Computational Ways to Enhance Protein Inhibitor Design

Robert L. Jernigan,[1] Kannan Sankar,[1] Kejue Jia,[1] Eshel Faraggi,[2,3] and Andrzej Kloczkowski[4,5]

[1] Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA

[2] Research and Information Systems, LLC, Indianapolis, Indiana, 46240, USA

[3] Department of Physics, Indiana University Purdue University Indianapolis, Indianapolis, Indiana 46202, USA

[4] Battelle Center for Mathematical Medicine, Nationwide Children's Hospital, Columbus Ohio, 43205 USA.

[5] Department of Pediatrics, The Ohio State University, Columbus, Ohio, USA.

Abstract: Two new computational approaches are described to aid the design of new peptide-based drugs by evaluating ensembles of protein structures from their dynamics, and the assessment of the structures using empirical contact potential. This builds on the concept that conformational variability can aid the binding process and, for disordered proteins, can even facilitate the binding of more diverse ligands. This latter consideration should mean that such a design process should be less restrictive so that multiple designed inhibitors might be effective. The example chosen here focuses on proteins/peptides binding to hemagglutinin (HA) to block the large-scale conformational change occurring during its activation. Variability in the conformations are considered from sets of experimental structures, or as an alternative their simple computed dynamics, and the set of designed peptides/small proteins from Baker (1) designed to bind to hemagglutinin, is the large set considered are assessed with the new empirical contact potentials.

Introduction

Influenza infection is a widespread cause of major medical concern because of the rapid viral evolution, that causes both occasional pandemics and more frequent problems almost every year. It has been estimated that the annual outbreaks by influenza A and B viruses over the past 100 years have had an even greater impact than all other past pandemics combined (2-4). The extremely high mutation rate of the virus means that any given vaccine soon becomes outdated. Thus, vaccination offers limited protection, especially when facing highly virulent nature and rapid evolution of influenza (5). Although some effective anti-influenza drugs have been developed, drug resistance usually appears rapidly.

Hemagglutinin (HA) is a major surface glycoprotein of this virus that is involved in four of the most important aspects of influenza infection: (a) it is the target of antibodies that neutralize infectivity, (b) it undergoes antigenic drift to escape neutralization, (c) it binds to cell-surface receptors to initiate infection, and (d) it mediates the fusion of viral and host membranes essential for viral entry. The large-scale conformational changes in HA are critical for the steps in which the virus inserts itself into the host cells by fusing to the host membrane and the residues involved in this process are highly conserved across different types and subtypes during antigenic drift. These residues can serve as important targets for developing broad-reacting antiviral inhibitors (6-9). Based on a set of crystal structures of the HA-antibody complex showing the conformational changes to HA during the essential activation steps, David Baker and his colleagues designed a novel HA inhibitor for Group 1 of type A virus (1).

Influenza HA is a homo-trimeric protein where each monomer contains two disulfide-bonded polypeptides, HA1 and HA2. HA1 is responsible for attaching to host cell-surface receptors while HA2 mediates the fusion of the influenza envelope with the endosomal membrane, thus allowing the entry of influenza RNA into the host cell. The pre- (2) and post-fusion structures (3) of HA1 are essentially the same, while those of HA2 (2;4;5) are drastically different, see **Fig. 1**.

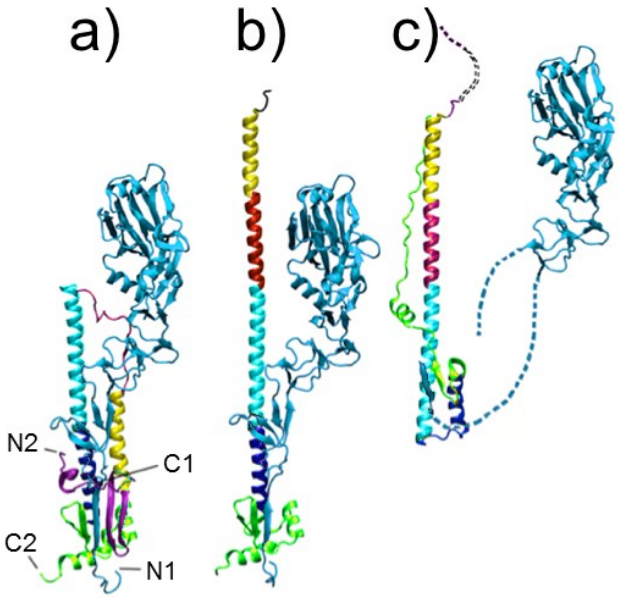


Figure 1. **The HA in the pre- (a), intermediate (b) and post-fusion (c) states**. The termini of HA1 (light blue) and HA2 are labeled as N1, C1, N2, and C2, respectively in (a). The dotted lines in (c) indicate unresolved regions. The structures have all been aligned on the cyan helix, which is the only region in that domain that does not change during the transition.

The structural change in HA2 includes a partial unfolding of the long α-helix into a loop (dark blue) and the folding of an inter-helix loop (in red) into a part of the long α-helix, thus delivering both N- (blue) and C-terminal (pink) fragments to the same end of the molecule upon the fusion of viral and endosomal membranes.

The protein gp41 of HIV-1 is the membrane fusion protein, is similar to HA2 of HA (6). In that case, peptides derived from the C-terminal region of gp41 corresponding to the outer-layer helices, referred to as C-peptides, were found to inhibit HIV-1 infection with IC_{50} in the nanomolar range (7-9). C-peptides are believed to act by binding to the exposed surface of the N-terminal central three-helical bundle in a transient pre-fusion gp41 intermediate, thereby blocking membrane fusion. One such L-peptide, T-20/ Enfuvirtide with 36-residues, was approved few years ago as a drug by the Food and Drug Administration (10); it shows high efficacy in suppressing resistant HIV-1 strains. Moreover, efforts to target a prominent pocket on the surface of the central three-helical bundle have led to the discovery of small, cyclic D-peptides that inhibit HIV-1 infection, thereby validating the pocket as a potential target for small-molecule HIV-1 fusion inhibitors (11).

To evade host antibody recognition, the HA protein on the surface of influenza virus, primarily on the globular domain, must constantly mutate. This interferes in important ways with any vaccine, and reduces the vaccine's efficiency and useful lifetime. However, no matter how much the influenza virus mutates, it must maintain the ability to induce membrane fusion to ensure its propagation. Thus, the stem domain that is primarily responsible for inducing membrane fusion is the most conserved. Ian Wilson's group identified antibodies that broadly neutralize influenza A virus Group 1 (12) (**Fig. 2a**), Group 2 (13), Group 1 and 2 (14)

and influenza type A and B viruses (15), see **Fig. 2b**. All these antibodies recognize epitopes located in the stem domain. David Baker's group designed small proteins against influenza A virus Group 1 (1) (**Fig. 2a**). In addition, they identified a conserved patch on the surface of the central helical bundle in the low-pH postfusion state (**Fig. 2c**). These three interfaces may all serve as useful targets for developing inhibitors against influenza virus.

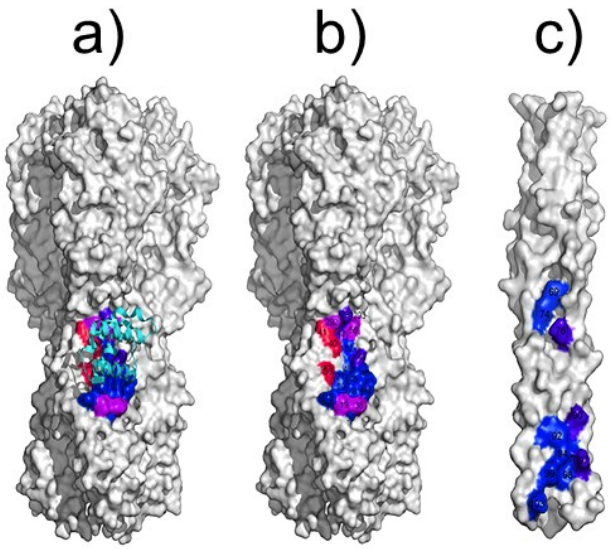


Figure 2. Interfaces on HA that could likely be targeted in inhibitor design. a) The interface conserved among Group 1 influenza A virus (in colors). b) The interface conserved among all influenza A and B viruses. c) The conserved patch on the surface of the central helical bundle at low-pH on HA.

Molecular recognition in general and protein-protein interactions in particular are essential in almost every aspect of biological function. Moreover, proteins that bind other proteins with high affinity and high specificity have numerous applications for diagnostics and therapeutics. Currently, antibodies are by far the most commonly used proteins for both detection and therapeutic intervention. However, antibodies are large proteins that are expensive to produce and difficult to deliver. Thus, it would be important progress for biomedicine to be able to design novel protein-binding modules at will.

The set of 88 proteins that were designed and tested by Baker and colleagues provides an excellent test set for use in the present study. Below we consider the dynamics of the structure in two different ways, from a set of experimental structures and from computed dynamics. Then we apply new knowledge-based free energies to rank the different designs, specifically predicting which designs are likely to bind. – Baker and colleagues were not able to do this without experimental testing. These are empirical free energy contact potentials developed by Jernigan, Kloczkowski and Faraggi that have proven to be highly successfully in blind-tests at past CASP experiments. In the present paper, we aim to make some suggestions for new ways to sample conformations of a target protein and how to assess the designed structures.

Multiple Experimental Structures Capture the Important Functional Motions within a Hemagglutinin Structure Set

The 43 structures of hemagglutinin listed in Table 1 were collected from the PDB with a BLAST search, retaining only those structures present as trimeric complexes of the HA1 and HA2 subunits. The individual

subunits were extracted separately and aligned. This yields a total of 129 structures of the HA1 + HA2 monomers that were superimposed onto the central structure (PDB: 1mqm) using the Combinatorial Extension (CE) algorithm, and these have a continuous distribution of RMSDs from 0 to 3.3 Å.

Table 1. The PDB identifiers of the 43 structures of hemagglutinin used here for extracting dynamics

1HGD	2HMG	3FKU	4BGZ	4KPQ
1MQL	2IBX	3HMG	4BH1	4KPS
1MQM	2WR7	3LZG	4DJ6	5HMG
1MQN	2WRB	3M5G	4EDB	
1RD8	2WRD	3M6S	4F23	
1RUY	2WRE	3S11	4F3Z	
1RUZ	2WRF	3SM5	4FIU	
1RV0	2WRG	3UBE	4GXX	
1RVX	2WRH	3VUN	4JTX	
2FK0	3EYM	3ZTJ	4KDM	

After these have been superimposed, then the covariances in for all pairs of positions were computed. Then, Principal Component Analysis is performed on this dataset. The input is the set of all of the structures in the set (16;17). From these data, the average position of each point in the reference structure is computed as $\langle x_i \rangle$ and the covariances for each pair of points i and j was computed according to $c_{ij} = \left\langle \left(x_i - \langle x_i \rangle \right) \left(x_j - \langle x_j \rangle \right) \right\rangle$, where brackets < > indicate averages over the set of structures. The covariance matrix *C* can be decomposed as $C = P \Delta P^T$, where the eigenvectors *P* represent the principal components (PCs) and the eigenvalues are the elements of the diagonal matrix Δ. The eigenvalues are sorted in order. Each eigenvalue is directly proportional to the amount of the total variance it captures. The results of this analysis are shown in **Fig. 3** for the set of coarse-grained hemagglutinins, which shows how truly limited the characteristic motions are within the structure set. Clearly it does not require many of these characteristic motions to capture nearly all of the overall motions.

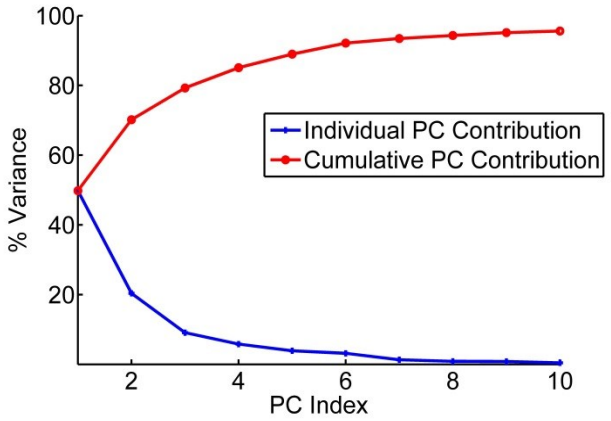
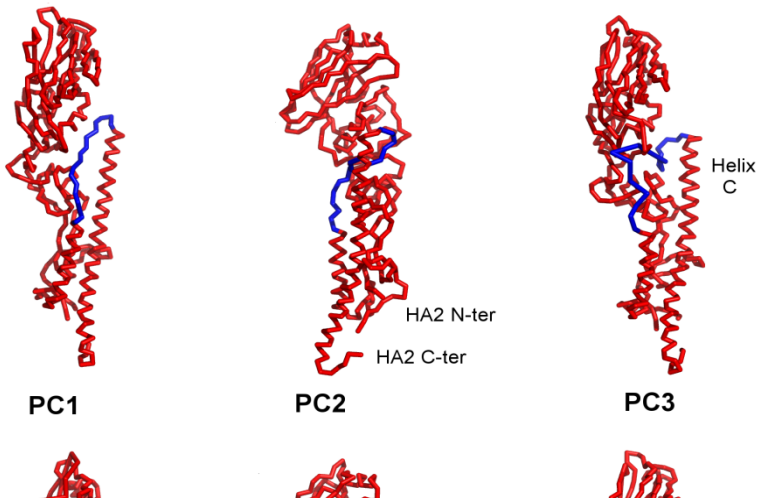
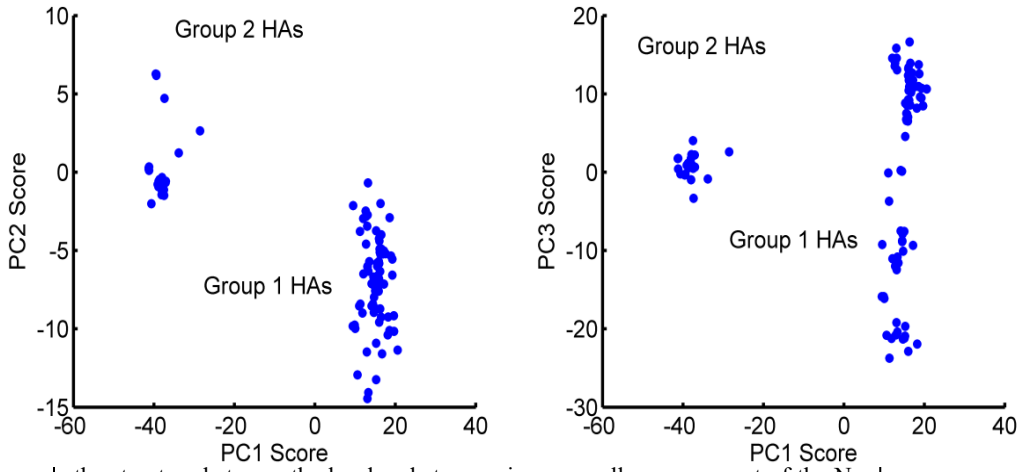


Figure 3. Principal Component Contributions to the Total Motions of Hemagglutinin. Percent of variance explained by each individual PC is shown in blue and the cumulative contribution of each PC to the total variance/motion in red. **The first 5 PCs account for 90% of the total motions present in the set of 43 structures.**

Characterization of the Global Motions in Hemagglutinins

Based on their sequences, HAs have been subdivided into two main groups: Group 1 (H1, H2, H5, H6, H8, H9, H11, H12, H13 and H16) and Group 2 (H3, H4, H7, H10, H14 and H15) (18). Interestingly the first three PCs separately cluster into these two major groups, with minor exceptions. The distribution of the experimental structures over the PCs are shown in **Fig. 4 for pairs of PCs**. This distinctive clustering can be seen clearly.





the structure between the head and stem regions as well as movement of the N-

Figure 4. Distribution of the 129 HA monomeric structures projected onto the first 3 PCs. (A) PC2-PC1 space and (B) PC3-PC1 space. PC1 separates the Group 1 and Group 2 hemagglutinins into different clusters. Outliers H13 and H16 have been eliminated from the figure for the sake of clarity. PC1 has a major gap between the two groups of clustered structures. PC3 appears to be populated in two somewhat similar clusters, with Group 1 showing a particularly wide range of PC3 values.

Different conformations can bind to different partners and thus include dynamics in the process will improve the probability of success in computational protein design. When the PCs are visualized on the structures it can be seen that the first three PCs represent primarily motions in the B-loop (blue) that is involved in the large-scale transition. PC1, PC2 and PC3 can be interpreted as primarily involving conformations changes in the C-terminus, the central and N-terminus parts of the B-loop (see **Fig. 5**). Interestingly the B-loop is a region with a strong tendency to form a coiled-coil and is implicated in the formation of the pre-hairpin intermediate in the 'spring loaded mechanism' of HA action (19;20). The PC3 motion also clearly demonstrates the shift in the loop necessary for it to position itself at the top of helix C. In addition, PC2 captures a hinge motion in the head of HA with respect to the stem as well as well as a motion at the N-terminus of HA2 (fusion peptide) that is subsequently exposed for insertion into the membrane during fusion. These computed structures show a high level of variability of conformations particularly for the B-loop, which relate well to the known conformational transition, even though the full extent of motions is not shown in Fig. 5. As shown in Table 2, these PCs provide a useful representation of changes present in the ensemble of structures.

Anisotropic Network Models (ANM) Can Substitute, If Insufficient Numbers of Experimental Structures Are Available

Elastic Network Models of proteins, such as the Gaussian Network Model (GNM) and Anisotropic Network Models (ANM) of proteins as developed by Tirion (21), Bahar, Erman and Jernigan (22-30) computationally yield information about protein fluctuation dynamics, the directions of motions of the residues and atoms around their equilibrium positions. This information has already been used by Bahar, Jernigan, Kloczkowski and many others with significant success (23;31-33) to explain functional motions and mechanisms in proteins, nucleic acids and large biological assemblies, such as the ribosome. ANM could be used as an alternative to calculate the normal modes from a single structure when insufficient numbers of experimental structures or structures having sufficient variability are not available to perform PC analysis, then normal modes from the elastic network models could also be used to compute entropies (34). (But, as we show below, contact entropies are simpler and provide significant gains.) In ANM, the potential energy *V* is a function of the displacement vector *D* of each point in the structure $V = \frac{\gamma}{2} DHD^T$, where γ is the spring constant for all closely interacting points in a structure (here we used a cutoff distance of 13 Å between alpha-carbons for coarse-grained models retaining only C^{α} atoms) to establish the spring connections between residues), and *H* is the Hessian matrix containing the second derivatives of the energy, with respect to each of the coordinates x, y, z. For a structure with *n* residues, the Hessian matrix *H* contains $n \times n$ super-elements each of size 3×3. The Hessian matrix *H* can be decomposed (22) as $H = M\Lambda M^T$, where Λ is a diagonal

matrix comprising the eigenvalues with the eigenvectors forming the columns of the matrix *M*. This decomposition generates 3*n*-6 normal modes (the first 6 modes account for the rigid body translations and rotations of the system) reflecting the vibrational fluctuations, so singular value decomposition is utilized.

Comparing Directions of Motions Using Overlaps

The alignment between the directions of a given experimental PC and a given computed normal mode can be measured by comparing the directions of motion in their overlap, as defined by Tama and Sanejouand (35) : $O_{ij} = \frac{\left|P_i \cdot M_j\right|}{\left\|P_i\right\|\left\|M_j\right\|}$, where P_i is the *i*th PC for and M_j is the *j*th normal mode. A perfect match yields an overlap value of 1, meaning these motions are in the same direction. We also define the cumulative overlap (CO) between the first *k* vectors M_j and P_i as $CO(k) = \left(\sum_{j=1}^{k} O_{ij}^2\right)^{\frac{1}{2}}$.

The high overlaps between the two methods ensures the reliability of the computed dynamics. The 1st, 2nd and 3rd PCs have good overlaps of 0.57, 0.43 and 0.34 with the 3rd, 2nd and 1st individual modes, respectively. We compare the first three PC's from the X-ray set with the first 20 normal modes from the elastic network models, and these are relatively high between all three PCs of the X-ray hemagglutinin and the set of normal modes for the computed normal modes (see **Table 2**).

Table 2. Cumulative overlaps between computed ANM modes and PCs from the set of experimental hemagglutinin structures

	CO		
	3 Modes	**6 Modes**	**20 Modes**
PC1	0.60	**0.66**	**0.71**
PC2	**0.50**	**0.57**	**0.65**
PC3	0.40	0.44	**0.60**

Strategies for Generating and Ranking an Ensemble of Structures and Identifying a Structure Module Targeted for Inhibitor Design

Identifying the most conformationally variable part of the structure is the aim here. These are the parts of a structure that should be the most useful to use for inhibitor design. These parts can be identified simply by computing the changes in all internal distances over the

Figure 6. Examples of the diversity of conformations from the first 3 PCs for the B loop (blue) of hemagglutinin. The PCs can be used to generate an ensemble of conformations. Each of three shows a conformation generated from one PC. The motions showed that this loop is the most flexible part of the structure and possesses an extremely diverse set of conformations.

ensemble. Examples of such potential binding parts to target have been extracted from the ensemble of sampled conformations for HA generated by utilizing combinations of the first several PCs (**Fig. 6)**. This highly variable segment should be susceptible to binding by a broader range of ligands.

Assessing Peptide/Protein Designs with New Empirical Contact Potentials

Here we present new strategies for the assessment of bound ligand structures by taking as our target the designed small proteins from David Baker and his colleagues that were targeted to bind hemagglutinin (1;36). This provides an interesting relatively large dataset, which we can use to test our assessment method. The Baker designs, originating from small, monomeric proteins in the PDB having between 80 and 250 residues, were targeted against a hydrophobic region on the 'stem' of hemagglutinin. Of the 88 designs that they tested, only *two* were reported to have detectable binding affinity for hemagglutinin (these affinity were subsequently improved in rounds of randomization and selection).

Four-body Coarse-Grained Contact Potentials (37-38). Four-body potentials were developed by Kloczkowski and Jernigan to account for the cooperative interactions in proteins; they take into account the coarse-grained contact interactions together with the extent of solvent exposure, and thus provide a more detailed and more cooperative representation of protein interaction energies than do pairwise potentials. Capturing this cooperativity is considered to be critical for evaluating densely packed protein structures. These potentials are highly empirical and are based simply on the observed frequency of occurrences of different types of amino acids in closely interacting quartets of amino acid types within a large set of protein structures. We have found that these four-body contact potentials can discriminate well between native structures and partially unfolded or deliberately misfolded structures. These have also included short-range backbone energies (39). We tested these optimized potentials at CASP9 as the prediction group 4_BODY_POTENTIALS from Iowa State University. There were 110 other human prediction groups participating in CASP9 competition, and 140 prediction servers. According to Nick Grishin, the assessor of free modeling techniques at CASP9, 4_BODY_POTENTIALS was one of most successful groups in free modeling at that time, ranking third, according to the averaged z_{score} both for best models, and top models. Free modeling is the most difficult and most challenge in protein structure prediction, when the sequence of the protein has only a low sequence similarity in comparison to any known protein structures. This success at CASP9 demonstrates clearly that the cooperative multibody interactions are an appropriate tool for assessing predicted structures, and we apply them here to Baker's hemagglutinin inhibitor structures. Later we have added in electrostatic interactions and these were tested at the subsequent CASP10.

Including Entropies in the Inhibitor Assessments

The Elastic Network Models (ENM) have proven themselves to be highly useful in representing the global motions for a wide variety of diverse protein structures (22;24-27;31-33;39-73). Since they have proven to be so successful in capturing the global most import important motions of protein structures, it is reasonable to expect that they should be able also to estimate the conformational entropies of structures. We employ the Elastic Network Model to compute the motions of protein structures and then these motions are then used directly to approximate the entropy of a conformation (34;74) . We previously (74) used vibrational entropies based on the frequencies of the normal modes, but more recently have found significant gains by utilizing the mean square fluctuations computed from the ENM as a direct measure of entropy: $\Delta S \propto \Gamma^{-1} = \sum_{i=2}^{N} \frac{1}{\lambda_i}\left(Q_i Q_i^T\right)$ where Q is a normal mode vector, λ the corresponding square frequency, Γ the system's Hessian, and Γ^{-1} its pseudo-inverse. We obtain the Free Energy changes from $\Delta G = \Delta E - T\Delta S$ by simply combining the four-body potential with the ENM-based entropy (34). The excellent blind-tested performance of our method in CASP experiments shows that our methodology is an outstanding tool for assessing protein designs, such as the ones from Baker's hemagglutinin inhibitor designs.

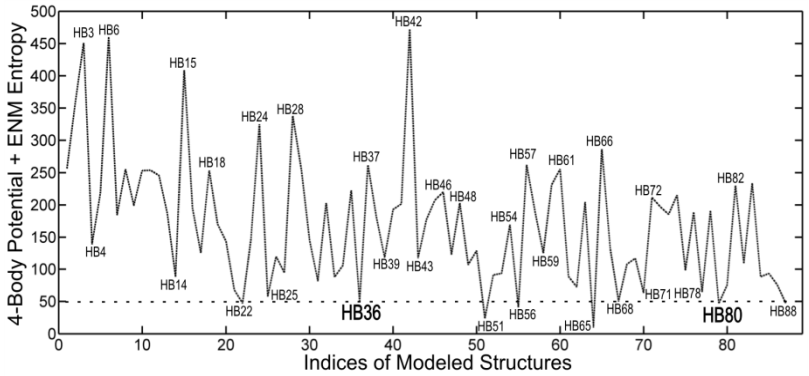


Figure 7. Ranking by Coarse-Grained Free Energies of inhibitor proteins designed against hemagglutinin. Free energies are given on the ordinate axis (arbitrary scale), and the different structures (from the pdb) are indicated along the abscissa. The 8 top-ranked structures with favorable free energies can be seen to be HB22, HB36, HB51, HB56, HB65, HB68, HB80, and HB88**.** This demonstrates the utility of the coarse-grained free energies to computationally screen for favorable structures. The two structures HB36 and HB80 were experimentally shown to be functional.

These New Free Energies Successfully Select Native-like Poses in Protein-Protein Docking

We have applied this method to the set of 89 inhibitor proteins designed against hemagglutinin by David Baker's group, and we find that it provides a useful screen for that set of structures. Structures having the lowest energies indicate stable favorable conformations. However, stable structures are not always functional. In this case, we tested a set of eight structures at local minimums of the energy landscape ranked by their energies. From these, it was reported that two of them were found to be functional. (see **Fig. 7).**

Discussion

Here we have outlined a simple new way to use protein dynamics for peptide/protein design studies. This approach serves to identify those specific regions in the structure having particularly wide-ranging conformational variability, which could be of particular importance for targeting computational design efforts. Specifically, the highly variable segments should be able to bind to a particularly wide range of diverse ligands. Such variable conformations are well known to be important for the promiscuous binding exhibited by disordered proteins and using this approach should have some advantage. Using such more localized protein targets might be an important new approach for targeted computational design. Another advantage of this is that more exhaustive computations can be carried out for smaller targets.

Application of the potentials described above to assess structural designs would allow ranking of sets of designed inhibitor proteins. The differences in rankings should allow to conclude the extent to which the large-scale backbone fluctuations identified in the dynamics could be utilized in the design process. This would require a significantly larger effort than has been presented here. Of course, the potentials themselves are empirical and could be modified to reflect the data from the experimental studies on the designed molecules for the specific class of targeted protein, which is one of a major advantages of the adaptability of the empirical potentials in any particular application.

Our approach can be extended by detailed analysis of allosteric sites that are important for drug design. Most drugs are designed to bind directly to the primary active sites named orthosteric sites, to inhibit or modify the function of the protein. Binding of a drug to the active site prevents binding to a virus or other disease-related agent and most drugs are designed to fit into the primary active sites. However, adverse side effects of a drug may occur because many enzymes or receptors with related functions may have similarities in their active sites.

A new approach to drug design is based on secondary binding site effects. In this approach, small molecule drugs are designed to bind at secondary binding sites called allosteric sites (75-81). A potential drug - an allosteric modulator binds to an allosteric site and remotely modifies the conformation of the primary binding site of the protein. Allosteric sites are controlled by intrinsic protein dynamics, and the approach proposed here could also be applied to these allosteric sites.

Acknowledgements: This research was supported by NSF grant DBI-1661391, and NIH grant R01GM127701.

References

1. Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch EM, Wilson IA, Baker D. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science 2011;332:816-21.
2. Wilson IA, Skehel JJ, Wiley DC. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature 1981;289:366-73.
3. Bizebard T, Gigant B, Rigolet P, Rasmussen B, Diat O, Bosecke P, Wharton SA, Skehel JJ, Knossow M. Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. Nature 1995;376:92-4.
4. Bullough PA, Hughson FM, Skehel JJ, Wiley DC. Structure of influenza haemagglutinin at the pH of membrane fusion. Nature 1994;371:37-43.
5. Chen J, Skehel JJ, Wiley DC. N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA(2) subunit to form an N cap that terminates the triple-stranded coiled coil. Proc.Natl.Acad.Sci U.S.A 1999;96:8967-72.
6. Skehel JJ, Wiley DC. Coiled coils in both intracellular vesicle and viral membrane fusion. Cell 1998;95:871-4.
7. Wild CT, Shugars DC, Greenwell TK, McDanal CB, Matthews TJ. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. Proc.Natl.Acad.Sci U.S.A 1994;91:9770-4.
8. Jiang S, Lin K, Strick N, Neurath AR. HIV-1 inhibition by a peptide. Nature 1993;365:113.
9. Chan DC, Chutkowski CT, Kim PS. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. Proc.Natl.Acad.Sci U.S.A 1998;95:15613-7.
10. FDA notifications. FDA approves Fuzeon, the first fusion inhibitor. AIDS Alert 2003;18(78).
11. Eckert DM, Malashkevich VN, Hong LH, Carr PA, Kim PS. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell 1999;99:103-15.
12. Ekiert DC, Bhabha G, Elsliger MA, Friesen RH, Jongeneelen M, Throsby M, Goudsmit J, Wilson IA. Antibody recognition of a highly conserved influenza virus epitope. Science 2009;324:246-51.
13. Murphy BR, Webster RG. Orthomyxoviruses. In: Knipe DM, Howley PM, Williams L, editors. Fields Virology. Boston: Philadelphia; 2001.
14. Ekiert DC, Kashyap AK, Steel J, Rubrum A, Bhabha G, Khayat R, Lee JH, Dillon MA, O'Neil RE, Faynboym AM, et al. Cross-neutralization of influenza A viruses mediated by a single antibody loop. Nature 2012;489:526-32.
15. Dreyfus C, Laursen NS, Kwaks T, Zuijdgeest D, Khayat R, Ekiert DC, Lee JH, Metlagel Z, Bujny MV, Jongeneelen M, et al. Highly conserved protective epitopes on influenza B viruses. Science 2012;337:1343-8.
16. Teodoro ML, Philips Jr. GN, Kavraki LE. Understanding protein flexibility through dimensionality reduction. J.Comput.Biol. 2003;10:617-34.
17. Teodoro ML, Philips Jr. GN, Kavraki LE. A dimensionality reduction approach to modeling protein flexibility. Int.Conf.Comput.Mole.Biol.(RECOMB) 2002;299-308.
18. Air GM. Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. Proc.Natl.Acad.Sci U.S.A 1981;78:7639-43.
19. Xu R, Wilson IA. Structural characterization of an early fusion intermediate of influenza virus hemagglutinin. J Virol. 2011;85:5172-82.
20. Carr CM, Kim PS. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. Cell 1993 May 21;73(4):823-32.

21. Tirion MM. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. Phys.Rev.Lett. 1996;77:1905-8.
22. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys.J. 2001;80:505-15.
23. Bahar I, Jernigan RL. Cooperative structural transitions induced by non-homogeneous intramolecular interactions in compact globular proteins. Biophys J. 1994;66:467-81.
24. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Fold.Des 1997;):173-81.
25. Sen TZ, Feng Y, Garcia JV, Kloczkowski A, Jernigan RL. The Extent of Cooperativity of Protein Motions Observed with Elastic Network Models Is Similar for Atomic and Coarser-Grained Models. J Chem Theory.Comput. 2006;2:696-704.
26. Demirel MC, Atilgan AR, Jernigan RL, Erman B, Bahar I. Identificatioin of kinetically hot residues in proteins. Protein Sci. 1998;7:871-84.
27. Bahar I, Jernigan RL. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase-bound forms. J.Mol.Biol. 1998;281:2522-32.
28. Doruker P, Jernigan RL, Bahar I. Dynamics of large proteins through hierarchical levels of coarse-grained structures. J.Comput.Chem. 2002;23:119-27.
29. Doruker P, Jernigan RL. Functional motions can be extracted from on-lattice construction of protein structures. Proteins 2003 Nov 1;53(2):174-81.
30. Doruker P, Jernigan RL, Navizet I, Hernandez R. Important fluctuation dynamics of large protein structures are preserved upon renormalization. Int J Quantum Chem 2002;90:822-37.
31. Isin B, Tirupula KC, Oltvai ZN, Klein-Seetharaman J, Bahar I. Identification of motions in membrane proteins by elastic network models and their experimental validation. Methods Mol Biol. 2012;914:285-317.
32. Keskin O, Bahar I, Flatow D, Covell DG, Jernigan RL. Molecular mechanisms of chaperonin GroEL-GroES function. Biochemistry 2002;41:491-501.
33. Keskin O, Durell SR, Bahar I, Jernigan RL, Covell DG. Relating molecular flexibility to function: a case study of tubulin. Biophys.J. 2002;83:663-80.
34. Zimmermann MT, Leelananda SP, Kloczkowski A, Jernigan RL. Combining statistical potentials with dynamics-based entropies improves selection from protein decoys and docking poses. J Phys.Chem B 2012;116:6725-31.
35. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. Protein Eng 2001;14:1-6.
36. Fleishman SJ, Baker D. Role of the biomolecular energy gap in protein design, structure, and evolution. Cell 2012;149:262-73.
37. Feng Y, Kloczkowski A, Jernigan RL. Four-body contact potentials derived from two protein datasets to discriminate native structures from decoys. Proteins 2007;68:57-66.
38. Feng Y, Kloczkowski A, Jernigan RL. Potentials 'R' Us web-server for protein energy estimations with coarse-grained knowledge-based potentials. BMC.Bioinformatics 2010;11:92.
39. Bahar I, Kaplan M, Jernigan RL. Short-range conformational energies, secondary structure propensities, and recognition of correct sequence-structure matches. Proteins 1997;29:292-308.
40. Kim MH, Seo S, Jeong JI, Kim BJ, Liu WK, Lim BS, Choi JB, Kim MK. A mass weighted chemical elastic network model elucidates closed form domain motions in proteins. Protein Sci. 2013;22:605-13.
41. Karaca E, Bonvin AM. A multidomain flexible docking approach to deal with large conformational changes in the modeling of biomolecular complexes. Structure 2011;19:555-65.

42. Zheng W. Accurate flexible fitting of high-resolution protein structures into cryo-electron microscopy maps using coarse-grained pseudo-energy minimization. Biophys J 2011;100:478-88.
43. Wieninger SA, Serpersu EH, Ullmann GM. ATP binding enables broad antibiotic selectivity of aminoglycoside phosphotransferase(3')-IIIa: an elastic network analysis. J Mol Biol. 2011;409:450-65.
44. Yang L, Song G, Carriquiry A, Jernigan RL. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. Structure. 2008;16:321-30.
45. Bahar I, Rader AJ. Coarse-grained normal mode analysis in structural biology. Curr.Opin.Struct.Biol. 2005;15:586-92.
46. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. J.Mol.Biol. 1999;285:1023-37.
47. Duttmann M, Mittnenzweig M, Togashi Y, Yanagida T, Mikhailov AS. Complex intramolecular mechanics of G-actin--an elastic network study. PLoS ONE 2012;7:e45859.
48. Bahar I, Jernigan RL. Cooperative fluctuations and subunit communication in tryptophan synthase. Biochemistry 1999;38:3478-90.
49. Bahar I, Wallqvist A, Covell DG, Jernigan RL. Correlation between native-state hydrogen exchange and cooperative residue fluctuations from a simple model. Biochemistry 1998;37:1067-75.
50. Zheng W, Auerbach A. Decrypting the sequence of structural events during the gating transition of pentameric ligand-gated ion channels based on an interpolated elastic network model. PLoS Comput.Biol. 2011;7:e1001046.
51. May ER, Brooks CL, III. Determination of viral capsid elastic properties from equilibrium thermal fluctuations. Phys.Rev.Lett. 2011;106:188101.
52. Martin DR, Ozkan SB, Matyushov DV. Dissipative electro-elastic network model of protein electrostatics. Phys.Biol. 2012;9:036004.
53. Bahar I, Erman B, Haliloglu T, Jernigan RL. Efficient characterization of collective motions and interresidue correlations in proteins by low-resolution simulations. Biochemistry 1997;36:13512-23.
54. Jernigan RL, Yang L, Song G, Doruker P. Elastic Network Models of Coarse-Grained Proteins Are Effective for Studying the Structural Control Exerted over Their Dynamics. In "Coarse-Graining of Condensed Phase and Biomolecular Systems" G Voth (Ed) Taylor and Francis Group LLC. CRC Press 2008;237-54.
55. Sanejouand YH. Elastic network models: theoretical and empirical foundations. Methods Mol Biol. 2013;924:601-16.
56. Gniewek P, Kolinski A, Jernigan RL, Kloczkowski A. Elastic network normal modes provide a basis for protein structure refinement. J Chem Phys. 2012;136:195101.
57. Keskin O, Bahar I, Badretdinov AY, Ptitsyn OB, Jernigan RL. Empirical solvent-mediated potentials hold for both intra-molecular and inter-molecular inter-residue interactions. Protein Sci. 1998;7:2578-86.
58. Yang L, Song G, Jernigan RL. How well can we understand large-scale protein motions using normal modes of elastic network models? Biophys.J 2007;93:920-9.
59. Zimmermann MT, Skliros A, Kloczkowski A, Jernigan RL. Immunoglobulin structure exhibits control over CDR motion. Immunome.Res 2011;7:5.
60. Bahar I, Jernigan RL. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. J.Mol.Biol. 1997;26:195-214.
61. Zimmermann MT, Kloczkowski A, Jernigan RL. MAVENs: motion analysis and visualization of elastic networks and structural ensembles. BMC Bioinformatics. 2011;12:264.

62. Globisch C, Krishnamani V, Deserno M, Peter C. Optimization of an elastic network augmented coarse grained model to study CCMV capsid deformation. PLoS ONE 2013;8:e60582.
63. Jernigan RL, Kloczkowski A. Packing regularities in biological structures relate to their dynamics. Methods Mol Biol. 2007;350:251-76.
64. Zhu F, Hummer G. Pore opening and closing of a pentameric ligand-gated ion channel. Proc.Natl.Acad.Sci U.S.A 2010;107:19814-9.
65. Bakan A, Meireles LM, Bahar I. ProDy: protein dynamics inferred from theory and experiments. Bioinformatics. 2011;27:1575-7.
66. Yang L, Song G, Jernigan RL. Protein elastic network models and the ranges of cooperativity. Proc.Natl.Acad.Sci.U.S.A 2009;106:12347-52.
67. Keskin O, Jernigan RL, Bahar I. Proteins with similar architecture exhibit similar large-scale dynamic behavior. Biophys.J. 2000;78:2093-106.
68. Jernigan RL, Demirel MC, Bahar I. Relating Structure to Function through the Dominant Modes of Motion of DNA Topoisomerase II . Int J Quant Chem (B Pullman Memorial Volume) 1999;vol.175:301-12.
69. Jernigan RL, Bahar I, Covell DG, Atilgan AR, Erman B, Flatow DT. Relating the Structure of HIV-1 Reverse Transcriptase to Its Processing Step. J.Biomol.Struct.Dyn. 2000;17(sup1):49-55.
70. Dasgupta B, Nakamura H, Kinjo AR. Rigid-body motions of interacting proteins dominate multispecific binding of ubiquitin in a shape-dependent manner. Proteins 2014;82:77-89
71. Ruvinsky AM, Kirys T, Tuzikov AV, Vakser IA. Structure fluctuations and conformational changes in protein binding. J Bioinform.Comput.Biol. 2012;10:1241002.
72. Peng C, Head-Gordon T. The dynamical mechanism of auto-inhibition of AMP-activated protein kinase. PLoS Comput.Biol. 2011;7:e1002082.
73. Uyar A, Kurkcuoglu O, Nilsson L, Doruker P. The elastic network model reveals a consistent picture on intrinsic functional dynamics of type II restriction endonucleases. Phys.Biol. 2011;8:056001.
74. Zimmermann MT, Leelananda SP, Gniewek P, Feng Y, Jernigan RL, Kloczkowski A. Free energies for coarse-grained proteins by integrating multibody statistical contact potentials with entropies from elastic network models. J Struct.Funct.Genomics 2011;12:137-47.
75. Guarnera E, Berezovsky IN. Allosteric sites: remote control in regulation of protein activity. Curr Opin Struct.Biol. 2016;37:1-8.
76. Guarnera E, Berezovsky IN. Allosteric drugs and mutations: chances, challenges, and necessity. Curr.Opin.Struct.Biol. 2020; 62:149-157.
77. Wodak SJ, et al. Allostery in Its Many Disguises: From Theory to Applications. Structure 2019;27:566-578, 2019.
78. Schueler-Furman O, Wodak SJ. Computational approaches to investigating allostery. Curr.Opin.Struct. Biol. 2016;41:159-171.
79. Tsai CJ, Nussinov R. A unified view of "how allostery works". PLoSComput.Biol. 2014;10:e1003394.
80. Dokholyan NV. Controlling Allosteric Networks in Proteins. Chem.Rev. 2016;116:6463-87.
81. Zhang Y, Doruker P, Kaynak B, Zhang S, Krieger J, Li H, Bahar I. Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. Curr.Opin.Struct.Biol. 2020;62:14-21.