PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

PCCP

ARTICLE

Predicting long term cooperativity and specific modulators of receptor interactions in human transferrin from dynamics within a Single Microstate

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Haleh Abdizadeh[a] and Canan Atilgan[a*]

Transferrin (Tf) is an essential transport protein circulating iron in the blood and delivering to tissues. It displays highly pH dependent cooperativity between the two lobes each carrying an iron, and forms a tight complex with the receptor during endocytosis and recycling back to the serum. We explore short-term dynamics within selected microstates of human Tf to identify functional information relevant to long-term dynamics. While the variance-covariance matrix delineates cooperativity between the domains of Tf at serum pH which is lost at endosomal pH, its decomposition does not bring about additional information. We employ perturbation-response scanning (PRS) to extract essential components that contribute to a pre-selected conformational change. Since large-scale motions may require key residues to mediate correlated motions between different regions of the protein, we use PRS to predict those involved in the conformational transitions between the iron bound and free hTf. Physiological and endosomal conditions are mimicked to identify critical residues for holo→apo and apo→holo transitions. Iron binding motions are mainly orchestrated by residues at the synergistic anion uptake sites, a finding also corroborated by additional molecular dynamics simulations where these sites are perturbed by docking the anion. Iron release is not readily accessible at serum pH, while at endosomal pH single residue perturbations on any residue encourage the large transition that involves a complex twisting of the two domains relative to each other, simultaneously opening both lobes. pH dependent change in the dynamics is traced to the altered electrostatic potential distribution along the surface. Examination of local dynamics in hTf-receptor pair reveals cooperativity in the quaternary structure and explains resistance to iron release in the complex. Meanwhile, analysis of hTf complex with a bacterial receptor that has evolved to sequester iron identifies two regions contacting rapidly evolving residues that mechanically manipulate dissociation from the pathogen.

Introduction

Local fluctuations on sub-micro second time scales lead to the structural stability and diverse functionality observed in proteins on millisecond and slower time scales [1, 2]. Functionality of the proteins, in return, is based on a particular tertiary structure and the ability to undergo conformational switching as a response to environmental stimuli such as ionic strength or pH changes, mutations and interactions with other biomolecules[1, 3]. Therefore, unraveling the mechanistic communication involved during dynamic interconversion between multiple functional states on long time scales may help explain molecular and chemical processes underlying protein function[4-6]. Different, poorly populated intermediate conformational sub-states between initial and target structures are closely spaced in their energetic basins and many of them are accessible under varying external conditions. Redistribution of the population regulates protein function and is crucial for the survival of organisms [7].

With the advent of coarse-grained methods, allostery, with its broad physical definition has been investigated by a wealth of techniques and allosteric regulation and its effects on conformational distribution have been quantified. Normal mode analysis (NMA) and protein network constructing models such as anisotropic network model (ANM) [8] and Gaussian network model (GNM) [9] have been very influential in extracting residue clusters correlating protein motions to functionality. ANM and GNM are examples of network based models which are convenient tools for identifying intermediate processes leading to protein folding and function [10-12]. Other examples applying harmonic approximation include calculation of ligand binding effects on protein reaction rates and conformational space distribution before and after complex formation [13]. These approaches provide an enhanced accuracy

[a,*] Faculty of Engineering and Natural Sciences, Sabanci University, Tuzla, Istanbul 34956, Turkey, Email: canan@sabanciuniv.edu

beyond that of classic NMA analyses to identify allosteric regulatory sites [14].

We have introduced Perturbation-Response Scanning (PRS) as an approach to determine single residues playing active roles in manipulating the conformational change in proteins [4, 15-17]. It requires two distinct conformations of a protein, determined, e.g. by x-ray crystallography, as input; it relies on linear response theory (LRT) to relate fictitiously introduced external forces acting on a protein to the perturbed positions of the residues [18]. In PRS, one performs a residue-by-residue scanning of one of the conformations by applying forces in a variety of directions and then records the subset of residues/forces that invoke conformational changes closest to that known due to the two endpoint conformers. One can thus map the regions on the protein surface whose perturbation leads to the expected conformational change. Besides mapping active site residues that are prime regions for e.g. drug targeting, this approach also has the potential of pointing out allosteric locations. For example, we have studied the proteins calmodulin [4] and ferric binding protein (FBP) [15] via PRS. By mutating those residues that were implicated in allosteric communication, we later verified through classical molecular dynamics (MD) simulations that they affect the conformation distributions [19, 20]. In another study, we have performed PRS on subtilisin in complex with its inhibitor to pinpoint hot residues involved in catalytic mechanism and stability of the enzyme [17]. PRS has also been used in the conformation generation step of a flexible docking scheme for exploring protein-ligand interactions [21]. In a similar spirit, a method based on distance constraint model to impose constraints on the torsional degrees of freedom of the protein to mimic a hypothetical ligand-binding situation has been successful in predicting protein functions [22, 23].

Here, we study human serum transferrin (hTf) from the transferrin family of glycoproteins, a major iron transport system in vertebrates and invertebrates. Mutants or protonation events at distinct sites might help overcome energy barriers in conformational shifts necessary for metal release/uptake processes of hTf. Ligand binding and release in hTf requires significant conformational changes that regulate its function [24]. hTf is a 679 residue protein that displays a remarkable affinity toward iron (K_d~10^{-22} M [25]) and yet readily releases it at the reduced pH of the endosome in the presence of its receptor [26]. We shall focus on the two reported states of the full length hTf as our model system [27, 28] as well as its states bound to receptors from human and bacteria [27, 29].

hTf folds into the N-lobe and C-lobe domains which bear equal metal binding configurations but overall structural dissimilarities resulting in asymmetrical iron binding/release kinetics 30. Modifications in the hTf conformation after iron exchange may be important for the recognition of the protein by its receptor (TfR) prior to iron delivery via receptor mediated endocytosis 31. TfR interacts strongly (Kd~0.5 μM) and rapidly (50 μs) with the C-lobe of mono- and diferric hTf via its helical domain 32-34 whereas the protease-like domain has weaker interactions (Kd~2.3 nM) with the N-lobe via slow kinetics (8000 s) [29, 33, 34]. The lower pH of the endosome, presence of salts and chelators, and interaction with TfR activate iron release from hTf. Protonation of carbonate is a prerequisite to iron release and iron loss from the N-lobe occurs 17-fold faster in the absence of TfR [33, 35]. Further protonation reactions trigger conformational changes leading to iron loss from the C-lobe. However, the reverse iron release path is also probable 31. After equilibration in the endosome, apo-hTf remains bound to TfR; together they go back to the extracellular membrane where hTf finally dissociates from TfR and is recycled in neutral bulk medium. Subsequent to the iron release, it is reduced to Fe^{2+} by members of the STEAP family of metalloreductases [36, 37]. An alternative pathway might be provided by TfR to drive the reduction of Fe^{3+} in transferrin. Dhungana et al. suggest that physiological reduction of Fe^{3+} might occur while iron is still bound to hTf in the presence of TfR [38, 39]. Therein, this redox potential shift is attributed to the different binding affinity of TfR towards apo and holo hTf in the transferrin cycle.

hTf is also the target of a highly developed receptor mediated iron uptake system used by pathogenic bacteria [40]. The direct iron piracy process initially involves binding of an outer membrane protein called transferrin binding protein A (TbpA) and the co-receptor TbpB to hTf [27]. Our aim is to effectively represent the mechanistic coupling between different conformations of hTf at the atomic level. We examine the dynamics within selected local minima to decipher kinetically hot residues involved in the shift of hTf between its apo and holo forms, and the factors contributing to the pH sensitivity of this shift. Furthermore, since conformational changes of hTf seem to be crucial for recognition by receptors, we also study the hTf-TfR and hTf-TbpA complexes to unravel the molecular basis of hTf function.

Experimental

Materials

Protein structures. hTf structures we use have been described in detail elsewhere [41]. Briefly, the overall structure of bilobal hTf exhibits two homologous halves termed N- and C-lobe [42]. It is further divided into N_1-(1-95 and 247-331), N_2-(96-246), C_1-(339-425 and 573-679) and C_2-subdomains (426-572) (Fig. 1b). Intra/inter-lobe interactions and an unstructured peptide linker (332-338) stabilize the tertiary structure of the protein. X-ray crystal structures of hTf display two different conformations, the iron-free fully open apo [28] and the fully closed iron-loaded holo conformation with two Fe^{+3} ions located in a cleft formed between the subdomains [27]. These structures are denoted by A and H, respectively. Further details are given in the supplementary text S1. We also study the complexes of hTf with its cell surface receptor in human (TfR) and a bacterial protein TbpA that has evolved to sequester iron from hTf. The former has the PDB code 1SUV [29] and the RMSD with iron loaded, unbound hTf is 4.7 Å [27]. The TbpA bound structure has PDB code 3V8X [27] and lacks iron and

synergistic anion. Its RMSD from free apo and holo hTf is 3.0 and 6.9 Å, respectively.

Methods

MD simulations. In previous studies, we have shown that the ideal length of MD simulation to properly represent the C matrix (see eq. 1, below) of a protein to be used in LRT is on the order of 20 ns [43]. Four sets of MD simulations for holo and apo systems have been carried out. We study holo structure at two different pH values. We label the holo form at serum (~7.4) and endosomal pH (~5.6) as $\mathbf{H^s}$ and $\mathbf{H^e}$, respectively. Apo system, labeled **A**, and apo hTf in the presence of carbonate anion is labeled as $\mathbf{A^C}$, are only simulated at serum pH. We have at least two replicas of each MD run, simulation lengths ranging from 50 ns to 500 ns (text S2), and the total MD simulation time is 1.72 μs.

To test the effect of carbonate binding on the hTf conformation ($\mathbf{A^C}$ system), we dock carbonate based-on the coordination of equivalent residues of the bacterial apo-FBP which does have its synergistic anion phosphate in the crystal structure (PDB code 1D9V [44]). We use comparison of the carbonate-binding sites on either lobe of the holo structure from [27] to the holo FBP in determining equivalent sites as well as experimental papers that determine carbonate anchoring residues [45, 46]. Accordingly, we choose residues 95, 120, 124-129 and 188 in the N-lobe and residues 426, 452,456-459 and 517 in the C-lobe as our primary carbonate-binding location for the docking procedure. We have used AutoDock 4.2 for this purpose [46, 47].

Further details on the MD simulations, AutoDock parameters and selection of charge states are given in supplementary text S2. We note that in reference [41], we have verified that our choice of protonation states reflect the experimentally observed dynamics of hTf.

PRS method. We analyze the conformational change between two different conformers of hTf by applying PRS to two available crystal structures. We examine the inter-conversion propensity between two microstates of the protein by employing fictitious external forces. Herein, different microstates are defined as distinct states separated by a well-defined energy barrier. The detailed theory of PRS has been laid-out in previous studies [15, 16], and is summarized in supplementary text S3. In brief, the ligand free state of a protein may be described by a perturbation of the Hamiltonian of the bound state, or vice versa. Under LRT, the shift in the coordinates due to unbinding is approximated by [15, 48]:

$$\Delta\mathbf{R}_1 = \langle\mathbf{R}\rangle_1 - \langle\mathbf{R}\rangle_0 \cong \frac{1}{3k_BT}\langle\Delta\mathbf{R}\Delta\mathbf{R}^T\rangle_0\Delta\mathbf{F} = \frac{1}{3k_BT}\mathbf{C}\Delta\mathbf{F} \quad (1)$$

where the subscripts 1 and 0 denote perturbed and unperturbed configurations of the protein, k_B is the Boltzmann constant, T is temperature. **ΔF** vector contains the components of the externally inserted force vectors on the selected residues. PRS requires the construction of the variance-covariance matrix, **C** (eqn (1)) which may be obtained from either the atomic coordinate trajectories of MD simulations of suitable length [43], or by imposing the approximation of harmonic springs between pairs of interacting atoms. In this work, the former has been done for the isolated hTf structures, while the latter is employed for the prohibitively large receptor-hTf complexes. PRS technique relies on repeating eqn (1) by scanning the residues of the protein one-by-one and focusing further on those perturbations that best mimic the conformational change **ΔR**.

To assess the quality of the predicted displacements of all residues resulting from a force applied on selected residue *i*, we use the correlation coefficient between the predicted and experimental displacements, averaged over all the affected residues, *k*:

$$C_i = \frac{\sum_{k=1}^{N}[(\Delta R_k)^i - (\overline{\Delta R})^i](\Delta S_k - \overline{\Delta S})}{(N-1)\sigma_R\sigma_S} \quad (2)$$

where the overbar indicates the average, ΔS_k are the displacements between the initial and the target forms obtained from the PDB structures, σ_S and σ_R are the corresponding root mean squared values [49]. A value close to 1 implies good agreement with the experimental conformational change, while a value of zero indicates lack of correlation between experimental and theoretical findings.

Modal analysis. In NMA (or ANM when the **C** matrix is constructed via a cut-off distance), a direct analysis of the C matrix is done, as opposed to PRS which uses it as a Kernel that filters displacements due to inserted perturbations (eqn (1)). **C** is decomposed as $\mathbf{C^{-1}=U\Lambda^{-1}U^T}$, where $\mathbf{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues of **C**, and **U** is the orthonormal matrix whose columns, **u***j*, are its eigenvectors. To assess the quality of the modes obtained by **C**, we use the correlation by replacing the displacement vector upon perturbation, **ΔR**, with the normal vector, **u***j* in eqn (2) [49].

Results and Discussion

Proteins of the transferrin family utilize conformational shifts between different states occurring before or after iron capture/release to function. It has been well established that chelators [50], salt [51], pH [52] and presence of TfR [33] affect the functioning of hTf in different states. The coupling between local and global motions governs the equilibrated sequence of events in iron regulation. For example, iron release from holo hTf, whose iron contacting residues and overall structure is displayed in Fig. 1, is accompanied by a major conformational change. It involves the two jaws enclosing the binding site moving apart through a hinge motion, accompanied by a 55-65° reorientation of the two subdomains of either lobe [53].

Since it is a computationally prohibitive task to simulate the long-term dynamics of the different conformers of hTf in various environments, we delineate the extent to which computationally inexpensive methods may uncover the long-term response of the protein to environmental changes. Point mutations introduced on either lobe of hTf have been shown to have unique consequences on the function of hTf (these include iron coordinating residues, see e.g. [54, 55], second shell

residues, see e.g. [56] and residues at the hTf-TfR interface [57]). Here we seek regions potentially controlling iron binding/release without *a priori* assumptions on their location on the protein.

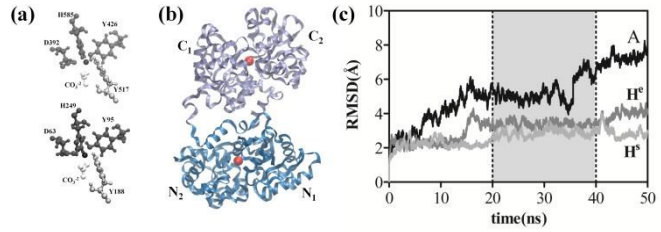

Fig. 1 (a) Iron (black spheres) coordinating residues in the N- and C-lobes. Residues in gray coordinate iron from N_1/C_1 and residues in silver are the sole coordinating residues from N_2/C_2. Carbonate is displayed in white. (b) Overall fold of holo hTf; N domain in blue, C-domain in gray. Subdomains are labeled. (c) RMSD of the three states of hTf; A, H^e and H^s are in black, gray and silver, respectively. The 20 to 40 ns portion of MD trajectory used for PRS analysis is shaded in gray. We find that the A system displays large fluctuations throughout the MD simulations.

There is a rich literature on the biochemistry of hTf. In particular, binding and release of iron and other metals as well as a variety of synergistic anions in the presence or absence of its receptor have been quantified (see [41] and references cited therein). Herein, we first study hTf in the absence of TfR. Conveniently, the methods used in this study require short enough simulations that may be achieved in moderate computational times. We first build the variance-covariance (**C**) matrices from the 20 to 40 ns piece of the MD trajectories (eqn (1)). We note that we have generated longer and multiple trajectories for each of the systems studied here, and we have repeated the analyses described in this work on all equilibrated 20 ns pieces. We have verified that the main results we present are independent of the choice of the piece. The only requirement is that the **C** matrix represents the same microstate within the time window of observations; we have previously shown that 20 ns is adequate for this purpose [43]. Too long time windows might allow sampling neighboring microstates in which case the harmonic assumption imposed on the $<\Delta\mathbf{R}\Delta\mathbf{R}^T>_0$ matrix in eqn (1) fails. We verify that after 20 ns the MD simulations sample the conformational substates around an equilibrium structure. The RMSD of the trajectory for the three systems are displayed in Fig. 1c; note that although we have produced 300 and 200 ns trajectories of the holo, endosomal pH ($\mathbf{H^e}$) 200 ns ones for apo (**A**), and 60 ns trajectories for the holo, serum pH ($\mathbf{H^s}$) system, we present the representative initial 50 ns portion of one of these in Fig. 1c to focus on the 20–40 ns region of interest. The RMSD profiles for the 200 ns trajectories of the $\mathbf{H^s}$, $\mathbf{H^e}$ and **A** systems may be viewed in ref. [41] and that of the $\mathbf{A^C}$ system is displayed in Fig. 2a. We find that the $\mathbf{H^s}$ system remains in the initial state throughout the simulation. $\mathbf{H^e}$ has a jump at 15 ns and then equilibrates. While the **A** form displays larger RMSD which may implicate a non-equilibrated structure, in a recent study we have shown that due to the absence of iron, it has an intrinsic propensity to sample a single wide microstate sweeping a large RMSD between 2–8 Å. The slowest mode that accounts for nearly 50% of all motions in each of four different 50 ns chunks of this system displays the same motion; see Fig. S4 in ref. [41].

Apo→holo transition requires perturbing synergistic anion binding sites as an intermediate step. The formation of a stable hTf-iron complex involves a sequence of events [24]. Thus, the iron bound state of hTf may not be assumed to be the result of a direct transition between apo and holo forms. Nevertheless, we study the A → H transition by applying the PRS analysis to these end points. Note that in this transition we cannot specify the difference between serum versus endosomal pH structures, since their average conformation is the same.

We first note that there is no single dominant mode of motion that describes the majority of the observed conformational change that involves a complex twisting of the two domains relative to each other, accompanied by closing of both lobes. The total correlations (see Modal Analysis section under Methods) between the displacements defined by the eigenvectors for the top five modes and the A → H transition is 0.49, 0.56, 0.25, 0.29 and 0.38. We shall see below, however, that select single residue perturbations provide a much better descriptor of the conformational change than any single mode.

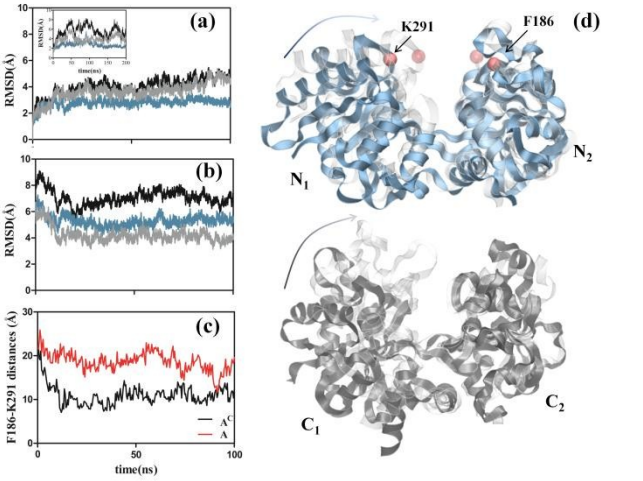

Fig. 2 (a) RMSD of the $\mathbf{A^C}$ system from its initial structure. RMSD of the protein, the N-lobe and the C-lobe is displayed in black, blue and gray, respectively. N-lobe is equilibrated in the first 5 ns and does not display large fluctuations. However, RMSD of the C-lobe increases up to ca. 5 Å and affects the overall structure. In the inset, RMSD of the apo hTf in the absence of the carbonate anion (**A** system) is shown. Equilibrated fluctuations of the apo hTf in the presence of carbonate anion in comparison to the large fluctuation in its absence show how carbonate stabilizes the apo form of hTf. **(b)** RMSD of $\mathbf{A^C}$ system from the holo hTf structure. Color coding is

the same as in (a). The initial large RMSD decreases with time, demonstrating the closing effect of carbonate binding to the apo conformation. (c) Distance profile between residues F186 and K291, for the $\mathbf{A}^{\mathbf{C}}$ (black) and **A** (red) systems, respectively. Upon entry of the synergistic anion, the distance between the tip residues residing on the loops enclosing the cavity decreases from 18.2±2.1 Å to 10.6±1.6 Å (averaged after the first 20 ns of each trajectory). (d) Superposition of the initial (translucent) and final (transparent) structure of the $\mathbf{A}^{\mathbf{C}}$ system at 100 ns for the N-domain (blue) and C-domain (gray); best fitting of the structures is to the N_2 and C_2 subdomains, respectively, so that the closing of the other subdomain over the cavity (shown by the arrows) is better visualized. C_α atoms of F186 and K291 whose separation is monitored in part (c) are displayed as red spheres.

In Table 1, we list the residues that are accentuated by the PRS analysis. These are the single residues whose perturbation selects the coordinate change towards the target conformation. As selection criteria of residues that are effective in directing the protein towards alternative conformations, we take several approaches depending on the distribution of the maximum of the C_i values, C_i^{max}, obtained from the 200 perturbations and calculated through eqn (2). We first list C_i^{max} in ascending order: (i) If there is a sharp decrease in the correlation values, we list the top residues until that gap. (ii) If there is a smooth decrease in the correlation values, we list the top 10 residues. We also check the location of the residues that do not survive this selection cutoff. In this work, we have found that the next 10 residues are spatial neighbors of the listed ones. Therefore, the interpretation of the results are not sensitive to the cutoff point. (iii) If there are too many residues with $C_i^{max} > 0.8$, we do not list any residues as "identified." Conversely, we interpret this behavior indicative of a natural propensity of the protein to interconvert to the target conformation. (iv) If all residues have $C_i^{max} < 0.6$, we interpret this behavior as indicative of resistance to interconvert to the target conformation. However, we still list the top residues according to criterion (i) or (ii). As an example, in Fig. S1, we display the resulting maximum correlation coefficients, C_i^{max}, between the calculated and experimental data for every residue of **A** → **H** and $\mathbf{H}^{\mathbf{s}}$ → **A** transitions (eqn (2)); these exemplify cases (i) and (iv), respectively. Note that each point on these figures is the result of the comparison of the displacements of the 674 residues in response to a perturbation applied at the selected residue, *i*, such as that obtained from the correlation between the experimental displacements and predicted displacements upon force application.

Of the 16 residues listed in Table 1 and having significant C_i^{max} values, five (D63, S189, H451, T452, G516) are in carbonate binding regions. Thus, perturbing these residues (e.g. by associating with carbonate) might induce displacements that better overlap with the fully iron bound conformation, in conformity with experiments that show the first step towards iron binding is carbonate capture [58]. Thus, the PRS analysis from the 50 ns trajectory lets us make the prediction that perturbations due to the entrance of carbonate into the binding cavity are expected to lead to conformational changes encouraging iron uptake from the medium.

We test this idea by showing that docking carbonate anion to the apo protein followed by all atom MD simulations drives the completely open cleft of each lobe in hTf towards an intermediate with a more compact configuration. In the docking procedure (see Methods) residues 95, 120, 124-127 and 188 in the N-lobe and 452, 456-459 and 517 in the C-lobe are used as contacts; among them are the PRS identifies residues T452 (a direct contact) and S189/H451/G516 (residues neighboring the binding site). Following 100 ns of MD simulations of this $\mathbf{A}^{\mathbf{C}}$ system, we observe that a partially closed conformation is promoted. Such an intermediate is expected since until iron enters the binding site, the iron-binding cleft will not close tightly. We note that both carbonates remain in their respective clefts throughout the 100 ns. In Fig. 2, the RMSD profiles of the $\mathbf{A}^{\mathbf{C}}$ system in comparison to its original structure (Fig. 2a) and to holo hTf (Fig. 2b) are displayed. In the inset to Fig. 2a, the RMSD of the **A** system is also shown from a 200 ns trajectory. We have previously presented results for this simulation and we have verified that it displays large fluctuations, but does not lose its average conformation [41]. In fact, four separate 50 ns chunks of this trajectory as well as an independent 50 ns run have the same average conformation which is at most 2.8 Å from the crystal structure that has 2.7 Å resolution. Each trajectory displays similar variance-covariance matrices with a dominant hinge-bending as its slowest mode of motion. Therefore, at equilibrium, apo hTf in the absence of the synergistic anion displays an extraordinarily wide microstate.

The motions of apo hTf upon the addition of carbonate are reduced, accompanied by a closing over the iron binding cavity (Fig. 2). The presence of carbonate has a stabilizing effect on the system with dampened fluctuations and a net change towards an alternative, intermediate structure, suitable for iron binding. The average RMSD for the last 10 ns of MD simulations are 4.8±0.2, 2.8±0.1 and 4.7±0.2 Å for the full structure, N-lobe and C-lobe, respectively. Fig. 2b displays the RMSD of the $\mathbf{A}^{\mathbf{C}}$ system with respect to holo hTf where the difference between $\mathbf{A}^{\mathbf{C}}$ and $\mathbf{H}^{\mathbf{s}}$ decreases due to the presence of carbonate. The average RMSD during the last 10 ns of MD simulations decreases from initial values of 8.2, 6.4 and 6.6 Å to 6.8±0.3, 5.4±0.2 and 3.8±0.2 Å for the full structure, N-lobe and C-lobe, respectively.

We find that the reduction in RMSD is not merely due to local arrangements accompanying carbonate docking, but are rather global changes that require large motions of the loops closing over the iron binding cleft; e.g. as monitored by the distance between F186 and K291 located on the lids of each subdomains of the N-lobe (Fig. 2c). It reduces from the initial value of 21.5 Å to an average of 18.2±2.1 Å in **A** while it attains an average value of 10.6±1.6 Å in $\mathbf{A}^{\mathbf{C}}$ upon carbonate binding.

Table 1. Summary of PRS results

Initial → target structure	PDB entries	Maximum correlation (eq. 2)	Identified residues	Notes
A → H	2HAV:B→3V83:B	0.75	K4 T5 R7 A59 V60 T61 L62 D63 S189 G198 D201 V251 H451 T452 L497 G516	Residues with C_i^{max} = [0.70-0.75]
H^s → A	3V83:B→2HAV:B	0.52	C498 C523 V525 E526 P549 W550 N553	Residues with C_i^{max} = 0.52
H^e → A	3V83:B→2HAV:B	0.89	All residues	All residues have $C_i^{max} > 0.85$
hTf–TbpA → H	3V8X:B →3V83:B	0.77	S390 V429 A430 R568 K569 V587 R602 Q603 Q605 H606 G609 L621 F631 R632	Residues with C_i^{max} = [0.73-0.77]

The closing motion of carbonate on **A** is presented in more detail by superposition of the $\mathbf{A}^{\mathbf{C}}$ structure at 100 ns on the initial structure (Fig. 2d). In particular, residues 159-189 and 274-295 in the N-lobe and amino acid sequences 343-369, 488-515 and 610-633 in the C-lobe display large motions resulting in the partially closed structure. These residues also have been shown to be highly fluctuating in the **H** system and mark the iron exit pathway [41]. Thus, we show that carbonate binding may act as a perturbation to shift the equilibrium towards the holo form. Presence of carbonate restricts the wide microstate sampled by the **A** system. Iron capture at this step might result in further structural changes to attain full closure of the subdomains.

We have checked the coordination of carbonate at the end of 100 ns. Residues we have originally specified as contact points remain within 6 Å; in addition, iron binding residues 392 and 585 as well as 85 which was identified as an anchor for carbonate in iron bound form [59] is also within 6 Å of the anion following the conformational change, implying the intermediate observed in the simulations is a likely scenario in the iron binding steps of hTf. Thus, carbonate capturing manipulates the overall structure of apo hTf, more so on the C-lobe than the N-lobe since the RMSD profile of the former is more similar to that of the holo form (Fig. 2b). This is consistent with the experimental observations that initial iron binding occurs on the C-lobe and further promotes the binding of the second iron on the N-lobe [60].

D63 mutants affect iron binding propensity. While most residues implicated as controlling the apo→holo transition interact with the synergistic anion in the holo form, the top residues L62 and D63 with C_i^{max}= 0.75 residing on a β-strand with indices 59–63 on the N_1 subdomain (listed in Table 1 along with V251 in direct contact with this strand) are not involved in synergistic ion coordination, but only iron coordination. Direct experimental evidence of their influence on the iron binding propensity of apo hTf has been well provided. This highly conserved sequence [61] is situated in the vicinity of N-lobe binding site and also displays low thermal fluctuations. Of main interest is D63, the only iron-coordinating residue from N_1 subdomain, which coordinates iron through the carboxylate oxygen. D63 provides structure-stabilizing effects in closed form by bridging N_1 and N_2 subdomains [62]. Experimentally, this residue has been perturbed through a series of mutations. D63S mutant shows that shorter serine side chain along with the reduced charge of mutant makes iron coordination less favorable [63]. D63E mutant has the same overall fold as the native holo hTf despite the weakened iron binding affinity; although it maintains iron coordination, the structure of metal binding site is altered [64]. The distorted coordination of iron is followed by the accommodation of the bulk methylene group of the glutamic acid side chain in the binding region and results in the loss of hydrogen bonding with neighboring residues [64].

While iron is initially coordinated to H249 and Y95, domain closure and retaining a closed conformation is maintained by D63. In both D63S and D63E mutants, Fe-H249 and Fe-Y95 distances change and N-lobe gets trapped in a semi-closed form [62]. The ability of D63E mutant to retain iron at lower pH also reduces [64]. Furthermore, the family of mutants, D63S, D63N, D63E and D63A produce altered interaction with synergistic anion and accelerate iron loss to chelators [55]. Thus, for mutants of D63, tight iron binding and apo → holo transition do not take place in a closed, compact form.

As a final note, residues K4, T5 and R7 also emerge as possible remote hot residues affecting apo → holo transition (Table 1). Presence of charged residues identified by PRS elucidates the importance of N-terminus in stabilizing hTf dynamics and reorienting it towards another conformation with fully engaged irons. Thus, we propose these N-terminus residues to be studied via mutagenesis to determine their effect on iron binding propensity of hTf. We next turn our attention to the reverse problem of dissociation of iron and investigate which structural elements facilitate iron release from holo hTf at endosomal pH while preventing it at serum pH.

The two domains of holo hTf are perfectly correlated at serum pH, resisting iron release. The residue-residue correlation map at serum pH, calculated through eqn (S2) is displayed in Fig. 3a. Residues K88-P91, S501-L503 and a region centered on N611-T613 are correlated with the whole protein structure (marked by the red strips on the figure and indicated by the arrows). All these residues are located on unstructured loops. The cooperativity between the two lobes is also manifested in the slowest three modes at this pH,

displayed in Fig. 4a. The total contribution of these modes to the overall motion is 67 % of all modes, computed from their eigenvalues, λ_i, as $(\lambda_i/\Sigma\,\lambda_i) \times 100$ (47, 14, and 6% for the respective modes). There is no domain opening motion in any of these three modes. They are all circular motions around the N- or the C-domains in counter directions, appearing roughly along the x-z, y-z and x-y, planes, respectively. By inspection of these motions, iron loss at serum pH is not expected to occur. In fact, the partially open state of hTf observed at serum pH [65] is not implicated by these motions, and protonation of liganding tyrosines appear to be a prerequisite for lobe-opening to be observed on time scales accessible by MD simulations as we have shown in recent work via two independent MD simulations of length 500 and 150 ns [41]. While the modes of motion provide information on the general propensity for conformational change in a given structure, understanding molecular details of local and long-range interactions that redistribute the conformations introduce another challenge. Iron release pathway also involves multiple steps consisting of interaction with TfR, protonation steps and interaction with metalloreductases. We explore the first step of dissociation by employing the PRS approach.

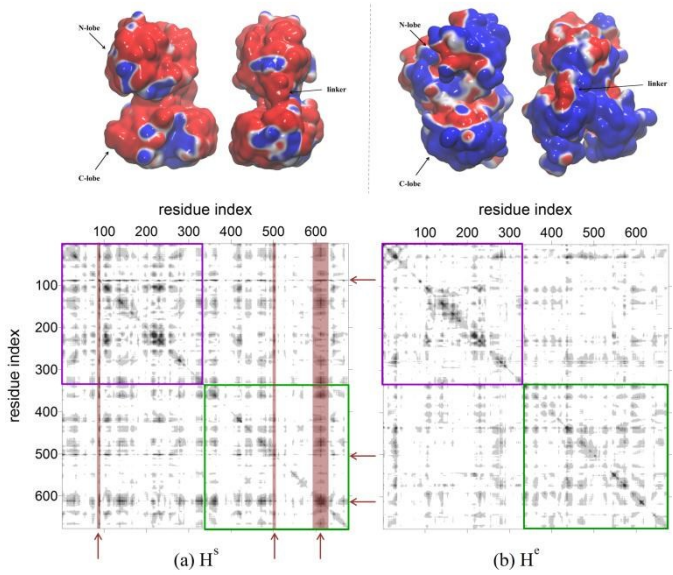

Fig. 3. Electrostatics isocontours (top row) and cross-correlation maps (eqn (S2), bottom row) for holo hTf, at (a) serum pH, $\mathbf{H^s}$ and (b) endosomal pH, $\mathbf{H^e}$. Maps are normalized so that the largest correlation is 1 (black). The regions corresponding to N-lobe – N-lobe interactions are boxed in purple and that for the intra-C-lobe are boxed in green. The N-lobe – C-lobe cross correlations appear in the remaining regions and display nearly complete lack of cooperativity between the two subdomains at endosomal pH, while pointing out residues organizing cooperativity at serum pH. Residues acting cooperatively through all regions in the latter are highlighted by red strips, also indicated by the arrows on the symmetric map. Above each system, $\mathbf{H^s}$ and $\mathbf{H^e}$, snapshots of electrostatics isocontours of hTf drawn at $\pm 0.5 k_B T/e$ are displayed. Blue is positive and red is negative. For each case, hTf is shown in two orientations: On the left, front views of the protein are displayed where the linker is placed at the back; on the right, a 180° rotation along the vertical axis is displayed with the linker in view. At serum pH, protein surface is predominantly negatively charged. At endosomal pH the two lobes reorient and the subdomains of either lobe are opened to expose the iron binding cleft. The charge distribution on the surface is now predominantly positive, especially on the C-lobe.

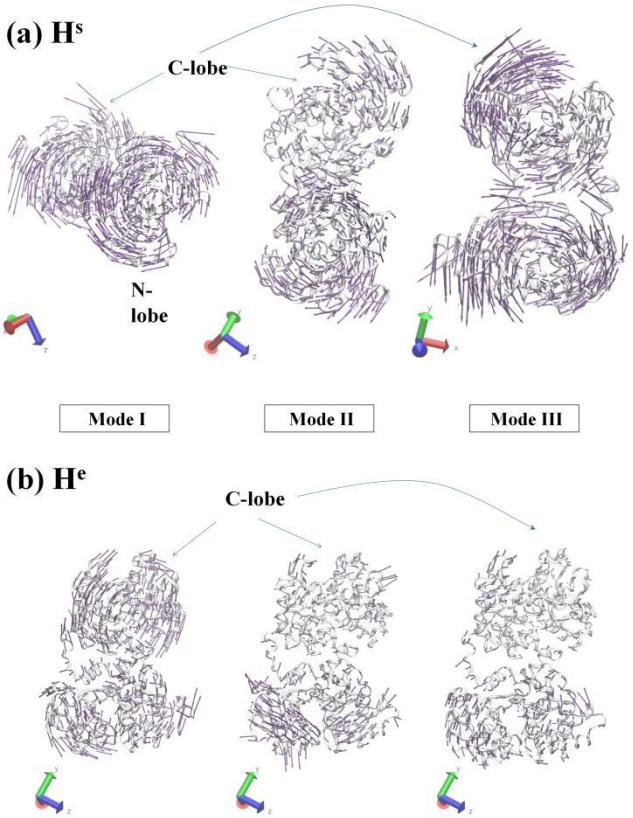

Fig. 4. (a) Slowest three modes of the **C** matrices obtained for holo hTf at serum pH ($\mathbf{H^s}$). Protein structure is reoriented to display the mode shapes indicated by the arrows. The motions within each lobe are unified for the whole domain while the N- and the C-lobes have opposite sense of direction in all three modes. There are no sub-domain separation motions, consistent with the slow release kinetics at serum pH. (b) Slowest three modes of the **C** matrices obtained for holo hTf at endosomal pH ($\mathbf{H^e}$). Protein structure is displayed in the same direction for all three cases. Each lobe moves independent of the other, In the slowest mode of motion, both lobes are mobile and correlated within the domain, but there is lack of correlated motions between the two domains. In the next two modes, N-lobe motions are dominant and there is little to no motion in the C-lobe. Mode II displays hinge-like opening around the iron binding cleft of the N-lobe, consistent with iron removal at this pH.

Residues identified by PRS in $\mathbf{H^s} \rightarrow \mathbf{A}$ transition (Table 1) display low

conformational transition propensity, measured by their correlation coefficient (~0.52). Such a low value is attributed to the rare event of releasing iron at serum pH. In serum environment at 37°C, iron release rates of 54.7×10^{-3} min^{-1} and 40.4×10^{-3} min^{-1}, respectively from the N- and the C-lobes, have been measured [66]. Thus, with such slow processes compared to fast endocytosis (~minute range [67]), we do not expect the release process to occur. The top residues selected according to PRS are all confined to the C-lobe (those having C_i^{max}= 0.52 are listed in Table 1). While these are top residues, even when the hTf-TfR complex is formed as the first step towards iron release process in vivo, the structure will have little propensity to open at this pH, despite having stronger interactions along the C-lobe interface [57].

Endosomal pH leads to a loss of cooperativity between the two domains of holo hTf and readily promotes iron release. The residue-residue correlations obtained from MD simulations at endosomal pH are displayed in Fig. 3b. In stark contrast to the case at serum pH, here only very weak correlations between the two lobes are observed and the two domains act independent of each other. Thus, at the lower pH (incorporated in our simulations via protonation of selected histidines and neutralization of iron coordinating tyrosines; see text S2), the surface loops whose regions are shaded on Fig. 3a "soften" and lose their ability to orchestrate correlations across domains.

The top three modes, contributing 60 % of the motion, are displayed in Fig. 4b, with the respective contributions being 39, 13 and 8 %. In the most collective mode (I), each lobe displays correlations within their own domain, but the relative orientation of the motion is orthogonal across the domains (roughly in the y-z plane in the N-lobe and x-y plane in the C-lobe). In modes II and III, dominant motions appear on the N-lobe which is consistent with the experimentally observed primary iron release from the N-lobe in the absence of TfR. Moreover, mode II is a hinge-like opening motion around the iron binding cleft of the N-lobe, hinting that there is intrinsic propensity for iron dissociation at endosomal pH.

The $H^e \rightarrow A$ transition in the absence of TfR is also monitored by PRS. At this pH, iron release from the N-lobe is faster (17.7±2.2 min^{-1}) and has precedence relative to the slower release (0.65±0.06 min^{-1}) from the C-lobe [31]. While there is asymmetry in the release rates from the two lobes, they are both three orders of magnitudes larger than at serum pH. We find that this conformational change may be induced by non-specific perturbations arriving on any of the residues. We make this conclusion because all residues have at least one perturbation for which the correlation is greater than 0.85. Thus, at endosomal pH, the protein has the propensity to open irrespective of the location of perturbation. The large number of residues with high correlations obtained by the PRS approach parallels what we had observed for the PRS of apo → holo transition in bacterial FBP [15] whereby the apo form has an intrinsic capability to close, corroborated by the extreme association constants on the order of 10^{17}–10^{22} M^{-1} depending on the measurement conditions [38, 39].

To establish how point charge alterations affect the overall response of the protein and interpret the mechanistic basis of pH affects iron affinity, we monitor the electrostatic potential distribution along the protein surface using the APBS package set to the default values [68]. While lowering the pH significantly alters this distribution, single point mutants or single protonation events might also change it [15]. In Fig. 4, top row, these surfaces are displayed for the $\mathbf{H^s}$ and $\mathbf{H^e}$ systems. The $\mathbf{H^s}$ system is the initial conformation taken from the PDB although we check that repeating these calculations on four other snapshots collected at equal spacing along the trajectory yields the same picture. For $\mathbf{H^e}$ system, the displayed conformation is that attained at the end of the 200 ns MD simulation; again we verify that conformations selected from 50 ns apart trajectory pieces yield similar distributions even before complete equilibration is attained at 150 ns. The main effect of lowering pH on the electrostatic distribution of the system is the engorged positively charged surface on the protein, particularly on the C-domain. Thus, the uniformly distributed positive charge of the acidic endosome effectively acts as a forcing agent on the protein surface at lower pH. As a result, the repulsion between the subdomains of each lobe drives the iron-binding cleft to open and promotes iron release, more so on the C-lobe than the N-lobe.

hTf – receptor complexes. We next study the effect of the interacting with receptors on the functioning of hTf by analyzing complex structures. To utilize iron in the cells, TfR recruits hTf on the cell surface. In bacteria, a highly sophisticated mechanism for sequestering iron from humans has been developed [40] whereby TbpA is the protein responsible for hijacking iron from hTf in the first step of the process.

We model the hTf – TfR complex with ANM using the complex structure (PDB code: 1SUV) obtained by electron microscopy [29, 69]. The slowest mode that contributes 49 % of the motion is displayed in Fig. 5a. TfR interacts with both lobes, although the C-lobe interface is larger. The complex acts unified, with both domains of hTf closing up on the receptor. This observation is in conformity with the fact that the complex remains bound through endocytosis and the release of iron in the endosome, until they reach cell surface [31]. Furthermore, the cooperativity between the two lobes that prevents iron release in isolated hTf at serum pH is maintained in the presence of its receptor (compare Fig. 4a and 5a). While the resolution of this structure (7.5 Å) allows for making general comments on the overall motion, we refrain from using it in the PRS analysis which requires more specific atomic detail. We then model the hTf – TbpA complex by ANM, using the PDB structure 3V8X [27] (2.6 Å resolution). The overall conformation of hTf interacting with TbpA adopts an intermediate between that of holo and apo form and exposes iron to the bacterial receptor. TbpA exclusively interacts with the C-lobe residues. The slowest two modes contribute equally by ~13 % in this case, and they are displayed in Fig. 5b and c.

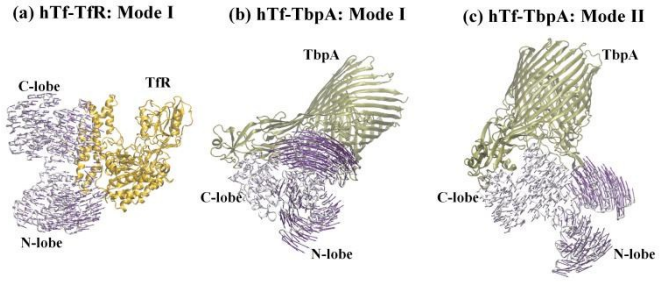

Fig. 5. Complexes of hTf with receptor from human and bacteria. (a) hTf-TfR complex. Slowest mode of motion of the complex (contributing 49% of the motion) is displayed by the arrows. TfR (yellow) interacts with both lobes, although the C-lobe interface has more cohesive interactions. (b-c) hTf-TbpA complex. The two slowest modes of motion contributing ~13 % each have mainly N-lobe motions; these lack domain separation consistent with iron removal by TbpA only from the C-lobe.

Since the hTf structure is held at the C-lobe, the dominant motions occur on the N-lobe, with shear-like motions in mode I (Fig. 5b) and rotational motion of N-lobe in mode II (Fig. 5c). As TbpA sequesters iron from the C-lobe with the aid of a plug loop, hinge-like motions on the N-lobe exposing that iron is not expected to occur.

PRS analysis results are mapped onto this structure in Fig. 6 (also see Table 1) where the identified residues are colored in orange. In a recent study, Barber and Elde have determined sites on hTf are subject to pathogen-driven evolution by cloning and sequencing transferrin from 21 different monkey species for phylogenetic analysis [70]. Single substitutions in hTf at rapidly evolving sites to counteract bacterial iron piracy determined in that study are highlighted in blue (434-436, 439, 576, 591-593, 614, 617, 619, 625) and the plug loop in TbpA is shown in red. We thus make the observation that a patch of residues directly interacting with the plug loop (S390 V587 R602 Q603 Q605 H606 G609 L621 F631 R632) may be perturbed to access the TbpA-free conformation. Another patch that is ~30 Å from the plug loop may also be perturbed to achieve the same conformational change (V429 A430 R568 K569), implicating this as an allosteric regulator of conformational change. During evolution, no direct substitutions in residues of these patches are observed, possibly because the rigidity of the regions must be maintained for mechanical manipulation of the conformations. Based on our observations, we propose that point mutations occurring on residues directly contacting them might offer a perfect solution during the persistent evolutionary conflicts between pathogens and primates.

Conclusion

We have applied computationally simple and efficient methods sampling local states of hTf to extract residues that might exhibit

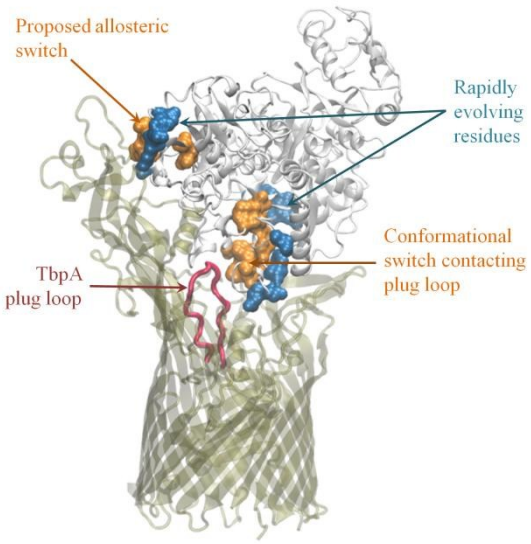

Fig. 6. Residues emerging as conformational switch from TbpA bound to free conformation of hTf appear as two patches that display large C_i values (orange). The direct switch contacts the plug loop in the pathogenic protein responsible for sequestering iron from the C-lobe of hTf is (red). The proposed allosteric switch is 30 Å away from the plug loop. Both switches are complemented by residues harboring rapidly evolving residues (blue) [70].

unique contributions to iron metabolism. Combination of the PRS technique with MD and ANM provides a set of powerful tools unraveling novel local and remote modulators that are effective on the ion binding site conformations. The bilobal structure of hTf features an excellent model to study how the two lobes accomplish orchestrated motions necessary for different ligand binding/release dynamics observed in various environments.

Biologically relevant processes are only partially sampled on the time scales accessible by conventional MD simulations. While the correlation matrices obtained from short time scale dynamics (20 ns in this case) and even network models (for the complexes) provide information on the propensity of the occurrence of certain conformational changes, they do not directly identify residues that are located on mechanically adept regions to trigger an estimated change. In the PRS approach, relatively short lengths of equilibrated substates sampled by MD simulations are employed to decipher the local dynamical contributions leading to global rearrangements of the protein structure towards pre-selected conformations [43]. PRS method is influential in that, it selects residues whose perturbation is expected to shift the conformational surface towards the alternate state while eliminating heuristic biases for focusing on regions that are directly in the active site of a protein.

Via PRS, we have been able to determine sites on the apo form whose perturbation is expected to shift the conformation

distribution towards holo-like structures. We have then been able to test this idea with conventional MD simulations where these locations are perturbed via synergistic anion docking, which is the natural ligand expected to enter the iron binding cavity and prepare it for iron uptake. The simulations show that such a local perturbation leads to global motions on both lobes (Fig. 2).

For the holo form, PRS is also able to explain the lack of iron release propensity at serum pH and its ease of release at endosomal pH. For these cases, there are no single residues whose perturbation is expected to lead to the conformational change. We find that electrostatic surface distributions drastically change at the two pH values which clearly explain the observed phenomena. We note that, in this study we pursue the idea of proton driven Fe^{+3} dissociation. However, experimental data supports that in the presence of the hTfR, there is a shift of the Fe^{+3}/Fe^{+2} redox potential in the positive direction making reduction of bound Fe^{+2} a viable initial step in the mechanistic pathway for Fe release in the endosome. Since atomistic modeling and MD simulations in the presence of the receptor are too expensive for current resources, and that we utilize a coarse grained approach to modelling receptor-hTf complex, we cannot explore this possibility in the current study.

This study also offers a possible mechanism on how single residue substitutions may be drastically effective under the positive selection in evolutionary path of hTf [70], despite the conjecture that single-point mutations would not be resistant toward TbpA binding due to the large buried interface (~2500 $Å^2$ [27]). Herein, residues that undergo positive selection to avoid iron sequestration by pathogens are shown to be in direct contact with two patches of residues that are mechanically capable of steering hTf away from binding the pathogenic protein. Therefore, we propose target locations on the surface of hTf for which ligands may be designed to modify the hTf-TbpA interface and prevent the iron scavenging process without altering the functioning of the protein.

Acknowledgments.

The authors acknowledge stimulating discussions with Ali Rana Atilgan.

This work was supported by the Scientific and Technological Research Council of Turkey (grant numbers 110T624 and 113Z408) and CMST COST action CM1306.

References

1. S. T. Whitten, B. Garcia-Moreno and V. J. Hilser, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, 102, 4282-4287.
2. M. Grossman, N. Sela-Passwell and I. Sagi, *Curr. Opin. Struct. Biol.*, 2011, 21, 678-685.
3. J. L. England, *Structure*, 2011, 19, 967-975.
4. A. R. Atilgan, A. O. Aykut and C. Atilgan, *J. Chem. Phys.*, 2011, 135, 155102.
5. A. Ghosh, R. Sakaguchi, C. Liu, S. Vishveshwara and Y.-M. Hou, *J. Biol. Chem.*, 2011, 286, 37721-37731.
6. A. Sethi, J. Eargle, A. A. Black and Z. Luthey-Schulten, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, 106, 6620-6625.
7. K. Gunasekaran, B. Y. Ma and R. Nussinov, *Proteins: Struct. Funct. Bioinform.*, 2004, 57, 433-443.
8. A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin and I. Bahar, *Biophys. J.*, 2001, 80, 505-515.
9. I. Bahar, A. R. Atilgan and B. Erman, *Fold. Des.*, 1997, 2, 173-181.
10. V. Tozzini, *Acc. Chem. Res.*, 2010, 43, 220-230.
11. M. Orozco, *Chem. Soc. Rev.*, 2014, 43, 5051-5066.
12. H. Dietz, F. Berkemeier, M. Bertz and M. Rief, *Proceedings of the National Academy of Sciences*, 2006, 103, 12724-12728.
13. D. Ming and M. E. Wall, *Proteins: Structure, Function, and Bioinformatics*, 2005, 59, 697-707.
14. D. Ming and M. E. Wall, *Phys. Rev. Lett.*, 2005, 95, 198103.
15. C. Atilgan and A. R. Atilgan, *PLoS Comput. Biol.*, 2009, 5, e1000544.
16. C. Atilgan, Z. N. Gerek, S. B. Ozkan and A. R. Atilgan, *Biophys. J.*, 2010, 99, 933-943.
17. H. Abdizadeh, G. Guven, A. R. Atilgan and C. Atilgan, *Journal of Enzyme Inhibition and Medicinal Chemistry, in press*, 2015.
18. M. Ikeguchi, J. Ueno, M. Sato and A. Kidera, *Phys. Rev. Lett.*, 2005, 94.
19. A. O. Aykut, A. R. Atilgan and C. Atilgan, *PLoS Comput. Biol.*, 2013, 9, e1003366.
20. G. Guven, A. R. Atilgan and C. Atilgan, *The Journal of Physical Chemistry B* 2014, 118, 11677-11687.
21. A. Bolia, Z. N. Gerek and S. B. Ozkan, *J. Chem. Inf. Model.*, 2014, 54, 913-925.
22. J. M. Mottonen, D. J. Jacobs and D. R. Livesay, *Biophys. J.*, 2010, 99, 2245-2254.
23. D. Jacobs, D. Livesay, J. Mottonen, O. Vorov, A. Istomin and D. Verma, in *Allostery*, ed. A. W. Fenton, Springer New York, 2012, vol. 796, ch. 15, pp. 279-304.
24. J.-M. El Hage Chahine, M. Hémadi and N.-T. Ha-Duong, *Biochim. Biophys. Acta Biochim Biophys Acta*, 2012, 1820, 334-347.
25. P. Aisen and I. Listowsky, *Annu. Rev. Biochem.*, 1980, 49, 357-393.
26. A. Dautry-Varsat, A. Ciechanover and H. F. Lodish, *Proc. Natl. Acad. Sci. USA*, 1983, 80, 2258-2262.
27. N. Noinaj, N. C. Easley, M. Oke, N. Mizuno, J. Gumbart, E. Boura, A. N. Steere, O. Zak, P. Aisen, E. Tajkhorshid, R. W. Evans, A. R. Gorringe, A. B. Mason, A. C. Steven and S. K. Buchanan, *Nature*, 2012, 483, 53-U92.
28. J. Wally, P. J. Halbrooks, C. Vonrhein, M. A. Rould, S. J. Everse, A. B. Mason and S. K. Buchanan, *J. Biol. Chem.*, 2006, 281, 24934-24944.
29. Y. Cheng, O. Zak, P. Alsen, S. C. Harrison and T. Walz, *Cell*, 2004, 116, 565-576.
30. J. Williams and K. Moreton, *Biochem. J.*, 1980, 185, 483-488.
31. A. N. Steere, S. L. Byrne, N. D. Chasteen and A. B. Mason, *Biochim. Biophys. Acta*, 2012, 1820, 326-333.
32. R. Liu, J.-Q. Guan, O. Zak, P. Aisen and M. R. Chance, *Biochemistry*, 2003, 42, 12447-12454.
33. M. Hémadi, N. T. Ha-Duong and J. M. El Hage Chahine, *J. Mol. Biol.*, 2006, 358, 1125-1136.
34. M. Hémadi, P. H. Kahn, G. Miquel and J.-M. El Hage Chahine, *Biochemistry*, 2004, 43, 1736-1745.

35. J. M. El Hage Chahine and R. Pakdaman, *Eur. J. Biochem.*, 1995, 230, 1102-1110.
36. N. C. Andrews, *The International Journal of Biochemistry & Cell Biology*, 1999, 31, 991-994.
37. R. S. Ohgami, D. R. Campagna, E. L. Greer, B. Antiochos, A. McDonald, J. Chen, J. J. Sharp, Y. Fujiwara, J. E. Barker and M. D. Fleming, *Nat. Genet.*, 2005, 37, 1264-1269.
38. S. Dhungana, C. H. Taboy, D. S. Anderson, K. G. Vaughan, P. Aisen, T. A. Mietzner and A. L. Crumbliss, *Proc Natl Acad Sci U S A*, 2003, 100, 3659-3664.
39. J. J. Heymann, K. D. Weaver, T. A. Mietzner and A. L. Crumbliss, *Journal of the American Chemical Society*, 2007, 129, 9704-9712.
40. C. J. P. Siburt, T. A. Mietzner and A. L. Crumbliss, *Biochimica Et Biophysica Acta (BBA)-General Subjects*, 2012, 1820, 379-392.
41. H. Abdizadeh, A. Atilgan and C. Atilgan, *JBIC Journal of Biological Inorganic Chemistry*, 2015, 20, 705-718.
42. R. T. MacGillivray, E. Mendez, S. K. Sinha, M. R. Sutton, J. Lineback-Zins and K. Brew, *Proc. Natl. Acad. Sci. USA*, 1982, 79, 2504-2508.
43. C. Atilgan, O. B. Okan and A. R. Atilgan, *Annual Review of Biophysics*, 2012, 41, 205-225.
44. C. M. Bruns, D. S. Anderson, K. G. Vaughan, P. A. Williams, A. J. Nowalk, D. E. McRee and T. A. Mietzner, *Biochemistry*, 2001, 40, 15631-15637.
45. T. E. Adams, A. B. Mason, Q.-Y. He, P. J. Halbrooks, S. K. Briggs, V. C. Smith, R. T. A. MacGillivray and S. J. Everse, *J. Biol. Chem.*, 2003, 278, 6027-6033.
46. W. R. Harris, *Biochim. Biophys. Acta*, 2012, 1820, 348-361.
47. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, 30, 2785-2791.
48. L. S. Yilmaz and A. R. Atilgan, *J. Chem. Phys.*, 2000, 113, 4454-4464.
49. F. Tama and Y.-H. Sanejouand, *Protein Eng.*, 2001, 14, 1-6.
50. R. Pakdaman, F. B. Abdallah and J.-M. El Hage Chahine, *J. Mol. Biol.*, 1999, 293, 1273-1284.
51. S. L. Byrne, N. D. Chasteen, A. N. Steere and A. B. Mason, *J. Mol. Biol.*, 2010, 396, 130-140.
52. A. N. Steere, S. L. Byrne, N. D. Chasteen, V. C. Smith, R. T. A. MacGillivray and A. B. Mason, *J. Biol. Inorg. Chem.*, 2010, 15, 1341-1352.
53. P. D. Jeffrey, M. C. Bewley, R. T. A. MacGillivray, A. B. Mason, R. C. Woodworth and E. N. Baker, *Biochemistry*, 1998, 37, 13978-13986.
54. R. T. A. MacGillivray, M. C. Bewley, C. A. Smith, Q.-Y. He, A. B. Mason, R. C. Woodworth and E. N. Baker, *Biochemistry*, 2000, 39, 1211-1216.
55. Q.-Y. He, A. B. Mason, R. C. Woodworth, B. M. Tam, T. Wadsworth and R. T. A. MacGillivray, *Biochemistry*, 1997, 36, 5522-5528.
56. D. Nurizzo, H. M. Baker, Q.-Y. He, R. T. A. MacGillivray, A. B. Mason, R. C. Woodworth and E. N. Baker, *Biochemistry*, 2001, 40, 1616-1623.
57. A. N. Steere, B. F. Miller, S. E. Roberts, S. L. Byrne, N. D. Chasteen, V. C. Smith, R. T. A. MacGillivray and A. B. Mason, *Biochemistry*, 2012, 51, 686-694.
58. L. Bellounis, R. Pakdaman and J. M. El Hage Chahine, *J. Phys. Org. Chem.*, 1996, 9, 111-118.
59. Q. Y. He, A. B. Mason, R. C. Woodworth, B. M. Tam, R. T. A. MacGillivray, J. K. Grady and N. D. Chasteen, *J. Biol. Chem.*, 1998, 273, 17018-17024.
60. R. Pakdaman and J. M. El Hage Chahine, *Eur. J. Biochem.*, 1996, 236, 922-931.
61. G. Celniker, G. Nimrod, H. Ashkenazy, F. Glaser, E. Martz, I. Mayrose, T. Pupko and N. Ben-Tal, *Isr. J. Chem.*, 2013, 53, 199-206.
62. J. G. Grossmann, A. B. Mason, R. C. Woodworth, M. Neu, P. F. Lindley and S. S. Hasnain, *J. Mol. Biol.*, 1993, 231, 554-558.
63. J. G. Grossmann, J. B. Crawley, R. W. Strange, K. J. Patel, L. M. Murphy, M. Neu, R. W. Evans and S. S. Hasnain, *J. Mol. Biol.*, 1998, 279, 461-472.
64. H. M. Baker, Q.-Y. He, S. K. Briggs, A. B. Mason and E. N. Baker, *Biochemistry*, 2003, 42, 7084-7089.
65. N. Yang, H. Zhang, M. Wang, Q. Hao and H. Sun, *Sci. Rep.*, 2012, 2, 999.
66. P. J. Halbrooks, Q. Y. He, S. K. Briggs, S. J. Everse, V. C. Smith, R. T. A. MacGillivray and A. B. Mason, *Biochemistry*, 2003, 42, 3701-3707.
67. D. Sheff, L. Pelletier, C. B. O'Connell, G. Warren and I. Mellman, *J. Cell Biol.*, 2002, 156, 797-804.
68. N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proceedings of the National Academy of Sciences*, 2001, 98, 10037-10041.
69. C. M. Lawrence, S. Ray, M. Babyonyshev, R. Galluser, D. W. Borhani and S. C. Harrison, *Science*, 1999, 286, 779-782.
70. M. F. Barber and N. C. Elde, *Science*, 2014, 346, 1362-1366.