Conformational thermodynamics of biomolecular complexes: The histogram-based method

Amit Das1#, Samapan Sikdar1, Mahua Ghosh*1 and J. Chakrabarti*1,2

1Department of Chemical, Biological And Macromolecular Sciences, S. N. Bose National Center for Basic Sciences, Sector III, Block JD, Salt lake City Kolkata 700098, INDIA

2Also at Unit of Nanoscience and Technology-II and The Thematic Unit of Excellence on Computational Materials Science, S. N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt Lake, Kolkata-700098, INDIA

#Presently at National Centre for Biological Sciences (TIFR), Bellary Road, Bangalore 560 065, INDIA

*Email: jaydeb@bose.res.in and mahuaghosh@bose.res.in

Abstract. Conformational changes in biomacromolecules govern majority of biological processes. Complete characterization of conformational contributions to thermodynamics of complexation of biomacromolecules has been challenging. Although, advances in NMR relaxation experiments and several computational studies have revealed important aspects of conformational entropy changes, efficient and large-scale estimations still remain an intriguing facet. Recent histogram-based method (HBM) offers a simple yet rigorous route to estimate both conformational entropy and free energy changes from same set of histograms in an efficient manner. The HBM utilizes the power of histograms which can be generated as accurately as desired from an arbitrarily large sample space from atomistic simulation trajectories. Here we discuss some recent applications of the HBM, using dihedral angles of amino acid residues as conformational variables, which provide good measure of conformational thermodynamics of several protein-peptide complexes, obtained from NMR, metal-ion binding to an important metalloprotein, interfacial changes in protein-protein complex and insight to protein function, coupled with conformational changes. We conclude the paper with a few future directions worth pursuing.
1. Introduction

Bio-macromolecules are large flexible molecules which exist in numerous possible conformations in biological environments. Fluctuations among these conformations are of fundamental importance in their structure and function. One classic example of such conformational fluctuation is the folding of proteins into its native conformation [1]. The biomacromolecules need to adopt suitable conformations in general to participate in molecular recognition, signal transduction, gene expression and so on [2,3]. Furthermore, subtle conformational changes of different biomacromolecules have been utilized in various nano-biotechnological applications [4]. Thus, microscopic understanding of conformational fluctuations in biomacromolecules is important from both biological and technological standpoints.

The most important aspect is the thermodynamic characterization of conformational fluctuations. Experimentally, the isothermal titration calorimetry (ITC) measurements yield the binding enthalpy and the binding free energy for bio-macromolecular complexes [5] but not the thermodynamics of the conformational changes. The NMR relaxation experiments provide measure of the conformational entropy costs in bio-macromolecular complexes from the bond order parameter [2]. Such studies indicate that the conformational entropy changes constitute an important contribution of the thermodynamic changes in the complex. The computer simulations [6,7] often estimate the conformational entropy from the normal modes associated with atomic vibrations. However, the atomic Cartesian coordinates often are not suitable to capture all possible bond rotations, thus providing poor estimates [8] of conformational entropy. The approaches based on purely statistical scoring functions from the crystal structure databases [9] are devoid of microscopic details. The protein dihedral angles, defined by the angle between different atomic planes in the molecule, constitute a convenient set of microscopic conformational variable to define a conformational state of a protein. Multidimensional histograms of the dihedral angle are shown to provide the estimation of the conformational entropy [10] for small molecules. Such studies further show that nearly ~80% of the conformational entropy for different small molecules, like hydrogen peroxide, methanol and some alkanes, is accounted for by neglecting all sorts of correlations among the dihedral angles [11]. In biomacromolecules the dihedral correlations have been generally found [12,13] to be short-ranged. These observations highlight the importance of completely reduced one-dimensional histograms of the independent dihedral angles [14,15]. Accurate and efficient estimation of conformational free energy is challenging as well. The free energy difference between two conformational states of small biomolecules has been measured by the population ratios of the two states via UV resonance Raman measurements [16]. Computational methodologies for conformational free energies are of two types: (1) estimating the absolute free energies of a conformational state [17]; and (2) evaluating the free energy differences between two states [18-20]. However, these approaches are inefficient to study medium to large bio-macromolecular systems.

It is thus important to have efficient and accurate method for evaluating the conformational thermodynamics for large bio-macromolecule, like proteins. If the dihedral correlations can be ignored, the histogram-based method (HBM) [21] utilizing the one-dimensional histograms of the dihedral angles to extract the thermodynamics of conformational changes, including the entropy and free energy, by simply adding all the dihedral angle contributions of a protein. The basic idea can be understood as follows: Suppose that the normalized histograms of a given dihedral angle are computed in two conformational states. One can compare the entropies in the two states using the Gibb’s formula and the free energy by comparing the peaks, relating the logarithm of the ratio to the free energy. Thus, unlike the existing methodologies, the HBM can yield both the information simultaneously from a single set of simulations. Moreover, the HBM is capable of yielding the conformational thermodynamics data even for the individual residues.

The histograms of dihedral fluctuations about different conformational states of a protein can be sampled from the equilibrium trajectories from all-atom MD simulations. In this article we focus on
(a couple of interesting aspects of the thermodynamics of conformational changes in a protein: (1) The conformational entropy estimated from the one-dimensional histograms can reproduce the conformational entropy obtained for large protein molecules from NMR experiments. This validates the HBM as a viable technique to estimate the conformational thermodynamics from microscopic conformational variables for bio-macromolecules. (2) The changes in the conformational thermodynamics are different for different residues. Even if a given conformational state is stabilized in thermodynamic sense, all the residues are not necessarily stabilized. The destabilized residues participate in binding events which gives clue for functional activities of different residues in different conformational states. The paper is organized as follows: We discuss the theoretical basis of the HBM in Sec. 2. We discuss the extraction of conformational thermodynamics for a system where the experimental data is available in Sec 3. We illustrate the method for a couple of cases, like ion binding to a protein in Sec. 4 and protein-protein interface in Sec. 5. We show the application of HBM in understanding protein function in Sec 6 and conclude with some future directions in Sec. 7.

2. The histogram-based method (HBM)

Let us consider a system with a set of conformational variables \(\{\xi_i\}\). The normalized probability distribution at a temperature \(T\), is given by:

\[
P(\xi_i) = \frac{1}{Z} \exp[-\mathcal{H}(\xi_i)/k_B T] \tag{1}
\]

where \(k_B\) is the Boltzmann constant, \(\mathcal{H}(\xi_i)\), the Hamiltonian and \(Z\), the partition function in the conformational space of the system which normalizes the probability distribution. Let us fix one of the conformational variables \(\xi_i = \xi\). By integrating over the other variables in Eq. 1:

\[
P(\xi = \xi) = (1/Z) \left[ \exp[-\mathcal{H}(\xi_i)/k_B T] \delta(\xi_i - \xi) \right] \int d\xi_i = (1/Z) \exp[-G(\xi)/k_B T] \tag{2}
\]

we get the effective free energy \(G(\xi)\) or the potential of mean force associated with a given conformational variable \(\xi\).

Let us specifically consider the case of a protein complexed with a ligand molecule. For dihedral angles \(\theta\) of protein \(p\) and dihedral angles \(\tau\) of ligand, one expresses \(\xi = (\theta, \tau)\). Subsequently, the subscripts \(p+l\), \(p\) and \(l\) indicate quantities associated with the complex, the protein and the ligand respectively, while the superscripts \(c\) and \(f\) denote the bound and free state respectively. Then Eq. 2 defines the effective free energies as:

\[
P^c_{p+l}(\xi) = \frac{1}{Z} \exp[-G^c_{p+l}(\xi)/k_B T],
\]

\[
P^f_p(\theta) = \frac{1}{Z} \exp[-G^f_p(\theta)/k_B T]
\]

and \(P^f_l(\tau) = \frac{1}{Z} \exp[-G^f_l(\tau)/k_B T]\). \tag{3}

Therefore, the free energy change for \(\xi\) due to complexation is

\[
\Delta G_{\text{conf}}(\xi) = -k_B T \ln \left[ \frac{P^c_{p+l}(\xi)}{P^f_p(\theta) P^f_l(\tau)} \right] \tag{4}
\]
If the conformational variables are independent, and $P_{c,i}^\xi = P_p^\theta \cdot P_l^\tau$. Equation 4 then gives

$$
\Delta G_{\text{conf}}(\xi) = -k_B T \ln \left[ \frac{P_p^\theta}{P_p^\theta} \right] - k_B T \ln \left[ \frac{P_l^\tau}{P_l^\tau} \right],
$$

(5)

which indicates that the thermodynamics is given separately in terms of the individual dihedrals. Summing over all the dihedral angles of protein and ligand gives the total conformational free energy change from Eq.5:

$$
\Delta G_{\text{conf}} = -k_B T \sum_{\theta} \ln \left[ \frac{P_p^\theta}{P_p^\theta} \right] - k_B T \sum_{\tau} \ln \left[ \frac{P_l^\tau}{P_l^\tau} \right] = \Delta G_{\text{prot}}^\text{conf} + \Delta G_{\text{lig}}^\text{conf}.
$$

(6)

Eqs. (5) and (6) may be simplified if the probability distributions are sharply peaked. Let us denote the normalized histograms $H_p^\theta$ for a protein dihedral, and $H_l^\tau$ for a ligand dihedral in the bound and the free states, respectively. The peak of the histogram defines the equilibrium value of the relevant dihedrals. Then the equilibrium conformational free energy cost associated with any protein dihedral $\theta$ is

$$
\Delta G_{\text{conf}}^\text{eq}(\theta) = -k_B T \ln \left[ \frac{H_p^\theta}{H_{p,\text{max}}^\theta} \right],
$$

(7)

where the subscript ‘max’ denotes the maximum of histogram. For multi-modal histograms the free energies can be estimated by taking average, weighted by the maximum values of the peaks. For a particular dihedral $\xi$ with multi-modal histograms in both free and complex states, the free energy cost is

$$
\sum_{i,j} c_{ij} \Delta G_{ij} = -k_B T \ln \left( \frac{H_c^\xi}{H_{c,\text{max}}^\xi} \right),
$$

the respective weights.

The conformational entropy for a particular dihedral $\xi$ can be estimated directly using the Gibbs entropy formula [22]:

$$
S_{\text{conf}}(\xi) = -k_B \sum H(\xi) \ln H(\xi),
$$

(8)

The HBM can be extended to estimate the free energy and entropy costs associated with the changes of distribution of any quantity. For instance, the interfacial water molecules undergo changes in the distribution over the interface compared to the free components. Let us denote the distribution of the interfacial water ($iw$) molecules within 5 Å of the interfacial residues in the complex by $g^c(r)$, and those within 5 Å of the binding regions of the free components by $g^f(r)$ respectively. Here $r$ is the separation between the O-atoms of two water molecules. The free energy cost is given by straightforward extension of Eq. 7:
\[
\Delta G_{iw} = -k_BT \ln \left[ g_{\max}^{r}(r) / \left( g_{N,max}^{f}(r) g_{\max}^{f}(r) \right) \right]
\]  

(9)

where the peak values in the respective \( g(r) \) are used. The corresponding entropy cost \( (T\Delta S_{iw}) \) is obtained using the Gibbs formula as in Eq. 8:

\[
\Delta S_{iw} = -k_B \sum_i g_i^{r} \ln g_i^{r} - \sum_i g_i^{f} \ln g_i^{f} + \sum_i g_i^{f} \ln g_i^{f}
\]  

(10)

where the sum has been taken over non-zero bins marked by the index \( i \).

3. Application to Calmodulin-Peptide complexes

The HBM has been applied to the Calmodulin-peptide complexes. Calmodulin (CaM) is an ubiquitous eukaryotic protein which in calcium (Ca\(^{2+}\)) loaded (holo-) state acts as a primary mediator of activities of various cellular proteins responding to changes in intracellular Ca\(^{2+}\) levels [27]. CaM belongs to the family of proteins containing EF-hand [28], a helix-loop-helix structural motif that binds to Ca\(^{2+}\) ions via the loop. CaM has two globular domains: N-terminal (N-domain) and the C-terminal domain (C-domain), each domain having two EF-hands, I and II in the N-domain and III and IV in the C-domain (Figure 1a). Two domains are connected via a linker region consisting of 29 residues (residues 68-92). Metal-free (apo) CaM, upon Ca\(^{2+}\)-saturation, undergoes subtle conformational changes [29,30] in the metal-bound (holo) form (Figure 1a) with the linker becoming a long helix and thus exposing the target-binding hydrophobic faces of CaM for target peptides [5]. Moreover, the linker gets deformed to wrap around the peptides (Figure 1b) which are CaM-binding sequences of large number of proteins including several regulatory enzymes e.g. protein kinases, phosphodiesterases, cyclases etc. Five peptides have been considered: CaM-binding sequences of smooth muscle myosin light chain kinase (smMLCK) [31], the neuronal and endothelial nitric oxide synthases (nNOS and eNOS respectively) [32], the calmodulin kinase I (CaMKI)[5] and the calmodulin kinase kinase (CaMKK) [33]. For all these CaM-peptide complexes the ITC data and conformational entropy changes \( \Delta S_{\text{conf}} \) measured via NMR relaxation experiments are known [2]. According to observations, the total changes in conformational entropy \( \Delta S_{\text{conf}}^{\text{tot}} \) is linearly correlated with the total binding entropy \( \Delta S_{\text{bind}}^{\text{tot}} \) for the complexes. Recent work [2] also shows that \( \Delta S_{\text{conf}}^{\text{tot}} \) for CaM and the target peptides are linearly related to \( \left< \Delta S_{Me}^2 \right> \), the average changes in residue-weighted methyl group order-parameters \( S_{Me}^2 \), which describes the rotational freedom of the methyl group.

Figure 1: Binding events of CaM. (a) Apo- to holo- transition of CaM using cartoon representations. The EF-hand loops are coloured: I (purple), II (yellow), III (pink) and IV (cyan). The part of linker that undergo loop to helix transformation is marked by circle. (b) CaM bound to a target peptide (blue), smMLCK.
The dihedral angles have been calculated from the atomic coordinates in all atom simulation trajectories. The all-atom MD simulations have been performed for the free protein, metal bound protein, free peptide and the complex in explicitly electroneutral solvent. Simulations are done with the NAMD program [23] at 308K and 1 atm pressure in isothermal-isobaric (NPT) ensemble under standard protocols [24], using the CHARMM27 force-field [25], periodic boundary conditions and 1 femtosecond time-step. The number of total particles including water, pressure and temperature are kept fixed for each case to make the simulated ensembles equivalent. The initial configurations are chosen from following protein data bank (PDB) entries: 1CDL (smMLCK), 1NIW (eNOS), 2O60 (nNOS), 2L7L (CaMKI), 1CKK (CaMKK), 1CLL (metal bound CaM) and 1CFD (metal free CaM). The peptide coordinates taken out from the complex structures are the initial configurations of the free peptide simulations. The 50 ns long simulations capture most of the protein motions and peptide motions relevant for binding. The equilibration, which takes about 10 ns, is ensured in any run by monitoring the root mean square deviation. The dihedral angles (backbone dihedrals $\phi$, $\psi$ and the side-chain dihedrals $\chi_1$, $\chi_2$, $\chi_3$, $\chi_4$ and $\chi_5$) are calculated from angle between the relevant atomic planes [26].

Figures 2a-c show the equilibrium correlations (Eq. 5) among different dihedrals. The correlation data among different dihedrals in CaM (Figure 2a) and peptides (Figure 2b) in various complexes and the cross-correlations between CaM and peptide dihedrals (Figure 2c) indicate only short-ranged correlations as found in earlier studies [13]. In absence of significant correlations, the histograms of the individual dihedrals are taken for the calculation of thermodynamics. Figure 2d shows three set of histograms in free and complex states for the dihedral angle $\phi$ of CaM-residue Glu45 in smMLCK-complex calculated from 1000 equilibrated configurations from different parts of the trajectory. All the backbone dihedrals (Figure 2e) exhibit sharp single-peaked histograms with the maxima around the equilibrium values in the initial configuration (PDB coordinates). Multi-modal histograms have been mostly observed for the side-chain dihedrals, as illustrated for $\chi_2$ of CaM residue Leu69 (Figure 2f) and for $\chi_2$ of CaM residue Asp118 (Figure 2g), indicating different rotameric states. We use Eqs. 8 and 9 to compute the changes in the conformational thermodynamics of the complex with respect to the free components. The similarities of these histograms in Figure 2d indicate the convergence of thermodynamic quantities calculated based on them, for instance, conformational entropies (Figure 3). Due to equivalence of the samples, the thermodynamic quantities are computed via a flat average over the entire equilibrium trajectory.
**Figure 2:** Equilibrium dihedral correlations and histograms. (a-c) $\xi - \xi'$ correlations $C_{\xi\xi'}(s)$ (Eq. 5) where $\xi$ and $\xi'$ are dihedral angles from two residues and $s$ is the difference between their locations. All correlation data have been normalized to $C_{\xi\xi'}(s = 0)$ (a) Dihedrals of CaM in different complexes (b) Dihedrals of peptides in different complexes (c) Cross-correlations among side-chain dihedrals $\xi$ from protein and $\xi'$ from peptide. Representative histograms of (d) $\phi$ of Glu45 of CaM in free and bound form in smMLCK-complex. The inset shows the near-peak region for the free case. In each case three convergent histograms are shown in solid, dashed and dotted lines sampled from different parts of the MD-trajectory; (e) $\psi$ of Lys802 of smMLCK peptide in free (dotted) and CaM-bound form (solid). Multi-modal histograms are shown for (f) $\chi_2$ of CaM residue Leu69 in free and bound form in CaMKK complex and (g) $\chi_2$ of CaM residue Asp118 in free and bound form in CaMKI complex.

**Figure 3:** The convergence of simulated conformational entropy is shown for the smMLCK complex chosen as a representative case. The convergence demonstrates the robustness of the HBM. The errors in the data are very small: 0.05 kJ K\(^{-1}\) mol\(^{-1}\) for free CaM, 0.04 kJ K\(^{-1}\) mol\(^{-1}\) for CaM in the complex and nearly zero for the peptide.
In Figure 4a, the total conformational entropies $\Delta S_{\text{conf}}^{\text{tot}}$ of the complex obtained from HBM [21] has been compared with the NMR results [2]. Both the theoretical and experimental $\Delta S_{\text{conf}}^{\text{tot}}$ are plotted against the corresponding $\Delta S_{\text{bind}}^{\text{tot}}$ from ITC [2], showing linear correlations between $\Delta S_{\text{conf}}^{\text{tot}}$ and $\Delta S_{\text{bind}}^{\text{tot}}$. The HBM provide an overall good estimate for all cases, with the correlation line (slope $m = 1.3$ and linear correlation coefficient $R = 0.95$) [21] agreeing quite well with the experimental data ($m = 0.95$ and $R = 0.75$) [2]. The $\Delta S_{\text{conf}}$ (in kJ K$^{-1}$ mol$^{-1}$) for all the side-chain dihedrals in the complexes showing multi-peak histograms contribute more than 60% of the total conformational entropy stabilizations of the complexes. This indicates the importance of the redistributions of populations among various side-chain rotamers in the binding. In Figure 4b, the contributions of the CaM and peptide are reported separately in the complexes. The CaM contributions show good agreement between the HBM and the experimental data. The best fit theoretical line (dashed, $m = 1.1$, $R = 0.88$) almost quantitatively matches the experimental correlation line (solid, $m = 1.0$, $R = 0.94$). CaM in CaMKI has been observed to be most ordered and least ordered in nNOS which support the same experimental observations [2]. For the peptide, similar ordering has been observed for all the cases which explain the experimental trend [2].

**Figure 4:** Comparison of theoretical and experimental (NMR [2]) conformational entropy data for CaM-peptide binding plotted against the experimental $\Delta S_{\text{bind}}^{\text{tot}}$ from ITC measurements [2]: (a) Calculated (circles) and experimental (triangles) $\Delta S_{\text{conf}}^{\text{tot}}$. (b) The conformational entropy contributions of the components obtained from HBM: CaM (open squares) and peptide (open diamonds) contributions. The corresponding experimental data [2] for CaM contributions (filled squares) and peptides (filled diamonds) are also shown. The lines represent the best linear fits to the respective data.
The total conformational free energy changes $\Delta G_{\text{conf}}^{\text{tot}}$ and the contributions of the components are shown in Figure 5a along with the experimental total binding free energy cost $[2] \Delta G_{\text{bind}}^{\text{tot}}$. The $-\Delta G_{\text{bind}}^{\text{tot}}$ values fall in a very narrow window (45-52 kJ mol$^{-1}$). The estimated $-\Delta G_{\text{conf}}^{\text{tot}}$ lie in a similar range (27-47 kJ mol$^{-1}$) except CaMKK where the extent of stabilization is nearly double. Although the $\Delta G_{\text{bind}}^{\text{tot}}$-data seem to indicate that all the peptides exhibit similar affinity to CaM, the conformational contributions of the components bring out a different picture. The protein is conformationally most stabilized in CaMKK-complex, while least stabilized in eNOS-complex. This separates out CaMKK from the others which may be a signature of its opposite binding orientation compared to other four, namely, the N-terminal of four of the peptides interact with the C-terminal of CaM, except CaMKK for where the C-terminal of the peptide interacts with the N-terminal of the protein [33].

A negative change in conformational free energy $\Delta G_{i}^{\text{conf}}$ indicates conformational stability in the complex with respect to the free components, whereas that in $T \Delta S_{i}^{\text{conf}}$ represents conformational ordering. The individual residues undergo different degree of changes in the complex with respect to the free states. The thermodynamic changes at each protein residue are shown by surface representations in Figure 5b-c for the most and least stabilized complexes, both free-energetically and entropically. In Figure 5b, CaMKK and eNOS complexes are shown, where the green residues are stabilized, the red residues are destabilized and the white residues undergo a marginal change in conformational free energy. Figure 5c shows the entropy changes for CaMKI and nNOS complexes where the green residues are ordered while the red residues are disordered in the bound protein compared to the free-state.

Figure 5: (a) Calculated $\Delta G_{\text{conf}}^{\text{tot}}$ and the individual protein and peptide contributions along with the experimental $\Delta G_{\text{bind}}^{\text{tot}}$. All data are in kJ mol$^{-1}$. Surface representations of CaM-peptide complexes showing the residue-wise (b) $\Delta G_{\text{conf}}$ of CaM for the cases where the protein is most stabilized (CaMKK) and least stabilized (eNOS); (c) $\Delta S_{\text{conf}}$ of CaM in most ordered (CaMKI) and disordered (nNOS) complexes. The stabilized residues are green and destabilized ones are red. The residues undergoing minor changes are white. Peptide is in violet cartoon representation.
The HBM is applied to make predictions on the conformational thermodynamics of binding of a target peptide from death associated protein kinase (DAPK2) [34] to CaM (PDB ID 1ZUZ) and known ITC data (\( \Delta G_{\text{bind}}^{\text{tot}} = -39.5 \text{ kJ mol}^{-1} \) and \( \Delta S_{\text{bind}}^{\text{tot}} = -0.28 \text{ kJ K}^{-1} \text{ mol}^{-1} \) at 300K). However, no experimental data is known regarding the conformational entropies. Here \( \Delta S_{\text{conf}}^{\text{tot}} = -1.29 \text{ kJ K}^{-1} \text{ mol}^{-1} \) using the Gibbs formula where the CaM and peptide contributions are \(-0.98\) and \(-0.31 \text{ kJ K}^{-1} \text{ mol}^{-1}\) respectively. The \( \Delta S_{\text{conf}}^{\text{tot}} \) value follow the same linear scaling with \( \Delta S_{\text{Me}}^{2} \) as the other complexes as shown by the closed circle in Fig.4a. We get \( \Delta G_{\text{conf}}^{\text{tot}} = -38.1 \text{ kJ mol}^{-1}\) with CaM and peptide contributions being \(-15.4\) and \(-22.7 \text{ kJ mol}^{-1}\), respectively. These free energy values are very similar to the case of smMLCK. Residues 77-83 constitute the maximum destabilized PBR in DAPK2-complex: \( \Delta G_{\text{conf}} = +21.2 \text{ kJ mol}^{-1} \) and \( \Delta S_{\text{conf}} = +0.02 \text{ kJ K}^{-1} \text{ mol}^{-1} \).

The agreement between the results on \( \Delta S_{\text{conf}}^{\text{tot}} \) from HBM and those of Marlow et al.[2] has got a strong implication. Marlow et al. [2] connects the NMR data on \( \langle \Delta S_{\text{Me}}^{2} \rangle \), the average changes of residue weighted \( S_{\text{Me}}^{2} \), to \( \Delta S_{\text{conf}}^{\text{tot}} \) for several CaM-peptide complexes. By definition, 
\[
\langle \Delta S_{\text{Me}}^{2} \rangle = n_{\text{CaM}} \langle \Delta S_{\text{Me}}^{2} \rangle_{\text{CaM}} + n_{\text{pep}} \langle \Delta S_{\text{Me}}^{2} \rangle_{\text{pep}}
\]
where \( n_{\text{CaM}} \) and \( n_{\text{pep}} \) are the numbers of residues in CaM and peptide respectively. 
\[
\langle \Delta S_{\text{Me}}^{2} \rangle_{\text{CaM}} = \left\langle S_{\text{Me}}^{2} \right\rangle_{\text{CaM}} - \left\langle S_{\text{Me}}^{2} \right\rangle_{f}\text{CaM}
\]
and 
\[
\langle \Delta S_{\text{Me}}^{2} \rangle_{\text{pep}} = \left\langle S_{\text{Me}}^{2} \right\rangle_{\text{pep}} - \left\langle S_{\text{Me}}^{2} \right\rangle_{f}\text{pep}
\]
where the average is taken over the available methyl groups in the respective system. A microscopic justification for the linear connection described by Marlow et al. [2] between \( \Delta S_{\text{conf}}^{\text{tot}} \) and \( \langle \Delta S_{\text{Me}}^{2} \rangle \) is provided by the HBM. Figure 6 shows the linearity between the estimated \( \Delta S_{\text{conf}}^{\text{tot}} \) from HBM and \( \langle \Delta S_{\text{Me}}^{2} \rangle \) directly obtained from simulations. The slope (\( m = -0.050 \)) of this linear fit matches quite well with those reported in Marlow et al. (\( m \sim -0.04 \)) [2]. CaMKK has been an outlier in both the calculated [21] as well as the experimental data [2]. The conformational enthalpy changes \( \Delta H_{\text{conf}}^{\text{tot}} (= \Delta G_{\text{conf}}^{\text{tot}} + T \Delta S_{\text{conf}}^{\text{tot}}) \) should also have the same linearity with \( \langle \Delta S_{\text{Me}}^{2} \rangle \) as \( \Delta S_{\text{conf}}^{\text{tot}} \) for thermodynamic consistency. This is shown also in Figure 6 via the plot of \( \Delta H_{\text{conf}}^{\text{tot}} / T \) values of the complexes, excluding the outlier CaMKK.

![Figure 6: Plot of \( -\Delta S_{\text{conf}}^{\text{tot}} \) (circles) and \( -\Delta H_{\text{conf}}^{\text{tot}} / T \) (squares) vs. \( \langle \Delta S_{\text{Me}}^{2} \rangle \) from HBM excluding CaMKK, along with the best fit lines (dotted for \( -\Delta S_{\text{conf}}^{\text{tot}} \) and dashed for \( -\Delta H_{\text{conf}}^{\text{tot}} / T \)).](image-url)
4. Conformational thermodynamics for Ca$^{2+}$-ion binding to CaM

The HBM has been applied to study the Ca$^{2+}$-ion binding to apo-CaM [35]. Although some experimental [36-38] and theoretical studies [39,40] exist, the conformational thermodynamics of CaM upon Ca$^{2+}$-ion binding has remained largely unexplored. The conformational thermodynamic changes have been calculated for the holo-state with respect to the apo-state. Here, $\Delta G_{\text{conf}}^{\text{tot}} = -54.9$ kJ mol$^{-1}$ for binding of four Ca$^{2+}$, nearly 30% of the experimentally measured binding free energy [36]. The corresponding $T\Delta S_{\text{conf}}^{\text{tot}} = -128.7$ kJ mol$^{-1}$ indicating substantial drop in conformational flexibility due to Ca$^{2+}$ binding. In absence of any experimental data, an alternative verification of the estimated $T\Delta S_{\text{conf}}^{\text{tot}}$ can be obtained from $\langle \Delta S_{\text{Me}}^2 \rangle$ calculated from the simulations. Using the theoretical value of the slope ($-0.05$) of the linear dependence between $T\Delta S_{\text{conf}}^{\text{tot}}$ and $\langle \Delta S_{\text{Me}}^2 \rangle$ [21], $T\Delta S_{\text{conf}}^{\text{tot}} = -169$ kJ mol$^{-1}$ from estimated $\langle \Delta S_{\text{Me}}^2 \rangle (=10.95)$ which is close to the estimated $T\Delta S_{\text{conf}}^{\text{tot}}$ from dihedral histograms.

Figure 7 shows the contributions of different type of residues in these loops towards the respective $T\Delta S_{\text{conf}}^{L}$ (Figure 7a) and $\Delta G_{\text{conf}}^{L}$ (Figure 7b). The trends are very similar in both the panels for all the loops. In all cases except loop III, the main change comes from the acidic residues, nearly all coordinating to Ca$^{2+}$. However, the ordering and stabilization of the acidic residues in loop III is nearly half of those observed in loops I and IV. The basic residues in loops I are ordered and stabilized, while that in loop III remains almost unchanged. The residues with hydrophobic side-chains and those with polar side-chains order and stabilize amply in loops I, II and IV.

Figure 7: Data for the four EF-hand loops. (a) $T\Delta S_{\text{conf}}$ and (b) $\Delta G_{\text{conf}}$ contributions of different types of residues in the loops. The loop descriptions are: loop I - residues 20-31 (amino acid sequence DGDGTITITKKE), loop II - residues 56-67 (sequence DADGNTIDFPE), loop III - residues 93-104 (sequence DKDNGYISAAE) and loop IV - residues 129-140 (sequence DIDGDGQVNYEE).

The most striking conformational change upon Ca$^{2+}$ binding to apo-CaM occurs in the linker (residues 64-92), connecting the two domains, where a loop (the encircled region in Figure 1a) becomes helical in holo-state helping CaM to expose its hydrophobic surface for target-binding. The metal-induced changes of the linker helix in CaM are quite substantial: $T\Delta S_{\text{conf}} = -44.2$ and $\Delta G_{\text{conf}} = -13.3$ kJ mol$^{-1}$ for all 29 residues in the linker. These values account for ~34% of $T\Delta S_{\text{conf}}^{\text{tot}}$ and ~25% of $\Delta G_{\text{conf}}^{\text{tot}}$ for the whole protein. The majority of the changes in linker helix come from the loop
region that becomes helical upon Ca$^{2+}$-binding. $T\Delta S_{\text{conf}} = -20.5$ and $\Delta G_{\text{conf}} = -10.2$ kJ mol$^{-1}$ for these six residues 76-81 account for $\sim46\%$ of $T\Delta S_{\text{conf}}$ and $\sim77\%$ of $\Delta G_{\text{conf}}$ estimated for linker helix.

5. Thermodynamics of conformational changes at protein-protein interface

Recent experiments with bio-macromolecular complexes suggest that structural modifications at the interfaces are vital for stability of the complexes. Although several qualitative aspects about such interfaces are known from structural data, quantification of the interfacial changes is lacking. The thermodynamic changes at the interface in the complex have been estimated using the HBM [41]. A structurally well-characterized protein-protein complex [42], comprising of an enzyme Nuclease A, NucA [43] and its specific inhibitor NuiA [44], a protein has been considered. NucA, a sugar non-specific nuclease from <i>Anabaena sp.</i>, is one of the most active nucleases known. These nucleases, having the capacity to hydrolyze nucleic acids without any base-preference, are important for their roles in host cellular mechanisms [45-54]. These enzymes require divalent metal-ion cofactor (like Mn$^{2+}$, Mg$^{2+}$) at the active-site to function. NuiA, an intrinsic inhibitor, deactivates NucA specifically via formation of a 1:1 complex with a high stability constant ($>10^9$ M$^{-1}$) [47, 48]. The crystal structure of NucA-NuiA complex is particularly important, since isolation of any substrate-bound NucA has been unsuccessful so far due to its high activity. This complex throws light on the plausible nuclease activity of NucA. The crystal structure reveals the conformational changes of NucA and NuiA compared to their free forms and indicate that the electrostatic interactions between basic and the acidic residues dominate the interfacial changes [42, 53]. The crystal structure also indicates the presence of water molecule mediated interactions at the interface. A quantitative analysis of the contributions of water molecules to the interfacial thermodynamic changes is important.

The HBM has been applied on the dihedral angles calculated from the simulated conformations of the entire complex and the free components. Figs. 7(a) and (b) show the conformational entropy and free energy changes of the different types of residues in the interface of the complex with respect to those in the free components. The basic residues of NucA and acidic residues of NuiA undergo maximum stabilization and ordering out of all interfacial residues. This quantitatively confirms the suggestion of the crystal structure data the enzyme-inhibitor interaction is predominantly electrostatic. The stabilizations and ordering of the polar side-chain containing residues at the interface are also significant.
Figure 7: Thermodynamic changes at interface. (a) The conformational entropy and (b) conformational free energy changes of different types of the interfacial residues of NucA and NuiA are shown using a bar diagram.

The HBM has also been extended to the distributions of water molecules present in the NucA-NuiA interface as indicated by Eqs. 9 and 10, using the pair distribution function $g(r)$ [55] for the water molecules which is the probability of finding the centers of mass of a pair of molecules at given separation $r$. The $g(r)$ for the interfacial water molecules is shown in Figure 8 along with those for the waters present around the binding surfaces of free NucA and free NuiA. The gross behavior of these water distributions is the same in all three cases, except the differences in the peak value. The maximum peak value is observed in the complex, indicating high localization of the water molecules at the interface due to tight inhibitor-binding. Such localization leads to a substantial drop of entropy ($T\Delta S_{w} \approx -10.2$ kJ mol$^{-1}$ for about 300 interfacial waters) with respect to the free components. However, the free energy cost turns out to be negligible ($\Delta G_{w} \approx -0.6$ kJ mol$^{-1}$), indicating entropy dominated contribution of water to the interfacial thermodynamics.
6. Prediction of protein function

The connection between conformational states and the functionalities of metalloproteins has been illustrated through microscopic conformational thermodynamics [56]. A specific instance is considered where metal ions induce conformational changes in \(\alpha\)-Lactalbumin (aLA), a Ca\(^{2+}\) ion binding metalloprotein, expressed in the mammary glands during lactation [57]. It participates in a variety of cellular functions, like lactose synthesis [58] and fatty acid binding [59].

As per the crystal structure data [60], the bovine aLA comprises of an \(\alpha\)-helical domain and a small \(\beta\)-sheet domain separated by a cleft [61]. The metal ion free native-apo structure of aLA is stable under neutral pH and physiological salt concentration. Calorimetric measurements [62, 63] suggest that apo-aLA binds to physiologically abundant Ca\(^{2+}\) at the metal binding loop connecting the two domains, resulting in metal ion bound holo structure. Fluorescence and Circular Dichroism studies show that the metal ions provide stability to native-apo aLA against heat or other denaturing agents [64-67] along with change in conformational state by narrowing of the interfacial cleft. The Ca\(^{2+}\)-aLA forms complex with \(\beta\)1,4-galactosyltransferase (\(\beta\)4GalT), to promote lactose synthesis [58]. The crystal structure of aLA complexed with \(\beta\)4GalT has been reported [58, 68]. Another interesting property of aLA is its ability to form molten globule (MG) conformation [61] in acidic pH upon removal of Ca\(^{2+}\). The MG state has fluctuating tertiary structure and binds to various fatty acids, like oleic acid. It has been identified that the oleic acid is the cytotoxic factor of XAMLET, whereas the MG-aLA acts as its carrier [69, 70]. The binding sites for fatty acid are largely unknown, since the MG state cannot be crystallized. Although the native-aLA cannot bind to oleic acid, it is still inconclusive whether divalent metal ion bound native Ca\(^{2+}\)-aLA can bind to oleic acid [70-74]. Since at physiological condition, the pH is neutral and there is abundance of Ca\(^{2+}\), it is of paramount interest to understand the role of these ions in fatty acid binding to aLA, particularly in the context of cytotoxic activities.

The histograms of dihedral angles of the protein, generated from all-atom molecular dynamics simulations, are used to compare the conformational thermodynamics between the holo Ca\(^{2+}\)-aLA and native apo-aLA. We relate the thermodynamics to the functionality of the protein: The residues having instability in conformational free energy (\(\Delta G_{\text{conf}}^{\text{inst}} > 0\)) in the holo form compared to the native-apo state of the protein, are likely to participate in further binding event to attain thermodynamic stability by reducing the free energy. The reduction in free energy is further aided by disordering (\(T\Delta S_{\text{conf}}^{\text{inst}} > 0\)). We test this approach for the case of \(\beta\)4GalT binding to Ca\(^{2+}\)-aLA where the binding residues are known from crystal structure data. Among the binding residues, the C-terminal residues like aspartate (D) 116, glutamine (Q) 117, tryptophan (W) 118 and leucine (L) 119 are destabilized and disordered and can dock \(\beta\)4GalT onto Ca\(^{2+}\)-aLA.

![Figure 8: The oxygen-oxygen g(r) of the interfacial water molecules in the complex and the waters around the binding regions of NucA and NuiA.](image-url)
Figure 9: The conformational thermodynamic changes (a) at the β4GalT binding residues and (b) the basic and the hydrophobic residues which can aid in oleic acid binding to Ca\(^{2+}\)-aLA compared to the native state.

Fatty acids being amphipathic in nature with negatively charged carboxylate (COO\(^-\)) head-groups and long hydrophobic tail, the complexation of fatty acids with aLA is driven by both electrostatic and hydrophobic interactions [69, 75]. We apply similar analysis to oleic acid binding and predict that the Ca\(^{2+}\)-aLA complex can bind to oleic acid through the basic histidine (H) 32 of A2 helix and the hydrophobic residues, namely, isoleucine (I) 59, W60 and I95 of the interfacial cleft. We find by docking analysis that the negatively charged head group of oleic acid is stabilized through electrostatic interactions with the side chain imidazole ring of the basic H32 (separation ~ 5.0 Å). The long hydrophobic tail of oleic acid extends towards the interfacial cleft. This long tail is further stabilized through hydrophobic interactions with W104, I95, I59 and W60. The functional sites are often conserved through evolution in a protein family. Multiple sequence alignment of thirty-six different species of aLA shows that the basic H32 along with the hydrophobic residues like W60, I95 and W104 are highly conserved across the aLA sequences.

7. Conclusion
To summarize, the thermodynamic changes in biomacromolecular conformations can be extracted from the distributions of the dihedral angles using the HBM from a single set of simulations in the relevant conformational states. This method is sufficiently fast and accurate to be applicable large macromolecular complex like protein-protein, protein-DNA or protein-ligand complexes. The detailed thermodynamic analysis of the binding regions would enable us to identify the prime spots of binding. The functional activities coupled to the protein conformations can be extracted from the destabilized residues due to the changes in conformation. The HBM may be useful to: (1) compare the relative stability and ordering between different states and thus selecting thermodynamically plausible state to supplement homology modelling in structurally unknown protein[76]; (2) provide functional insight following changes in conformational states and hence, aiding in rational drug designing [77] and a variety of nano-biotechnological applications [78]. The dynamical information extracted from the time series of the dihedral fluctuations may provide useful insight to dynamic phenomena associated with bio-macromolecular interactions, including signal transduction [79].
Acknowledgements
AD thanks Council of Scientific and Industrial Research, India and SS thanks the University Grants
Commission, India for research fellowships. MG and JC thanks the Department of Science and
Technology, India for funding.

References
[1] Thirumalai D, Liu Z, O’Brien EP, Reddy G, *Curr Opin Struct Biol*, 2013, 23, 22–29.
[2] Marlow MS, Dogan J, Frederic KK, Valentine KG, Wand AJ, *Nature Chem Biol*, 2010, 6, 352-358.
[3] Soukup JK, Soukup GA, *Curr Opin Struct Biol*, 2004, 14, 344–349.
[4] Hall WP, Modica J, Anker J, Lin Y, Mrksich M, et al., *Nano Lett*, 2011, 11, 1098–1105.
[5] Broxha RD, Lopez MM, Vogel HJ, Makhadadze GI, *J Biol Chem*, 2001, 276, 14083–14091.
[6] Schafer H, Mark AE, Gunsteren WFv, *J Chem Phys*, 2000, 113, 7809-7817.
[7] Baron R, Hu’enberger PH, McCammon JA, *J Chem Theory Comput*, 2009, 5, 3150–3160.
[8] Chang C-E, Chen W, Gilson MK, *J Chem Theory Comput*, 2009, 5, 3150–3160.
[9] Gromiha MM, Fukui K, *J Biomol Struct Dyn*, 2011, 51, 721–729.
[10] Trbovic N, Cho J-H, Abel R, Friesner RA, Rance M, et al., *J Am Chem Soc*, 2009, 131, 615–622.
[11] Killian BJ, Kravitz JY, Gilson MK, *J Chem Phys*, 2007, 127, 024107.
[12] Mendez R, Bastolla U, *Phys Rev Lett*, 2010, 104, 228103.
[13] Gilson MK, *Chem Phys Lett*, 2009, 131, 14610–14611.
[14] Li D-W, Showalter SA, Bru’schweiler R, *Phys Rev Lett*, 2010, 102, 118108.
[15] Ytreberg FM, Zuckerman DM, *J Chem Phys*, 2006, 124, 104105.
[16] Cecchini M, Krivov SV, Spichyta M, Karplus M, *J Phys Chem B*, 2009, 113, 9728–9740.
[17] Park S, Lau AY, Roux B, *J Chem Phys*, 2008, 129, 134102.
[18] Spichyta M, Cecchinini M, Karplus M, *J Phys Chem Lett*, 2010, 1, 1922–1926.
[19] Das A, Chakrabarti I, Ghosh M, *Biophys J*, 2013, 104, 1274-1284.
[20] Reif F, *Fundamentals of statistical and thermal physics*, 1965 (Auckland: McGraw-Hill).
[21] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al., *J Comp Chem*, 2005, 26, 1781-1802.
[22] Humphrey W, Dalke A, Schulten K, *J Mol Graphics*, 1996, 14, 33-38.
[23] Brooks BR, III CLB, Mackerrall AD, Nilsson L, Petrella RJ, et al., *J Comp Chem*, 2009, 30, 1545-1615.
[24] Branden C, Tooze J, *Introduction to Protein Structure*, 1999 (New York: Garland Publishing, Inc.).
[25] Babu YS, Sack JS, Greenhough TJ, Bugg CE, Means AR, et al., *Nature*, 1985, 315, 37-40.
[26] Nisell MR, Thulin E, Fagan PA, Forsén S, Chazin WJ, *Proc Natl Acad Sci U S A*, 2001, 98, 198–205.
[27] Kuboniwa H, Tjandra N, Grzesiek S, Ren H, Klee C, et al., *Nature Struct Mol Biol*, 1995, 2, 768-776.
[28] Zhang M, Tanaka T, Ikura M, *Nature Struct Mol Biol*, 1995, 2, 758 - 767.
[29] Wintride PL, Privalov PL, *J Mol Biol*, 1997, 266, 1050-1062.
[30] Zhang M, Vogel HJ, *J Biol Chem*, 1994, 269, 981-985.
[31] Osawa M, Tokumitsu H, Windell MB, Kurihara H, Orita M, et al., *Nature Struct Mol Biol*, 1999, 6, 819 - 824.
[32] Kuczerka K, Kursula P, *J Biomol Struct Dyn*, 2012, 30, 45-61.
[33] Das A, Chakrabarti I, Ghosh M, *Chem Phys Lett*, 2013, 581, 91-95.
[34] Linse S, Helmersson A, Forsen S, *J Biol Chem*, 1991, 266, 8050-8054.
[37] Wu G, Gao Z, Dong A, Yu S, *Int J Biol Macromol*, 2012, **50**, 1011–1017.

[38] Ye Y, Lee H-W, Yang W, Shealy S, Yang JJ, *J Am Chem Soc*, 2005, **127**, 3743-3750.

[39] Kobayashi C, Takada S., *Biophys J*, 2006, **90**, 3043–3051.

[40] Lepsik M, Field MJ, *J Phys Chem B*, 2007, **111**, 10012-10022.

[41] Das, A., Chakrabarti, J. and Ghosh M, *Mol Biosyst*, 2014, **10**, 437-445.

[42] Ghosh, M.; Meiss, G.; Pingoud, A.; London, R.E and Pedersen, L.C, *J. Biol. Chem.*, 2007, **282**, 5682–5690.

[43] Muro-Pastor, A. M., Flores, E., Herrero, A. & Wolk, C. P., *Mol. Microbiol.*, 1992, **6**, 3021-3030.

[44] Muro-Pastor, A. M., Herrero, A., Flores, E., *J. Mol. Biol.*, 1997, **268**, 589-598.

[45] Meiss, G., Franke, I., Gimadutdinow, O., Urbanke, C., and Pingoud, A., *Eur. J. Biochem.*, 1998, **251**, 924–934.

[46] Rangarajan, E.S. and Shankar, V., *FEMS Microbiol. Rev.*, 2001, **25**, 924–934.

[47] Hsia, K.C., Li, C.L. and Yuan, H.S., *Curr. Opin. Struct. Biol.*, 2005, **15**, 126-134.

[48] Fox K.R. and Waring, M.J., *Biochim. Biophys. Acta.*, 1987, **909**, 145-155.

[49] Krupp G. and Gross, H.J., *Nucleic Acids Res.*, 1979, **6**, 3481-3490.

[50] Varecha, M., Amrichova, J., Zimmermann, M., Ulman, V., Lukasova, E. and Kozubek, M., *Apoptosis* 2007, **12**, 1155–1171.

[51] Korn, C.; Meiss, G.; Gast, F.U.; Gimadutdinow, O.; Urbanke, C. and Pingoud, A., *Gene*, 2000, **253**, 221–229.

[52] Meiss, G., Gimadutdinow, O., Haberland, B., Pingoud, A., *J. Mol. Biol.*, 2000, **297**, 521-534.

[53] Ghosh, M.; Meiss, G.; Pingoud, A.; London, R.E and Pedersen, L.C, *J. Biol. Chem.*, 2005, **280**, 27990–27997.

[54] Kirby, T.W.; Mueller, G.A.; DeRose, E.F.; Lebetkin, M.S.; Meiss, G.; Pingoud, A. and London, R.E, *J. Mol. Biol.*, 2002, **320**, 771–782.

[55] Hansen, J. P.; and McDonald, I. R. *Theory of Simple Liquids*, 2006 (Academic: San Diego).

[56] Sikdar, S.; Chakrabarti, J.; Ghosh, M. *Mol Biosyst.*, 2014, **10**, 3280-3289.

[57] Anderson, P. J.; Brooks, C. L.; Berliner, L. J. *Biochemistry*, 1997, **36**, 11648-11654.

[58] Ramakrishnan, B.; Shah, P. S.; Qasba, P. K. *J Biol Chem*, 2001, **276**, 37665-37671.

[59] Cawthern, K. M.; Narayan, M.; Chaudhuri, D.; Permyakov, E. A.; Berliner, L. J. *J Biol Chem*, 1997, **272**, 30812-30816.

[60] Ostrovsky, A. V.; Kalinichenko, L. P.; Emelyanenko, V. I.; Klimanov, A. V.; Permyakov, E. A. *Biophys Chem*, 1988, **30**, 105-112.

[61] Permyakov, E. A.; Kalinichenko, L. P.; Morozova, L. A.; Yarmolenko, V. V.; Burstein, E. A. *Biochem Biophys Res Commun*, 1981, **102**, 1-7.

[62] Hiraoka, Y.; Sugai, S. *Int J Pept Protein Res*, 1984, **23**, 535-542.

[63] Hiraoka, Y.; Sugai, S. *Int J Pept Protein Res*, 1985, **26**, 252-261.

[64] Spolaore, B.; Pinato, O.; Canton, M.; Zambonin, M.; Polverino de Laureto, P.; Fontana, A. *Biochemistry*, 2010, **49**, 8658-8667.

[65] Svennson, M.; Fast, J.; Mossberg, A. K.; Duringer, C.; Gustafsson, L.; Hallgren, O.; Brooks, C. L.; Berliner, L.; Linse, S.; Svanborg, C. *Protein Sci*, 2003, **12**, 2794-2804.

[66] Stanciuc, N.; Aprodu, I.; Raseanu, G.; Bahrim, G. *Eur Food Res Technol*, 2013, **236**, 257-266.

[67] Barbana, C.; Perez, M. D.; Sanchez, L.; Dulgalarondo, M.; Chobert, J. M.; Haertle, T. *Int Dairy J*, 2006, **16**, 18-25.
[73] Knyazeva, E. L.; Grishchenko, V. M.; Fadeev, R. S.; Akatov, V. S.; Permyakov, S. E.; Permyakov, E. A. Biochemistry, 2008, 47, 13127-13137.

[74] Svanborg, C.; Agerstam, H.; Aronson, A.; Bjerkvig, R.; Düringer, C.; Fischer, W.; Gustafsson, L.; Hallgren, O.; Leijonhuvud, I.; Linse, S.; Mossberg, A. K.; Nilsson, H.; Petersson, J.; Svensson, M. Adv Cancer Res, 2003, 88, 1-29.

[75] Tolin, S.; De Franceschi, G.; Spolaore, B.; Frare, E.; Canton, M.; Polverino de Laureto, P.; Fontana, A. FEBS J, 2010, 277, 163-173.

[76] Punta, M.; Coggill, P. C.; Eberhardt, R. Y.; Mistry, J.; Tate, J.; Boursnell, C.; Pang, N.; Forslund, K.; Cerić, G.; Clements, J.; Heger, A.; Holm, L.; Sonnhammer, E. L.; Eddy, S. R.; Bateman, A.; Finn, R. D. Nucleic Acids Res, 2012, 40, D290.

[77] Mavromoustakos, T.; Durdagi, S.; Koukoulitsa, C.; Šimcic, M.; Papadopoulos, M. G.; Hodoscek, M.; Grđadolnik, S. G. Curr Med Chem, 2011, 18, 2517-2530.

[78] Salata, O. J Nanobiotechnology, 2004, 2, 3.

[79] Berg J.M., Tymoczko J.L., Stryer L., Signal-Transduction Pathways: An Introduction to Information Metabolism, Biochemistry 2002 (New York: W H Freeman) Chapter 15