DEVELOPMENT OF A NEW META-SCORE FOR PROTEIN STRUCTURE PREDICTION FROM SEVEN ALL-ATOM DISTANCE DEPENDENT POTENTIALS USING SUPPORT VECTOR REGRESSION

MATSUYUKI SHIROTA[1,2] mshirota@hgc.jp
TAKASHI ISHIDA[1] t-ishida@hgc.jp
KENGO KINOSHITA[1,2] kinosita@hgc.jp

[1] *Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan*
[2] *Institute for Bioinformatics Research and Development, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

An accurate scoring function is required for protein structure prediction. The scoring function should distinguish the native structure among model structures (decoys) and it also should have correlation with the quality of the decoys. However, we had observed the trade-off between the two requirements for seven all-atom distance dependent potentials in the previous study, where the native structure could be discriminated by examining the fine atomic details, whereas the correlation could be improved by examining coarse-grained interactions. To overcome this problem, in this study, we tried to make an improved scoring function by combining the seven potentials. First, the seven potentials were normalized by the expected energy values of the native and reference states of the target protein. Second, the relationship between the seven normalized energies and the quality (GDT_TS) of the structure were learned using support vector regression with the decoy sets of CASP6 as the training set. Then the meta-score was obtained as the predicted GDT_TS and it was tested with the decoys of the CASP7 experiment. The meta-score showed improvement in correlations with the GDT_TS and in the Z-score of the native structure. It also showed comparable performances in the ΔGDT and enrichment criteria, with the best component potentials. The meta-score could be also used as the absolute quality of the structures. Our study suggests the benefit of combining several different scoring functions for model evaluation.

Keywords: protein structure prediction; statistical potential; support vector regression; quality assessment

1. Introduction

Predicting the native structure of a protein molecule from its amino acid sequence is still a very challenging task in structural bioinformatics. The process of structure prediction consists of the following two steps; (i) enumerating a lot of possible model structures, or

decoys, of the target protein and (ii) selecting the most likely structure among them. For the success of the second process, an accurate scoring function of protein structures is now required. There have been a large number of attempts for the development of the evaluation function of a protein structure. A large part of the studies relied on the analysis of known protein structures in the Protein Data Bank [2] to find the most appropriate structure of the target protein. They are called knowledge-based potentials or statistical potentials [8, 13, 17, 23, 26]. Statistical potentials are constructed by counting the occurrence of protein conformations in the PDB, comparing the observed frequency with the expected one and calculating the energy of the conformation by the log-ratio of the observed and expected occurrences by assuming of the Boltzmann distribution. Various structural elements can be used for statistical potentials, such as burial and exposure of residues [7, 19], secondary structures [3], amino acid or atomic contacts [9, 25], distances between amino acids or atoms [8, 12-14, 18, 26]. In spite of many efforts that have been made, a community-wide experiment of protein structure prediction (CASP) showed that it is still difficult to pick up a few suitable models from the sampled structures [10].

There will be two requirements for a scoring function of model structures. One is to evaluate the native structure as the lowest energy among the non-native structures. This is expected from the thermodynamic hypothesis of protein folding which states that the native state of a protein molecule is the one with the lowest free energy of the system [1]. The other requirement is the correlation between the energy and the quality of the structures, which originates from the funnel-like potential energy landscape of the protein molecule [20].

Although it is desirable to fulfill these two requirements, it may be difficult to optimize the performances of native discrimination and correlation simultaneously. This is because the native structure can be discriminated by examining the fine atomic details such as Van der Waals interactions, whereas the near-native decoys and bad ones can be separated by examining coarse-grained interactions such as hydrophobicity.

We had observed that there was no single potential that could work best in these two performances among seven all-atom distance dependent statistical potentials with different reference states [15]. To overcome the trade-off, in this study, we tried to make an improved scoring function that satisfied these two requirements. The new scoring function, the meta-score, was developed as a predictor of the structural similarity to the native structure (GDT_TS [22]) by combining the energy values of the seven potentials with a support vector regression (SVR). The SVR was trained with the CASP6 decoy sets and tested with the CASP7 decoy sets.

2. Materials and Methods

2.1. Decoy Sets

The performance of the potentials was evaluated using the decoy sets of the 6th and 7th CASP (Critical Assessment on the techniques for protein Structure Prediction) experiments [10, 11]. All of the native and model structures were obtained via the CASP

web site (http://predictioncenter.org/). We considered protein domains as the unit of folding and employed the definitions of the CASP organizers. The CASP6 and CASP7 decoy sets included 90 and 122 domains, respectively. The model structures with any missing atoms, for which the corresponding atoms were observed in the native structure, were removed from the decoy set.

2.2. Quality of the Structure

We defined the quality of a model structure by its structural similarity to the native structure, measured by the GDT_TS [22]. GDT_TS is the average of the fraction of matched residues between the native and model structures within various distance thresholds after global structural superimposition. This score was scaled within [0,1], so that it approached 1 for good decoys and 0 for bad ones. The GDT_TSs of all the native structures are 1. We used TM-score program [24] to calculate the GDT_TS.

2.3. Component Statistical Potentials

All the seven all-atom distance dependent potentials have been described previously, and they were classified according to the hydrophobic and attractive characteristics [15]. A hydrophobic potential evaluates the clustering of atoms of hydrophobic residues favorably, whereas an attractive potential prefers compact globular structures. The seven potentials were called RAPDF [13], AKBP [8], DFIRE [26], SHELL, BALL, FAR and DS [5]. Briefly, RAPDF is a potential without hydrophobic or attractive characteristics. SHELL, AKBP and FAR are the potentials in which hydrophobic characteristics are added to RAPDF. BALL is a potential which consist of RAPDF + attractive characteristics and DFIRE has both the hydrophobic and attractive characteristics added to RAPDF. DS has a similar characteristic to RAPDF. These characteristics of potentials were shown to influence on their performance. The hydrophobic characteristic improved the correlation with the GDT_TS, but decreased the ability to discriminate the native structures. An attractive potential worked well for decoys relatively similar to the native structure, but poor for decoys whose similarities to the native structures were low.

The seven potentials were calculated from the statistics of the non-redundant set of 1,788 SCOP domains. We considered 167 atom types, by treating all of the non-hydrogen atoms to have different types when they are in different amino acid residues [13]. We used the distance intervals described by Zhou and Zhou [26] to count the frequency at each distance r. The width of the bin was 2 Å for r <2 Å, 0.5 Å for 2 < r < 8 Å, and 1 Å for 8 < r < 15 Å. The atom pairs with more than 15Å separation were not considered in the calculation of potentials.

Using these potentials, the raw energy of the model structure was calculated by

$$E^{\text{model}} = \sum_{ab} \sum_{r} n_{ab}^{\text{model}} \times E_{ab}(r)$$

, where n_{ab}^{model} is the occurrence of atom pair *ab* at a distance *r* in the model structure. $E_{ab}(r)$ is the atom pair interaction energy of the statistical potentials. The atom pair type *ab* takes 167 x 167 types and the distance *r* takes 3.5 Å < *r* < 15 Å. For AKBP and FAR potentials which have the strongest hydrophobic characteristics, the distance of energy evaluation was shortened to 3.5 Å < *r* < 6.5 Å to reduce the hydrophobicity [15] and they were called AKBP' and FAR', respectively.

2.4. Normalization of the All-Atom Distance Dependent Potentials

The potential energies of the model structures calculated above could be compared with each other, only if all the structures had the same amino acid sequence, because the numbers and the compositions of the atom pairs in a structure vary with its sequence. This fact causes some problems in the regression of the quality of structures from the energies, because different models for various targets with similar GDT_TSs can have quite diverse raw potential values, which are not suitable for regression. The relationship between the energies and the GDT_TS would be different for each target. Therefore, the energy values should be normalized according to the amino acid sequence for regression.

For this purpose, each of the potentials was normalized as,

$$E^{\text{normalized}} = \left(E^{\text{model}} - \widetilde{E}^{\text{reference}}\right) / \left(\widetilde{E}^{\text{native}} - \widetilde{E}^{\text{reference}}\right)$$

, where $\widetilde{E}^{\text{native}}$ and $\widetilde{E}^{\text{reference}}$ are the estimated potential values for the native and reference state structures, respectively, of the same protein. The normalized energies are expected to have a constant relationship to the quality of the structure for different proteins. The details about the normalization will be described elsewhere.

2.5. Development of the Meta-Score

We developed a new meta-score of protein structures by combining the information from seven normalized energies of the all-atom distance dependent statistical potentials. This meta-score can be regarded as a predicted GDT-TS of a structure estimated by a regression from the seven normalized energies. The underlying assumption for this attempt is that the relationship between the normalized energies and the quality (GDT_TS) would be universal for proteins of various amino acid sequences. This assumption can be reasonable from the viewpoint that both the normalized energies and the GDT_TS indicate the energetic and structural distance to the native structure, respectively, between 0 and 1, regardless of amino acid sequence.

For the purpose, we employed the support vector regression (SVR) algorithm, which was implemented using the libsvm package [4] with the regression mode. SVR can solve nonlinear regression problems using the kernel functions. In this study, the Gaussian radial basis function (RBF) was selected as the kernel function. We adjusted two parameters, the width γ of the Gaussian function and the regularization parameter *C*. γ controls how peaked Gaussian functions were and *C* controls the cost for the regression errors. These parameters were optimized by the leave-one-out jackknife method using

CASP6 decoy sets. The meta-scores were trained with various width γ and the regularization parameter C using all but one decoy sets of CASP6, and the resulting meta-scores were subjected to evaluate the model structures of the remaining decoy set. Then the correlation coefficients between the meta-score and the GDT_TS were calculated. All the 90 decoy sets were evaluated in this manner and the parameter set, γ and C, that yielded the highest averaged correlation coefficient was selected. After this parameter optimization, a meta-score was trained using all of the 90 CASP6 decoy sets of 17,184 structures with the optimized parameters. The meta-score was tested by the decoys of CASP7.

2.6. Assessment of Potentials

Five criteria were used for the assessment of the potentials or the meta-score (in the following, the potential values and the meta-scores are simply referred to as scores) as in the previous study [15]: (1) PCC, the Pearson's correlation coefficient between the score and the GDT_TS of the decoys; (2) Z-score, the difference between the native score and the average of the decoy scores in units of their standard deviation; (3) ΔGDT, the difference between the GDT_TS of the best decoy and that of the best-scored decoy; and (4, 5), 10% and 20% enrichments; x% enrichment is the relative frequency of the most accurate x% models in the GDT_TS among the x% best scored models and x%.

2.7. Statistical Significance of the Difference in Performance

These five assessment measures were calculated for each target domain of CASP7, using each of the scores, and they were averaged for 104 template-based modeling (TBM) targets and 18 free modeling (FM) targets. The statistical significance of the difference in performance between the scores for the TBM and FM categories was assessed by the Wilcoxon signed rank test [16] with a threshold of $p<0.01$. More precisely, for a single decoy set k, the difference in performance between two scores X and Y was determined by using a bootstrap approach, as follows. Half of the decoys were sampled randomly 100 times, and the performances were calculated for both X and Y. Then, we define the difference in performance for the decoy set k to be $d_k(X,Y) = (m_k(X) - m_k(Y))/(\sigma_k(X) + \sigma_k(Y))$, where $m_k(X)$ and $\sigma_k(X)$ are the mean and the standard deviation of the performance of X. Finally, the distribution $d_k(X,Y)$ ($k=1,\ldots, n$, where n is the number of decoy sets) was evaluated by the Wilcoxon signed rank test.

3. Results and Discussion

3.1. Performances for the Training Set

Table 1 shows the assessment results for the seven all-atom distance dependent statistical potentials and the meta-score averaged over 90 CASP6 decoy sets. Among the seven

component potentials, DFIRE showed the largest PCC, whereas RAPDF showed the largest Z-score and BALL showed the best ΔGDT and the largest 10% and 20% enrichment values.

These results of the seven potentials for the training set were basically the same as the results reported in the previous article [15]. A non-hydrophobic potential, RAPDF, has the highest Z-score of the native structure, but smaller PCC with the GDT_TS than the hydrophobic potentials, AKBP', SHELL, FAR'. The attractive characteristics are usually effective as observed for DFIRE and BALL. DS has a similar characteristic with, but a little worse performance than RAPDF. In short, seven potentials had their own advantages and disadvantages in model evaluation and none of them outperformed others in all the assessment criteria.

On the other hand, the assessments of the meta-score with the training CASP6 decoy sets showed improvements overall the seven potentials in both PCC and Z-score. The improvements were also observed in ΔGDT and enrichment criteria. Thus, the meta-score has combined the strong points of the component potentials, although these results are for the training sets.

Table 1. Results of the assessment of the seven normalized energies (columns 2-8) and the meta-score (column 9) with the CASP6 decoy sets. The meta-score was calculated using the Jackknife-method as described in the text. The rows 2-6 correspond to the five assessment criteria. PCC: Pearson's correlation coefficient between score and the GDT_TS, Z-score: the deviation of the native energy from the mean energy of the decoys in units of standard deviation, ΔGDT: the difference between the best GDT_TS among the decoys and the GDT_TS of the best-scored decoy, Enr 10%, 20%: 10% and 20% enrichments respectively. These assessment values were calculated for all of the 90 decoy sets of CASP6 and the averages were listed.

	RAPDF	AKBP'	DFIRE	SHELL	BALL	FAR'	DS	Meta
PCC	0.38	0.41	0.56	0.47	0.53	0.43	0.37	0.57
Z	3.02	2.77	2.47	2.88	2.93	2.74	2.79	3.34
ΔGDT	0.13	0.12	0.15	0.11	0.11	0.12	0.14	0.10
Enr10	2.53	2.64	3.00	2.73	3.11	2.65	2.36	3.11
Enr20	1.92	2.03	2.24	2.09	2.25	2.06	1.82	2.30

3.2. Performances for the Test Set

Table 2 shows the assessment results with the test decoy sets of CASP7. These decoy sets were classified as either a) TBM targets or b) FM targets.

For both categories, the meta-score showed higher correlation with GDT_TS and better Z-score than the seven potentials. The meta-score displayed significantly better PCC values than all potentials but DFIRE for TMB targets and SHELL for FM targets. Significant differences in Z-score values were observed between the meta-score and the potentials except for RAPDF for TBM modeling targets and RAPDF, AKBP' and SHELL for FM targets by the same test. These results indicated that the meta-score could satisfy the two requirements, the correlation and native discrimination, and overcame the trade-off of the seven component potentials.

The meta-score demonstrated the smallest ΔGDT for TBM targets. DFIRE and BALL showed similar performances with the meta-score, indicating that the attractive characteristics played an important role in choosing the best model from decoys. DFIRE slightly outperformed the meta-score in ΔGDT for FM targets, although the difference was not significant due to small sample size (18 targets).

In 10% and 20% enrichment criteria, the meta-score was not the best score either for TBM targets or for FM targets. The results of enrichment assessment revealed the difference between the TBM and FM decoy sets. The two attractive potentials, DFIRE and BALL, showed good performances for TBM targets but they performed worse than the non-attractive five potentials for FM targets. The best result for FM targets was achieved by a hydrophobic potential SHELL. Despite these differences, the meta-score showed comparable performance to the best potentials for both categories (BALL for TBM and SHELL for FM).

The performance of these scores was also compared with the previously reported scoring functions, DOPE [14] and Verify3D [3]. DOPE is another type of all-atom distance dependent statistical potential based on a reference state corresponding to non-interacting atoms in a homogeneous sphere, with the radius dependent on a sample native structure. Verify3D assesses the compatibility of each residue with its environment. The performances of these two potentials were comparable to the seven component potentials in the five assessment criteria for both TBM targets and FM targets. The meta-score significantly outperformed these two scores in all five criteria for TBM targets. The meta-score also outperformed these two scores for FM targets although the significance was not clear because of the small number of targets.

Table 2. Assessments with the CASP7 decoy sets (test set) averaged over a) 104 TBM targets and b) 18 FM targets. The columns represent the seven potentials (columns 2-8), the meta-score (column 9), DOPE [14] (column 10) and Verify3D [3] (column 11). The rows 2-6 correspond to the five assessment criteria. The results for the seven potentials were taken from the Table III of our previous paper [15] except for AKBP' and FAR'. An asterisk represents the significant difference between the meta-score and the score in the assessment of the same row. The highest performance column in each row is underlined.

a) TBM targets

	RAPDF	AKBP'	DFIRE	SHELL	BALL	FAR'	DS	Meta	DOPE	V3D
PCC	0.46*	0.48*	0.60	0.52*	0.58*	0.49*	0.44*	0.60	0.43*	0.48*
Z	2.83	2.74*	2.32*	2.64*	2.81*	2.72*	2.64*	3.11	2.53*	2.34*
ΔGDT	0.08*	0.07*	0.07	0.07*	0.05	0.07*	0.09*	0.05	0.06*	0.09*
Enr10	3.03*	2.85*	3.70	3.14*	3.81	2.88*	2.93*	3.70	2.76*	2.64*
Enr20	2.21*	2.21*	2.52	2.31*	2.54	2.24*	2.16*	2.53	2.10*	2.07*

b) FM targets

	RAPDF	AKBP'	DFIRE	SHELL	BALL	FAR'	DS	Meta	DOPE	V3D
PCC	0.47*	0.55*	0.55*	0.58	0.54*	0.56*	0.43*	0.61	0.44*	0.54*
Z	2.28	2.16	2.03*	2.25	2.47*	2.14*	2.13*	2.78	2.36	2.38*
ΔGDT	0.10	0.10	0.09	0.10	0.10	0.10	0.11	0.10	0.10	0.10
Enr10	3.21	3.64	2.97	3.77	2.84*	3.62	3.11	3.55	2.75*	3.14
Enr20	2.22	2.40	2.16	2.49	2.15*	2.40	2.04*	2.44	2.09*	2.23

3.3. Evaluation of the Meta-Score as an Absolute Quality Score for Protein Structures

Since a meta-score was constructed as a predictor of the GDT-TS of each structure, it can also be regarded as the absolute quality score. We checked the ability of the meta-score to approximate the true GDT_TS.

Figure 1 shows the scatter plot of the predicted and true GDT_TS of all the native and decoy structures for 122 targets, totaling 33,758 structures. The overall correlation coefficient and the root mean squared error (RMSE) between the meta-score and the true GDT-TS were 0.84 and 0.142, respectively.

There have been some reports on estimating the absolute quality of models. Wan et al [21] developed a method to evaluate the absolute quality of a model structure as a predicted GDT_TS value using the identities between the observed and predicted structural features of various kinds and support vector regression. They reported the overall correlation coefficient and the RMSE for the CASP7 decoys 0.87 and 0.123, respectively. Eramian et al [6] reported a prediction of NO3.5Å score (the fraction of Cα atoms in a model that are within 3.5Å of the corresponding atom in the native structure after superposition) of their 580,317 comparative modeling structures using support vector regression which utilized various scoring functions. They reported that the PCC between the observed and predicted NO3.5Å was 0.84. Although the direct comparison to these reports was not appropriate due to the different model structures used in these studies (the decoy set of Wan et al. was not completely the same as that of ours because their method evaluated the whole chains instead of the domains), our meta-score was almost comparable to these results.

We created a predictor of the absolute quality of decoys from the component potentials. The normalization of the component potentials was the key to absolute quality assessment. Employing other statistical potentials in their normalized form would help improve the performance for model evaluation and increase the accuracy of the prediction of the GDT_TS.

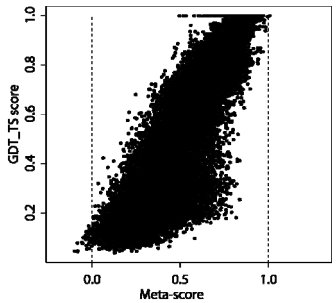

Figure 1. The meta-score and the GDT_TS score of the CASP7 native and model structures.

4. Conclusion

We have developed a new meta-score which combined the normalized energies of the seven all-atom distance dependent statistical potentials with support vector regression. The meta-score performed better than or comparable to the best of the component potentials in five assessment criteria. Therefore, the meta-score could combine the advantages of the component potentials. In addition, it could be used as the absolute quality score of the model structures, which had comparable accuracy with the previous results.

Acknowledgments

This work was partially supported by a Grant-in Aid for Scientific Research on the Priority Area "Transportsome" from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by a Grant-in-aid from the Institute for Bioinformatics Research and Development, Japan Science and Technology Corporation (BIRD-JST) to KK. Computation time was provided by the Super Computer System, Human Genome Center, Institute of Medical Science, The University of Tokyo.

References

[1] Anfinsen, C. B., Principles that govern the folding of protein chains, *Science*, 181(96):223-30, 1973

[2] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E., The Protein Data Bank, *Nucleic Acids Res*, 28(1):235-42, 2000

[3] Bowie, J. U., Luthy, R. and Eisenberg, D., A method to identify protein sequences that fold into a known three-dimensional structure, *Science*, 253(5016):164-70, 1991

[4] Chang, C.-C. and Lin, C.-J., LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2001

[5] DeBolt, S. E. and Skolnick, J., Evaluation of atomic level mean force potentials via inverse folding and inverse refinement of protein structures: atomic burial position and pairwise non-bonded interactions, *Protein Eng*, 9(8):637-55, 1996

[6] Eramian, D., Eswar, N., Shen, M. Y. and Sali, A., How well can the accuracy of comparative protein structure models be predicted?, *Protein Sci*, 17(11):1881-93, 2008

[7] Jones, D. T., Taylor, W. R. and Thornton, J. M., A new approach to protein fold recognition, *Nature*, 358(6381):86-9, 1992

[8] Lu, H. and Skolnick, J., A distance-dependent atomic knowledge-based potential for improved protein structure selection, *Proteins*, 44(3):223-32, 2001

[9] Miyazawa, S. and Jernigan, R. L., Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. , *Macromolecules*, 18:534-552, 1985

[10] Moult, J., Fidelis, K., Kryshtafovych, A., Rost, B., Hubbard, T. and Tramontano, A., Critical assessment of methods of protein structure prediction-Round VII, *Proteins*, 69 Suppl 8:3-9, 2007
[11] Moult, J., Fidelis, K., Rost, B., Hubbard, T. and Tramontano, A., Critical assessment of methods of protein structure prediction (CASP)--round 6, *Proteins*, 61 Suppl 7:3-7, 2005
[12] Rykunov, D. and Fiser, A., Effects of amino acid composition, finite size of proteins, and sparse statistics on distance-dependent statistical pair potentials, *Proteins*, 67(3):559-68, 2007
[13] Samudrala, R. and Moult, J., An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction, *J Mol Biol*, 275(5):895-916, 1998
[14] Shen, M. Y. and Sali, A., Statistical potential for assessment and prediction of protein structures, *Protein Sci*, 15(11):2507-24, 2006
[15] Shirota, M., Ishida, T. and Kinosita, K., Analyses on hydrophobicity and attractiveness of all-atom-distance-dependent potentials, *Protein Sci*, 18(9):1906-1915, 2009
[16] Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, 1956
[17] Simons, K. T., Kooperberg, C., Huang, E. and Baker, D., Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions, *J Mol Biol*, 268(1):209-25, 1997
[18] Sippl, M. J., Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins, *J Mol Biol*, 213(4):859-83, 1990
[19] Sippl, M. J., Recognition of errors in three-dimensional structures of proteins, *Proteins*, 17(4):355-62, 1993
[20] Skolnick, J., In quest of an empirical potential for protein structure prediction, *Curr Opin Struct Biol*, 16(2):166-71, 2006
[21] Wang, Z., Tegge, A. N. and Cheng, J., Evaluating the absolute quality of a single protein model using structural features and support vector machines, *Proteins*, 75(3):638-47, 2009
[22] Zemla, A., LGA: A method for finding 3D similarities in protein structures, *Nucleic Acids Res*, 31(13):3370-4, 2003
[23] Zhang, Y., Kolinski, A. and Skolnick, J., TOUCHSTONE II: a new approach to ab initio protein structure prediction, *Biophys J*, 85(2):1145-64, 2003
[24] Zhang, Y. and Skolnick, J., Scoring function for automated assessment of protein structure template quality, *Proteins*, 57(4):702-10, 2004
[25] Zhang, C., Vasmatzis, G., Cornette, J. L. and DeLisi, C., Determination of atomic desolvation energies from the structures of crystallized proteins, *J Mol Biol*, 267(3):707-26, 1997
[26] Zhou, H. and 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):2714-26, 2002