Evaluation of the Stability of Folding Nucleus upon Mutation

Mathieu Lonquety[1,2], Zoé Lacroix[1,3], and Jacques Chomilier[2]

[1] Scientific Data Management Laboratory, Arizona State University
Tempe AZ 85281-5706, USA
[2] IMPMC, Université Pierre et Marie Curie, UMR 7590 CNRS
140 rue de Lourmel, 75015 Paris, France
[3] Pharmaceutical Genomics Division, Translational Genomics Research Institute
13400 E Shea Blvd, Scottsdale AZ 85259, USA

Abstract. The development of a method that accurately predicts protein folding nucleus is critical at least on two points. On one hand, they can participate to misfolded proteins and therefore they are related to several amyloid diseases. On the other hand, as they constitute structural anchors, their prediction from the sequence can be valuable to improve database screening algorithms. The concept of Most Interacting Residues (MIR) aims at predicting the amino acids more likely to initiate protein folding. An alternative approach describes a protein 3D structure as a series of Tightened End Fragments (TEF). Their spatially close ends have been shown to be mainly located in the folding nucleus. While the current sequence-driven approach seems to capture all MIR, the structure-driven method partially fails to predict known folding. We present a stability-based analysis of protein folding to increase the recall and precision of these two methods.

Results: Prediction of the folding nucleus by MIR algorithm is in agreement with mutation stability prediction.

Availability: The database is available at:
http://bioinformatics.eas.asu.edu/Stability/index.php. The MIR calculation program is available at:
http://bioserv.rpbs.univ-paris-diderot.fr/cgi-bin/MIR and the TEF program at:
http://bioserv.rpbs.univ-paris-diderot.fr/TEF.

Contact: jacques.chomilier@impmc.jussieu.fr

Keywords: Protein folding, folding nucleus, structure stability, point mutations.

1 Introduction

Structural bioinformatics has been particularly productive for the past decade partially thanks to the contribution of physics disciplines. One of its main focuses is the study of protein folding and, in particular, the prediction of folding nuclei. Modeling and predicting protein folding mechanisms is critical because

M. Chetty, A. Ngom, and S. Ahmad (Eds.): PRIB 2008, LNBI 5265, pp. 54–65, 2008.

a misfolded protein may result in the formation of aggregates that may play a role in most misfolding diseases such as amyloid ones [1,2,3,4].

The folding nucleus model [5,6,7,8,9] is based on the assumption that protein folding begins with just a few amino acids that strongly interact with each other. These strong interactions initiate the folding that is completed by a successive folding of the remaining parts of the structure to constitute a compact globule. Within this model, a precise and accurate prediction of the main amino acids responsible for initiating the folding provides enough constraints to simulate the whole folding mechanism. For that purpose, Papandreou *et al.* developed an algorithm devoted to the search of the Most Interacting Residues (MIR) [10]. The current algorithm has a good recall however its precision needs improvement as several studies consider that the minimal number of amino acids needed to initiate a folding process is significantly less than the 15 % found in average with the MIR prediction algorithm [8,11,12,13].

Proteins can be described as a succession of Tightened End Fragments or TEF [14,15,16,17] which spatially close ends (lower than 10 Å) are deeply buried in the cores of globular domains. Their ending positions could represent the folding nucleus while TEF would correspond to the final fold for these portions. Indeed, we have previously demonstrated that the TEF ends correspond statistically to hydrophobic residues highly conserved in multiple alignments of proteins of common function [17]. These particular positions have been called topohydrophobic, and they are clearly related to amino acids belonging to the folding nucleus [18]. They are derived from multiple alignments of distantly related sequences, typically less than 30 % identity. It constitutes a limitation of the prediction process since most of the available algorithms for multiple alignments of highly divergent sequences produce controversial results [19]. We have shown that MIR and tophohydrophobic positions match in two thirds of the cases which confirms a reasonable recall of the MIR prediction algorithm. In other words, one has a mean to predict, from the single information of the sequence, positions (MIR) including the folding nucleus.

In this paper we present a stability-based analysis that was conducted to better characterize MIRs. The expected results were to improve the precision of the MIR method by refining the algorithm with constraints related to the prediction of the stability changes induced by point mutations. We assume that the folding nucleus is the deep core of the structure and thus should be very sensitive to point mutations. For example, if a keystone substitutes another one with a different shape, the vaulting will collapse almost every time.

2 Material and Methods

2.1 MIR Prediction Algorithm

A Monte Carlo algorithm is used to simulate the early steps of protein folding on a (2,1,0) lattice. An amino acid is randomly selected and displaced to a new available position on the lattice. The energy of both initial and final conformations is computed from the Miyazawa and Jernigan potential of mean force [20]

and the Metropolis criterion is then applied [21,10]. The starting point is the protein structure in a random coil conformation and the simulation is typically conducted on 10^6 Monte Carlo steps.

This simulation is repeated 100 times with different initial conformations. The number of first neighbors is recorded after each series of 10 Monte Carlo steps, and at the end of the process, an average Number of Contact Neighbors (NCN) is calculated for each amino acid of the sequence. Actually, amino acids surrounded by many others play a role in the compactness of the protein and thus are called Most Interacting Residues (MIR). In contrast, the ones with few neighbors are called Less Interacting Residues (LIR).

2.2 TEF Assignment

Along the backbone of a protein, some pairs of amino acids can be very close in several places, with a typical distance between their alpha carbons below 10 Å. The histogram of the sequence separation between these "contact" amino acids is not smooth, and presents a maximum around 25 amino acids [15]. These sequence fragments were initially called closed loops [14].

Later on, it has been shown that the ends of these closed loops are mainly occupied by hydrophobic amino acids. A thorough analysis demonstrated that these hydrophobic amino acids were highly conserved among structures of the same family, although containing distantly related sequences: these positions were called topohydrophobic [22].

The concept of TEF emerged from the junction between closed loops and topohydrophobic positions mainly located at their ends.

2.3 Free Energy Calculation

Gibbs free energy change due to mutation is a good approximation to characterize the stability of a given structure. It consists of a succession of energetic terms that attempt to capture all the properties and forces that drive the conformation of a protein. In our study we focus on the difference of these energies for the wild type structure ΔG_{wild} and for the mutant structure ΔG_{mutant}. Considering that in the literature various stability prediction methods use different nomenclature, $\Delta\Delta G$ is defined as follows:

$$\Delta\Delta G = \Delta G_{\text{mutant}} - \Delta G_{\text{wild}} \ . \tag{1}$$

The unit is kcal/mol. $\Delta\Delta G$ describes whether it costs more in energy to have the mutated amino acid or the wild type one. For example, if $\Delta\Delta G < 0$ then it costs more in energy to have the wild type structure than the mutant one thus the mutation is more favorable to the structure stability. Conversely, if $\Delta\Delta G > 0$, the mutant structure ΔG is higher than the wild type one thus the mutation is less favorable to the structure stability.

2.4 Stability Analysis

Many methods have been implemented to predict stability changes induced by point mutations. MUpro [23] and I-Mutant (sequence version) [24] both predict stability changes on a protein sequence whereas DFIRE [25], I-Mutant (sequence + structure version) [24] and PoPMuSiC [26] use protein sequence and structure to predict these changes. Other methods exist but have been rejected due to some restrictions: CUPSAT[1] [27] was not available in a standalone version and the current version of FoldX[2] [28] only computes mutations to Alanine. To avoid biases from one or the other method, we present a comprehensive analysis with five existing tools (the two versions of I-Mutant are considered as two different tools).

We use the Protherm database [29] that collects thermodynamic data published in the scientific literature and thus includes measured values of $\Delta\Delta G$ to compare our prediction to experimental data. It is available at the following URL: http://gibk26.bse.kyutech.ac.jp/jouhou/Protherm/protherm.html.

2.5 Data Set

Our analysis was conducted on a dataset published by the Protein Folding Fragments European consortium that can be found at the URL: http://bioserv.rpbs.univ-paris-diderot.fr/PFF/. MIR predictions and TEF calculations were already performed on the selected 116 protein sequences.

The experimental dataset consisted of 116 protein sequences for a total of 15,183 amino acids. Each sequence was processed with each of the five stability prediction tools, for each amino acid, and for each of the 19 possible mutations. We computed 1,442,385 different $\Delta\Delta G$ values. In order to manage and publish our produced data and results in a more efficient way than output flat files, a database was created and is available to the community at http://bioinformatics.eas.asu.edu/Stability/ where more information about the data can be found.

3 Results

3.1 MIR and TEF

The MIR concept aims at characterizing the main amino acids involved in the early steps of the protein folding process. The TEF method splits a structure into fragments with spatially close ends that interact with each other. A previous study [10] has demonstrated that TEF ends (within a range of ± 5 positions) correspond to MIR in 57 % of the cases. As we are looking at coherent methods to determine the folding nucleus, we observe that MIR over predict the TEF ends, therefore we hypothesize that restricting the MIR to the ones in agreement with TEF ends would capture the expected amino acids responsible for protein folding.

[1] CUPSAT is available at http://cupsat.tu-bs.de/.
[2] FoldX is available at http://foldx.crg.es//.

Then, separation between "good" MIR and "bad" MIR emerges, and we define them as TEF related and TEF independent MIR. The TEF independent MIR are expected to be the noise in the MIR prediction algorithm. The TEF related ones are those in a ± 3 amino acids window around a TEF end. As the experimental validation of folding nucleus is rather difficult, one way to validate MIR prediction (i.e., TEF related MIR are the nucleus residues) is the comparison with other structural data. Indeed, there is nowadays no experimental technique able to determine which residues constitute the folding nucleus.

The Φ value experimental determination [30] attests whether one amino acid is in the folding nucleus or not, but all the mutants need to be constructed to validate the hypothesis. Nevertheless in some cases, such as CI2 for instance, a low Φ value can be obtained for a residue attributed to the folding nucleus by convergent experiments [31]. We propose here to add constraints derived from energy stability evaluation in order to increase the agreement between prediction and experiment; these constraints are restrictred to thermodynamics experiments as they are not supposed to be suspicious.

3.2 Stability Changes upon Point Mutations

Because they are structurally compulsory for the complete folding, we assume that folding nucleus positions are very sensitive to mutation, in the sense that a mutation would destabilize the protein. We thus decide to verify this assumption by computing stability changes upon point mutations for all the sequences which already have been processed for the MIR and TEF predictions (See Material and Methods section). There exist numerous software devoted to this task among which we focused on: DFIRE, two versions of I-Mutant, MUpro, and PoPMuSiC.

We first start by calculating $\Delta\Delta G$ resulting from mutations at each position for the five tools. We retrieve all experimental values for proteins either present in our database and in Protherm [29]. 1409 different mutations with their experimental $\Delta\Delta G$ were gathered. A correlation then appear between experimental $\Delta\Delta G$ and predicted $\Delta\Delta G$. The two versions of I-Mutant obtain the best score (represented in Fig. 1) with 0.96 correlation coefficient, just followed by MUpro with 0.86. The remaining tools PoPMuSiC and DFIRE show an average correlation with 0.53 and 0.48 respectively. The goal of these correlations is to verify that the tools used in this work are accurate and truthfully. The excellent correlation for I-Mutant and MUpro can be explained by the fact that both software used data extracted from the Protherm database as a training set for their algorithm. No other experimental data was available for this study.

Three tools can be considered as efficient and two others have to be carefully apprehended. Nevertheless, for a better overview of the stability changes concept, all five tools are kept in this study and a comparison between the two types of MIR and their relative stability changes upon mutation can be performed. We then verify if stability upon mutation would allow to discriminate among the two types of MIR, according to their location relative to TEF ends. Stability prediction is characterized by $\Delta\Delta G$ on each amino acid and for each possible mutation. We compute a stability score to compare with MIR prediction as

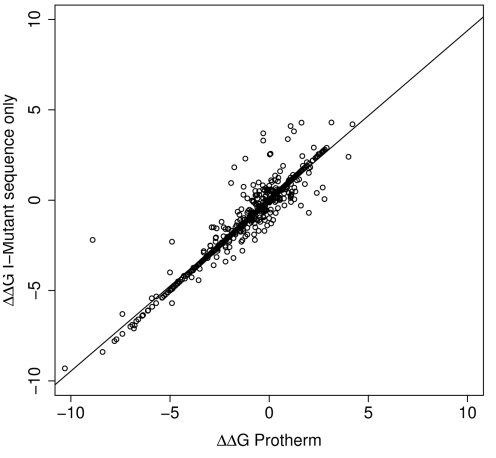

Fig. 1. Graph of $\Delta\Delta G$ (kcal/mol) predicted by I-Mutant (sequence only version) as a function of the experimental one, taken from the Protherm database. The line represents the correlation between both experiment and prediction.

follows. To synthesize the mean stability change tendency for a given amino acid, the 19 $\Delta\Delta G$ have been summed in one value. Actually, to normalize this result, instead of summing the $\Delta\Delta G$, a score has been given to each $\Delta\Delta G$. If $\Delta\Delta G < 0$, the mutation is considered as stabilizing and it is granted a value of $+1$. Conversely, if $\Delta\Delta G > 0$, the mutation is considered as destabilizing and the value is -1. This procedure produces a score in the range of $[-19,+19]$ which reflects the global stability change for an amino acid upon its mutation. The lower the score, the more sensitive to mutation, i.e., the native residue is the most stable. This stability score is computed for each amino acid of all the sequences in the data set and for the five different tools. Moreover, a consensus tool has been created which corresponds to the mean of the five programs. Graphs are plotted, upon request on the server, to get an overview of the stability score over a whole sequence. One example is given in Fig. 2 where the stability scores along a whole sequence have been represented for the five tools in the case of the engrailed homeodomain (PDB code: 1enh). Stability scores curves have also been smoothened for an easier interpretation.

We observe that along the homeodomain sequence, stability changes score ranges from -19 to $+15$. One can notice that most values are under 0 which means that there are more positions destabilized by mutations than stabilized ones. This observation is in agreement with the principles of Evolution which tend to favor stable protein structures.

It is thus possible to detect the most sensitive positions to a mutation. Indeed, the minima of stability scores are positions for which mutations induce the most destabilizing changes, regarding free energy, along the structure. We can locate these positions and compare them with the MIR.

To bear evidence of an eventual co-localisation of MIR and stable positions, we compute the distance in sequence that separates each MIR (TEF related and TEF independent) from the nearest minimum of stability score. These differences of positions have been computed for all the sequences of the database and for each tool. Figure 3 shows an example of these deltas of positions for the consensus tool. It appears that there is no distinction between the two classes of MIR as both have their respective peak centered on the same position.

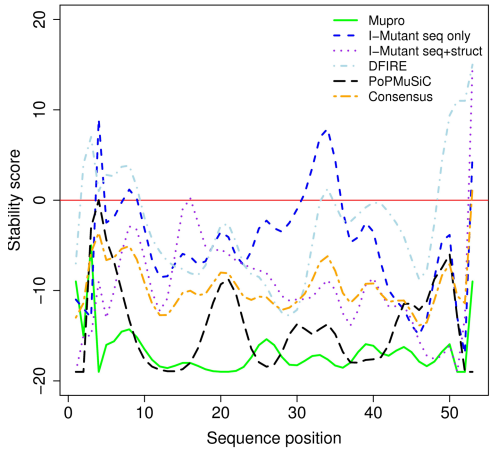

Fig. 2. Representation of the stability scores for each amino acid of the 1enh sequence. The five lines represent each one a different tool. The consensus graph is also represented.

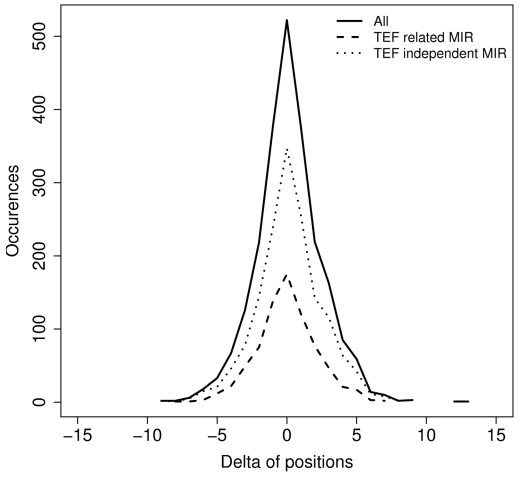

Fig. 3. The origin of the abscissa corresponds to the position of each MIR (TEF related and independent) and one calculates the distance to the closest minimum of stability scores on the whole dataset and for the consensus tool.

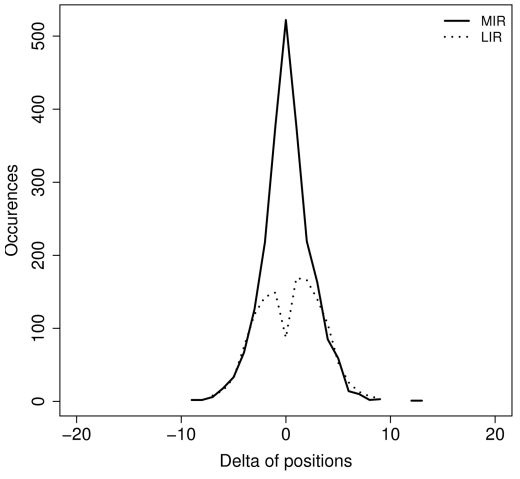

Fig. 4. Sequence separation (Delta) between MIR, LIR, and minima of stability scores on the whole data set and for the consensus tool. The origin is taken as for Fig. 3.

However, MIR have been shown to statistically match the topohydrophobic positions, corresponding in several experiments to the folding nucleus [18]. The conservation of the folding nucleus among species is still under a strong debate. If we assume, in agreement with Shakhnovich [31], that the folding nucleus is the subject of an additional evolution pressure, considering simulations on distantly related sequences of the same fold, instead of single one, may help in a better definition of the folding nucleus.

Two hypotheses emerge: the MIR algorithm is not accurate enough or the TEF assignment has to be improved. A direct comparison of each of these methods with minimua of protein stability scores has been processed.

The determination of MIR accuracy relies on their good match with minima of stability scores but also on the comparison between their antagonists i.e., LIR for Less Interacting Residues and stability scores under the same protocol. These amino acids are the ones with the smallest number of contact neighbors and are thus assumed to be mainly located at the interface of the protein and the solvent. Results are shown in Fig. 4. As already seen in Fig. 3, MIR are statistically located at minima of stability scores with a peak for a delta of 0. If one takes a window centered on 0, on the $[-1, +1]$ range, 55 % of the MIR correspond to a minimum of stability score. For the LIR, we observe that there is a clear minimum on the 0 position and two peaks centered on the -2 and +2 positions. The conclusion is that the MIR concept is in good agreement with the concept of sensitivity to the structure stability. This prediction method succeeds in correctly locating the most stable residues, that can be either located at the ends of TEF or elsewhere, because both classes of MIR match the location of the lowest scores.

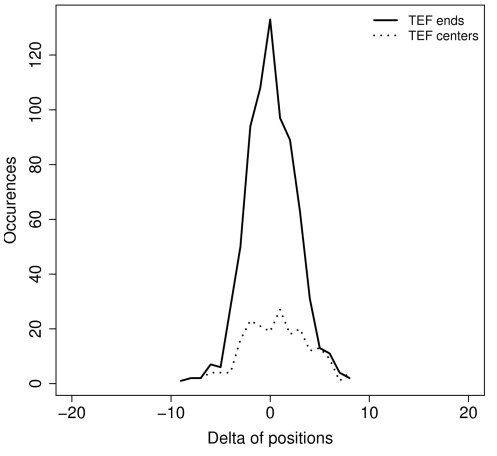

Fig. 5. Sequence separation (Delta) between TEF ends, TEF centers, and minima of stability scores on the whole data set and for the consensus tool.

We then compare specific TEF positions to the positions with the lowest stability scores. We consider on one hand TEF ends, as they are assumed to be in the folding nucleus, and TEF centers on the other hand. Figure 5 represents the distance in terms of amino acids (delta) between the TEF ends/TEF centers and minima of stability score on the other side. The results are compatible with the ones presented for Fig. 4 with MIR and LIR. TEF ends match the positions where stability is the highest (lowest scores), while TEF centers do not. Therefore, one can conclude that MIR predictions capture some physics of the folding process, by finding residues forming the core, evaluated here on the basis of the most stable positions toward mutation, independently of their location relative to the TEF.

The relative efficiency of this method has been confirmed by the calculation of solvent accessible surface for all amino acids and for each sequence of the database. The mean value is of $53\,\text{Å}^2$ for one amino acid among all the sequences of the dataset. If we now consider amino acids which are characterized as MIR, this mean drops to $33\,\text{Å}^2$. For the LIR, we obtain a rise to $64\,\text{Å}^2$. This observation also gives another evidence of the efficiency of the MIR method as low solvent accessible surface induces that the considered amino acid is buried inside the globular domain.

For the results observed for the TEF the conclusion is less evident. TEF ends are centered on the positions of the highest stability, but TEF centers graph is more ambiguous as there is kind of a plateau in the range $[-3, +3]$. It thus means that TEF ends are quite in agreement with stability scores minima but TEF centers do not show any tendency to be reluctant to stability scores minima.

4 Conclusion

Protein folding is nowadays one of the biggest challenges in structural bioinformatics. The MIR method is devoted to the prediction of the residues forming the

folding nucleus of proteins. Some refinement has been proposed to improve the accuracy of the current algorithm such as the use of additional input based on topology and stability. The structural analysis of proteins in terms of TEF was a relevant choice as it captures ends of fragments buried in the core of the protein. MIR are constituted of two families, the ones present at the ends of TEF and the other ones found elsewhere. It was hypothesized that folding nucleus would preferentially be located at TEF ends. We checked the relevance of this separation by comparing the presence of MIR with positions known to be stable upon point mutation. We actually evidenced that both classes of MIR are highly stable positions with respect to mutations. This result may be interpreted in the following way: if we admit the assumption that most stable positions toward mutation are indicative of the inclusion in the folding nucleus, then MIR is a rather satisfactory method to predict this nucleus. In addition, we assume that split of the protein structures into TEF should be improved, and in particular, one might think of secondary contacts, i.e. two residues located in the middle of a TEF, and close from one each other.

Although we probably overestimate the number of amino acids involved in the folding nucleus, our approach might be a help for selecting positions susceptible of experimental muations in order to perform Φ_F determination. A long term application of this prediction nucleus algorithm is its inclusion in database screening tools, in order to give a stronger weight once a residue has been postulated as belonging to the nucleus. One might guess that this would help in retrieving more distantly related sequences than present methods.

Acknowledgement. RPBS resources have been used in order to produce all these results. The authors would like to thank Dr. Nikolaos Papandreou for his contribution to the MIR calculation program. Many thanks for the authors of the different software devoted to stability changes and especially to Dr. Jean-Marc Kwasigroch for his help on using the PoPMuSiC software. Dr. Pierre Tufféry is acknowledged for his help in implementing the MIR and TEF programs on the RPBS server.

This research was partially supported by the National Science Foundation[3] (grants IIS 0431174, IIS 0551444, and IIS 0612273) and by an invitation of the Université Pierre et Marie Curie. J.C. and M.L. benefited of a EU grant QLG2-2002-01298.

References

1. Brockwell, D.J., Smith, D.A., Radford, S.E.: Protein folding mechanisms: new methods and emerging ideas. Curr. Opin. Struct. Biol. 10(1), 16–25 (2000)
2. Steward, R.E., MacArthur, M.W., Laskowski, R.A., Thornton, J.M.: Molecular basis of inherited diseases: a structural perspective. Trends Genet. 19(9), 505–513 (2003)

[3] Any opinion, finding, and conclusion or recommendation expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

3. Mogensen, J.E., Ipsen, H., Holm, J., Otzen, D.E.: Elimination of a misfolded folding intermediate by a single point mutation. Biochemistry 43(12), 3357–3367 (2004)
4. Cerdà-Costa, N., Esteras-Chopo, A., Avilés, F.X., Serrano, L., Villegas, V.: Early kinetics of amyloid fibril formation reveals conformational reorganisation of initial aggregates. J. Mol. Biol. 366(4), 1351–1363 (2007)
5. Wetlaufer, D.B.: Nucleation, rapid folding, and globular intrachain regions in proteins. Proc. Natl. Acad. Sci. USA 70(3), 697–701 (1973)
6. Abkevich, V.I., Gutin, A.M., Shakhnovich, E.I.: Specific nucleus as the transition state for protein folding: evidence from the lattice model. Biochemistry 33(33), 10026–10036 (1994)
7. Fersht, A.R.: Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. Proc. Natl. Acad. Sci. USA 92(24), 10869–10873 (1995)
8. Shakhnovich, E., Abkevich, V., Ptitsyn, O.: Conserved residues and the mechanism of protein folding. Nature 379(6560), 96–98 (1996)
9. Fersht, A.R.: Nucleation mechanisms in protein folding. Curr. Opin. Struct. Biol. 7(1), 3–9 (1997)
10. Papandreou, N., Berezovsky, I.N., Lopes, A., Eliopoulos, E., Chomilier, J.: Universal positions in globular proteins. Eur. J. Biochem. 271(23-24), 4762–4768 (2004)
11. Sacile, R., Ruggiero, C.: Hunting for "key residues" in the modeling of globular protein folding: an artificial neural network-based approach. IEEE Trans Nanobioscience 1(2), 85–91 (2002)
12. Religa, T.L., Markson, J.S., Mayor, U., Freund, S.M.V., Fersht, A.R.: Solution structure of a protein denatured state and folding intermediate. Nature 437(7061), 1053–1056 (2005)
13. Alexander, P.A., He, Y., Chen, Y., Orban, J., Bryan, P.N.: The design and characterization of two proteins with 88% sequence identity but different structure and function. Proc. Natl. Acad. Sci. 104(29), 11961–11963 (1968)
14. Ittah, V., Haas, E.: Nonlocal interactions stabilize long range loops in the initial folding intermediates of reduced bovine pancreatic trypsin inhibitor. Biochemistry 34(13), 4493–4506 (1995)
15. Berezovsky, I.N., Grosberg, A.Y., Trifonov, E.N.: Closed loops of nearly standard size: common basic element of protein structure. FEBS 466(2-3), 283–286 (2000)
16. Berezovsky, I.N., Kirzhner, V.M., Kirzhner, A., Trifonov, E.N.: Protein folding: looping from hydrophobic nuclei. Proteins 45(4), 346–350 (2001)
17. Lamarine, M., Mornon, J.P., Berezovsky, N., Chomilier, J.: Distribution of tightened end fragments of globular proteins statistically matches that of topohydrophobic positions: towards an efficient punctuation of protein folding? Cell Mol. Life Sci. 58(3), 492–498 (2001)
18. Poupon, A., Mornon, J.P.: Predicting the protein folding nucleus from sequences [correction of a sequence]. FEBS Lett. 452(3), 283–289 (1999)
19. Baussand, J., Deremble, C., Carbone, A.: Periodic distributions of hydrophobic amino acids allows the definition of fundamental building blocks to align distantly related proteins. Proteins 67(3), 695–708 (2007)
20. Miyazawa, S., Jernigan, R.L.: Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. J. Mol. Biol. 256(3), 623–644 (1996)
21. Chomilier, J., Lamarine, M., Mornon, J.P., Torres, J.H., Eliopoulos, E., Papandreou, N.: Analysis of fragments induced by simulated lattice protein folding. C. R. Biol. 327(5), 431–443 (2004)

22. Poupon, A., Mornon, J.P.: Populations of hydrophobic amino acids within protein globular domains: identification of conserved "topohydrophobic" positions. Proteins 33(3), 329–342 (1998)
23. Cheng, J., Randall, A., Baldi, P.: Prediction of protein stability changes for single-site mutations using support vector machines. Proteins 62(4), 1125–1132 (2006)
24. Capriotti, E., Fariselli, P., Casadio, R.: I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 33(Web Server issue), W306–W310 (2005)
25. Zhou, H., Zhou, Y.: Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. Protein Sci. 11(11), 2714–2726 (2002)
26. Gilis, D., Rooman, M.: PoPMuSiC, an algorithm for predicting protein mutant stability changes: application to prion proteins. Protein Eng. 13(12), 849–856 (2000)
27. Parthiban, V., Gromiha, M.M., Schomburg, D.: CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res. 34(Web Server issue), W239–W242 (2006)
28. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., Serrano, L.: The FoldX web server: an online force field. Nucleic Acids Res. 33(Web Server issue), W382–W388 (2005)
29. Kumar, M.D.S., Bava, K.A., Gromiha, M.M., Prabakaran, P., Kitajima, K., Uedaira, H., Sarai, A.: ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. Nucleic Acids Res. 34(Database issue), D204–D206 (2006)
30. Fersht, A.R., Daggett, V.: Protein folding and unfolding at atomic resolution. Cell 108(4), 573–582 (2002)
31. Shakhnovich, E.: Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. Chem. Rev. 106(5), 1559–1588 (2006)