Mathematical Biosciences and Engineering

MBE, 19(12): 13252–13275.
DOI: 10.3934/mbe.2022620
Received: 10 May 2022
Revised: 17 August 2022
Accepted: 21 August 2022
Published: 09 September 2022

http://www.aimspress.com/journal/MBE

*Research article*

# Modeling ligand-macromolecular interactions as eigenvalue-based transition-state dissociation constants may offer insights into biochemical function of the resulting complexes

**Siddhartha Kundu***

Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110029, India

* **Correspondence:** Email: siddhartha_kundu@yahoo.co.in, siddhartha_kundu@aiims.edu.

**Abstract:** A ligand when bound to a macromolecule (protein, DNA, RNA) will influence the biochemical function of that macromolecule. This observation is empirical and attributable to the association of the ligand with the amino acids/nucleotides that comprise the macromolecule. The binding affinity is a measure of the strength-of-association of a macromolecule for its ligand and is numerically characterized by the association/dissociation constant. However, despite being widely used, a mathematically rigorous explanation by which the association/dissociation constant can influence the biochemistry and molecular biology of the resulting complex is not available. Here, the ligand-macromolecular complex is modeled as a homo- or hetero-dimer with a finite and equal number of atoms/residues per monomer. The pairwise interactions are numeric, empirically motivated and are randomly chosen from a standard uniform distribution. The transition-state dissociation constants are the strictly positive real part of all complex eigenvalues of this interaction matrix, belong to the open interval $(0,1)$, and form a sequence whose terms are finite, monotonic, non-increasing and convergent. The theoretical results are rigorous, presented as theorems, lemmas and corollaries and are complemented by numerical studies. An inferential analysis of the clinical outcomes of amino acid substitutions of selected enzyme homodimers is also presented. These findings are extendible to higher-order complexes such as those likely to occur *in vivo*. The study also presents a schema by which a ligand can be annotated and partitioned into high- and low-affinity variants. The influence of the transition-state dissociation constants on the biochemistry and molecular biology of non-haem iron (II)- and 2-oxoglutarate-dependent dioxygenases (catalysis) and major histocompatibility complex (I) mediated export of high-affinity peptides (non-enzymatic association/dissociation) are examined as special cases.

**Keywords:** dissociation constant; eigenvalues; interaction matrix; model of ligand-macromolecular complex as a homo- or hetero-dimer; Finite, monotonic, non-decreasing, bounded and convergent sequences; iron (II)- and 2-oxoglutarate-dependent catalysis; major histocompatibility complex (I) mediated export of high-affinity peptides

## 1. Introduction

Ligands, are biochemical modifiers of macromolecular structure and can impact biological function. These can be co-factors (iron, cobalt, copper, zinc), co-enzymes such as Nicotinamide- and Flavin- Adenine Dinucleotides (NAD, FAD) and full-length molecules with short binding sites [1,2]. Ligands, unlike substrates/co-substrates are either reversibly altered or not at all. The biochemical role of ligands, *in vivo*, is complex and can influence both, enzyme-mediated substrate catalysis and non-enzymatic association and dissociation interactions. Whilst, the interaction with competitive inhibitors, co-factors or co-enzymes involves definitive and direct modifications to the active site residues, the effect of a ligand can be allosteric and indirect [3–5]. The latter involves both long-distance conformational changes and non-covalent interactions (hydrogen, Van der Waals, hydrophobic, electrostatic) [6–11]. Empirical data suggests that the binding affinity or the strength-of-association of a macromolecule for its ligand is a critical determinant of function [7–11]. For example, 2,3-Bisphophoglycerate is a potent modifier of Hemoglobin function and does so by shifting the oxygen-dissociation curve to the right. In its absence Hemoglobin retains high affinity for molecular oxygen (left-shift of the oxygen dissociation curve), an undesirable effect on its role as a transporter [10,11]. Similarly, Ascorbic acid maintains iron in its reduced state in the gastrointestinal tract and as part of the active site of non-haem iron (II)- and 2-oxoglutarate-dependent dioxygenases [12,13]. Deficiency of Ascorbic acid is implicated in tardy iron absorption in the ileum as well as a range of collagen disorders such as scurvy (Prolyl- and Lysyl-hydroxylases) [14,15]. Conversely, proteins which are modified such as those that may originate from missense or amino acid substitutions and are secondary to genomic variants such as single nucleotide polymorphisms (SNPs) and insertions-deletions (indels), will also result in several clinical outcomes [16,17]. Here, too, enzyme catalysis is directly affected if these are present at the active site or is impacted indirectly (folding, stability, complex formation) when present elsewhere [16–18].

There is a large volume of literature which describes macromolecules in terms of either residues (amino acids, nucleotides) interacting or an all atom-based interaction matrix. The proponent of the 2D approach is the Gaussian network model (GNM), while the Anisotropic network model (ANM) is representative of the 3D approach [19–22]. The fundamental premise of both these approaches is the elastic network model (ENM) [19]. Here, an atom or residue is modeled as an elastic mass and the interaction between a pair of atoms/residues is dependent on the selection of a pre-determined cut-off distance [22]. The force constant, although, not a parameter by definition, has also been studied and shows good correlation with B-factor data [23]. A major application of these studies is normal mode analysis (NMA), which has been used to glean valuable insights into the structural dynamics of the investigated macromolecule and into B-factor distribution [23,24]. Despite this success, there are significant limitations of this approach, including inadequate descriptors for the type of interactions computed by the Hessian matrix and data points that are dependent on a preselected cut-off distance (5–10 Ang, GNM; 10–15 Ang, ANM) [20,22,25]. Parameter-free

versions of the ENM (pfENM), GNM (pfGNM) and ANM (pfANM), to resolve the latter, have been described and compared to establish B-factor distribution (isotropic, anisotropic) [25]. Additionally, many of these studies have focused on inferring biophysical characteristics such as cross-correlational fluctuations and mean square displacements. From a functional standpoint, however, it is not clear whether these data can be utilized to derive/study parameters such as the Michaelis-Menten constant (Km) or the association/dissociation (Ka/Kd). Since, these depend on the presence of an organic or inorganic modifier, i.e., ligand/substrate/co-factor/co-substrate, in addition to the modeled macromolecule, its exclusion is another major lacuna of these studies.

Despite the availability of clinical, empirical, analytical and computational data, a mathematically rigorous explanation for the heterogeneity in biochemical function, for a ligand-macromolecular complex, is missing. The work presented models a ligand and macromolecule as a homo- or hetero-dimer and subsumes a finite and equal number of atoms/residues per monomer. The pairwise interactions of the resulting square matrix will be chosen randomly from a standard uniform distribution. The resulting eigenvalues will be analyzed and modeled in accordance with known literature on biochemical reactions to generate biologically viable and usable dissociation constants. The theoretical results will be complemented by numerical studies where applicable. Additionally, and through various theorems, lemmas and corollaries, a schema to partition ligands into high- and low-affinity variants will also be discussed. The suitability of the transition-state dissociation constants as a model for ligand-macromolecular interactions will be inferentially assessed by analyzing the clinical outcomes of amino acid substitutions of selected enzyme homodimers. The relevance of the model to biochemical function will be discussed by examining the ligand-macromolecular complex for known ligands of non-haem iron (II)- and 2-oxoglutarate-dependent dioxygenases (Fe2OG) and the major histocompatibility complex (I) (MHC1).

The general outline of the manuscript includes an initial section where the system to be modeled, rationale for this study, and formal definitions are introduced (Section 2). The model is analyzed, formulated and presented as theorems, lemmas and corollaries (Section 3). This section also includes a numerical study to demonstrate and validate the theoretical assertions made. The biological relevance of these findings are discussed with an analysis of clinical outcomes of enzyme sequence variants and case studies of enzyme- and non-enzymatic complex formation (Section 4). A brief conclusion that summarizes the presented study, limitations and future directions is included at the end of the manuscript (Section 5). Details of all proofs are included after the conclusions (Section 6).

## 2. Model description and definitions

Consider the generic interaction between macromolecule $(c)$ and ligand $(\mu)$,

$$c\mu \underset{r_b}{\overset{r_f}{\rightleftarrows}} c + \mu \quad RXN(1)$$

We can represent this interaction/reaction, at a steady state, with the rate equations [26],

$$R_d(c\mu) = r_f.[c\mu]^{L_{c\mu}\geq 0} \tag{1}$$

$$R_a(c\mu) = r_b.[c]^{L_c\geq 0}[\mu]^{L_\mu\geq 0} \tag{2}$$

At steady state,

$$R_d(c\mu) = R_a(c\mu) \tag{3}$$

$$\Rightarrow \frac{r_f}{r_b} = \frac{[c]^{Lc\geq 0}.[\mu]^{L\mu\geq 0}}{[c\mu]^{Lc\mu\geq 0}} \tag{4}$$

$$\frac{r_f}{r_b} = Kd(c\mu) \tag{5}$$

Here,

$$c := Macromolecule$$
$$\mu := Ligand$$
$$[.] := Molar\ concentration\ of\ reactant\ in\ standard\ form\ (M)$$
$$R_a(c\mu) := Rate\ of\ association\ of\ complex\ (M\ s^{-1})$$
$$R_d(c\mu) := Rate\ of\ dissociation\ of\ complex\ (M\ s^{-1})$$
$$r_f := Rate\ constants\ of\ forward\ reaction\ (s^{-1})$$
$$r_b := Rate\ constants\ of\ reverse\ reaction\ (M^{-1}\ s^{-1})$$
$$L := Stoichiometry\ of\ reactant(s)$$
$$Kd(c\mu) := Dissociation\ constant\ for\ ligand - macromolecular\ complex\ (M)$$

It is clear that a ligand-macromolecular complex may exist in one of three distinct states. These include: a) perfect association, b) perfect dissociation and c) an intermediate- or transition-state; and can be represented in terms of the dissociation constant,

***Case (1)*** $Perfect\ association$ *Def. (1)*

$$\frac{1}{r_b} \to \infty; r_f \to 0 \tag{6,7}$$

$$\Rightarrow Kd(c\mu) \approx 0 \tag{8}$$

***Case (2)*** $Perfect\ disassociation$ *Def. (2)*

$$r_f \ggg r_b \tag{9}$$

$$\Rightarrow Kd(c\mu) \geq 1 \tag{10}$$

***Case (3)*** $Transient - state\ disassociation\ constant$ *Def. (3)*

$$r_f \lesseqgtr r_b \tag{11}$$

$$\Rightarrow Kd(c\mu) \in \mathbb{R} \cap (0,1) \tag{12}$$

Consider an atom/residue-based representation (amino acids/nucleotides) of a generic set of monomeric macromolecules, $\boldsymbol{\mathcal{C}} = \{protein, DNA, RNA\}$, with $z = 1,2,\ldots,Z$ components each with $i$-indexed ($i = 1,2,\ldots,I$) c-atoms/residues,

$$\mathbf{c} \equiv \mathcal{C}_z \in \boldsymbol{\mathcal{C}} | \mathbf{c} = [c_1\ c_2\ \ldots.\ c_{i=I}]^T, I \in \mathbb{N} \tag{13}$$

The analogous model of a monomer ligand, $\boldsymbol{\mathcal{L}} = \{small\ molecule, peptide, oligonucleotide\}$, with $j$-indexed $(j = 1,2,\ldots,J)$ $\mu$-atoms/residues is,

$$\boldsymbol{\mu} \in \boldsymbol{\mathcal{L}} | \boldsymbol{\mu} = [\mu_1\ \mu_2\ \ldots.\ \mu_{j=J}], J \in \mathbb{N} \tag{14}$$

It is also assumed that the ligand-macromolecular complex is a homo- or hetero-dimer with an equal number of atoms/residues $(I = J)$ per monomer. The interaction matrix is,

$$\langle \boldsymbol{c}|\boldsymbol{\mu}\rangle = [c_1\, c_2..c_{i=I}]^T \times [\mu_1\, \mu_2..\mu_{j=J}] = \boldsymbol{\mathcal{C}_{z\mu}} = (c_i\mu_j) \subset \mathbb{R}^{I\times J} \tag{15}$$

The numerical values of this matrix are chosen randomly from the standard uniform distribution,

$$\boldsymbol{\mathcal{C}}_{z\mu} = (c_i\mu_j) \in \mathcal{U}_{[0,1]} \tag{16}$$

The rationale for this choice is that each pairwise interaction is subsumed to be a function of an arbitrary number of non-bonded interactions (long- and short-range) and is therefore, unique. Clearly, this implies the existence of $\{I, J\}$-linear independent vectors,

$$rank(\boldsymbol{\mathcal{C}}_{z\mu}) = \{I, J\} \tag{17}$$

Since $\boldsymbol{\mathcal{C}_{z\mu}}$ is diagonalizable there exists a diagonal matrix, $\boldsymbol{\mathcal{KC}}_{z\mu}$,

$$\boldsymbol{\mathcal{KC}}_{z\mu} = \boldsymbol{X^{-1}\mathcal{C}}_{z\mu}\boldsymbol{X} \tag{18}$$

$$z_{i=j} = z \in diag(\boldsymbol{\mathcal{KC}}_{z\mu}) \subset \mathbb{C} \tag{19}$$

$\boldsymbol{\mathcal{C}_{z\mu}}$, is non-symmetric the computed eigenvalues of the modeled ligand-macromolecular interaction matrix can have positive and negative real parts,

$$\{\alpha_{ij} = Re(z) \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R}|i = j, z \in \mathbb{C}\} \tag{20}$$

$$\{\alpha_{ij} = Re(z) \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R}|i = j, z \in \mathbb{C}\} \tag{20.1}$$

A further subdivision can be made in accordance with established literature,

$$\{\alpha_{ij} = Re(z) \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R}|i = j, z \in \mathbb{C}\} \;\; \omega_\infty = Kd(c\mu) \geq 1 \tag{21}$$

$$Perfect\ association \equiv Reverse \;\; \omega_0 = Kd(c\mu) = 0 \tag{22}$$

This selection generates the set,

$$\omega = \alpha_{ij} \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R} \cap (0,1) \tag{23}$$

$$\omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \qquad \textit{Def. (4)}$$

$$\#\boldsymbol{Kd}(\boldsymbol{c\mu}) = A \tag{23.1}$$

We can combine these to get a preliminary definition of the transition-state disassociation constants. These are the strictly positive real part of all complex eigenvalues that characterize a ligand-macromolecular complex with an equal number of atoms/residues per monomer and belong to the open interval (0,1),

$$\{\omega = \mathcal{R}\mathrm{e(z)} \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R} \cap (0,1)|z \in \mathbb{C}\} \qquad \textit{Def. (5a)}$$

## 3. Results

### *3.1. Mathematical analysis of transition-state dissociation constants of the modeled ligand-macromolecular complex*

Whilst, the states of perfect association and dissociation are key determinants of whether a reaction will occur or not, the transition-state dissociation constants may offer insights into the origins of threshold values, feedback mechanisms and other regulatory checkpoints. However, in order to ascribe biological relevance to these findings we must establish various bounds which can then be utilized to assess and thence assay the function of a ligand-macromolecular complex.

***Theorem 1*** **(T1)**: The linear map between the transition-state dissociation constants and the eigenvalues that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecular complex is the injection,

$$g: \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \mapsto \boldsymbol{Kd}\left(\boldsymbol{\mathcal{C}_{z\mu}}\right) \tag{24}$$

***Theorem 2*** **(T2):** The transition-state dissociation constants that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecular complex is a monotonic and non-increasing sequence,

$$\{\omega_a\}_{a\leq a+1} |\ \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{25}$$

***Theorem 3*** **(T3):** The transition-state dissociation constants that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecule complex are monotonic, bounded and therefore, convergent,

$$\lim_{a\to\infty} \{\omega_a\}_{a\leq a+1} = \{0,1\} |\ \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{26}$$

***Corollary 1*** **(C1):** The transition-state dissociation constants that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecule complex is a sequence with defined greatest-lower and least-upper -bounds,

$$inf\{\omega_a\}_{a\leq a+1} < \{\omega_a\}_{a\leq a+1} < sup\{\omega_a\}_{a\leq a+1} \tag{27}$$

***Corollary 2*** **(C2;** without proof**):** The cardinality of the set of transition-state dissociation constants that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecule complex is finite,

$$\#\boldsymbol{Kd}(\boldsymbol{c\mu}) = A < \#\boldsymbol{Kd}\left(\boldsymbol{\mathcal{C}_{z\mu}}\right) = \{I, J\} \tag{28}$$

Using **T1–T3** and **C1**, **C2** we can refine our definition of the transition-state dissociation constants for the modeled ligand-macromolecular complex,

$$\{\omega_a\}_{a\leq a+1} |\ \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}); a = 1,2 \ldots A \qquad \textit{Def. (5b)}$$

where,

$$\omega = \mathcal{R}\mathrm{e}(\mathrm{z}) \in \boldsymbol{Kd}\left(\boldsymbol{\mathcal{C}_{z\mu}}\right) \subset \mathbb{R} \cap (0,1) | z \in \mathbb{C}$$

### *3.2. Annotation of a ligand on the basis of the transition-state dissociation constants of the modeled ligand-macromolecular complex*

It is clear from the above results that the eigenvalue-based transition-state dissociation constants are continuous and can potentially model the multiplicity of intermediate- or transient-states that a ligand-macromolecular complex may adopt. It should therefore, be possible to partition the transition-state dissociation constants into functionally distinct subsets and will be characteristic for a specific ligand-macromolecular complex.

***Theorem 4*** **(T4):** The transition-state dissociation constants that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecule complex is a proper subset of the complete set of the real part of all complex eigenvalues that comprise the interaction matrix,

$$\boldsymbol{Kd}(\boldsymbol{c\mu}) \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{29}$$

***Corollary 3*** **(C3):** The distribution of the transition-state dissociation constants that characterize the interactions of the homo- or hetero-dimer form of the modeled ligand-macromolecule complex will result in a schema by which we can annotate the ligand as a high $(\mu_{high})$- or low $(\mu_{low})$-affinity variant,

$$\mu = \{\mu_{high}, \mu_{low}\} \tag{30}$$

### *3.3. Modeling the transition-state dissociation constants of higher-order ligand-macromolecular complexes*

Biologically relevant macromolecular complexes are characterized by heterogeneity and high-order $(Z \geq 2)$. This multimer-form of a macromolecule is formed around a primary molecule and its interactions. These may be protein-protein, DNA/RNA-protein or DNA-RNA-protein.

The multimer-form of ligand-macromolecular complex is easily modeled using the mathematical framework defined earlier. Here, the ligand-macromolecular complex is considered as a set of interacting monomers and the binding to a ligand occurs via a single unique monomer,

$$\{\mathcal{C}_z \in \boldsymbol{\mathcal{C}} | \mathcal{C}_1 = \mathcal{C}_2 \ldots = \mathcal{C}_{z=Z} | Z \geq 2\} \qquad \textit{Def. (6a)}$$

where,

$$\mathcal{C}_z \equiv \mathcal{C}_{z\mu} \equiv \mu\mathcal{C}_z \qquad \textit{Def. (6b)}$$

Here,

$$\mu := Ligand$$
$$\mathcal{C}_z := Unique\ monomer\ of\ a\ macromolecule\ that\ associates\ with\ a\ ligand$$
$$\mu\mathcal{C}_z := Ligand - macromolecular\ complex$$

On the basis of these definitions we can re-index the remaining monomers, i.e., after excluding the unique monomer that binds to the ligand,

$$\{\mathcal{C}_y \in \widetilde{\boldsymbol{\mathcal{C}}} = \boldsymbol{\mathcal{C}} \backslash \{\mathcal{C}_z\} | y = 1,2 \ldots Y; Y = Z - 1\} \qquad \textit{Def. (7)}$$

We now derive an expression for the multimer (higher-order)-form of a ligand-macromolecular complex.

***Theorem 5 (T5)***: The multimer-form of a ligand-macromolecular complex comprising identical monomer units and with an arbitrary unit associating with a ligand is,

$$\prod_{y=1}^{y=Y} \mathcal{C}_z.\mathcal{C}_y = \mathcal{C}_z.\mathcal{C}_Y^y | y = 1,2 \dots Y; Y = Z-1; Z \geq 2; z = 1,2 \dots Z \tag{31}$$

Rewriting, this result in terms of the definition of the multimer form of a ligand-macromolecular complex,

$$\mathcal{C}_z.\mathcal{C}^Y \equiv \mu\mathcal{C}_z.\mathcal{C}^Y \qquad \textit{Def. (8)}$$

***Theorem 6*** **(T6):** The linear map between the transition-state dissociation constants that characterize the interactions of the monomer- and multimer-forms of a ligand-macromolecular complex is a bijection,

$$h^{-1} \circ h: \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \leftrightarrow u \in \boldsymbol{Kd}(\mu\boldsymbol{\mathcal{C}_z}) \tag{32}$$

***Theorem 7*** **(T7)**: The linear map between the transition-state dissociation constants and the eigenvalues that characterize the multimer-form of a ligand-macromolecular complex is a composition and an injection,

$$g \circ h^{-1} \circ h: u \in \boldsymbol{Kd}(\mu\boldsymbol{\mathcal{C}_z}.\mathcal{C}^Y) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{33}$$

*3.4. Numerical studies to establish biological relevance of the transition-state dissociation constants*

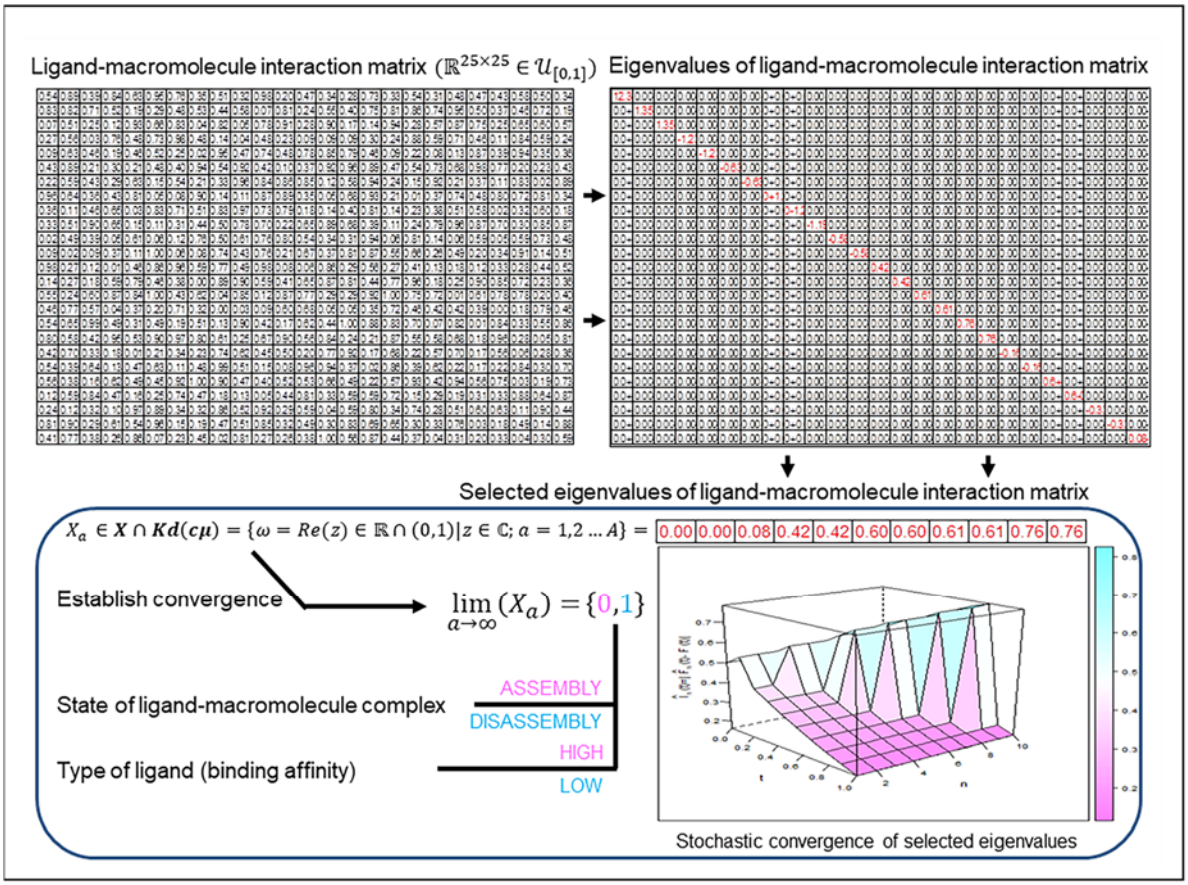


**Figure 1.** Numerical studies of ligand-macromolecular interactions as eigenvalue-based transition-state dissociation constants.

The aforementioned theoretical results establish the mathematical rigor behind the definition and development of the transition-state dissociation constants as a model for ligand-macromolecular interactions (T1–T7, C1–C3). These assertions are complemented and numerically validated in R-4.1.2. Here, the R-packages, "ConvergenceConcepts" and "pracma" are utilized to investigate and analyze the stochastic convergence of the eigenvalues generated by the interaction matrix of a ligand-macromolecular complex (Supplementary Text 1) [27]. The R-scripts to establish convergence along with data processing are developed in-house (Supplementary Text 2). The stepwise algorithm to compute and numerically validate the transition-state dissociation constants is presented (Figure 1).

**Step 1**: A ligand and macromolecule with an equal number of atoms/residues ($n = 25$) is chosen. Whilst, the complex can be modeled as a perfect homodimer, imperfect forms such as alternatively spliced isoenzymes are prevalent and commonly observed. Alternatively, the ligand can be modeled as a different macromolecule altogether.

**Step 2**: Populate the square interaction matrix with values randomly chosen from a uniform distribution, $\mathcal{U}_{[0,1]}$. These will represent one of three potential states for each interacting pair of atoms/residues of the modeled ligand-macromolecular complex (association, complete disassembly, transition-state).

**Step 3**: Compute the complex eigenvalues of this matrix and extract the real part of each.

**Step 4**: Form a sequence of the subset comprising those values that are strictly positive and belong to the open interval $(0,1)$.

**Step 5**: Establish the stochastic convergence in distribution and/or probability of the terms of this sequence to the expected upper (tsup)- and lower (tinf)-bounds, i.e., 0 and 1.

***Step 5.1***: Construct a sequence of random numbers, **X**, whose elements are uniquely mapped to the eigenvalue-based transition-state dissociation constants and represent intermediate- or transition-states of the modeled ligand-macromolecular complex,

$$\{X_a \in \boldsymbol{X} \cap \boldsymbol{Kd}(\boldsymbol{c\mu}) \subset \mathbb{R} \cap (0,1) | a = 1,2 \dots A\} \qquad \textit{Def. (9)}$$

Here,

$$A = \#\left(\boldsymbol{X} \cap \boldsymbol{Kd}(\boldsymbol{c\mu})\right) \tag{34}$$

***Step 5.2***: Establish convergence of this set of random numbers. Here, weak convergence will suffice (distribution, probability),

$$\lim_{a\to\infty}(X_a) \to \{tinf, tsup\} = \{0,1\} \tag{35}$$

The parameters to accomplish this numerically are,

$$nmax := Number\ of\ values\ to\ analyse$$
$$M := Number\ of\ paths$$
$$\varepsilon := Threshold\ value$$
$$tinf := Lower\ limit\ of\ interval\ to\ establish\ convergence$$
$$tsup := Upper\ limit\ of\ interval\ to\ establish\ convergence$$

The values of these parameters for the numerically studied example are,

$$nmax = A = 11 \tag{35.1}$$

$$M = 500 \tag{35.2}$$

$$\varepsilon = 0.01 \tag{35.3}$$

$$tinf = 0 \tag{35.4}$$

$$tsup = 1 \tag{35.5}$$

## 4. Discussion

The eigenvalue-based model of transition-state dissociation constants of a ligand-macromolecular complex asserts that there are several intermediate- or transition-states of a complex and that each of these has the potential to modify the biochemical process that the complex participates in.

### *4.1. Transition-state dissociation constants as a model of biochemical function for ligand-macromolecular complexes*

Ligand-macromolecular complexes, *in vivo*, possess a finite and in most cases, an incomparable number of atoms/residues. The theoretical results establish definition(s), bounds and metrics to assess biochemical function for both, monomer (**T1–T4, C1–C3**)- and multimer (**T5–T7**)-forms. The numerical data suggests that the set of transition-state dissociation constants can be finite, converge and retain statistical relevance (Figure 1).

***Proposition*** **(P):** The transition-state dissociation constants for the monomer ($z = Z = 1$)- and multimer ($Z \geq 2$)-forms of a ligand-macromolecular complex with a finite number of atoms/residues of each $(I, J)$ per monomer,

$$\left\{u \in \boldsymbol{Kd}(\boldsymbol{\mu \mathcal{C}_z}.\boldsymbol{\mathcal{C}^Y}) | \mathcal{C}_z \in \boldsymbol{\mathcal{C}}, \mu \in \boldsymbol{\mathcal{L}}; z = 1,2 \ldots Z; A = \{I, J\}\right\} \qquad \textit{Def. (10)}$$

is the finite set,

$$\boldsymbol{Kd}(\boldsymbol{\mu \mathcal{C}_z}.\boldsymbol{\mathcal{C}^Y}) \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}})$$

Here,

$$I = \#\mathcal{C}_z | \mathcal{C}_z \in \boldsymbol{\mathcal{C}} \tag{36}$$

$$J = \#\mu | \mu \in \boldsymbol{\mathcal{L}} \tag{37}$$

where,

$$\boldsymbol{Kd}(.) := Set\ of\ constrained\ eigenvalue - based\ transition - state\ dissociation\ constants$$

$$I := Finite\ number\ of\ atoms\ or\ residues\ of\ macromolecule$$

$$J := Finite\ number\ of\ atoms\ or\ residues\ of\ ligand$$

$$\mathcal{C}_{\mu z}.\mathcal{C}^Y := Multimer\ form\ of\ ligand - macromolecular\ complex$$

Enzyme-mediated catalysis, or lack thereof, results in metabolic enzyme disorders and may be inherited (inborn errors of metabolism) or acquired [28]. In order to assess the biomedical relevance

of modeling ligand-macromolecule interactions as transition-state dissociation constants, the clinical outcomes of amino acid substitutions of selected enzyme homo- or hetero-dimers are examined (Table 1). These outcomes, i.e., benign, likely benign, pathologic, likely pathologic, conflicting, uncertain significance, are defined in accordance with the prevalent nomenclature of the ClinVar database [16]. Here, the data annotated as "uncertain significance" are those sequence variants with a high likelihood ($\approx 90-95\%$) of being "benign" or "pathogenic" [17]. This means they are likely to classified as "true positive", and if ignored will result in a "false negative". On the other hand, an outcome designated as with a "conflicting interpretation" is likely to be due to unresolved contradictory findings in the presence or absence of confounding factors. If we assume perfect contradiction, i.e., 50%, and couple this with the previous result, we get a $\approx 70-73\%$ possibility that the variant of interest is a "true positive". This means, that here too, if missed a "false negative" will result. The metric of choice is the Recall (R) percentage,

$$R = \frac{TP}{TP+FN} \times 100 \tag{38}$$

$$R := Recall$$

$$TP := Known\ positives\ (benign, likely\ benign, pathogenic, likely\ pathogenic) \qquad Def.\ (12)$$

$$FN := Likely\ positives\ (conflicting\ data, uncertain\ significance) \qquad Def.\ (13)$$

**Table 1.** Analysis of clinical outcomes of amino acid substitutions of selected enzyme homo- and hetero-dimer forms present in ClinVar.

| | Enzyme | EC | CV | SNP | M | Co | US | B | LB | P | LP | FN | TP | R (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Glucokinase | 2.7.1.2 | 778 | 619 | 370 | 49 | 157 | 4 | 4 | 65 | 113 | 206 | 186 | 47.45 |
| 2 | Pyruvate kinase | 2.7.1.40 | 1239 | 629 | 209 | 14 | 135 | 7 | 7 | 30 | 20 | 149 | 64 | 30.05 |
| 3 | Cathepsin A | 3.4.16.x | 1776 | 1270 | 501 | 20 | 382 | 26 | 23 | 26 | 34 | 402 | 109 | 21.33 |
| 4 | Pyruvate dehydrogenase | 1.2.4.1 | 2317 | 1591 | 584 | 25 | 391 | 35 | 46 | 62 | 43 | 416 | 186 | 30.9 |
| 5 | Phosphofructokinase 1 | 2.7.1.11 | 165 | 32 | 10 | 2 | 4 | 2 | 1 | 1 | 1 | 6 | 5 | 45.45 |
| 6 | Phosphofructokinase 2 | 2.7.1.105 | 206 | 76 | 23 | 1 | 17 | 1 | 1 | 2 | 1 | 18 | 5 | 21.74 |
| 7 | Cystathione beta-synthase | 4.2.1.22 | 930 | 744 | 255 | 21 | 172 | 3 | 4 | 46 | 32 | 193 | 85 | 30.58 |
| 8 | DNA topoisomerase II | 5.6.2.2 | 466 | 319 | 152 | 0 | 122 | 11 | 14 | 3 | 1 | 122 | 29 | 19.21 |
| 9 | Guanylate cyclase 1 | 4.6.1.2 | 1572 | 1117 | 576 | 29 | 417 | 11 | 10 | 34 | 59 | 446 | 114 | 20.36 |
| 10 | Phenylalanine hydroxylase | 1.14.16.1 | 1314 | 1102 | 631 | 5 | 182 | 0 | 3 | 161 | 254 | 187 | 418 | 69.09 |

Note: EC: Enzyme commission number; CV: Number of clinical variants; SNP: Single nucleotide polymorphisms; M: Missense mutations; Co: Conflicting data; US: Uncertain significance; B: Benign; LB: Likely; P: Pathogenic; LP: Likely pathogenic; FN: False negative (Co + US); TP: True positive (B + LB + P + LP); R: Recall ($\frac{TP}{TP+FN} \times 100$).

It is clear from these data that amino acid substitutions (nature, type), either alone or in combination, comprise distinct transition-states and are significant contributors to the biochemical function of each enzyme dimer ($Recall \approx 19-70\%$). Since each of these states will result in a distinct $Kd$, it is easily inferred that the transition-state dissociation constants, for a complex, may be more representative of biochemical function (**T1–T4, C1–C3**).

*4.2. Ligand-macromolecule biochemistry as a function of its transition-state dissociation constants*

The discussion, *vide supra*, presents and highlights the biomedical relevance of modeling ligand-macromolecular interactions as transition-state dissociation constants. The results for monomer- and multimer-forms of ligand-macromolecular complexes are mathematically rigorous and have been validated, *in silico*. The results are now examined in context of biochemical function (enzyme, non-enzyme) for selected cases.

***Case 1****: Oxygen sensitive variants of non-haem iron (II)- and 2-oxoglutarate-dependent dioxygenases*

The non-haem iron (II)- and 2-oxoglutarate-dependent dioxygenases $(EC\ 1.14.x.y)$, comprise a large superfamily of enzymes, are present in all kingdoms of life and is chemically diverse (variable reaction chemistry, multiple substrates) [12,13,29]. Clinically relevant members include Phytanoyl-CoA dioxygenase (PHYT), Lysine Hydroxylases, and the Proline 4-Hydroxylases (P4H) amongst several others [12–15,29]. These enzymes have important roles in Phytanic acid metabolism and collagen maturation, with sub-optimal activities contributing to diseases such as Refsum and the Ehlers-Danlos (ED)-syndrome [14,15,30]. Here, too, the outcomes (clinical, non-clinical) associated with substitution mutations for PHYT suggest that modeling ligand-macromolecular interactions as transition-state dissociation constants may be a better index of biochemical function (**T1**–**T4, C1**–**C3, P**) [31].

P4Hs, are classified as being either hypoxia-sensitive $(H-P4H \equiv HP4H; EC\ 1.14.11.29)$ or collagen transforming $(C-P4H \equiv CP4H; EC\ 1.14.11.2)$ [28]. These reactions may be written,

$$HP4H - HIF + O_2 + 2OG + Proline \underset{r_b}{\overset{r_f}{\rightleftarrows}} Hydroxyproline + SA + CO_2 \quad RXN(2)$$

$$CP4H + O_2 + 2OG + Proline \underset{r_b}{\overset{r_f}{\rightleftarrows}} Hydroxyproline + SA + CO_2 \quad RXN(3)$$

$$\begin{aligned} r_f, r_b &:= Rate\ constants\ for\ forward\ and\ backward\ reactions\ at\ steady\ state \\ HP4H &:= Hypoxia\ inducible\ factor-dependent\ Proline\ 4-Hydroxylase \\ HIF \equiv \mu_{high} &:= Hypoxia-inducible\ factor\ (high-affinity\ modifier) \\ 2OG &:= 2-oxoglutarate \\ SA &:= Succinic\ acid \\ CO_2 &:= Carbon\ dioxide \end{aligned}$$

The amino acid identity between $HP4H$ and $CP4H$ notwithstanding, there are significant differences between the molecular biology that they exhibit. This implies that despite the similarity of co-factor $(iron(II))$, substrate $(L-Proline)$ and co-substrate $(2-oxoglutarate)$, the binding affinities for molecular dioxygen vary considerably [7,32],

$$Km_{HP4H} = 0.1 - 0.76\ mM \quad (39)$$

$$Km_{CP4H} = 0.03 - 1.5\ mM \quad (40)$$

The turnover numbers for the cognate substrate, too, differ significantly [7],

$$Kcat_{HP4H} = 0.015 - 0.733\ s^{-1} \quad (41)$$

$$Kcat_{CP4H} = 0.0188 - 0.02\ s^{-1} \quad (42)$$

Clearly, a plausible explanation for these disparate empirical observations is the binding of the

hypoxia-inducible factors (HIF) to P4H. The hypoxia-inducible factors (HIFs), are a family $(n = 3)$ of transcription factors which sense hypoxia and trigger the upregulation of hypoxia-dependent genes [32–34]. Here, although hypoxia-inducible factor, is a full length protein, the actual binding site is the C-terminal end of HP4H [7,35].

$$Kd \approx 0.000016 - 0.023\, mM \tag{43}$$

Some of these observations may be inferred from the partitioning of the transition-state dissociation constants into distinct subsets (Table 2):

**Table 2.** Ligand-macromolecular biochemistry as a function of its transition-state dissociation constants.

| | Case 1 | Case 2 |
|---|---|---|
| Ligand $(\boldsymbol{\mu} \in \boldsymbol{\mathcal{L}})$ | Hypoxia-inducible factor | Peptide |
| Macromolecule $(\boldsymbol{c} \equiv \boldsymbol{\mathcal{C}_z} \in \boldsymbol{\mathcal{C}})$ | HP4H, CP4H | $M1\beta$ |
| Primary complex $\langle \boldsymbol{c}\vert\boldsymbol{\mu}\rangle$ | $\langle \boldsymbol{HP4H}\vert\boldsymbol{\mu}\rangle$, $\langle \boldsymbol{CP4H}\vert\boldsymbol{\mu}\rangle$ | $\langle \boldsymbol{M1\beta}\vert\boldsymbol{\mu}\rangle$ |
| High-affinity variant | $\boldsymbol{Kd}(\boldsymbol{HP4H}\vert\boldsymbol{\mu_{high}})$ | $\boldsymbol{Kd}(\boldsymbol{M1\beta}\vert\boldsymbol{\mu_{high}})$ |
| $:= \boldsymbol{Kd}(\boldsymbol{\mu_{high}\mathcal{C}_z})$ | $= \boldsymbol{Kd}(\boldsymbol{\mu_{high}HP4H})$ | $= \boldsymbol{Kd}(\boldsymbol{\mu_{high}M1\beta})$ |
| Low-affinity variant | $\boldsymbol{Kd}(\boldsymbol{CP4H}\vert\boldsymbol{\mu_{low}})$ | $\boldsymbol{Kd}(\boldsymbol{M1\beta}\vert\boldsymbol{\mu_{low}}) = \boldsymbol{Kd}(\boldsymbol{\mu_{low}M1\beta})$ |
| $:= \boldsymbol{Kd}(\boldsymbol{\mu_{high}\mathcal{C}_z})$ | $= \boldsymbol{Kd}(\boldsymbol{\mu_{low}CP4H})$ | |
| Higher-order complex $\langle \boldsymbol{c}\vert\boldsymbol{\mu}\rangle.\mathcal{C}^{\gamma}$ | --- | $\langle \boldsymbol{M1\beta}\vert\boldsymbol{\mu}\rangle.PLC$ |
| High-affinity variant | --- | $\boldsymbol{Kd}(\boldsymbol{M1\beta}\vert\boldsymbol{\mu_{high}}.\boldsymbol{PLC})$ |
| $:= \boldsymbol{Kd}(\boldsymbol{\mu_{high}\mathcal{C}_z}.\mathcal{C}^{\gamma})$ | | $= \boldsymbol{Kd}(\boldsymbol{\mu_{high}M1\beta}.\boldsymbol{PLC})$ |
| Low-affinity variant | --- | $\boldsymbol{Kd}(\boldsymbol{M1\beta}\vert\boldsymbol{\mu_{low}}.\boldsymbol{PLC})$ |
| $:= \boldsymbol{Kd}(\boldsymbol{\mu_{low}\mathcal{C}_z}.\mathcal{C}^{\gamma})$ | | $= \boldsymbol{Kd}(\boldsymbol{\mu_{low}M1\beta}.\boldsymbol{PLC})$ |
| Functional relevance | $R_{HP4H}(t) \gg R_{CP4H}(t)$ | $M1\beta\vert\mu_{high}.PLC \rightarrow \alpha M1\beta$ (Anterograde)<br>$M1\beta\vert\mu_{low}.PLC \rightarrow rM1\beta$ (Retrograde) |

Note: $\boldsymbol{\mathcal{C}_{z\mu}}$: All pairwise-atom/residue based square interaction matrix of ligand and macromolecule, $\langle \boldsymbol{c}\vert\boldsymbol{\mu}\rangle$; $\boldsymbol{\mathcal{K}\mathcal{C}_{z\mu}}$: Diagonal matrix of the interactions of a ligand and macromolecule; $z_{i=j} = z$: Set of eigenvalues of $\boldsymbol{\mathcal{K}\mathcal{C}_{z\mu}}$ where $z \in diag(\boldsymbol{\mathcal{K}\mathcal{C}_{z\mu}}) \subset \mathbb{C}$; $\boldsymbol{Kd}(\boldsymbol{c\mu})$: Set of eigenvalue-based transition-state dissociation constants for monomer- and multimer-forms of ligand-macromolecular complexes, $\{\omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) = \alpha_{i=j} = Re(z) \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \cap (0,1)\}$; $\mu_{high}, \mu_{low}$: High- and low-affinity variants of an arbitrary ligand, $\mu = \{\mu_{high}, \mu_{low}\} \in \boldsymbol{\mathcal{L}}$; $\boldsymbol{\mu\mathcal{C}_z}.\mathcal{C}^{\gamma}$: Higher-order complex of ligand and macromolecule; $\boldsymbol{Kinf}$: Subset of transition-state dissociation constants of ligand-macromolecular interaction; $\boldsymbol{Ksup}$:Subset of transition-state dissociation constants of ligand-macromolecular interaction; $R(t)$: Rate of reaction; HP4H: Hypoxia stimulated- and Collagen Proline 4-Hydroxylase; CP4H: $M1\beta$: Heterodimer of major histocompatibility complex I (MHC1) with beta-2 microglobulin; PLC: Peptide loading complex.

In particular, binding of HIF restricts the range of the binding affinity of HP4H for molecular oxygen significantly,

$$\frac{\Delta Km_{HP4H}}{\Delta Km_{CP4H}} \times 100 \approx 45\% \tag{44}$$

Here, the set of conformers of HP4H once bound to HIF ensures that the catalytic activity of

HP4H for HIF is significantly reduced in the presence of hypoxia. This will extend the half-life of HIF and facilitate transcription of HIF-responsive genes [36,37]. In contrast, CP4H exhibits no such differential activity. Furthermore, there is a significant variation in the catalytic activity (turnover number) of these enzymes for their cognate substrate (L-Proline),

$$\frac{max(Kcat_{HP4H})-min(Kcat_{HP4H})}{max(Kcat_{CP4H})-min(Kcat_{CP4H})} = \frac{\Delta Kcat_{HP4H}}{\Delta Kcat_{CP4H}} \approx 600 \quad (45)$$

These data suggest that a ligand when bound to a macromolecule can affect the rate at which the resulting complex assembles or disassembles and thereby influence biochemical function. Hence, partitioning the transition-state dissociation constants of the ligand-macromolecular complex into distinct subsets may offer valuable insights into the *in vivo* function of enzymes in physiological and pathological states (**T4, C2, C3**).

***Case 2:*** *Generic model of MHC1-mediated high-affinity peptide export*

The peptide loading complex (PLC) is a higher-order $(Z > 2; Tapasin, ERp57, MHC1)$ complex that assembles at the endoplasmic reticulum (ER)-membrane and functions to transport cytosolic peptides into the ER-lumen *en route* to the plasma membrane [38,39]. Whilst, regulation of this process, by Tapasin is well studied, the role of peptides and the possible mechanism(s) of action is unclear [40–43]. A low-affinity peptide-driven (LAPD)-model of the MHC1-mediated export of high-affinity peptides to the plasma membrane of nucleated cells has been proposed and investigated *in silico* [44]. A major proponent of this study was simulating the differential disassembly of PLC in response to peptides with varying affinities (high, low) for the MHC1-$\beta_2$-microglobulin [44]. In fact, data from the simulations suggested that low-affinity peptides may not only actively participate in the transport of high-affinity peptide export, but could also regulate the same [44]. Another interesting observation discussed was the role of low-affinity peptides in priming the MHC1-export apparatus, such that irrespective of the nature of the cellular insult (acute, chronic), export of high-affinity peptides was rapid, continuous and efficient [44].

Utilizing the partition schema for the transition-state dissociation constants from the current analysis (Table 2), we can model and rewrite the differential disassembly of the PLC,

$$\langle \boldsymbol{M1\beta|\mu_{high}} \rangle . PLC \underset{r_b}{\overset{r_f}{\rightleftarrows}} Tapasin - ERp57 + aM1\beta \quad RXN(4)$$

$$\langle \boldsymbol{M1\beta|\mu_{low}} \rangle . PLC \underset{r_b}{\overset{r_f}{\rightleftarrows}} Tapasin - ERp57 - \mu_{low} + rM1\beta \quad RXN(5)$$

$$\begin{aligned}
r_f, r_b &:= Rate\ constants\ for\ forward\ and\ backward\ reactions\ at\ steady\ state \\
M1\beta &:= Heterodimer\ of\ MHC1\ with\ beta-2\ microglobulin \\
\mu_{low} &:= Low-affinity\ peptide \\
\mu_{high} &:= High-affinity\ peptide \\
\langle \boldsymbol{M1\beta|\mu_{low}} \rangle &:= Set\ of\ peptides\ with\ low\ affinity\ for\ MHC1\ and\ in\ complex\ with\ MHC1 \\
\langle \boldsymbol{M1\beta|\mu_{high}} \rangle &:= Set\ of\ peptides\ with\ high\ affinity\ for\ MHC1\ and\ in\ complex\ with\ MHC1 \\
eM1\beta &:= Net\ exportable\ complex\ of\ high-affinity\ peptide\ with\ MHC1 \\
aM1\beta &:= Anterograde-derived\ exportable\ complex\ of\ high-affinity\ peptide\ with\ MHC1 \\
rM1\beta &:= Retrograde-derived\ exportable\ complex\ of\ high-affinity\ peptide\ with\ MHC1 \\
PLC &:= Peptide\ loading\ complex \\
ERp57 &:= Endoplasmic\ reticulum\ protein\ disulfide\ isomerase
\end{aligned}$$

The appropriate dissociation constants are,

$$Kd(RXN4) = Kd_{RXN4} = \frac{[Tapasin{-}ERp57][aM1\beta]}{[M1\beta|\mu_{high}]^{L_{M1\beta|\mu_{high}}{}^{\geq 0}}[PLC]^{L_{PLC}\geq 0}} \tag{46}$$

$$Kd(RXN5) = Kd_{RXN5} = \frac{[Tapasin{-}ERp57{-}\mu_{low}][rM1\beta]}{[M1\beta|\mu_{low}]^{L_{M1\beta|\mu_{low}}{}^{\geq 0}}[PLC]^{L_{PLC}\geq 0}} \tag{47}$$

Rewriting these equations in terms of the peptide-bound MHC1,

$$Kd_{RXN4} = \frac{\zeta}{[M1\beta|\mu_{high}]^{L_{M1\beta|\mu_{high}}{}^{\geq 0}}} \tag{48}$$

where,

$$\zeta = \frac{[Tapasin{-}ERp57][aM1\beta]}{[PLC]^{L_{PLC}\geq 0}} \tag{48.1}$$

$$Kd_{RXN5} = \frac{\zeta}{[M1\beta|\mu_{low}]^{L_{M1\beta|\mu_{low}}{}^{\geq 0}}} \tag{49}$$

where,

$$\zeta = \frac{[Tapasin{-}ERp57{-}\mu_{low}][rM1\beta]}{[PLC]^{L_{PLC}\geq 0}} \tag{49.1}$$

Clearly,

$$Kd_{RXN4} \propto \frac{1}{[M1\beta|\mu_{high}]^{L_{M1\beta|\mu_{high}}{}^{\geq 0}}} \tag{50}$$

$$Kd_{RXN5} \propto \frac{1}{[M1\beta|\mu_{low}]^{L_{M1\beta|\mu_{low}}{}^{\geq 0}}} \tag{51}$$

These results suggest that,

$$Kd_{RXN4} \simeq 1.0\ (Perfect\ disassociation)\ and\ [aM1\beta] \to \infty \tag{52}$$

$$Kd_{RXN5} \simeq 0.0\ (Perfect\ association)\ and\ [rM1\beta] \to 0 \tag{53}$$

and is in accordance with existing empirical and simulation data,

$$[\mu_{high}] \lll [\mu_{low}], [aM1\beta] \ggg [rM1\beta] \tag{54}$$

Here, the partitioning of transition-state dissociation constants into low- and high-affinity peptides for the MHC1 can provide valuable insights into the underlying molecular biology of MHC1-mediated high-affinity peptide transport under physiological and pathological conditions (**T5–T7, P**) [42–44].

## 5. Conclusions

The work presented models ligand-macromolecular interactions as eigenvalue-based transition-state disassociation constants. The interaction matrix is an all-atom/residue pairwise comparison between the ligand and macromolecule and comprises numerical values drawn randomly

from a standard uniform distribution. The transition-state dissociation constants are the strictly positive real part of all complex eigenvalues of this ligand-macromolecular interaction matrix, belong to the open interval (0,1) and form a sequence whose terms are finite, monotonic, non-increasing and convergent. The findings are rigorous, numerically robust and can be extended to higher-order complexes. The study, additionally, suggests a schema by which a ligand may be partitioned into high- and low-affinity variants. This study, although theoretical offers a plausible explanation into the underlying biochemistry (enzyme-mediated substrate catalysis, assembly/disassembly and inhibitor kinetics) of ligand-macromolecular complexes. Future investigations may include assigning weights to each interaction, investigating origins of co-operativity in enzyme catalysis and inhibitory kinetics amongst others.

## 6. Proofs

This section provides formal proofs for the included theorems, corollaries and proposition.

**Proof (T1)**:

From *Defs.* (4) and (5),

$$\text{For every } \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu})\ \exists g(\omega) \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) | g^{-1} \circ g(\omega) = \omega \tag{55}$$

Let,

$$\omega_x, \omega_y \in \boldsymbol{Kd}(\boldsymbol{c\mu}) | g(\omega_x), g(\omega_y) \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}); x \neq y; \{x, y\} \leq A$$

If,

$$g(\omega_x) = g(\omega_y) \tag{56}$$

then,

$$g^{-1} \circ g(\omega_x) = g^{-1} \circ g(\omega_y) \tag{57}$$

$$\Rightarrow \omega_x = \omega_y \tag{58}$$

If,

$$t = 0 | t \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{59}$$

From (55),

$$g^{-1} \circ g(t) = 0 \notin \boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{60}$$

Similarly, For,

$$t \geq 1 | t \in \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{61}$$

From (55),

$$g^{-1} \circ g(\omega) \geq 1 \notin \boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{62}$$

From (58), (60) and (62),

■

**Proof (T2):** (By induction)

For $a = 1$,

$$\omega_a = \max\left(\boldsymbol{Kd}(\boldsymbol{c\mu})\right) < 1 \quad \text{(By Def. (5))}, (63)$$

Assume $a = A - 1$,

$$\omega_a \geq \omega_{A-1} \quad (64)$$

Then $\exists a = A$,

$$\omega_a \geq \omega_{A-1} \geq \omega_A \quad (65)$$

For $a = 1$,

$$\omega_a = \min\left(\boldsymbol{Kd}(\boldsymbol{c\mu})\right) > 0 \quad \text{(By } \textit{Def.} \text{ (5))}, (66)$$

Assume $a = A - 1$,

$$\omega_a \leq \omega_{A-1} \quad (67)$$

$$\Rightarrow 0 < \{\omega_a\}_{a \leq a+1} \sim \boldsymbol{Kd}(\boldsymbol{c\mu}) < 1 \; \forall \omega$$

■

**Proof (T3):**

From (T2),

$$\{\omega_a\}_{a \leq a+1} < 1$$

Choose $\varepsilon \in \mathbb{R}_+, \varepsilon \to 0$,

$$\varepsilon^{a>A} < |\omega_{a>A} - 1| < \varepsilon$$

For every $a > A$,

$$\left|\lim_{a\to\infty} \omega_{a>A} - 1\right| < \varepsilon$$

$$\lim_{a\to\infty} |\omega_{a>A} - 1| < \varepsilon$$

$$\lim_{a\to\infty} \omega_{a>A} = 1 \quad (69)$$

Similarly,

$$\{\omega_a\}_{a \leq a+1} > 0$$

Choose $\varepsilon \in \mathbb{R}_+, \varepsilon \to 0$,

$$\frac{1}{\varepsilon} > |\omega_{a>A} - 0| > \varepsilon$$

For every $a > A$,

$$\left|\lim_{a\to\infty} \omega_{a>A} - 0\right| > \varepsilon$$

$$\lim_{a\to\infty} |\omega_{a>A} - 0| > \varepsilon$$

$$\lim_{a\to\infty} \omega_{a>A} = 0 \tag{70}$$

From (69) and (70),

$$\lim_{a\to\infty} \{\omega_a\}_{a\le a+1} = \{0,1\}$$

■

**Proof (C1):**

Assume,

$$sup\{\omega_a\}_{a\le a+1} = max\{\omega_a\}_{a\le a+1} = \omega_{a=1}$$

Choose $\varepsilon \in \mathbb{R}_+, \varepsilon \to 0$
then for any $a = 1,2..A$, we can find,

$$\varepsilon^A.\omega_{a=1} \lll \omega_{a=1} \tag{71}$$

$$\varepsilon^A.\omega_{a=1} \le \omega_{a=A} \tag{72}$$

$$\omega_{a=1} \le \omega_{a=A}.\left(\frac{1}{\varepsilon^A}\right) \tag{73}$$

Let $\delta \in \mathbb{R}_+, \delta \to 0$,

$$\omega_{a=1} - \delta < \omega_{a=A}.\left(\frac{1}{\varepsilon^A}\right) \tag{74}$$

Assume,

$$inf\{\omega_a\}_{a\le a+1} = min\{\omega_a\}_{a\le a+1} = \omega_{a=A}$$

Choose $\varepsilon \in \mathbb{R}_+, \varepsilon \to 0$
then for any $a = 1,2..A$, we can find,

$$\varepsilon^A.\omega_{a=A} < \omega_{a=A} \tag{75}$$

$$\omega_{a=A} \lll \omega_{a=A}.\left(\frac{1}{\varepsilon^A}\right) \tag{76}$$

$$\omega_{a=1} \le \omega_{a=1}.\left(\frac{1}{\varepsilon^A}\right) \tag{77}$$

Let $\delta \in \mathbb{R}_+, \delta \to 0$,

$$\omega_{a=1} < \omega_{a=1}.\left(\frac{1}{\varepsilon^A}\right) + \delta \tag{78}$$

From (74) and (78),

■

**Proof (T4):**

From (T2 and T3), (C1 and C2)

Case (1)

If

$$\{\omega_a\}_{a>A} = \boldsymbol{Ksup} \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}})$$

then,

$$\{\omega_a\}_{a\leq A} \in \boldsymbol{Kinf}\sim\boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{79}$$

Case (2)

If

$$\{\omega_a\}_{a>A} = \boldsymbol{Kinf} \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}})$$

then,

$$\{\omega_a\}_{a\leq A} \in \boldsymbol{Ksup}\sim\boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{80}$$

From (79) and (80),

$$\boldsymbol{Kinf} \cap \boldsymbol{Ksup} = \{\emptyset\} \tag{81}$$

Since,

$$\boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \supset (\boldsymbol{Kinf} \cup \boldsymbol{Ksup}) \cup (\boldsymbol{Kinf} \cap \boldsymbol{Ksup})$$

From (79)–(81),

$$\boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \supset \boldsymbol{Kinf} \cup \boldsymbol{Ksup} \tag{82}$$

$$\Rightarrow \boldsymbol{Kd}(\boldsymbol{c\mu}) \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}})$$

■

**Proof (C3):**

If,

$$\exists \mu \in \boldsymbol{\mathcal{L}} | \#\boldsymbol{Ksup} \ggg \#\boldsymbol{Kinf}\ \forall a$$

Then,

$$\mu \equiv \mu_{low} \tag{83}$$

Similarly, if,

$$\exists \mu \in \boldsymbol{\mathcal{L}} | \#\boldsymbol{Kinf} \ggg \#\boldsymbol{Ksup}\ \forall a$$

Then,

$$\mu \equiv \mu_{high} \tag{84}$$

From (83) and (84),

■

**Proof** (**T5**) (By induction)

Assume $z = 1; Z \geq 2,\ y = 1; Z = 2,$

$$\mathcal{C}_1.\prod_{y=1}^{y=Y=Z-1}\mathcal{C}_y = \mathcal{C}_1.\prod_{y=1}^{y=1}\mathcal{C}_y \tag{85}$$

$$= \mathcal{C}_1.\mathcal{C}_Y \tag{85.1}$$

$$= \mathcal{C}_1.\mathcal{C}^Y \tag{85.2}$$

Assume truth for $y = Y \gg 1,$

$$\mathcal{C}_1.\prod_{y=1}^{y=Y=Z-1}\mathcal{C}_y = \mathcal{C}_z.(\mathcal{C}_1 \ldots . \mathcal{C}_Y) \tag{86}$$

$$= \mathcal{C}_1.\mathcal{C}^Y \tag{86.1}$$

For $y = Y + 1$,

$$\mathcal{C}_1.\prod_{y=1}^{y=Y+1}\mathcal{C}_y = \mathcal{C}_1.\prod_{y=1}^{y=Y+1}\mathcal{C}_y \tag{87}$$

$$= \mathcal{C}_1.\left(\prod_{y=1}^{y=Y}\mathcal{C}_y\right).\left(\prod_{y=1}^{y=1}\mathcal{C}_y\right) \tag{87.1}$$

$$= \mathcal{C}_1.\mathcal{C}_Y^Y.\mathcal{C}_1^1 \tag{87.2}$$

$$= \mathcal{C}_1.\mathcal{C}^{Y+1} \tag{87.3}$$

From (85)–(87),

■

**Proof (T6):**

From *Defs. (6–8),*

$$\mu\mathcal{C}_z.\mathcal{C}^Y \equiv \mu\mathcal{C}_z \equiv \langle \boldsymbol{c}|\boldsymbol{\mu}\rangle \tag{88}$$

$$\Rightarrow \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y}) \sim \boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{89}$$

$$\Rightarrow \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y}) \cap \boldsymbol{Kd}(\boldsymbol{c\mu}) \tag{90}$$

From *Def. (5)*,

$$\omega_{a\in[1,A]} \in \left(\boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y}) \cap \boldsymbol{Kd}(\boldsymbol{c\mu})\right) \tag{91}$$

$$\left(\omega_{a\in[1,A]} \in \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y})\right) \cap \left(\omega_{a\in[1,A]} \in \boldsymbol{Kd}(\boldsymbol{c\mu})\right) \tag{92}$$

From (T2–T4), (C1–C3),

$$\forall a \qquad \left(\omega_{a=1,2\ldots A} \subseteq \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y})\right) \cap \left(\omega_{a=1,2\ldots A} \subseteq \boldsymbol{Kd}(\boldsymbol{c\mu})\right) \tag{93}$$

$$\Rightarrow \left(\{\omega_a\}_{a=1,2\ldots A} \sim \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y})\right) \cap \left(\{\omega_a\}_{a=1,2\ldots A} \sim \boldsymbol{Kd}(\boldsymbol{c\mu})\right) \tag{94}$$

Conversely, let,

$$u \in \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y}); \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu})$$

From (92) and (94),

$$\forall \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu})\ \exists u \in \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y}) | u = h(\omega) \tag{95}$$

$$\forall u \in \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y})\ \exists \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) | \omega = h^{-1}(u) = h^{-1}\left(h(\omega)\right) \tag{96}$$

From (95) and (96),

■

**Proof** (**T7**):

From (T6),

$$h^{-1} \circ h: \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu})$$

From (T1), *Def. (4)*

$$g: \omega \in \boldsymbol{Kd}(\boldsymbol{c\mu}) \mapsto \boldsymbol{Kd}(\boldsymbol{c\mu}) \cup \mathbb{R} = \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}})$$

Rewriting,

$$g: \omega \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{97}$$

$$g\left(h^{-1}(u)\right) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{98}$$

$$g \circ h^{-1}\left(h(\omega u)\right) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{99}$$

$$g \circ h^{-1} \circ h(\omega) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{100}$$

■

**Proof (P)**:

From (T1–T7) and *Defs. (4–8, 10),*

For $z = Z = 1$,

$$\mu\mathcal{C}_z.\mathcal{C}^Y = \mu\mathcal{C}_z \tag{101}$$

$$g \circ h^{-1}\left(u \in \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z})\right) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R} \tag{102}$$

$$g \circ h^{-1} \circ h(\omega) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{102.1}$$

$$\Rightarrow \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z}) \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{103}$$

For $Z \geq 2$,

$$g \circ h^{-1}\left(u \in \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y})\right) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \subset \mathbb{R} \tag{104}$$

$$g \circ h^{-1} \circ h(\omega) \mapsto \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{104.1}$$

$$\Rightarrow \boldsymbol{Kd}(\boldsymbol{\mu\mathcal{C}_z.\mathcal{C}^Y}) \subset \boldsymbol{Kd}(\boldsymbol{\mathcal{C}_{z\mu}}) \tag{105}$$

From (103) and (105),

■

## Acknowledgments

This work is funded by an early career intramural grant (Code No. A-766) awarded to SK by the All India Institute of Medical Sciences (AIIMS, New Delhi, INDIA).

## Conflict of interest

The author declare there is no conflict of interest.

## References

1. M. Su, Y. Ling, J. Yu, J. Wu, J. Xiao, Small proteins: untapped area of potential biological importance, *Front. Genet*., **4** (2013), 286. https://doi.org/10.3389/fgene.2013.00286

2. M. B. Pappalardi, D. E. McNulty, J. D. Martin, K. E. Fisher, Y. Jiang, M. C. Burns, et al., Biochemical characterization of human HIF hydroxylases using HIF protein substrates that contain all three hydroxylation sites, *Biochem. J.*, **436** (2011), 363–369. https://doi.org/10.1042/BJ20101201
3. L. Esposito, M. Ferrara, L. Tomasi, P. De Filippo, Hereditary methemoglobinemia caused by NADH methemoglobin reductase deficiency, *Pediatria (Napoli)*, **84** (1976), 411–422.
4. D. E. Koshland Jr., G. Nemethy, D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits, *Biochemistry*, **5** (1966), 365–385. https://doi.org/10.1021/bi00865a047
5. J. Monod, J. Wyman, J. P. Changeux, On the nature of allosteric transitions: A plausible model, *J. Mol. Biol.*, **12** (1965), 88–118. https://doi.org/10.1016/S0022-2836(65)80285-6
6. J. J. Hutton Jr., A. L. Trappel, S. Udenfriend, Requirements for alpha-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase, *Biochem. Biophys. Res. Commun.*, **24** (1966), 179–184. https://doi.org/10.1016/0006-291X(66)90716-9
7. S. Pektas, C. Y. Taabazuing, M. J. Knapp, Increased turnover at limiting $O_2$ concentrations by the $Thr^{387}$ → Ala variant of HIF-Prolyl Hydroxylase PHD2, *Biochemistry*, **54** (2015), 2851–2857. https://doi.org/10.1021/bi501540c
8. K. S. Hewitson, B. M. Lienard, M. A. McDonough, I. J. Clifton, D. Butler, A. S. Soares, et al., Structural and mechanistic studies on the inhibition of the hypoxia-inducible transcription factor hydroxylases by tricarboxylic acid cycle intermediates, *J. Biol. Chem.*, **282** (2007), 3293–301. https://doi.org/10.1074/jbc.M608337200
9. K. M. Paulsson, M. J. Kleijmeer, J. Griffith, M. Jevon, S. Chen, P. O. Anderson, et al., Association of tapasin and COPI provides a mechanism for the retrograde transport of major histocompatibility complex (MHC) class I molecules from the Golgi complex to the endoplasmic reticulum, *J. Biol. Chem.*, **277** (2002), 18266–18271. https://doi.org/10.1074/jbc.M201388200
10. R. Benesch, R. E. Benesch, The effect of organic phosphates from the human erythrocyte on the allosteric properties of haemoglobin, *Biochem. Biophys. Res. Commun.*, **26** (1967), 162–167. https://doi.org/10.1016/0006-291X(67)90228-8
11. P. J. Mulquiney, W. A. Bubb, P. W. Kuchel, Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: in vivo kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using 13C and 31P NMR, *Biochem. J.*, **342** (1999), 567–580. https://doi.org/10.1042/bj3420567
12. S. Martinez, R. P. Hausinger, Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-dependent Oxygenases, *J. Biol. Chem.*, **290** (2015), 20702–20711. https://doi.org/10.1074/jbc.R115.648691
13. I. J. Clifton, M. A. McDonough, D. Ehrismann, N. J. Kershaw, N. Granatino, C. J. Schofield, Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins, *J Inorg Biochem.*, **100** (2006), 644–669. https://doi.org/10.1016/j.jinorgbio.2006.01.024
14. K. L. Gorres, R. T. Raines, Prolyl 4-hydroxylase, *Crit. Rev. Biochem. Mol. Biol.*, **45** (2010), 106–124. https://doi.org/10.3109/10409231003627991
15. E. Hausmann, Cofactor requirements for the enzymatic hydroxylation of lysine in a polypeptide precursor of collagen, *Biochim. Biophys. Acta, Protein Struct.*, **133** (1967), 591–593. https://doi.org/10.1016/0005-2795(67)90566-1

16. M. J. Landrum, J. M. Lee, M. Benson, G. R. Brown, C. Chao, S. Chitipiralla, et al., ClinVar: improving access to variant interpretations and supporting evidence, *Nucleic Acids Res*., **46** (2018), D1062–D1067. https://doi.org/10.1093/nar/gkx1153
17. S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, et al., Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, *Genet. Med.*, **17** (2015), 405–424. https://doi.org/10.1038/gim.2015.30
18. S. Kundu, Mathematical model of a short translatable G-quadruplex and an assessment of its relevance to misfolding-induced proteostasis, *Math. Biosci. Eng.*, **17** (2020), 2470–2493. https://doi.org/10.3934/mbe.2020135
19. M. M. Tirion, Large amplitude elastic motions in proteins from a single-parameter, atomic analysis, *Phys. Rev. Lett.*, **77** (1996), 1905. https://doi.org/10.1103/PhysRevLett.77.1905
20. A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, *Biophys. J.*, **80** (2001), 505–515. https://doi.org/10.1016/S0006-3495(01)76033-X
21. P. Doruker, A. R. Atilgan, I. Bahar, Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor, *Proteins*, **40** (2000), 512–524. https://doi.org/10.1002/1097-0134(20000815)40:3<512::AID-PROT180>3.0.CO;2-M
22. I. Bahar, A. R. Atilgan, B. Erman, Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, *Fold. Des.*, **2** (1997), 173–181. https://doi.org/10.1016/S1359-0278(97)00024-2
23. K. Hinsen, Analysis of domain motions by approximate normal mode calculations, *Proteins*, **33** (1998), 417–429. https://doi.org/10.1002/(SICI)1097-0134(19981115)33:3<417::AID-PROT10>3.0.CO;2-8
24. S. Kundu, Insights into the mechanism(s) of digestion of crystalline cellulose by plant class C GH9 endoglucanases, *J. Mol. Model.*, **25** (2019), 240. https://doi.org/10.1007/s00894-019-4133-1
25. L. Yang, G. Song, R. L. Jernigan, Protein elastic network models and the ranges of cooperativity, *PNAS*, **106** (2009), 12347–12352. https://doi.org/10.1073/pnas.0902159106
26. X. Du, Y. Li, Y. L. Xia, S. Ai, J. Liang, P. Sang, et al., Insights into protein-ligand interactions: mechanisms, models, and methods, *Int. J. Mol. Sci.*, **17** (2016), 144. https://doi.org/10.3390/ijms17020144
27. P. L. de Micheaux, B. Liquet, Understanding convergence concepts: A visual-minded and graphical simulation-based approach, *Am. Stat.*, **63** (2009), 173–178. https://doi.org/10.1198/tas.2009.0032
28. S. Chaturvedi, A. K. Singh, A. K. Keshari, S. Maity, S. Sarkar, S. Saha, Human metabolic enzymes deficiency: A genetic mutation based approach, *Scientifica (Cairo)*, **2016** (2016), 9828672. https://doi.org/10.1155/2016/9828672
29. S. Kundu, Fe(2)OG: An integrated HMM profile-based web server to predict and analyze putative non-haem iron(II)- and 2-oxoglutarate-dependent dioxygenase function in protein sequences, *BMC Res. Notes*, **14** (2021), 80. https://doi.org/10.1186/s13104-021-05477-z
30. R. J. Wanders, J. C. Komen, Peroxisomes, Refsum's disease and the alpha- and omega-oxidation of phytanic acid, *Biochem. Soc. Trans.*, **35** (2007), 865–869. https://doi.org/10.1042/BST0350865

31. M. A. McDonough, K. L. Kavanagh, D. Butler, T. Searls, U. Oppermann, C. J. Schofield, Structure of human phytanoyl-CoA 2-hydroxylase identifies molecular mechanisms of Refsum disease, *J. Biol. Chem.,* **280** (2005), 41101–41110. https://doi.org/10.1074/jbc.M507528200
32. T. G. Smith, P. A. Robbins, P. J. Ratcliffe, The human side of hypoxia-inducible factor, *Br. J. Haematol.*, **141** (2008), 325–334. https://doi.org/10.1111/j.1365-2141.2008.07029.x
33. G. L. Wang, G. L. Semenza, Purification and characterization of hypoxia-inducible factor 1, *J. Biol. Chem.*, **270** (1995), 1230–1237. https://doi.org/10.1074/jbc.270.3.1230
34. S. E. Wilkins, M. I. Abboud, R. L. Hancock, C. J. Schofield, Targeting protein-protein interactions in the HIF system, *ChemMedChem*, **11** (2016), 773–786. https://doi.org/10.1002/cmdc.201600012
35. M. A. McDonough, V. Li, E. Flashman, R. Chowdhury, C. Mohr, B. M. R. Liénard, et al., Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2), *PNAS*, **103** (2006), 9814–9819. https://doi.org/10.1073/pnas.0601283103
36. P. H. Maxwell, M. S. Wiesener, G. W. Chang, S. C. Clifford, E. C. Vaux, M. E. Cockman, et al., The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis, *Nature*, **399** (1999), 271–275. https://doi.org/10.1038/20459
37. G. L. Semenza, Hydroxylation of HIF-1: oxygen sensing at the molecular level, *Physiology (Bethesda)*, **19** (2004), 176–182. https://doi.org/10.1152/physiol.00001.2004
38. D. R. Peaper, P. Cresswell, Regulation of MHC class I assembly and peptide binding, *Annu. Rev. Cell Dev. Biol.*, **24** (2008), 343–368. https://doi.org/10.1146/annurev.cellbio.24.110707.175347
39. E. W. Hewitt, The MHC class I antigen presentation pathway: strategies for viral immune evasion, *Immunology*, **110** (2003), 163–169. https://doi.org/10.1046/j.1365-2567.2003.01738.x
40. E. Rufer, R. M. Leonhardt, M. R. Knittler, Molecular architecture of the TAP-associated MHC class I peptide-loading complex, *J. Immunol.*, **179** (2007), 5717–5727. https://doi.org/10.4049/jimmunol.179.9.5717
41. A. Blees, D. Januliene, T. Hofmann, N. Koller, C. Schmidt, S. Trowitzsch, et al., Structure of the human MHC-I peptide-loading complex, *Nature*, **551** (2017), 525–528. https://doi.org/10.1038/nature24627
42. J. W. Yewdell, J. R. Bennink, Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses, *Annu. Rev. Immunol.*, **17** (1999), 51–88. https://doi.org/10.1146/annurev.immunol.17.1.51
43. P. V. Praveen, R. Yaneva, H. Kalbacher, S. Springer, Tapasin edits peptides on MHC class I molecules by accelerating peptide exchange, *Eur. J. Immunol.*, **40** (2010), 214–224. https://doi.org/10.1002/eji.200939342
44. S. Kundu, Mathematical modeling and stochastic simulations suggest that low-affinity peptides can bisect MHC1-mediated export of high-affinity peptides into “early”- and “late”-phases, *Heliyon*, **7** (2021), e07466. https://doi.org/10.1016/j.heliyon.2021.e07466



AIMS Press