upon galactose binding across the protein.

To understand the thermodynamic implications of the observed changes in dynamics, we employed the relationship between Lipari–Szabo order parameters (S^2) and conformational entropy derived by Yang and Kay (1996).

In addition to generalised order parameters, we observed significant increase in backbone dynamics in the complex as compared with the apo protein by several other measures. RMS deviations from the average structure were significantly larger for the galactose-ABP complex than for apo-ABP. Fluctuations of backbone heavy-atom positions across the trajectory were generally larger in the complex than in the apo protein.

As well as validating the experimental results, the MD simulations revealed details of dynamic changes in regions that could not be measured experimentally. Importantly, the simulations revealed complex changes in the dynamics of the ABP binding site: counter-intuitively, several residues in the binding site showed increase in flexibility upon binding, which was consistent with the trend seen throughout the rest of the protein. Other binding site residues displayed a decrease in flexibility, more in keeping with the intuitive expectation that ligand binding will reduce the conformational freedom of binding site residues (Figure 9).

The total entropy alters due to changes in the pico- to nanosecond motion upon galactose binding, estimated from MD simulations and NMR results, was 610(± 120) J/mol K, which gives $T\Delta S = 188(\pm 37)$ kJ/mol at 308 K. Clearly, this latter value is overestimated, as the

assumption of un-correlated motion is not likely to hold for all residues of the protein. It is evident, however, that the entropy change associated with changes in 'fast' dynamics contributes favourably to the free energy of binding. As the result, it allows the reduction of the unfavourable entropic contribution associated with ligand binding.
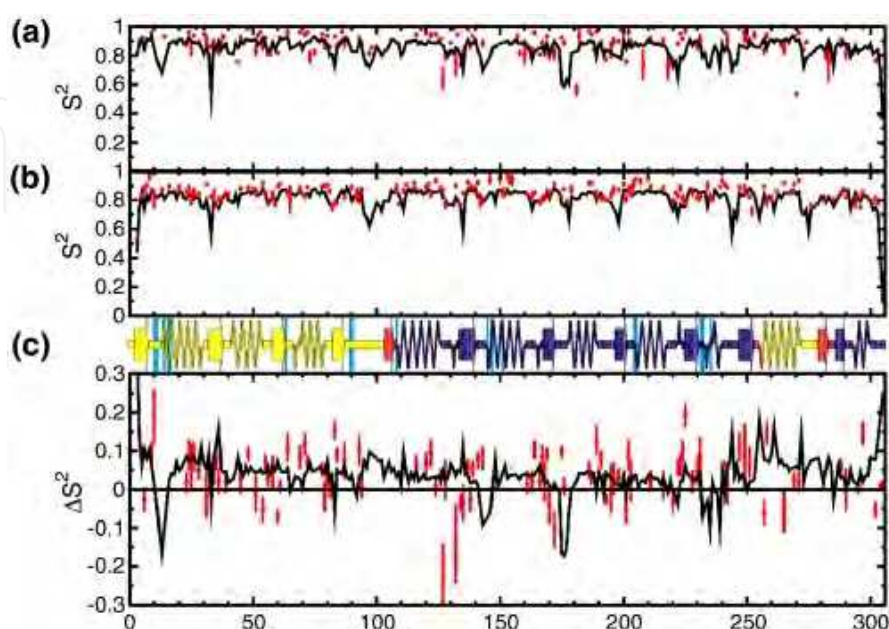

Fig. 9. Backbone amide order parameters for apo-ABP and galactose-ABP complex. S^2 order parameters were obtained by NMR relaxation measurements (red dots) and MD simulations (black lines). Data obtained for apo-ABP are showed in panel (a), while results for galactose-ABP complex are displayed in panel (b). Panel (c) shows changes in order parameters induced by galactose binding to ABP, with protein secondary structure elements displayed above. The N-domain of ABP is coloured is yellow, the C-domain is coloured blue, and the hinge region is coloured red. Residues interacting with the ligand are coloured cyan. This colouring scheme is consistent with colouring in Figure 10. (figure taken from MacRaild *et al.*, 2007)

4.2.3 Origins of entropic costs of binding

We investigated the origin of the large and unfavourable entropic contribution to the binding free energy of galactose-ABP, observed by ITC (Daranas *et al.*, 2004), in terms of the different contributions.

It is clear that formation of a ligand-protein complex will involve the loss of entropy associated with constraining the translational and rotational degrees of freedom of one binding partner with respect to the other. The magnitude of these unfavourable contributions to the ligand-protein interaction can be approximated. In the case of galactose-ABP interactions, we took an estimate of the loss of ligand translational and rotational entropy from the work by Turnbull *et al* (2004) and Lundquist and his coworkers (Lundquist and Toone, 2002), which gave at approximately 25 kJ/mol for the free energy penalty.

It is assumed that the bound ligand will experience a loss of entropy reflecting the loss of conformational flexibility of the ligand in solution. On the assumption that conformational degrees of freedom are substantially restricted upon ligand-protein binding, the entropic

penalty arising from loss in degrees of freedom of the galactose hydroxyl rotors alone is likely to be ~30 kJ/mol (Lundquist and Toone, 2002). Another contribution arises from the solvation effects. As calculated by means of QM/COSMO approach, desolvation energy of galactose is +87.6 kJ/mol (at 300 K), which is a significant unfavourable contribution (Bronowska and Homans, unpublished data).

It is known that the ABP binding site contains a significant number of tightly bound water molecules, which maintain the structure of the binding site and play a role in governing the specificity of ligand binding (Quiocho, 1993). Examination of the structures of ABP in complex with its ligands (galactose, fucose, and arabinose) revealed some 15 crystallographically resolved and structurally conserved water molecules within the binding site (MacRaild *et al.*, 2007). Dunitz (1994) estimated the maximal entropic cost of confinement of a single water molecule to 8 kJ/mol. Even if the cost of confinement of the water molecules in the binding site of ABP is lower than this maximal value, the overall cost of confining these water molecules in the binding site will be vast.

All these amount to unfavourable entropic contribution of galactose-ABP binding. However, the entropy of galactose-ABP interactions is much lower than the sum of these contributions: as measured by ITC the $T\Delta S^0$ amounts to –61 kJ/mol at 308 K (Daranas *et al.*, 2004). Observed discrepancy can be explained in terms of favourable contribution of protein dynamics to the entropy of galactose-ABP interactions, which was observed by NMR measurements and MD simulations.

4.3 Enthalpy-entropy compensation revisited: bovine carbonic anhydrase II

In studies on the thermodynamics of ligand-protein interactions, it is usually assumed that the bound ligand is fixed in the binding site. However, there is little direct experimental evidence for this assumption, and in the case of binding of p-substituted benzenesulfonamide inhibitors to bovine carbonic anhydrase II (BCA II), the observed thermodynamic binding signature assessed by ITC measurements leads indirectly to the conclusion that the bound ligands retain a considerable degree of flexibility (Krishnamurthy *et al.*, 2006).

BCAII and its binding to a large panel of ligands was reported in literature as a classic example of enthalpy-entropy compensation. Whitesides and his coworkers studied these interactions for a series of *p*-(glycine)*n*-substituted benzenesulfonamides (where n = 1-5) and found almost perfect enthalpy-entropy compensation across the series: as demonstrated by ITC measurements, the enthalpy of binding became less favorable and the entropy more favorable with increasing chain length. Changes in heat capacity were independent of chain length, which indicated that the observed changes in binding thermodynamic signatures across the series cannot be explained on the basis of the classical hydrophobic effect. In addition, strong evidence exists that these thermodynamic signatures are not driven by solvent effects (Syme *et al.*, 2010). To explain the observed data, a model was proposed, which assumed the increased mobility of ligand upon the chain length growth. Such increased flexibility of the bound ligand (favourable entropic contribution) would be compensated by a decreased number of direct ligand-protein contacts (unfavourable enthalpic contribution).

Such an increase in mobility of the ligand can be readily probed by ^{15}N NMR relaxation measurements and computational studies. Thus, we investigated these using a combination of the solution NMR and MD simulations. Two series of ligands were studied: The first

(series 1) consisted of six ligands with different chain lengths (n = 1-6), isotopically [15]N -labeled at the terminal amide, whereas the second (series 2) comprised six ligands with the same chain length (n = 6), but isotopically [15]N -labeled at a single amide at each position n. The ligands bound at the BCAII binding site are shown in Figure 10.

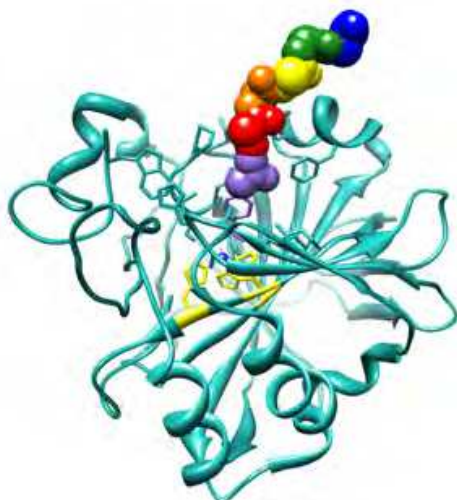Fig. 10. The binding of investigated benzenesulfonamides to BCAII. The protein backbone is coloured cyan, with three histidines (yellow) coordinating zinc atom (dark blue dot) displayed. The ligand is showed, with glycine side chains represented as VDW spheres and coloured as follows: n=1 – purple, n=2 – red, n=3 – orange, n=4 – bright yellow, n=5 – dark green, and n=6 – blue.

Contrary to expectation, we found that the observed thermodynamic binding signature could not be explained by residual ligand dynamics in the bound state, but rather results from the indirect influence of ligand chain length on protein dynamics.

Chemical shift changes on ligand binding were monitored by simple one-dimensional ^{1}H, ^{15}N HSQC spectra. It was possible, since each ligand was selectively [15]N -labeled. The results (Figure 11) indicated that residues proximal to the aromatic moiety of ligand show substantial changes in chemical shift, whereas residues more distant to the aromatic moiety show much smaller changes. The former is indicative of interactions with the protein, while the latter suggests less substantive interactions with the protein by these more distant residues.
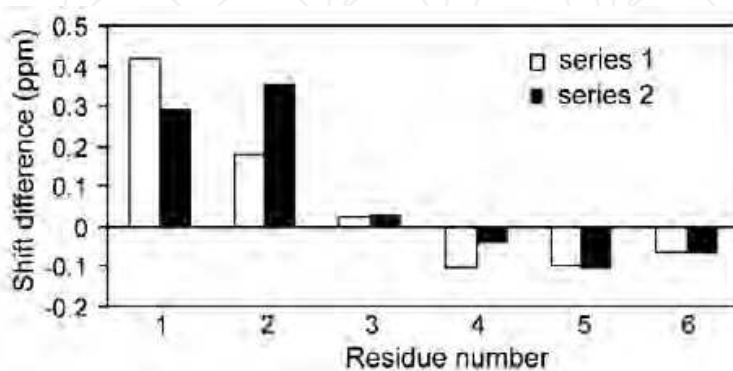
Fig. 11. Observed chemical shift differences between free ligands and ligand-BCAII complexes, for both series of ligands, plotted against the number of glycine residues (n from 1 to 6) in the side chain (Stoeckmann *et al.*, 2008).

In order to obtain a more detailed picture of ligand dynamics in the bound state, ^{15}N NMR relaxation data ($R1$, $R2$, NOE) were measured for each series of ligands. The results obtained showed that the ligand chain became more dynamic as n increased. However, there were substantial differences between series 1 and 2 ligands: series 1 ligands were much more dynamic for small n values, while in case of series 2 ligands, the three residues nearest the aromatic ring adopted slow dynamic motions, whereas the three residues distal to the aromatic ring adopted faster dynamics similar to series 1 ligands.

The relaxation data were analysed using the Lipari-Szabo model-free approach (1982) and the 'fast' dynamics was quantified by means of generalised order parameters. The entropic contributions arising from these 'fast' motions were assessed from S^2 parameters using the relation derived by Yang and Kay (1996).

The order parameters of all ligand Gly residues in both series were much smaller than those of backbone residues within protein, indicating a comparatively high mobility of the ligand chain. The first two residues of the ligand were relatively immobile adjacent to the aromatic ring and are engaged in direct interactions with the protein, whereas the last four Gly residues were significantly more mobile, with S^2 values indicating motions being unrestricted by BCAII protein. From S^2 values, the entropic contribution to ligand-BCAII binding arising from each Gly residue, could be obtained. These data are shown in Figure 12.
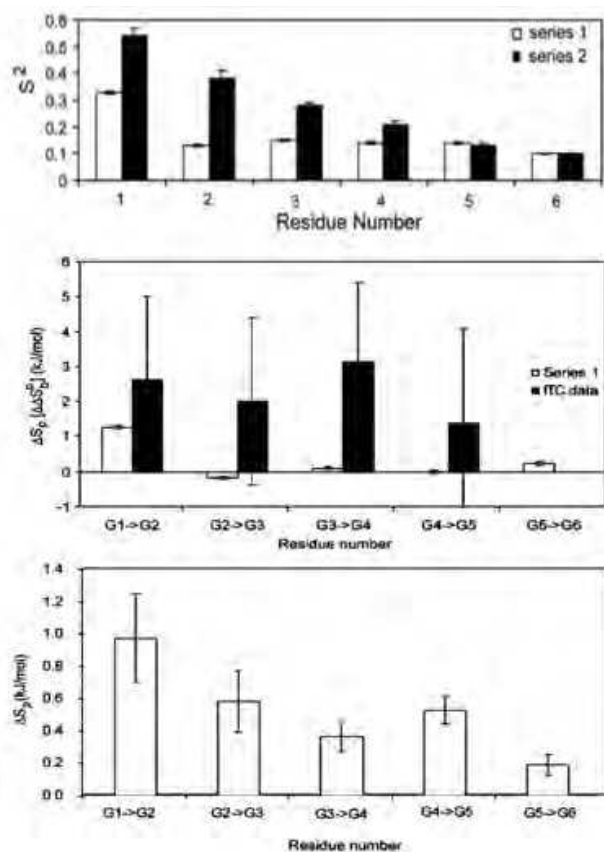
Fig. 12. Top panel: order parameters for backbone amide (N-H) bond vectors for ligand series 1 (white) and 2 (black), obtained from NMR relaxation measurements. Middle panel: Derived entropy differences for series 1 ligands (white) compared to entropies of binding obtained from ITC measurements (black) (Krishnamurthy *et al.*, 2006). Bottom panel: Entropy differences for series 2 ligands (per N-H bond vector), obtained by NMR.

Comparing these results with ITC data by Krishnamurthy *et al*. (2006), it is clear that a poor correlation exists between the change in ligand conformational entropy determined from NMR relaxation studies and the entropies of binding derived from ITC (Figure 14, middle panel). It indicates that a model based on increased dynamics of the ligand in the bound state is not a plausible explanation for the observed thermodynamic binding data. This is not entirely unexpected since the ITC values are global parameters, which include contributions not only from the ligand, but from protein and solvent as well. However, the role of solvation is unlikely to be the driving one in the case of ligand-BCAII binding – for three reasons. First, ΔC_p values for the interaction determined by ITC are independent of Gly chain length (Stoeckmann *et al*., 2008). Second, these values are fairly small: around 80 J/mol/K. Finally, ligands are not fully desolvated upon the binding event: more distal residues extend beyond the binding pocket and they interact with water molecules. The observed increase in entropy with respect to the ligand chain length is approximately linear, which argues against a significant solvation contribution.

It was hoped that assessment of the protein contribution would shed light on the observed binding signature. To achieve this, MD simulations of both series of ligands in complexes with BCAII were performed (Stoeckmann *et al*., 2008). In order to validate the methodology, generalised order parameters for ligand amide vectors were calculated from the trajectory and compared to NMR data. These MD trajectories were then used to probe the influence of ligand binding on protein dynamics. Specifically, S^2 values for backbone amide bond vectors, side chain terminal heavy-atom bond vectors, and corresponding conformational entropies were calculated for each complex with series 1 ligands.

The results obtained showed that the aromatic moiety became correspondingly more rigid with respect to series 1 ligand chain length. This was consistent with the NMR data showing that addition of successive glycine residues decreased the dynamics of the preceding units. Moreover, we observed the trend of increased dynamics of protein residue side chains with respect to ligand chain length (Table 2). This counter-intuitive observation that ligand binding increases protein dynamics has been observed in a number of ligand-protein systems, including ABP, which was described in the previous section of this chapter.

Residues	Gly2-Gly1	Gly3-Gly2	Gly4-Gly3	Gly5-Gly4	Gly6-Gly5
Biding site	4.37 ± 1.1	5.28 ± 1.2	4.33 ± 1.0	3.11 ± 1.0	6.04 ± 1.3
Whole protein	14.9 ± 1.7	4.6 ± 1.8	5.5 ± 2.2	9.9 ± 2.4	8.4 ± 2.5

Table 2. Differences in per-residue entropies quantified as TDS (in kJ/mol at temperature 300 K) for residues in the binding pocket of BCAII as well as for the whole BCAII protein. Displayed differences are result of changing side chain length of the ligand (Gly$_n$ – Gly$_{n-1}$).

Summarising, our results suggest that the enthalpy-entropy compensation observed for binding of ArGly$n$O- ligands to BCA II derives principally from an increase in protein dynamics, rather than ligand dynamics, with respect to the ligand chain length. Krishnamurthy and his coworkers showed that enthalpy-entropy compensation was observed for a range of BCAII ligands, whose structurally distinct chain types gave similar thermodynamic signatures (Krishnamurthy *et al*., 2006). This suggests that a common process is underway that is unlikely to be related to specific interactions between the chain

and the protein. In our study, we demonstrated an increase in protein dynamics upon binding longer-chained ligands. This observation provides an explanation for the enthalpy-entropy compensation across these structurally distinct ligands.

5. Conclusions

The notion of the binding event being the result of shape complementarity between ligand and protein binding site (key-and-lock model) has been a paradigm in the description of binding events and molecular recognition phenomena for a long time. The recent discovery of the important role played by protein dynamics and solvent effects, as well as the enthalpy-entropy compensation phenomenon, challenged this concept, and demanded the thorough examination of entropic contributions and solvent effects. Assessment of all these contributions to the thermodynamics of ligand-protein binding is a challenging task. Although understanding the role of each contribution and methods allowing for a complete dissection of thermodynamic contributions are tasks far from being completed, significant progress has been made in recent years. For instance, development of high-resolution heteronuclear NMR methods allowed for assessment of the contribution from protein degrees of freedom to the intrinsic entropy of binding. The usefulness of such approach has been demonstrated in the course of this chapter on several ligand-protein examples. In addition, progresses in the development of MD-related methodologies and advanced force fields enabled the application of the NMR-derived formalism on relevant time scales and the assessment of the intrinsic entropic contributions solely using computational methods. Development of QM methods allows the study of larger and larger systems, while advances in ITC calorimetry allow the use of very small amounts of reagents for a single experiment.

Despite this progress, much remains to be done. The enthalpy-entropy compensation phenomenon seems to be widespread among ligand-protein systems. It seems universal: binding restricts motions, while motions oppose tight confinement. However, our current knowledge about intrinsic protein dynamics is still insufficient to allow us to predict this phenomenon and hence to exploit it for the purposes of rational molecular design. Another challenge lies within the quantification of solvation contributions. There seem to be conflicting data regarding the contributions from confined water molecules. Their influence on binding can be favourable or unfavourable, enthalpy- or entropy- driven. Bound water molecules can be released upon ligand binding or – on the contrary – bind tighter (Poornima CS and Dean, 1995a-c). Their presence can make the protein structure more rigid (Mao *et al.*, 2000), or more flexible (Fischer and Verma, 1999). Finally, protein binding sites can be fully solvated prior to binding, or fully desolvated (Barratt *et al.*, 2006, Syme *et al.*, 2010). The only common feature that seems to exist is that the contribution of the solvation effects to the ligand-protein binding thermodynamics can be – and often is – significant.

Last but not least, intrinsic entropic contributions are notoriously difficult to quantify. A handful of experimental and theoretical methods can be employed to quantify these contributions, as have been described. However, all of these methods have their limitations, and one should be aware of these and of the assumptions that are being made. Theoretical results should be treated with caution, experimental data likewise, as they are based on many approximations and heavily dependent on the conditions applied. Care must be taken not to over-extrapolate data, and not fall the victim to confirmation bias.

Fundamentally, in order to predict free energy of binding accurately, it would be necessary to go beyond predicting a single 'dominant' conformation of the ligand-protein complex. It

should be emphasised that the overall shape of the free energy landscape controls the binding free energy. This shape is affected by the depth and width of the local minima, and the height and breadth of the energy barriers. The factors that shape that landscape include intrinsic entropic contributions of both interacting partners, ligand poses, protein conformations, solvent effects, and protonation states. Computational and experimental approaches combined together can provide insight into this crucial but otherwise hidden landscape, which is pivotal not only to understand the origin of each contribution and its role in the binding event, but which can allow a truly rational molecular design.

6. Acknowledgements

7. References

Alves, I.D.; Park, C.K.; Hruby, V.J. (2005). Plasmon resonance methods in GPCR signaling and other membrane events. Curr Protein Pept Sci. 6(4), pp. 293-312.

Axilrod, B.M.; & Teller, E. (1943). Interaction of the van der Waals' type between three atoms, J Chem Phys. 1943;11, pp.299–300. doi: 10.1063/1.1723844.

Baldus, M. (2006). Molecular interactions investigated by multi-dimensional solid-state NMR, Curr Opin Struct Biol. 2006 Oct;16(5), pp. 618-23.

Barratt, E.; Bingham, R.J.; Warner, D.J.; Laughton, C.A.; Phillips, S.E.; Homans, S.W. (2005). Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water, J Am Chem Soc. 2005 Aug 24;127(33), pp. 11827-34.

Barratt, E.; Bronowska, A.K.; Vondrásek, J.; Cerný, J.; Bingham, R.; Phillips, S.; Homans, S.W. (2006). Thermodynamic penalty arising from burial of a ligand polar group within a hydrophobic pocket of a protein receptor., J Mol Biol. 2006 Oct 6;362(5), pp. 994-1003.

Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges the RESP model. J. Phys. Chem. 1993;97, pp. 10269–10280.

Benjahad, A.; Guillemont, J.; Andries, K.; Nguyen, C.H.; Grierson, D.S. (2003). 3-Iodo-4-phenoxypyridinones (IOPY's), a new family of highly potent non-nucleoside inhibitors of HIV-1 reverse transcriptase. Bioorg. Med. Chem. Lett. 2003, 13, pp. 4309–4312.

Best, R.B.; & Vendruscolo, M. (2004). Determination of protein structures consistent with NMR order parameters, J Am Chem Soc. 2004 Jul 7;126(26), pp. 8090-1.

Bissantz, C.; Kuhn, B.; & Stahl, M. (2010). A medicinal chemist's guide to molecular interactions. J. Med. Chem. 2010;53(14), pp. 5061-5084. doi: 10.1021/jm100112j.

Boehm, H.J.; & Klebe, G. (1996).What can we learn from molecular recognition in protein−ligand complexes for the design of new drugs?. Angew. Chem., Int. Ed. 1996, 35, pp. 2588–2614.

Boehr, D.D.; Dyson, H.J.; & Wright, P.E. (2009). An NMR perspective on enzyme dynamics. Chem Rev, 2006; 106, pp. 3055-3079.

Brown, A. (2009). Analysis of Cooperativity by Isothermal Titration Calorimetry, Int. J. Mol. Sci. 2009, 10, pp. 3457 – 3477. doi:10.3390/ijms10083457

Burz, D.S.; Dutta, K.; Cowburn, D.; Shektman, A. (2006). In-cell NMR for protein-protein interactions (STINT-NMR), Nat Protoc. 2006;1(1), pp. 146-52.

Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. J. (1981) Solution Chem. 1981, 10, pp. 563–595.

Chakrabarti, P.; & Bhattacharyya, R. (2007). Geometry of nonbonded interactions involving planar groups in proteins. Prog. Biophys. Mol. Biol. 2007, 95, pp. 83–137.

Cieplak, P.; Cornell, W.D.; Bayly, C.; Kollman, P.A. (1995). Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins. J. Comput. Chem. 1995;16, pp. 1357–1377.

Clark, T.; Hennemann, M.; Murray, J.S.; Politzer, P. (2007). Halogen bonding: the sigma hole. J. Mol. Model. 2007, 13, pp. 291–296.

Connelly, P.R.; Thomson, J.A.; Fitzgibbon, M.J.; Bruzzese, F.J. (1993). Probing hydration contributions to the thermodynamics of ligand binding by proteins. Enthalpy and heat capacity changes of tacrolimus and rapamycin binding to FK506 binding protein in D2O and H2O, Biochemistry. 1993 Jun 1;32(21), pp. 5583-90.

Connelly, P.R.; Aldape, R.A.; Bruzzese, F.J.; Chambers, S.P.; Fitzgibbon, M.J.; Fleming, M.A.; Itoh, S.; Livingston, D.J.; Navia, M.A.; Thomson, J.A. et al. (1994). Enthalpy of hydrogen bond formation in a protein-ligand binding reaction, Proc Natl Acad Sci U S A. 1994 Mar 1;91(5), pp. 1964-8.

Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Kollman, P.A. (1993). Application of RESP charges to calculate conformational energies, hydrogen-bond energies, and free-energies of solvation. J. Am. Chem. Soc. 1993;115, pp. 9620–9631.

Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M. Jr.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, J. Am. Chem. Soc. 117, pp. 5179–5197.

Daranas, A.H.; Shimizu, H.; & Homans, S.W. (2004). Thermodynamics of binding of D-galactose and deoxy derivatives thereof to the L-arabinose-binding protein, J Am Chem Soc. 2004 Sep 29;126(38), pp. 11870-6.

Davis, A.M.; & Teague, S.J. (1999). Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. Angew. Chem., Int. Ed. 1999, 38, pp. 736–749.

Denisov, V.P.; Venu, K.; Peters, J.; Hoerlein, H.D.; Halle, B. (1997) Orientational disorder and entropy of water in protein cavities. J. Phys. Chem. B 1997, 101, pp. 9380–9389.

Desiraju, G.R. (2002) Hydrogen bridges in crystal engineering: interactions without borders. Acc. Chem. Res. 2002, 35, pp. 565–573.

Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U.J.; Akke, M. (2010). Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3, J Am Chem Soc. 2010 Oct 20;132(41), pp. 14577-89.

Dill, K.A. (1997). Additivity principles in biochemistry, J. Biol. Chem., 272, pp. 701-704.

Dixon, S.L. & K.M. Merz, Jr. (1996). Semiempirical Molecular Orbital Calculations with Linear System Size Scaling. J. Chem. Phys., 1996. 104: pp. 6643-6649.

Dixon, S.L. and K.M. Merz, Jr. (1997). Fast, Accurate Semiempirical Molecular Orbital Calculations for Macromolecules. J. Chem. Phys., 1997. 107(3): pp. 879-893.

Donchev, A.G. (2006).Many-body effects of dispersion interaction. J Chem Phys. 2006, pp. 125:074713. doi: 10.1063/1.2337283.

Dunitz, J.D. (1994).The entropic cost of bound water in crystals and biomolecules. Science 1994, 264, p. 670.

Dunitz, J.D. (1995). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions., Chem Biol, 1995;2(11), pp. 709-12.

Earl, D.J.; & Deem, M.W. (2005). Parallel tempering: theory, applications, and new perspectives, Phys Chem Chem Phys. 2005 Dec 7;7(23), pp. 3910-6

Ebbinghaus, S.; Kim, S.J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D.M.; Havenith, M. (2007). An extended dynamical hydration shell around proteins. Proc Natl Acad Sci U S A. 2007 Dec 26;104(52), pp- 20749-52.

Emsley, J. (1980). Very Strong Hydrogen Bonds. Chem Soc Rev, 1980;9(1), pp. 91-124.

Evans, D.A.; & Bronowska, A.K. (2010). Implications of fast-time scale dynamics of human DNA/RNA cytosine methyltransferases (DNMTs) for protein function, Theoretica Chimica Acta, Volume 125, Numbers 3-6, pp. 407- 418, doi: 10.1007/s00214-009-0681-2

Fanfrlík, J.; Bronowska, A.K.; Rezác, J.; Prenosil, O.; Konvalinka, J.; Hobza P. (2010). JA reliable docking/scoring scheme based on the semiempirical quantum mechanical PM6-DH2 method accurately covering dispersion and H- bonding: HIV-1 protease with 22 ligands, Phys Chem B. 2010 Oct 7;114(39), pp. 12666-78.

Fernandez, C.; & Wider, G. (2003). TROSY in NMR studies of the structure and function of large biological macromolecules, Curr Opin Struct Biol. 2003 Oct;13(5), pp. 570-80.

Finkelstein, A.V. (2007). Average and extreme multi-atom Van der Waals interactions: strong coupling of multi-atom Van der Waals interactions with covalent bonding. Chem Cent J. 2007, pp. 1:21. doi: 10.1186/1752-153X-1-21.

Finney, J.L.; & Soper, A.K.(1994). Solvent structure and perturbations in solutions of chemical and biological importance. Chem. Soc. Rev. 1994, 23, pp. 1–10.

Fischer, S.; & Verma, C.S. (1999). Binding of buried structural water increases the flexibility of proteins. Proc. Natl. Acad. Sci. U.S.A. 1999, 96, pp. 9613–9615.

Fischer, F.R; Wood, P.A.; Allen, F.H.; Diederich, F. (2008). Orthogonal dipolar interactions between amide carbonyl groups, Proc Natl Acad Sci U S A. 2008 Nov 11;105(45), pp. 17290-4.

Foloppe, N.; & Hubbard, R. (2006). Towards predictive ligand design with free-energy based computational methods?, Curr Med Chem. 2006;13(29), pp. 3583-608.

Ford, D.M. (2005). Enthalpy−entropy compensation is not a general feature of weak association. J. Am. Chem. Soc. 2005, 127, pp. 16167–16170.

Frauenfelder, H.; & Mezei, F. (2010). Neutron scattering and protein dynamics, Acta Crystallogr D Biol Crystallogr. 2010 Nov;66(Pt 11), pp. 1229-31.

Gilson, M.K.; & Zhou, H.X. (2007). Calculation of protein-ligand binding affinities, Annu Rev Biophys Biomol Struct. 2007;36, pp. 21-42.

Goodman, J.L.; Pagel, M.D.; & Stone, M.J. (2000). Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters, J Mol Biol. Jan 28;295(4), pp. 963-78.

Graf, C.; Stankiewicz, M.; Kramer, G.; Mayer, M.P. (2009). Spatially and kinetically resolved changes in the conformational dynamics of the Hsp90 chaperone machine, EMBO J. 2009 Mar 4;28(5), pp. 602-13.

Grover, J.R.; Walters, E.A.; & Hui, E.T. (1987). Dissociation energies of the benzene dimer and dimer cation. J. Phys. Chem. 1987, 91, pp. 3233–3237.

Harman, J.G. (2001). Allosteric regulation of the cAMP receptor protein, Biochim Biophys Acta. 2001 May 5;1547(1), pp. 1-17.

Heugen, U.; Schwaab, G.; Bründermann, E.; Heyden, M.; Yu, X.; Leitner, D.M.; Havenith, M. (2006). Solute-induced retardation of water dynamics probed directly by terahertz spectroscopy, Proc Natl Acad Sci U S A. 2006 Aug 15, 103(33), pp. 12301-6.

Hobza, P.; Selzle, H.L.; & Schlag, E.W. (1996). Potential Energy Surface for the Benzene Dimer. Results of ab initio CCSD(T) Calculations Show Two Nearly Isoenergetic Structures: T-Shaped and Parallel Displaced, J Phys Chem. 1996;100(48), pp. 18790-94.

Homans, S.W. (2005). Probing the binding entropy of ligand-protein interactions by NMR, Chembiochem. 2005 Sep;6(9), pp. 1585-91.

Homans, S.W. (2007). Dynamics and thermodynamics of ligand-protein interactions, Top Curr Chem 2007, 272, pp. 51-82.

Iltzsch, M.H.; Uber, S.S.; Tankersley, K.O.; el Kouni, M.H. (1995). Structure−activity relationship for the binding of nucleoside ligands to adenosine kinase from Toxoplasma gondii. Biochem. Pharmacol. 1995, 49, pp. 1501–1512.

Jelesarov, I.; & Bosshard, H.R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition, J Mol Recognition, 12(1), pp.3-18.

Ji, C.G.; & Zhang, J.Z. (2009). Electronic polarization is important in stabilizing the native structures of proteins, J Phys Chem B. 2009 Dec 10;113(49), pp. 16059-64.

Ji, C.G.; Mei, Y.; & Zhang, J.Z. (2008). Developing polarized protein-specific charges for protein dynamics: MD free energy calculation of pKa shifts for Asp26/Asp20 in thioredoxin, Biophys J. 2008 Aug;95(3), pp. 1080-1088.

Klamt, A.; & Schüürmann, G. (1993). COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradien, J.Chem.Soc.Perkin Trans. 2, p. 799-805.

Klepeis, J.L.; Lindorff-Larsen, K.; Dror, R.O.; Shaw, D:E. (2009). Long-timescale molecular dynamics simulations of protein structure and function, Curr Opin Struct Biol. 2009 Apr;19(2), pp. 120-7.

Knight, J.L.; & Brooks, C.L. III. (2009). Lambda-Dynamics free energy simulation methods. Journal of Computational Chemistry, 2009, pp. 1692-1700.

Krause, H.; Ernstberger, B.; Neusser, H.J. (1991). Binding energies of small benzene clusters. Chem. Phys. Lett. 1991, 184, pp. 411–417.

Krishnamurthy, V. M.; Bohall, B. R.; Semetey, V.; Whitesides, G. M. (2006). The paradoxical thermodynamic basis for the interaction of ethylene glycol, glycine, and sarcosine chains with bovine carbonic anhydrase II: an unexpected manifestation of enthalpy/entropy compensation. J. Am. Chem. Soc. 2006, 128, pp. 5802–5812.

Kuser, P.R.; Franzoni, L.; Ferrari, E.; Spisni, A.; Polikarpov, I. (2001). The X-ray structure of a recombinant major urinary protein at 1.75 A resolution. A comparative study of X-ray and NMR-derived structures, Acta Crystallogr D Biol Crystallogr. 2001 Dec;57(Pt 12), pp. 1863-9.

Laio, A.; & Gervasio, F.L. (2008). Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science, Rep. Prog. Phys. 71, p. 126601. doi: 10.1088/0034- 4885/71/12/126601

Lange, O.F.; Schäfer, L.V.; & Grübmüller, H. (2006). Flooding in GROMACS: accelerated barrier crossings in molecular dynamics, J Comput Chem. 2006 Nov 15;27(14), pp. 1693-702.

Langhorst, U.; Loris, R.; Denisov, V.P.; Doumen, J.; Roose, P.; Maes, D.; Halle, B.; Steyaert, J. (1999). Dissection of the structural and functional role of a conserved hydration site in RNase T1. Protein Sci. 1999 Apr;8(4), pp. 722-30.

Li, Z.; & Liu, W. (2010). Spin-adapted open-shell random phase approximation and time-dependent density functional theory. I. Theory, J Chem Phys. 2010 Aug 14;133(6), pp. 064106.

Lipari, G.; & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic reaonance relaxation in macromolecules. 1. Theory and range of validity, J. Am. Chem. Soc., 104, pp. 4546–4559.

Lu, Y. et al. (2009) Halogen bonding--a novel interaction for rational drug design? J Med Chem 52 (9), pp. 2854-2862.

Lundquist, J.J.; & Toone, E.J. (2002). The cluster glycoside effect, Chem. Rev. 2002;102, pp. 555–578.

MacRaild, C.A.; Daranas, A.H.; Bronowska, A.; Homans, S.W. (2007). Global changes in local protein dynamics reduce the entropic cost of carbohydrate binding in the arabinose-binding protein, J. Mol. Biol., 2007, 368(3), pp. 822- 832.

Makowski, L.; Gore, D.; Mandava, S.; Minh, D.; Park, S.; Rodi, D.J.; Fischetti RF. (2011). X-ray solution scattering studies of the structural diversity intrinsic to protein ensembles, Biopolymers. 2011 Apr 1. doi: 10.1002/bip.21631

Mao, Y.; Ratner, M.A.; & Jarrold, M.F. (2000). One water molecule stiffens a protein. J. Am. Chem. Soc. 2000, 122, pp. 2950–2951.

Mao, B.; Pear, M.R.; McCammon, J.A.; Quiocho F.A. (1982). Hinge-bending in L-arabinose-binding protein. The "Venus's- flytrap" model. J. Biol. Chem. 1982;257, pp. 1131–1133.

Matthews, B.W.; & Liu, L. (2009). A review about nothing: are apolar cavities in proteins really empty?. Protein Sci. 2009, 18, pp. 494–502.

Meyer, B.; & Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors, Angew Chem. Int. Ed. 2003, 42(8), pp. 864-90.

Michel, J.; Tirado-Rives, J.; & Jorgensen, W. L. (2009). Energetics of displacing water molecules from protein binding sites: consequences for ligand optimization. J. Am. Chem. Soc. 2009, 131, pp. 15403–15411.

Ming, D.; & Brüschweiler, R. (2004). Prediction of methyl-side chain dynamics in proteins, J Biomol NMR. 2004 Jul;29(3), pp. 363-8.

Mittermaier, A.; Davidson, A.R.; & Kay, L.E. (2003). Correlation between 2H NMR side-chain order parameters and sequence conservation in globular proteins, J Am Chem Soc. 2003 Jul 30;125(30), pp. 9004-5.

Nagendra, H.G.; Sukumar, N.; & Vijayan, M. (1998). Role of water in plasticity, stability, and action of proteins: the crystal structures of lysozyme at very low levels of hydration, Proteins. 1998 Aug 1;32(2), pp. 229-40.

Newcomer, M.E.; Lewis, B.A.; & Quiocho, F.A. (1981). The radius of gyration of L-arabinose-binding protein decreases upon binding of ligand. J. Biol. Chem. 1981;256, pp. 13218–13222.

Panigrahi, S.K.; & Desiraju, G.R.(2007). Strong and weak hydrogen bonds in the protein−ligand interface. Proteins 2007, 67, pp. 128–141.

Paulini, R.; Müller, K.; & Diederich, F. (2005). Orthogonal multipolar interactions in structural chemistry and biology. Angew. Chem., Int. Ed. 2005, 44, pp. 1788–1805.

Perozzo, R.; Folkers, G.; & Scapozza, L. (2004). Thermodynamics of protein-ligand interactions: history, presence, and future aspects, J. Recept. Signal Transduct. Res., 2004, 24(1-2), pp. 1-52.

Piela, L. (2007). Ideas of Quantum Chemistry, Elsevier, ISBN: 9780444522276.

Poornima, C.S.; & Dean, P.M. (1995). Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions., J Comput Aided Mol Des. 1995 Dec;9(6), pp. 500-12.

Poornima, C.S.; & Dean, P.M. (1995). Hydration in drug design. 2. Influence of local site surface shape on water binding. J. Comput.-Aided Mol. Des. 1995, 9, pp. 513–520.

Poornima, C.S.; & Dean, P.M. (1995). Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins., J Comput Aided Mol Des. 1995 Dec;9(6), pp. 521-31.

Popovych, N.; Sun, S., Ebright, R.H.; Kalodimos, C.G. (2006). Dynamically driven protein allostery, Nat Struct Mol Biol. 2006 Sep;13(9), pp. 831-8.

Quiocho, F.A. (1993). Probing the atomic interactions between proteins and carbohydrates. Biochem. Soc. Trans. 1993;21, pp. 442–448.

Raha, K.; & Merz, K.M. Jr. (2004). A quantum mechanics-based scoring function: study of zinc ion-mediated ligand binding, J Am Chem Soc. 2004 Feb 4;126(4), pp. 1020-1.

Raha, K.; Peters, M.B.; Wang, B.; Yu, N.; Wollacott, A.M.; Westerhoff, L.M.; Merz, K.M. Jr. (2007). The role of quantum mechanics in structure-based drug design, Drug Discov Today. 2007 Sep;12(17-18), pp. 725-31.

Rezac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P. (2009). Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes, J. Chem. Theory Comp. 2009, 5, pp. 1749-1760.

Ringer, A.L.; Senenko, A.; & Sherrill, C.D. (2007). Models of S/pi interactions in protein structures: Comparison of the H2S−benzene complex with PDB data. Protein Sci. 2007, 16, pp. 2216–2223.

Sadus, R.J. (1998). Exact calculation of the effect of three-body Axilrod-Teller interactions on vapour-liquid phase coexistence. Fluid Phase Equilibria. 1998;144, pp. 351–360. doi: 10.1016/S0378-3812(97)00279-3.

Scatena, L.F.; Brown, M.G.; & Richmond, G.L. (2001). Water at hydrophobic surfaces: weak hydrogen bonding and strong orientation effects. Science 2001, 292, pp. 908–912.

Schlitter, J. (1993). Estimation of absolute and relative entropies of macromolecules using the covariance matrix. Chem. Phys. Lett. 1993, 215, pp. 617–621.

Seeliger, D.; Haas, J.; & de Groot, B.L. (2007). Geometry-based sampling of conformational transitions in proteins, Structure, 2007;15(11), pp. 1482-92.

Senn, H.M.; & Thiel, W. (2009). QM/MM methods for biomolecular systems, Angew Chem Int Ed Engl. 2009;48(7), pp. 1198-229.

Sharp, K. (2001). Entropy−enthalpy compensation: fact or artifact?. Protein Sci. 2001, 10, pp. 661–667.

Shimokhina, N.; Bronowska, A.K.; & Homans, S.W. (2006). Contribution of Ligand Desolvation to Binding Thermodynamics in a Ligand–Protein Interaction, Angew Chem Int Ed Engl. 2006 Sep 25;45(38), pp. 6374-6.

Showalter, S.A.; & Brüschweiler, R. (2007). Validation of molecular dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: application to the AMBER99SB force field, J. Chem. Theory Comput., 2007, 3, pp. 961-975.

Sinnokrot, M.O.; & Sherrill, C.D. (2004). Substituent effects in pi−pi interactions: sandwich and T-shaped configurations. J. Am. Chem. Soc. 2004, 126, pp. 7690–7697.

Smith, J.C.; Merzel, F.; Bondar, A.N.; Tournier, A.; Fischer S. (2004). Structure, dynamics and reactions of protein hydration water., Philos Trans R Soc Lond B Biol Sci. 2004 August 29; 359(1448), pp. 1181–1190.

Stewart, J.J. (2009). Application of the PM6 method to modeling proteins, J Mol Model. 2009 Jul;15(7), pp. 765-805

Stoeckmann, H.; Bronowska, A.K.; Syme, N.R.; Thompson, G.S.; Kalverda, A.P.; Warriner, S.L.; Homans, S.W. (2008). Residual ligand entropy in the binding of p-substituted benzenesulfonamide ligands to bovine carbonic anhydrase II, J Am Chem Soc. 2008 Sep 17;130(37), pp. 12420-6.

Straatsma, T.P.; & Berendsen, H.J.C. (1988). Free energy of ionic hydration: Analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulations, J. Chem. Phys. 89, pp. 5876- 5886.

Timm, D.E.; Baker, L.J.; Mueller, H.; Zidek, L.; Novotny, M.V. (2001). Structural basis of pheromone binding to mouse major urinary protein (MUP-I), Protein Sci. 2001 May;10(5), pp. 997-1004.

Tochtrop, G.P.; Richter, K.; Tang, C.; Toner, J.J.; Covey, D.F.; Cistola, D.P. (2002). Energetics by NMR: site-specific binding in a positively cooperative system. Proc. Natl. Acad. Sci. U.S.A. 2002, 99, pp. 1847–1852.

Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanaba, K. (2000). Origin of the attraction and directionality of the NH/pi interaction: comparison with OH/pi and CH/pi interactions. J. Am. Chem. Soc. 2000, 122, pp. 11450– 11458.

Turk, J.A.; & Smithrud, D.B. (2001). Synthesis and physical properties of protein core mimetics, J. Org. Chem. 2001, 66, pp. 8328–8335.

Turnbull, W. B.; Precious, B. L.; & Homans, S. W. (2004). Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry, J. Am. Chem. Soc. 2004, 126, pp. 1047–1054.

Vallone, B.; Miele, A.E.; Vecchini, P.; Chiancone, E.; Brunori, M. (1998). Free energy of burying hydrophobic residues in the interface between protein subunits. Proc. Natl. Acad. Sci. U.S.A. 1998, 95, pp. 6103–6107.

Vyas, N.K.; Vyas, M.N.; & Quiocho, F.A. (1991). Comparison of the periplasmic receptors for L-arabinose, D-glucose/D- galactose, and D-ribose. Structural and Functional Similarity, J Biol Chem. 1991 Mar 15;266(8), pp. 5226-37.

Wand, A.J. (2001). Dynamic activation of protein function: a view emerging from NMR spectroscopy, Nat Struct Biol. 2001 Nov;8(11), pp. 926-31.

Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. (2004). Development and testing of a general amber force field. J Comp Chem. 2004;25, pp. 1157–1174. doi: 10.1002/jcc.20035.

Weiss, S. (2000). Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy, Nat Struct Biol. 2000 Sep;7(9), pp. 724-9.

Whitesides, G.M.; & Krishnamurthy, V.M. (2005). Designing ligands to bind proteins, Q. Rev. Biophys., 2005, 38, pp. 385- 395.

Williams, M.A.; & Ladbury, J.E. (2003). Hydrogen bonds in protein−ligand complexes. Methods Princ. Med. Chem. 2003, 19, pp. 137–161.

Williams, D.H.; Stephens, E.; & Zhou, M. (2003). Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes, J Mol Biol. 2003 May 30;329(2), pp. 389-99.

Xu, J.; Plaxco, K.W.; & Allen, J.S. (2006). Probing the collective vibrational dynamics of a protein in liquid water by terahertz absorption spectroscopy, Protein Sci. 2006 May; 15(5), pp. 1175–1181.

Yang, D.; & Kay, L.E. (1996). Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding, J. Mol. Biol., 1996, 263(2), pp. 369-82.

Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. (2007). Motifs for molecular recognition exploiting hydrophobic enclosure in protein−ligand binding. Proc. Natl. Acad. Sci. U.S.A. 2007, 104, pp. 808–813.

Zauhar, R.J.; Colbert, C.L.; Morgan, R.S.; Welsh, W.J. (2000). Evidence for a strong sulfur−aromatic interaction derived from crystallographic data. Biopolymers 2000, 53, pp. 233–248.

Zhou, P.; Tian, F.; Zou, J.; Shang, Z. (2010). Rediscovery of halogen bonds in protein-ligand complexes. Mini Rev Med Chem. 2010;10(4), pp. 309-314.
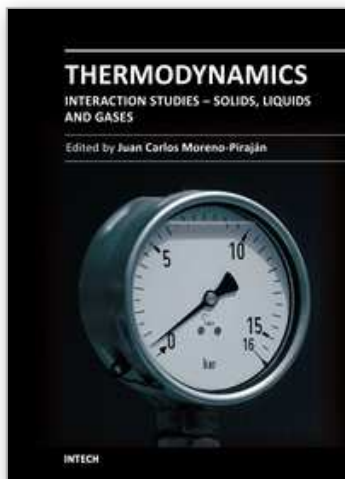
Thermodynamics - Interaction Studies - Solids, Liquids and Gases

Edited by Dr. Juan Carlos Moreno Pirajãn

Thermodynamics is one of the most exciting branches of physical chemistry which has greatly contributed to the modern science. Being concentrated on a wide range of applications of thermodynamics, this book gathers a series of contributions by the finest scientists in the world, gathered in an orderly manner. It can be used in post-graduate courses for students and as a reference book, as it is written in a language pleasing to the reader. It can also serve as a reference material for researchers to whom the thermodynamics is one of the area of interest.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds