4

Decomposition of Intramolecular Interactions Between Amino-Acids in Globular Proteins - A Consequence for Structural Classes of Proteins and Methods of Their Classification

Boris Fackovec[1,2] and Jiri Vondrasek[1]
[1]*Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague,*
[2]*Faculty of Natural Sciences, Charles University in Prague, Prague,*
Czech Republic

1. Introduction

An amino-acid in proteins shows two different, yet mutually dependent faces connected through the polymer character of a protein in the final product. They are the amino-acid side-chain and its corresponding backbone part. On the level of the side-chains, we often refer to specific structural arrangements such as hydrophobic cluster motifs, salt-bridge motifs or hydrogen-bond motifs characterizing various parts of a protein and usually assigned to a certain function. The backbone on the other hand offers limited, yet general structural motifs – α, β and random coil patterns. All of these mentioned amino-acid features contribute to the synergy demonstrated observably by protein stability and protein function.

Thermal stability is one of the most important features of the structure of a fully folded protein. It is defined as the difference in the Gibbs free energy between its native and denaturated states and as such is a function of temperature and implicitly a function of protein composition and the effect of the environment. Nevertheless, it is necessary to say that for this function we do not know yet the precise and general form which could be applicable for a large set of proteins. There have been many attempts to propose an intuitive, yet productive decomposition of Gibbs free stabilization energy (GFSE) into simple terms. One of the scenarios utilized for such purposes is that the total free energy is the sum of the free energies of various atomic groups and the hydrophobic effect. However, as the free energy is not additive and the fractionation of free energy to independent terms is difficult, this attempt has been quite unsuccessful.

The utilization of molecular modeling methodology and tools has opened a more systematic and perhaps more promising approach – the evaluation of the enthalpy term in the equation for Gibbs free energy with reasonable accuracy (Lazaridis, Archontis, & Karplus, 1995). The remaining entropy term could be obtained by fitting the corresponding analytical form to the experimental data. There are basically three different enthalpy contributions that we can separate. The first comes from the intramolecular interactions between the atoms of proteins, producing the largest stabilizing enthalpy contribution. The second comes from the interactions between the molecules of a solvent, and finally the third contribution is the result of the interactions between the atoms of the solute (protein) and the solvent.

It is commonly believed that the dominant force of protein folding and therefore the main stabilizing force of the native structure is the hydrophobic effect (Dill, 1990). However, it has been insightfully pointed out (Makhatadze et al. 1995) that a water environment destabilizes folded protein structures and the decomposition of enthalpy shows that the solvation models introduce significant errors. In these studies, it has been assumed that the denatured state of a protein can be identified with the fully unfolded state (Makhatadze et al. 1989), where residues do not interact with each other. Even in light of this hypothesis, the intramolecular interactions between amino-acids in a protein are expected to contribute significantly to its overall stability. However, the hypothesis has never been proved and the importance of the intramolecular interactions would be much higher if the unfolding were considered as "core melting" rather than "oil-droplet dissolution". Regardless of the denatured form, the intramolecular organization of a protein is the result of a subtle balance between the rigidity/flexibility of the protein backbone and the noncovalent interactions between protein's side-chains. This result in conformational unique and stable protein structures as well as the ratio between the importance of the backbone/side-chain contributions can vary for different proteins.

The main problem of the enthalpy (or the potential energy) approach is that we are unable to evaluate the enthalpy-entropy compensation; therefore, the theoretically determined enthalpy contribution should be adjusted in some other way. A realistic method is to correlate the calculated values with the experimental data obtained by microcalorimetry, where both the enthalpy and the entropy terms can be determined. On the level of particular amino-acids, we face the problem of their "denatured-state" definition for the reasonable decomposition of the free energy on individual amino-acids.

The dissection of the enthalpy contribution which the intra-molecular noncovalent interaction energy (part of the potential energy) is a component of seems to be a reasonable approach for the study of the role of the composing amino-acids in protein stabilization. We can decompose this energy into individual pairwise amino-acid contributions and determine their importance for protein stability. The evaluation of the interaction energy (of noncovalent origin) between biomolecules or between their parts is a traditional field of the symbiosis between experiment and theory, and the methodology is well described and highly developed (Müller-Dethlefs & P Hobza, 2000). The crucial condition for the success of the theoretical methodology is the accuracy of the methods utilized. Recently, it has been quite common to evaluate the potential energy of a protein at the suitable *ab initio* methodology level, but we are still severely limited by the size of the protein. Therefore, the Density Functional Theory methods (DFT) are the most utilized for such purposes (Riley, 2010). Unfortunately, the DFT methods fail to describe the noncovalent interactions reasonably mostly because of the missing electron correlation term. Even the new functionals recently introduced (Kolář, 2010) have failed to describe properly the noncovalent potential curve mostly in the repulsion and asymptotic regions. Such inaccuracies can be tolerated at the energy minima, but only a limited number of the interactions between amino-acids in proteins meet such a requirement. Therefore, only high-level *ab initio* methods can be utilized – at least for benchmark studies. As was shown on a set of representative interactions between amino-acid side-chains in proteins in 2009, empirical force fields (namely OPLS and AMBER) are suitable for the description of their interaction (Berka, 2009). Kolar (Kolář, 2010) tested the performance of the energy calculations using MM on a representative set, S22, and found quite satisfactory agreement between the empirical force fields and high-level *ab initio* methods. It was later shown that we can use the empirical force field with satisfactory accuracy also for the description of the intramolecular interaction-energy distributions for pairs of amino-acid side-chains (Berka, 2010). Still, one has to be aware

of the limitations of the force-field methods, namely for subtle cases of the interactions present in proteins. On the other hand, the utilization of empirical methods decreases the computational cost and provides an opportunity to investigate the trends presented in biomolecules if the highest accuracy is not the major issue.

The evaluation of the interaction energy between amino-acid residues resulted in the interaction energy matrix (IEM) concept being introduced in 2008 (Bendová-Biedermannová, 2008). The IEM approach was used to identify the key residues for protein stability in a model system – rubredoxin. The matrix carries information about the energy and the role of a residue in the protein structure, namely its interaction energy strength, which is more than the simple distance matrix concept. It also shows how much a certain residue is a hub within the context of the other interacting amino-acids. The IEM approach might also open new horizons for the investigations of proteins. The concept could be incorporated into the methods of protein-structure superpositions (similar to the DALI approach)(Holm & Sander, 1997) and can shed light on other protein-related issues – for example protein stability, folding kinetics, foldability and design.

The work presented in this study is based on the calculations of the amino-acid – amino-acid interaction energies (IEs) between all of the residues in approximately 1400 proteins to justify the roles of different amino-acids, their backbones and side-chains and their physical-chemical character for structural or stabilization preferences. We especially focused on the problem of how the interaction energy distributions are related to the secondary-structure content defined by the CATH (Orengo et al., 1997) and SCOP(Murzin, 1995) criteria.

2. Amino-acids in proteins and their distribution

2.1.1 Representative structure-set selection

All of the protein structures utilized in this study were obtained from the PDB database (download Jan 31, 2011). We selected only protein molecules with one chain, no ligands, resolved by the X-ray crystallography method at a minimum resolution of 2.0 A. We also omitted structures with a 70% sequence identity and higher. The database filter yielded 1531 structures. This number was slightly reduced by inconveniences with file processing to 1358. The characteristics of the set are illustrated in Figure 1 (size histogram, resolution histogram).

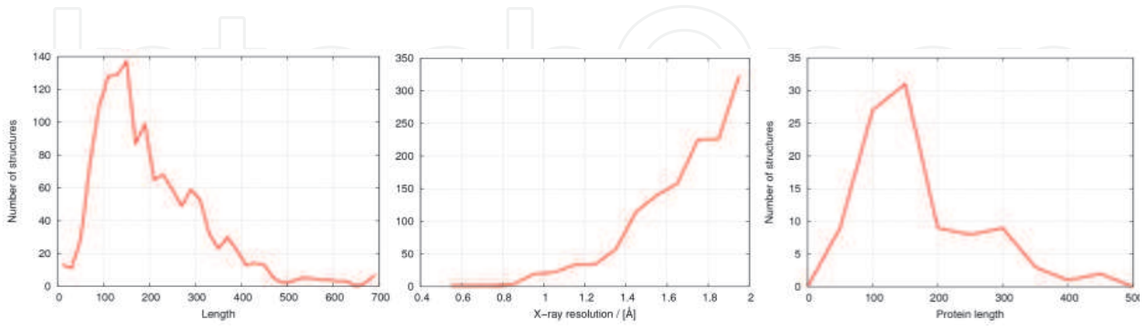


Fig. 1. a) Number of structures against protein length; binned by 20 AA; b) number of structures with a particular X-ray resolution; binned by 0.1 A; c) histogram of the sizes of the structures selected for secondary-structure studies.

Incomplete amino-acid side-chains (missing heavy atoms, disordered) were replaced by glycine in the cases where backbone atoms were available. Amino-acids with missing

backbone atoms would have discredited the whole set and were therefore omitted. The missing hydrogen atoms were added by the Xleap program from the AMBER (Case et al. 2010) simulation package for pH 7 and the parameters were assigned according to the OPLS FF (Jorgensen & Rives 1988). The ambiguity of protonation, mainly in the case of histidine, is discussed later. The structures were optimized using the GROMACS (Hess et al. 2008) molecular simulation package with the steepest descent algorithm being employed. The hydrogen atoms were optimized first and then the full optimization of the whole protein in the gas phase was performed.

To address the question of the residue selectivity for secondary structure motifs, the structures were classified according to the CATH and SCOP categories and four representative sets were selected. To prevent the interference of the size and secondary structure effect, we assured that the structure sets possess the same size distribution.

Hence, the structures pertaining to particular secondary-structure sets were binned according to their chain length (bin size 50, see Figure 2) and were randomly removed from the bins until the number of structures in the corresponding bins was the same for all the sets. This procedure resulted in four sets, each containing 99 structures.

2.1.2 The fragmentation of proteins

To differentiate between the particular types of interactions which every amino-acid can maintain, we assigned every atom of a residue to one of four attributes according to their occurrence in the backbone or to their occurrence in certain types of amino-acid side-chains. The attributes were as follows – BB – backbone atoms, CH – side-chain atoms of charged residues (asp, glu, lys, arg, his), PO – side-chain atoms of polar residues (asn, gln, thr, ser) and NP – side-chain atoms of nonpolar and aromatic residues (gly, ala, leu, ile, val, pro, cys, met, phe, tyr, trp). Such classification provides the lowest number of groups necessary to discern between interactions characterized by different distance dependencies and orders of magnitude (different physical characters). On the other hand, breaking residues into more parts is restrained by the resulting charges of the fragments which would introduce significant but artificial electrostatic energies. The OPLS force field guarantees that the backbone (which includes C_α) and side-chain fragments are neutral. The physical character of the interaction energies of the aromatic residues is close to those of nonpolar residues. Hence, taking into account digestibility of presented data, we decided not to increase the number of attributes.

2.1.3 The Interaction Energy Matrix (IEM) calculation

After all of the structural optimizations, the pairwise interaction energies for all of the residues at the OPLS level were calculated excluding those between backbones of adjacent amino-acid in primary structure which were set to zero. The interactions were calculated separately for the backbones and side-chains as the sum of the interatomic Lennard-Jones and Coulombic contributions in the gas phase (ε_r=1) using an in-house developed Python program utilizing the standard libraries. The classification of the amino-acid atoms in four groups resulted in ten types of mutual interactions – BB-BB, BB-CH, BB-PO, BB-NP, CH-CH, CH-PO, CH-NP, PO-PO, PO-NP, NP-NP – reflecting the attributes of the atoms involved. For example, CH-CH represents salt bridges and all of the interactions between the side-chains of charged residues regardless of their relative distance and charge sign.

Each type of interaction for one protein was represented by one interaction energy matrix, namely a NxN (where N denotes the number of residues) matrix containing the interaction energy between the atoms of residues i and j with particular attributes assigned. It is guaranteed that no interaction energy is counted twice, so the sum of all of the matrices provides the interaction energy between the corresponding residues.
In order to compare the residual energy content, we have introduced a residue interaction energy (RIE) characteristic for each residue. The RIE of a certain type is defined as the sum of all of the interactions the residue can maintain – the sum of all the numbers in a particular row (or column) in the IEM of that type. At the end, we have ten (NxN dimension, where N is the number of amino-acids) IEMs of different types in one protein. Most of the IEs are of course almost zero; some are set as zero by definition.

2.1.4 Representation of data – cumulative distribution functions and histograms

There are two main data representation schemes in this work. Those are as follows:
The distributions of RIEs of a certain type in one protein. For one specific type and one specific protein set (for example CH-CH in SCOP β), the following procedure was performed to acquire an average distribution representing the whole set. The non-zero RIEs calculated from appropriate IEM were sorted independently for each protein and the distributions were obtained as a plot of the RIE against the residue rank in the sorted list normalized to one. To enable the averaging of the distributions, we represented each one by 1001 equally distant (on the rank coordinate) points between 0 and 1 (instead of for example N in the case of RIE BB). The RIE for each point was obtained by linear interpolation using the nearest two points of the calculated distribution. The averaged distribution was obtained by averaging the RIEs of the corresponding points of the curves of all of the proteins pertaining to the set. The inverse of the averaged distribution is a quite smooth cumulative distribution function representing the average for the set.
The distributions of the RIEs of a certain type for a particular amino-acid were sampled from all of the 1358 proteins. The RIEs of a particular type and AA were sampled from all the proteins and binned to yield quite smooth histograms.

2.2 Secondary-structure dependence

The RIE distribution of a particular type in a protein describes the distribution of the energetic importance of the residues. An average distribution also characterizes the particular type of interaction in the ensemble – the fraction of the key residues, their importance, and the fraction of the residues with repulsive interactions. The magnitude interval of a distribution is a very important parameter. It contains information about the interaction strength in the native states of the proteins. Unfortunately, this information does not denote the contribution of particular interactions to stability as it lacks information on the denatured state.
The shape of the distribution determines the pressure exerted on a residue and might help estimate the actual contribution of the corresponding interactions to protein stability. It is not surprising that the BB RIEs correlate with the secondary structures as the classifications indirectly use the BB RIEs. However, the differences are smaller than one might expect. It is also clear that none of the interactions other than BB is affected by the secondary-structure content.
From Figure 2, it can be concluded that the difference between the CATH and SCOP classifications is more significant mainly in the case of α proteins. Figures 3 and 4 show all of the types of distributions for a nonpolar (ALA, Figure 3) and a polar (THR, Figure 4)

amino-acid. It is obvious that the BB RIE cumulative distributions are the only distributions to have their shape affected by the secondary-structure content and the particular AA RIE distributions show more than one peak. The distinctive peaks might be assigned to special structural features and their identification remains a task for future studies.

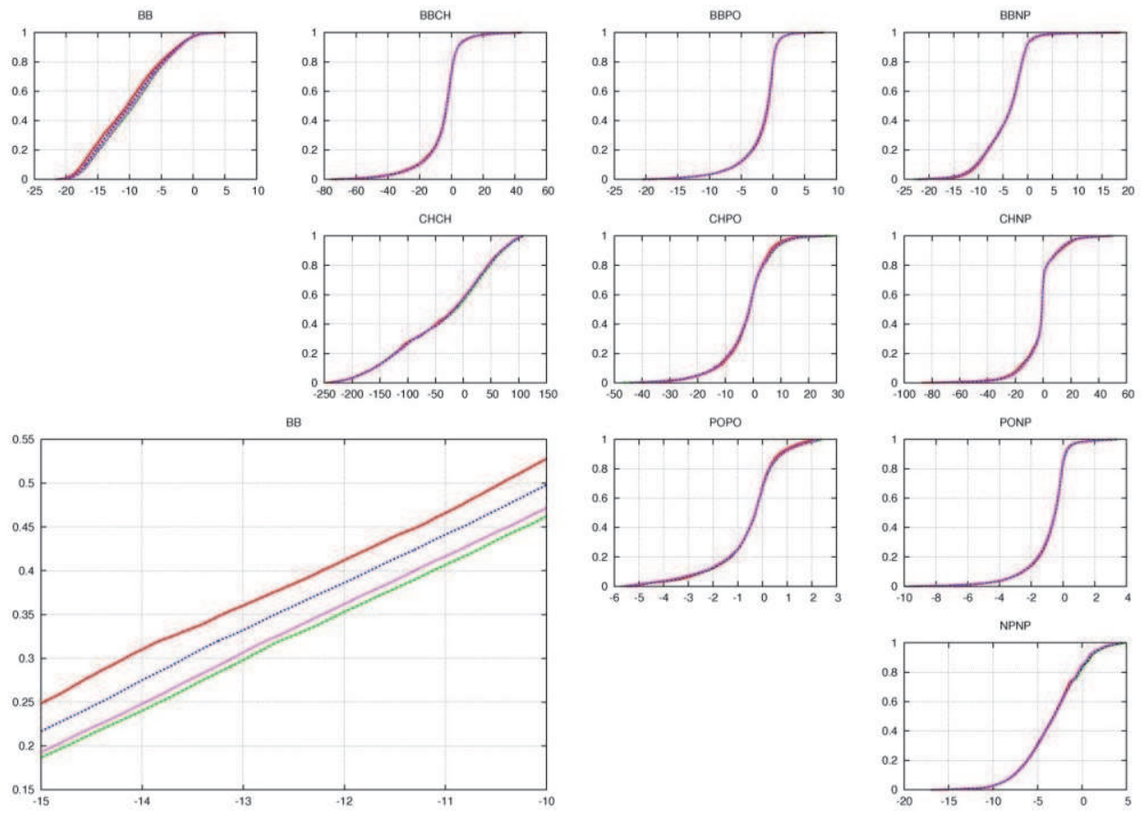


Fig. 2. The average RIE distributions of all ten types: a comparison of the secondary-structure classes. The colors of the lines correspond to the following structure sets: red – CATH α, blue – SCOP α, green – CATH β, magenta – SCOP β. The detail of the BB distribution in the bottom left corner is a zoom of the BB RIE distributions.

The fact that the CYS average NPNP RIE distribution is the only exception to the rule, because it has two peaks, can be explained by a different strength of the noncovalent interactions of the cystein SH group and cystine SS bridge.

The BB RIEs of particular AAs sampled through all of the structure sets are shown in Figure 5. There are remarkable differences between the shapes of the distributions corresponding to the α and β proteins as well as between the shapes of the distributions for particular AAs. Generally, the BB RIE distributions of the beta-structured proteins are shifted to a less attractive (less negative) noncovalent region.

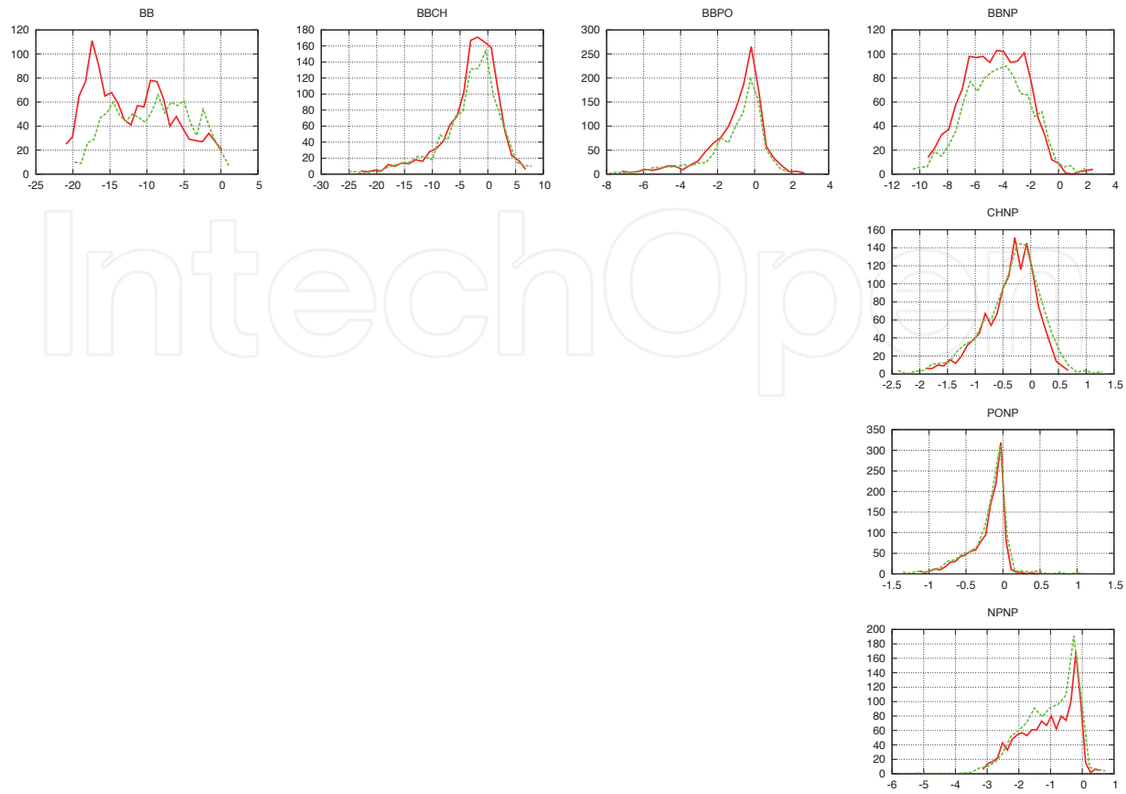


Fig. 3. All of the types of the RIE distributions of ALA. The red line corresponds to the CATH α set, the green line to the CATH β.

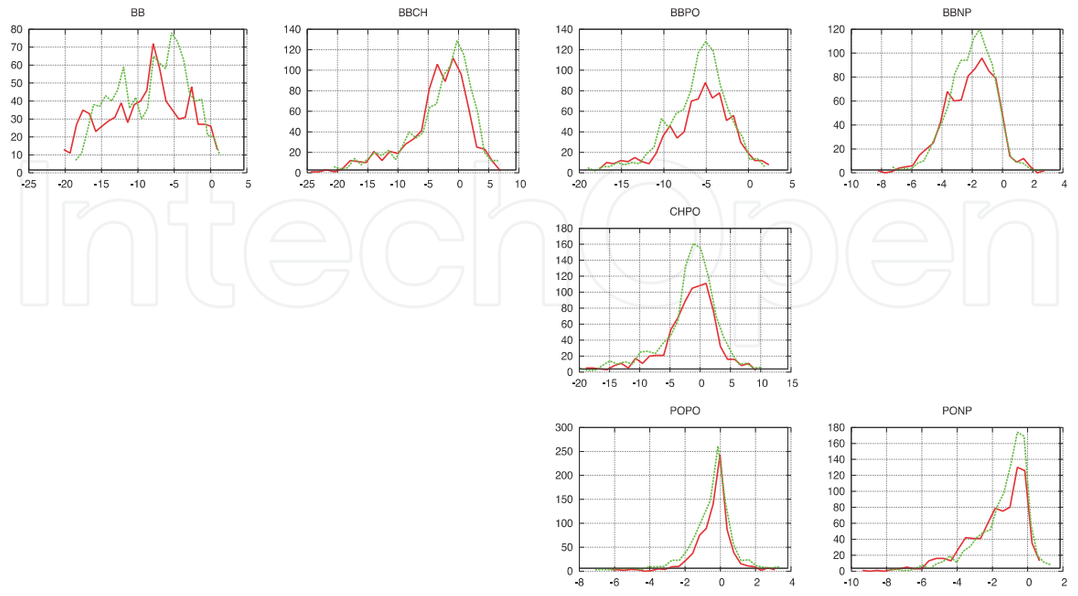


Fig. 4. All of the types of the RIE distributions of THR. The red line corresponds to the CATH α set, the green line to the CATH β.

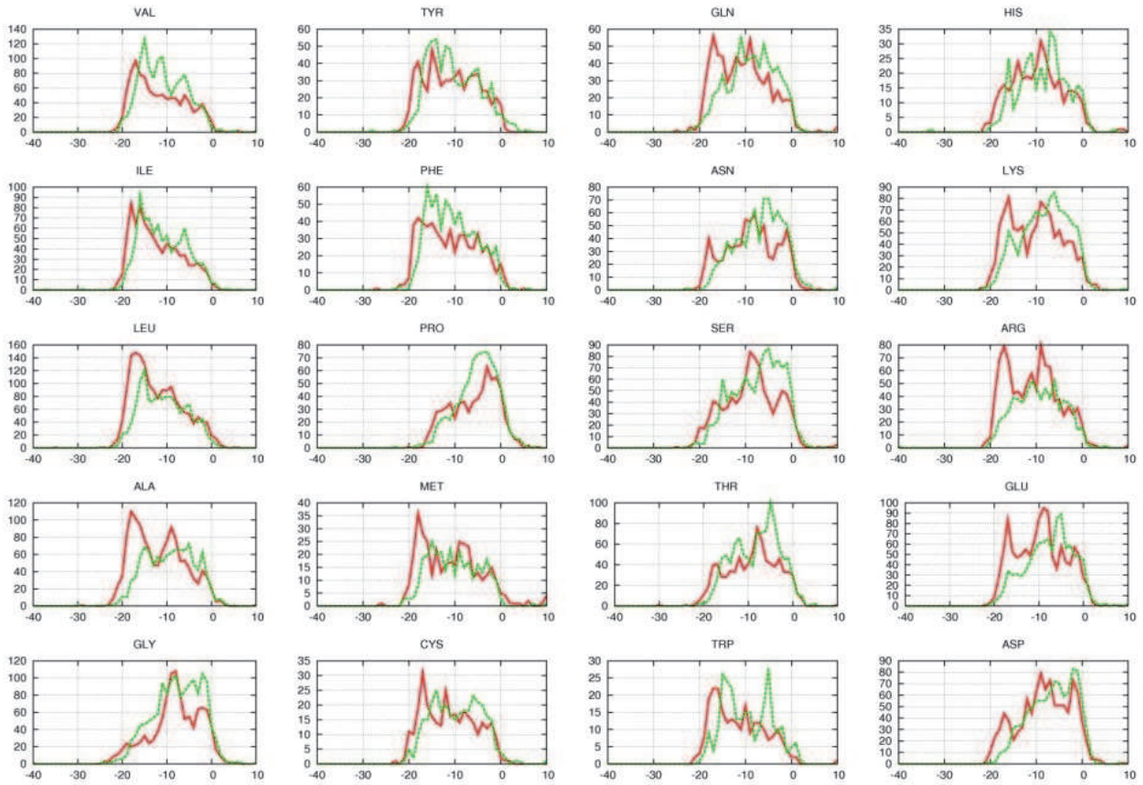


Fig. 5. The average BB RIE distributions for each AA. Sampled through proteins from the CATH α and CATH β sets. The red line corresponds to the CATH α set, the green line to the CATH β.

2.3 Size dependence

The proteins were selected based on their chain lengths up to fourteen groups regardless of their secondary-structure content. Their characteristics (chain-length range, average chain length, amino-acid type composition, number of proteins, number of residues of particular types, average surface area) are reviewed in Table 1.

min length	max length	average length	nonpolar resis	polar resis	charged resis	no. of structures	no. of residues
40	60	52.5	55.1	18.3	26.6	28	521
60	80	68.9	52.1	20.7	27.2	71	1342
80	100	90.7	53.0	18.5	28.5	105	1960
100	120	109.1	52.9	19.8	27.3	132	2492
120	140	129.1	53.9	18.8	27.3	125	2351
140	160	149.7	52.9	19.6	27.6	142	2655
160	180	169.6	53.8	19.5	26.6	85	1587
180	200	188.4	54.3	19.4	26.3	102	1909
200	220	209.9	53.1	19.9	26.9	65	1219
220	240	228.0	53.4	20.2	26.5	67	1253
240	260	249.3	54.3	18.9	26.8	57	1059
260	280	269.1	55.4	20.0	24.6	52	955
280	300	289.4	54.5	19.5	26.0	58	1061

Table 1. The characteristics of the structure sets used for the RIE-size dependence studies.

RIEs of a particular type were sampled from all of the proteins of a particular size group. The RIE averages were calculated separately for each interaction type of each size. The plots of the average RIEs against size are presented in four figures (Figures 6 to 9) in order to maintain the lucidity of the plots with lower magnitudes of average RIEs. The results reported in Figure 6 suggest that the RIE-size dependence varies significantly with the interaction type. On the one hand, the interaction of the polar residues with the backbone is almost independent of size. On the other hand, the interactions of the side-chains follow common rules, which are investigated later.

An interesting notion comes from a comparison of the magnitudes of the POPO and BBPO average RIEs. The lower RIE magnitudes in the case of POPO RIEs are probably caused by the lower probability of hydrogen-bond formation with polar side-chains in comparison with the backbone-polar side-chain because of the lower frequency of their occurrence.

A noticeable trend is the coupling of BBCH and CHPO interactions (see Figure 8). This binding may be ascribed to the same physical quality of these two types of interactions; they both represent charge–dipole interactions. The accuracy of the data can be estimated from the curve smoothness and is apparently lower in the case of charged residues. One possible reason for this trend is that the RIEs of charged residues are the products of a large compensation for the low amount of data.

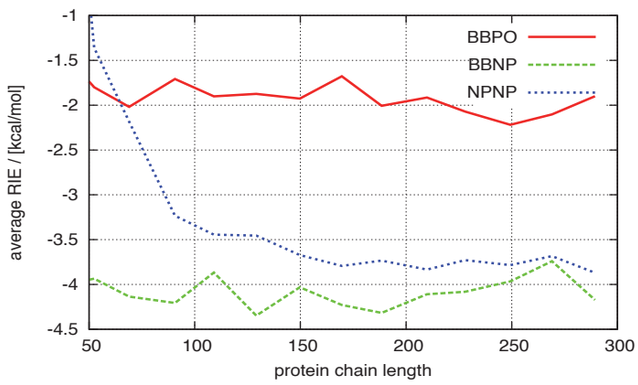


Fig. 6. The size dependence for BBPO, BBNP and NPNP interactions in the studied protein set. The NPNP differs significantly from the rest.

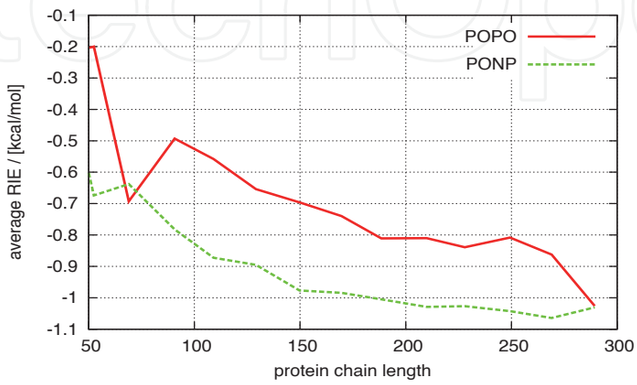


Fig. 7. The Size dependence for the POPO and PONP interactions in the studied protein set.

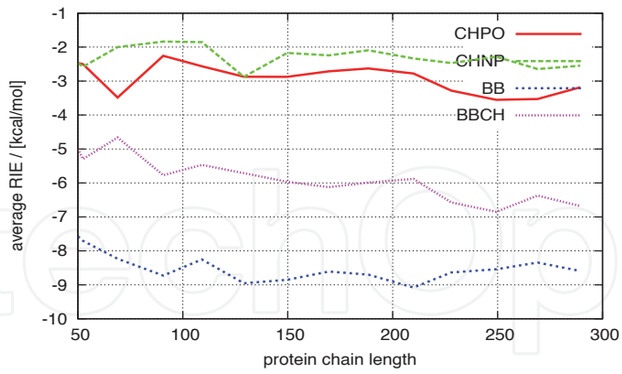


Fig. 8. The size dependence for the CHPO, CHNP, BB and BBCH interactions in the studied protein set.

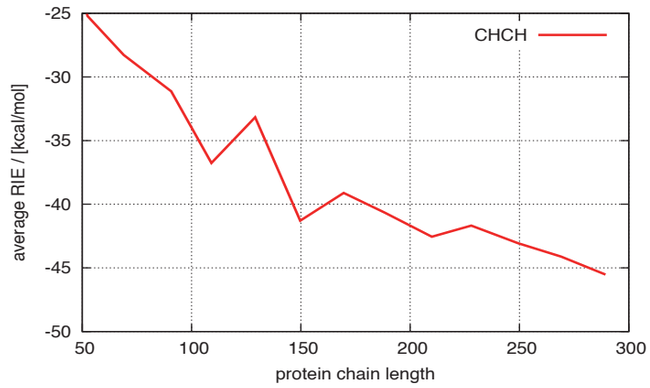


Fig. 9. The size dependence for the CHCH interactions in the studied protein set.

2.3.1 The model of size dependence for the interaction of nonpolar amino-acid side-chains

Two simple models were tested to explain the observed trends. In the first model, protein is assumed to be a sphere with nonpolar residues in its hydrophobic core and polar and charged residues forming its exterior shell. The size dependence of a NPNP RIE average is ascribed to the size dependence of the ratio between the core and surface residues. At infinite length, the NPNP RIE should reach its limit. The second model is more realistic in such a way that the core never behaves like a limitlessly increasing sphere and the volume occupied by the side-chains must reach its limit. This limits the NPNP RIE value which will not rise further with the increasing size of a protein and define a certain size of the most compact amino-acid arrangement.

2.3.1.1 The NPNP RIE Model 1

An average NP residue can be described by its characteristic length r, surface factor f_s (corresponding to its surface or interaction area $S_r = f_s r^2$), volume factor f_v (corresponding to its volume $V_r = f_v r^3$) and the NPNP RIE limit for the infinite bulk E_∞. As we are assuming that all of the nonpolar residues form the core which has a spherical shape, the core size is determined by the size of the protein and the ratio φ of nonpolar residues and all residues. A protein can

be described by its porosity ε (determining the ratio of the gap volume to the volume of the whole protein) and at least its length N. Assuming that all of these quantities except for N are constants, the volume of each protein can be expressed as $V_p = NV_r/(1-\varepsilon) = Nf_v r^3/(1-\varepsilon)$ and the core volume as $V_c = V_P\varphi = N\varphi f_v r^3/(1-\varepsilon)$. The interaction surface of the core residues can be considered as $S_i = N\varphi S_r$ and the core surface is $S_c = 4\pi r_c^2$. E can be calculated as

$$E = E_\infty \left(1 - kN^{-1/3}\right), \tag{1}$$

where

$$k = \frac{1}{f_S}\sqrt[3]{\frac{1024\pi f_v^2}{9\varphi\left(1-\varepsilon\right)^2}}.$$

The k and E_1 parameters were fitted to the calculated data using Equation (1). As can be seen in Figure 10, the fitted curve does not represent the data very well.

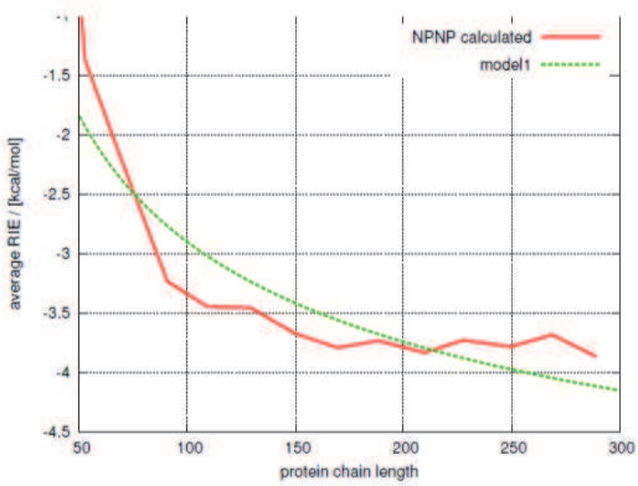


Fig. 10. The performance of Model 1

2.3.1.2 The NPNP RIE Model 2

The first model was extended by adding a new parameter, representing the domain size. The energy was represented by the following function:

$$E = \begin{cases} E = E_\infty \left(1 - kN^{-1/3}\right) & : N \le N_D \\ E_D & : N > N_D \end{cases}, \tag{2}$$

where N_D is the domain size and $E_D = E_\infty (1-kN_D^{-1/3})$ is NPNP RIE average at N_D. The parameters N_D, k and E_1 were fitted to the NPNP RIE averages. The agreement of the fitted curve with the data is satisfactory considering the simplicity of the model as one can see in Figure 11.

The coefficient k obtained by fitting the data is comparable to that obtained by a calculation using the estimated values of fv, fs, ε and the experimental value of φ. Other types of interactions seem to be unrelated to the domain size of a protein as there is no mechanism connected with size that we could follow.

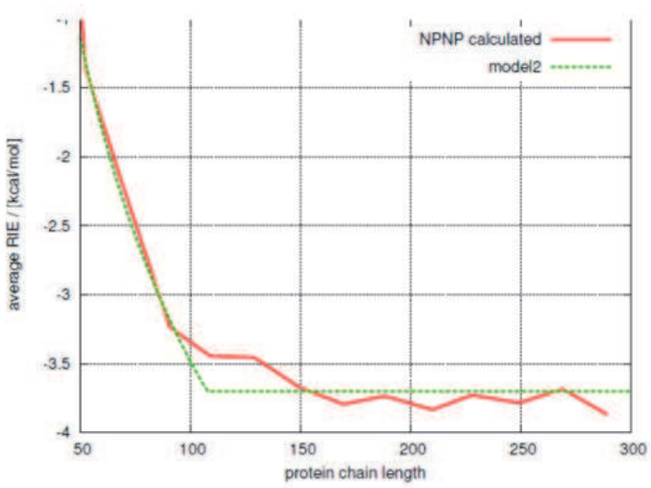


Fig. 11. The performance of Model 2

2.3.2 The reliability of the evaluated distributions

To adjust the reliability of our findings from computational point of view, we divided all of the proteins randomly into two groups. The distributions are indistinguishable, which proves that the distributions can be obtained by averaging even smaller sets of proteins. Additionally, we calculated the distributions using the OPLS force field in a Cα representation of the protein side-chains. Apparently (see Figure 12), the distributions for both FFs are the same. This not only proves that our results are robust against a FF parametrization error but also suggests that both FFs are within their limits equally good for RIE-distribution investigations.

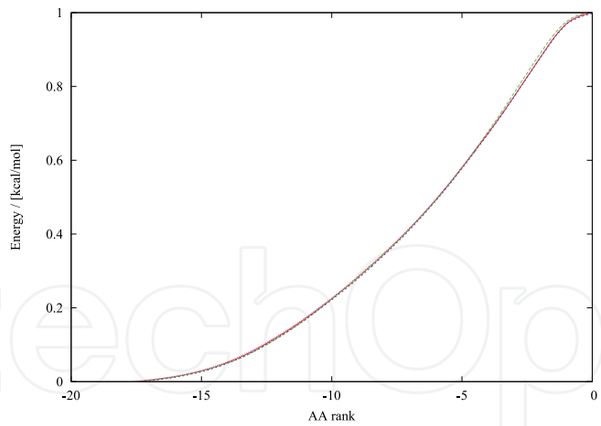


Fig. 12. A Comparison of the distributions obtained by averaging the distributions within the whole set using the OPLS Ca FF (dots) and Amber 03 C_a (full line) shows the robustness of the distributions against the FF used. The distributions obtained by averaging the distributions in two randomly chosen half-sets of structures calculated using the Amber C_a FF are indistinguishable, which proves that our set is sufficiently large.

3. Conclusion

RIE distributions in proteins, except for the BB RIE distributions, are not affected by secondary-structure content. The same applies for the distributions sampled for each amino-

acid separately. Hence, we can claim that the strength and selectivity of the SC-SC and SC-BB interaction do not correlate with the secondary-structure content.
The size dependence of the RIEs can be satisfactorily described by the second model proposed. Its three parameters can be fitted to the results obtained by FF calculations of a high number of protein structures. One of the parameters obtained by fitting to the NPNP RIE averages represents the optimum definition of the domain size in globular proteins. Although the models proposed apply for all types of NP and PO SC-SC interactions, the models fail in the description of the BB and CH interactions. Many interesting facts about the size dependence of the RIE averages were revealed. First, the BBCH and CHPO interactions seem to be bound by some as-yet unknown rule. Second, the PO interactions exhibit "strange" behavior at a protein chain length of approximately seventy residues. These findings need to be investigated more deeply.

4. Acknowledgment

This work was supported by Grant No. P208/10/0725 from the Czech Science Foundation and Grants LH11020 and LC512 from the Ministry of Education, Youth and Sports of the Czech Republic. It was also a part of research projects No. Z40550506 and No. SM6198959216.

5. References

Bendová-Biedermannová, L., Hobza, Pavel, & Vondrásek, J. (2008). Identifying stabilizing key residues in proteins using interresidue interaction energy matrix. *Proteins*, *72*(1), 402-13. doi: 10.1002/prot.21938.

Berka, K., Laskowski, R. a, Hobza, Pavel, & Vondrášek, J. (2010). Energy Matrix of Structurally Important Side-chain/Side-chain Interactions in Proteins. *Journal of Chemical Theory and Computation*, *6*(7), 2191-2203. doi: 10.1021/ct100007y.

Berka, K., Laskowski, R., Riley, K., & Hobza, Pavel. (2009). Representative amino-acid side-chain interactions in proteins. a comparison of highly accurate correlated ab initio quantum chemical and empirical potential. *Journal of Chemical*, *5*(4), 982-992. doi: 10.1021/ct800508v.

D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvai, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2010), *AMBER 11*, University of California, San Francisco.

Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, *29*(31), 7133–7155. ACS Publications.

Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* 2008;4(3):435-47

Holm, L., & Sander, C. (1997). Dali/FSSP classification of three-dimensional protein folds. *Nucleic acids research*, *25*(1), 231-4. Retrieved from

Jorgensen WL, TiradoRives J. The Opls Potential Functions for Proteins - Energy Minimizations for Crystals of Cyclic-Peptides and Crambin. *Journal of the American Chemical Society* 1988;110(6):1657-66

Kolář, M., Berka, K., Jurečka, K., Hobza, P. (2010). Comparative Study of Selected Wave Function and Density Functional Methods for Noncovalent Interaction Energy

Calculations Using the Extended S22 Data Set. *Journal of Chemical Theory and Computation*, *6*, 2365-2376.

Kolář, M., Berka, K., Jurecka, P., & Hobza, Pavel. (2010). On the Reliability of the AMBER Force Field and its Empirical Dispersion Contribution for the Description of Noncovalent Complexes. *Chemphyschem : a European journal of chemical physics and physical chemistry*, *11*(11), 2399-408. doi: 10.1002/cphc.201000109.

Lazaridis, T., Archontis, G., & Karplus, M. (1995). Enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics. *Advances in Protein Chemistry*, *47*, 231-306.

Makhatadze, G. I., & Khechinashvili, N. N., with Venyaminov SYu & Griko YuV. (1989). Heat capacity and conformation of proteins in the denatured state. *Journal of molecular biology*, *205*(4), 737-50..

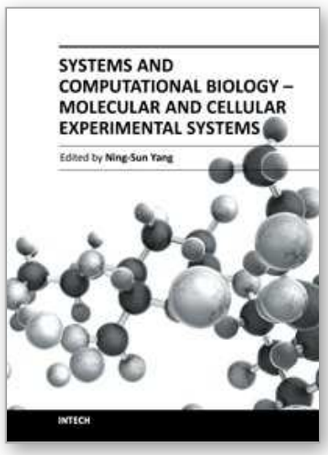
Makhatadze, G., & Privalov, P. (1995). Energetics of protein structure. *Advances in Protein Chemistry*, *47*, 307-425.

Müller-Dethlefs, K., & Hobza, P. (2000). Noncovalent interactions: a challenge for experiment and theory. *Chemical Reviews*, *100*(1), 143-167.

Murzin, A.G., Brenner, S.E., Hubbard, T., Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology*, *247*, 536-540..

Orengo, C. a, Michie, a D., Jones, S., Jones, D. T., Swindells, M. B., & Thornton, J. M. (1997). CATH--a hierarchic classification of protein domain structures. *Structure (London, England : 1993)*, *5*(8), 1093-108.

Riley, K. E., Pitoňák, M., Jurečka, P., & Hobza, Pavel. (2010). Stabilization and structure calculations for noncovalent interactions in extended molecular systems based on wave function and density functional theories. *Chemical Reviews*, *110*(9), 5023-63. doi: 10.1021/cr1000173.



Systems and Computational Biology - Molecular and Cellular Experimental Systems

Edited by Prof. Ning-Sun Yang

ISBN 978-953-307-280-7
Hard cover, 332 pages
Publisher InTech
Published online 15, September, 2011
Published in print edition September, 2011

Whereas some â€œmicroarrayâ€ or â€œbioinformaticsâ€ scientists among us may have been criticized as doing â€œcataloging researchâ€, the majority of us believe that we are sincerely exploring new scientific and technological systems to benefit human health, human food and animal feed production, and environmental protections. Indeed, we are humbled by the complexity, extent and beauty of cross-talks in various biological systems; on the other hand, we are becoming more educated and are able to start addressing honestly and skillfully the various important issues concerning translational medicine, global agriculture, and the environment. The two volumes of this book presents a series of high-quality research or review articles in a timely fashion to this emerging research field of our scientific community.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Boris Fackovec and Jiri Vondrasek (2011). Decomposition of Intramolecular Interactions Between Amino-Acids in Globular Proteins - A Consequence for Structural Classes of Proteins and Methods of Their Classification, Systems and Computational Biology - Molecular and Cellular Experimental Systems, Prof. Ning-Sun Yang (Ed.), ISBN: 978-953-307-280-7, InTech, Available from: http://www.intechopen.com/books/systems-and-computational-biology-molecular-and-cellular-experimental-systems/decomposition-of-intramolecular-interactions-between-amino-acids-in-globular-proteins-a-consequence-

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen