JOURNAL OF COMPUTATIONAL CHEMISTRY

Predicting Backbone Cα Angles and Dihedrals from Protein Sequences by Stacked Sparse Auto-Encoder Deep Neural Network

James Lyons,[a] Abdollah Dehzangi,[a,b] Rhys Heffernan,[a] Alok Sharma,[a,c] Kuldip Paliwal,[a] Abdul Sattar,[a,b] Yaoqi Zhou,*[d] and Yuedong Yang*[d]

Because a nearly constant distance between two neighbouring Cα atoms, local backbone structure of proteins can be represented accurately by the angle between $C\alpha_{i-1}$—$C\alpha_i$—$C\alpha_{i+1}$ (θ) and a dihedral angle rotated about the $C\alpha_i$—$C\alpha_{i+1}$ bond (τ). θ and τ angles, as the representative of structural properties of three to four amino-acid residues, offer a description of backbone conformations that is complementary to φ and ψ angles (single residue) and secondary structures (>3 residues). Here, we report the first machine-learning technique for sequence-based prediction of θ and τ angles. Predicted angles based on an independent test have a mean absolute error of 9° for θ and 34° for τ with a distribution on the θ-τ plane close to that of native values. The average root-mean-square distance of 10-residue fragment structures constructed from predicted θ and τ angles is only 1.9Å from their corresponding native structures. Predicted θ and τ angles are expected to be complementary to predicted ϕ and ψ angles and secondary structures for using in model validation and template-based as well as template-free structure prediction. The deep neural network learning technique is available as an on-line server called Structural Property prediction with Integrated DEep neuRal network (SPIDER) at http://sparks-lab.org. © 2014 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23718

Introduction

Template-based and template-free protein-structure prediction relies strongly on prediction of local backbone structures.[1,2] Protein local structure prediction is dominated by secondary structure prediction with its accuracy stagnant around 80% for more than a decade.[3,4] However, secondary structures are only a coarse-grained description of protein local structures in three states (helices, sheets, and coils) that are somewhat arbitrarily defined because helices and sheets are often not in their ideal shapes in protein structures. This arbitrariness has limited the theoretically achievable accuracy of three-state prediction to 88–90%.[4,5] Moreover, predicted coil residues do not have a well-defined structure.

An alternative approach to characterize the local backbone structure of a protein is to use three dihedral or rotational angles about the N—Cα bond (φ), the Cα—C bond (ψ), and the C—N bond (ω). A schematic illustration is shown in Figure 1. Because ω angles are restricted to 180° (the majority) or 0° due to rigid planar peptide bonds, two dihedral angles (ϕ and ψ) essentially determine the overall backbone structure. Unlike secondary structures, these dihedral angles (ϕ and ψ) can be predicted as continuous variables and their predicted accuracy has been improved over the years[6–8] so that it is closer to dihedral angles estimated according to NMR chemical shifts.[9] Predicted backbone dihedral angles were found to be more useful than predicted secondary structure as restrains for *ab initio* structure prediction.[9,10] It has also been utilized for improving sequence alignment,[11] secondary structure prediction,[3,12,13] and template-based structure prediction and fold recognition.[14–16] However, unlike the secondary structure of proteins, φ and ψ are limited to the conformation of a single residue.

Two different angles can also be used for representing protein backbones. As shown in Figure 1, they are the angle between $C\alpha_{i-1}$—$C\alpha_i$—$C\alpha_{i+1}$ (θ_i) and a dihedral angle rotated about the $C\alpha_i$—$C\alpha_{i+1}$ bond (τ_i). This two-angle representation is possible because neighbouring Cα atoms mostly have a fixed distance (3.8Å) due to the fixed plane in $C\alpha_{i-1}$—C—N—$C\alpha_i$. These two inter-residue angles (θ and τ) reflect the conformation of four connected, neighbouring residues that is longer than a single-residue conformation represented by φ and ψ angles. By comparison, a conformation represented by helical or sheet residues involves in an

[a] J. Lyons, A. Dehzangi, R. Heffernan, A. Sharma, K. Paliwal, A. Sattar
Institute for Integrated and Intelligent Systems, Griffith University, Brisbane, Australia
[b] A. Dehzangi, A. Sattar
National ICT Australia (NICTA), Brisbane, Australia
[c] A. Sharma
School of Engineering and Physics, University of the South Pacific, Private Mail Bag, Laucala Campus, Suva, Fiji
[d] Y. Zhou, Y. Yang
Institute for Glycomics and School of Information and Communication Technique, Griffith University, Parklands Dr. Southport, QLD 4222, Australia.
E-mail: yaoqi.zhou@griffith.edu.au or yuedong.yang@griffith.edu.au

Contract/grant sponsor: National Health and Medical Research Council of Australia (to Y.Z.); contract/grant number: 1059775

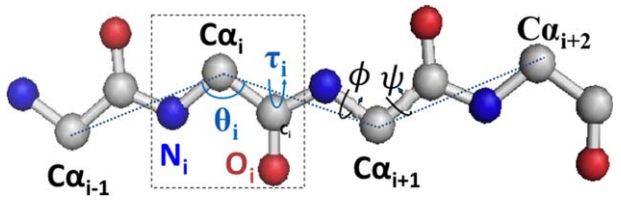

Figure 1. The schematic illustration of the protein backbone and associated angles. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

undefined number of residues (4 for 3_{10} helix, 5 for α-helix, and an undefined number of residues for sheet residues). Thus, secondary structure, ϕ/ψ and θ/τ provide complimentary local structural information along the backbone. Indeed, both predicted ϕ/ψ and secondary structure are useful for template-based structure prediction.[14]

In this article, we will develop the first machine-learning technique to predict θ and τ from protein sequences. This tool is needed not only because these two angles yield local structural information complementary to secondary structure and ϕ/ψ angles but also because they have been widely used in coarse-grained models for protein dynamics,[17] folding,[18] structure prediction,[19,20] conformational analysis,[21] and model validation.[22] That is, accurate prediction of θ and τ will be useful for template or template-free structure prediction as well as validation of predicted models. Using 4590 proteins for training and cross validation and 1199 proteins for an independent test, we have developed a deep-learning neural-network-based method that achieved θ and τ angles within 9 and 34 degrees, in average, of their native values.

Method

Datasets

In this study, we obtained a dataset of 5840 proteins with less than 25% sequence identity and X-ray resolution better than 2 Å from the protein sequence culling server PISCES.[23] After removing 51 proteins with obsolete IDs or missing data, the final dataset consists of 5789 proteins with 1,246,420 residues. We randomly selected 4590 proteins from this dataset for training and cross-validation (TR4590) and used the remaining 1199 proteins for an independent test (TS1199).

Deep neural-network learning

An Artificial Neural Network (ANN) consists of highly interconnected, multilayer processing units called neurons. Each neuron combines its inputs with a nonlinear sigmoid activation function to produce an output. Deep neural networks refer to feed-forward ANNs with three or more hidden layers. Multilayer networks were not widely used because of the difficulty to train neural-network weights. This has changed due to recent advances through unsupervised weight initialization, followed by fine-tuned supervised training.[24,25] In this study, unsupervised weight initialization was done by stacked sparse auto-encoder. A stacked auto-encoder treats each layer as an auto-encoder that maps the layer's inputs back to themselves.

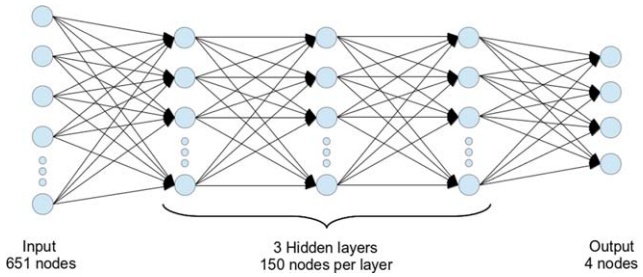

Figure 2. The general architecture of the stacked sparse auto-encoder deep neural network. Four output nodes are Sin(θ), Cos(θ), Sin(τ), and Cos(τ), respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

During training auto-encoders a sparsity penalty was utilized to prevent learning of the identity function.[26] Initialised weights were then refined by standard back propagation. The stacked sparse auto-encoder used in this study consists of three hidden layers with 150 hidden nodes in each layer (Fig. 2). The input data was normalised so that each feature is in the range of 0 to 1. For residues near the ends of a protein, the features of the amino acid residue at the other end of the protein were duplicated so that a full window could be used. The learning rate was initialised to start at 0.5 and was then decreased as training progressed. In this study, we used the deep neural network MATLAB toolbox implemented by Palm.[27]

Input features

Each amino acid was described by a vector of input features that include 20 values from the Position Specific Scoring Matrix generated by PSI-BLAST[28] with three iterations of searching against nonredundant sequence database with an *E*-value cut off of 0.001. We also used seven representative amino-acid properties: a steric parameter (graph shape index), hydrophobicity, volume, polarizability, isoelectric point, helix probability, and sheet probability.[29] In addition, we used predicted secondary structures (three probability values for helix, sheet, and coils) and predicted solvent accessible surface area (one value) from SPINE-X.[3] That is, this is a vector of 31 dimensions per amino acid residue. As before, we also used a window size of 21 amino acids (10 amino acids at each side of the target amino acid). This led to a total of 651 input features (21×31) for a given amino acid residue.

Output

Here, we attempt to predict two angles. One is θ, the angle between three consecutive Cα atoms of a protein backbone. The other one is τ, the dihedral angle between four consecutive Cα atoms of protein backbone. Two angles are predicted at the same time. To remove the effect of periodicity, we used four output nodes that correspond to Sin(θ), Cos(θ), Sin(τ), and Cos(τ), respectively. Predicted sine and cosine values were converted back to angles by using $\theta = \tan^{-1}[\sin(\theta)/\cos(\theta)]$ and $\tau = \tan^{-1}[\sin(\tau)/\cos(\tau)]$. Such transformation is widely used in signal processing and speech recognition.[30]

Table 1. Performance of θ and τ angle prediction based on the MAE as compared to θ and τ angles calculated from ϕ and ψ angles predicted by SPINE-X for two datasets (tenfold cross validation for TR4590 and independent test for TS1199).

MAE	TR4590(°)	TS1199(°)	TS1199(°)from predicted ϕ and ψ
θ−All	8.57 ± 0.01	8.6	9.6
θ−Helix	4.50 ± 0.02	4.5	4.5
θ−Sheet	10.45 ± 0.02	10.6	11.3
θ−Coil	11.437 ± 0.01	11.4	13.8
τ	33.4 ± 0.3	33.6	37.7
τ−Helix	17.1 ± 0.9	16.9	17.8
τ−Sheet	32.4 ± 0.1	33.1	39.1
τ−Coil	50.1 ± 0.3	50.2	56.4

Evaluation methods

We investigated the effectiveness of our proposed method using tenfold cross validation (TR4590) and independent test sets (TS1199). In tenfold cross validation, TR4590 was divided into 10 groups. Nine groups were used as a training dataset while the remaining group was used for test. This process was repeated 10 times until all the 10 groups were used once as the test dataset. In addition to tenfold cross validation, TR4590 was used as the training set and TS1199 was used as an independent test set. Comparison between tenfold cross validation and the test gives an indicator for the generality of the prediction tool. We evaluated the accuracy of our prediction by mean absolute error (MAE), the average absolute difference between predicted and experimentally determined angles. The periodicity of τ angles was taken care of by utilizing the smaller value of the absolute difference d_i $(=|\tau_i^{\mathrm{Pred}}-\tau_i^{\mathrm{Expt}}|)$ and $360-d_i$ for average.

Result

Table 1 compares the results of tenfold cross validation based on TR4590 and the independent test (TS1199). θ angles with a range of 0 to 180° were predicted significantly more accurate than τ angles with a range of −180° to 180°. The MAE is <9° for θ but 33–34° for τ. This level of accuracy can be compared to the baseline MAE values of 18.8° for θ and 86.2° for τ if θ and τ are assigned randomly according to their respective distributions. Accuracy for angles differs significantly in secondary structure types. The angles for helical residues have the highest accuracy (MAE<5° for θ and 17° for τ. The MAE for sheet residues is about twice larger than that for helical residues. Angles for coil residues have the largest error (τ in particular). Different levels of accuracy in different secondary structural types reflect the fact that helical structures are more locally stabilized than sheet structures while coil residues do not have a well-defined conformation. Similar trends were observed for prediction of backbone ϕ and ψ angles.[6–9] We also noted that MAEs from tenfold cross validation and from the independent test are essentially the same. This indicates the robustness of the method trained. Thus, here and hereafter, we will present the result based on the independent test only.

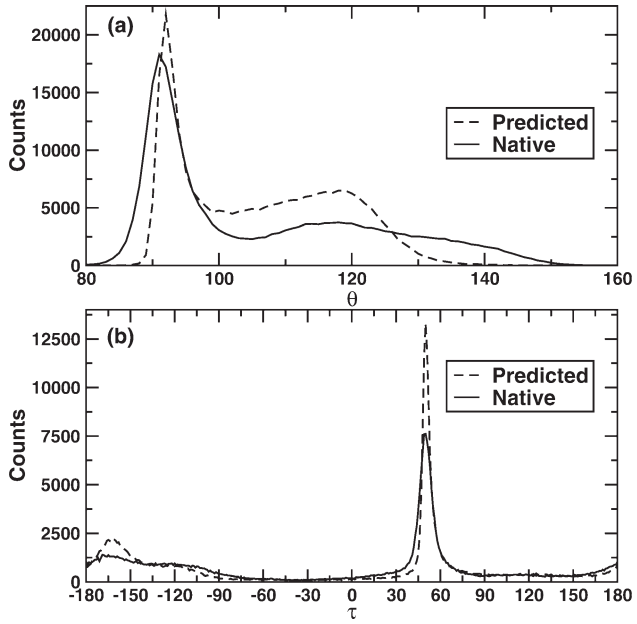

Figure 3. Predicted and actual distributions of θ (a) and τ (b) angles for the TS1199 dataset.

Actual and predicted distributions of θ and τ for TS1199 are shown in Figure 3. Predicted and actual distributions agree with each other very well. Both predicted and actual peaks for θ angles are located at 92° and 119°, respectively. Actual peaks for τ angles are also in good agreement with those predicted ones at 50° and −164°, respectively. Predicted peaks, however, are slightly narrowly than native peaks for all cases. Predicted and actual angle distributions also agree in a two-dimensional plane of θ and τ in Figure 4, the locations of three major populations were well captured by predicted distributions.

Table 2 lists the MAEs for 20 individual residue types along with their frequencies of occurrence in the TS1199 dataset. Glycine (G) has the largest MAE, corresponding to the fact that it is the most flexible residue due to lack of a side chain. Leucine (L), has the smallest MAE and interestingly also the most frequently occurred residue (9.2%). The angles for several other small hydrophobic residues [isoleucine (I), valine (V), and alanine (A)] are also in the pack of residues with smallest errors. There is no strong correlation between the MAE of an amino acid residue type and its frequency of occurrence.

In Figure 5, MAEs for predicted angles are shown as a function of relative solvent accessible surface area. MAEs for θ and τ have similar trend: two peaks separated by a valley (although in a smaller magnitude for θ). Both angles have the highest accuracy (the smallest error) at an intermediate range of solvent accessibility and the lowest accuracy (the largest error) at 90–100% solvent accessibility. The lowest accuracy at 90–100% solvent accessibility is likely due to the smallest number of residues at 90–100% solvent-accessible and 20% more coil residues in fully exposed residues.[3]

Figure 6 displays the fraction of proteins with more than a given fraction of correctly predicted angles (θ and τ). Here, a correct prediction is defined as 36° or less from the actual angle. We use 36° as a cut off value because it is relatively

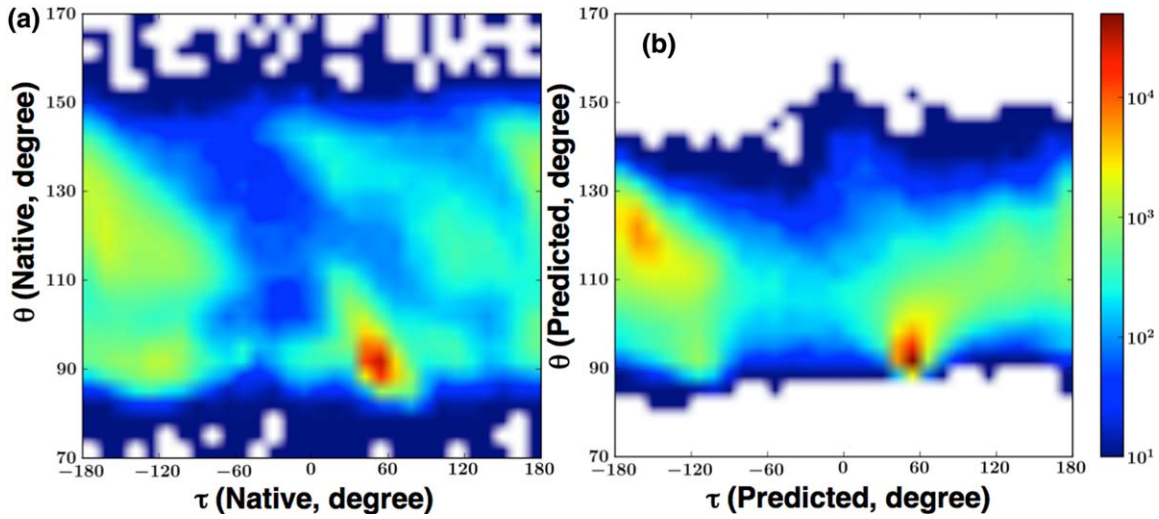

Figure 4. Actual (a) and predicted (b) distributions in the θ-τ plane for the TS1199 dataset.

easy for a conformational sampling technique to sample conformational changes within 36°. θ angles are always predicted within 36° for all residues in all proteins. 70% or more τ angles are predicted correctly for nearly 90% proteins. However, less than 10% proteins have 100% correctly predicted θ and τ.

θ and τ can also be calculated from backbone torsion angles ϕ and ψ by assuming $\omega = 180°$. Thus, it is of interest to compare the accuracy of θ and τ predicted in this work with those calculated from predicted ϕ and ψ. For the TS1199 dataset, we found that the MAE values for θ and τ derived from ϕ and ψ predicted by SPINE X[3] are 9.6° and 37.7°, respectively. Thus, the angles predicted in this work (MAE = 8.6° and 33.6°, respectively) are about 10% more accurate in θ or τ than those calculated from ϕ and ψ predicted by SPINE X. The largest improvement by direct prediction of θ or τ as shown in Table 1 is in coil residues. The MAE for a coil residue is reduced from 13.8° to 11.4° for θ and from 56.4° to 50.2° for τ.

Table 2. The MAEs of θ and τ prediction for 20 amino acid residue types along with their frequency of occurrence in the TS1199 dataset.

Amino acids	Frequency	Theta	Tau
A	8.3	7.5	28.5
C	1.4	10.1	38.1
D	5.9	8.4	38.9
E	6.7	7.1	29.7
F	4.0	9.3	34.2
G	7.2	12.3	51.5
H	2.3	9.6	37.9
I	5.6	7.2	26.2
K	5.8	7.8	30.9
L	9.2	6.9	25.9
M	2.1	7.9	29.2
N	4.4	9.0	41.0
P	4.6	8.5	33.5
Q	3.8	7.6	30.6
R	5.1	8.0	31.2
S	5.9	10.7	40.4
T	5.6	9.9	35.6
V	7.1	7.7	27.7
W	1.5	9.2	35.3
Y	3.6	9.3	34.5
Average		8.6	33.6

One application of predicted θ and τ is to utilize them for direct construction of local structures whose accuracies can be measured by the root-mean-square distance (RMSD) from their corresponding native conformations. Fragment structures of a length L are derived from predicted angles using the TS1199 dataset with a sliding window (1 to L, 2 to $L+1$, 3 to $L+2$, and etc.). For $L = 15$, a total of 229,681 fragments are constructed. Each fragment structure was built by using the standard Cα-Cα distance of 3.8 Å, and predicted θ and τ angles. We compared the accuracy of local structures from predicted θ and τ to those from ϕ and ψ predicted by SPINE X in Figure 7a. The RMSD between a native local structure (15 residue fragment) and its corresponding local

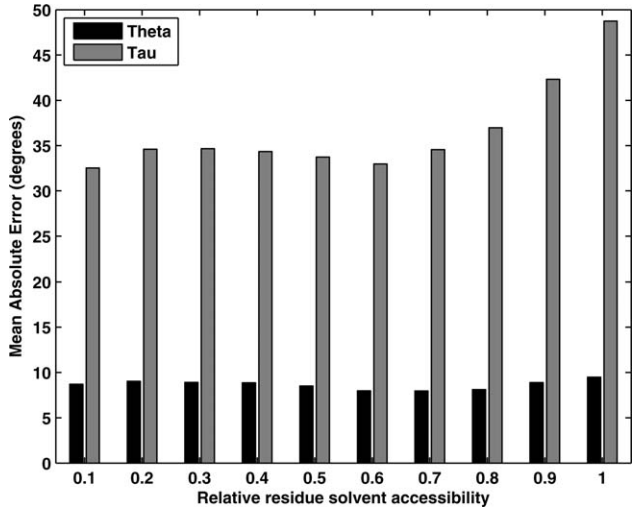

Figure 5. MAEs as a function of relative solvent accessibility for the TS1199 dataset.

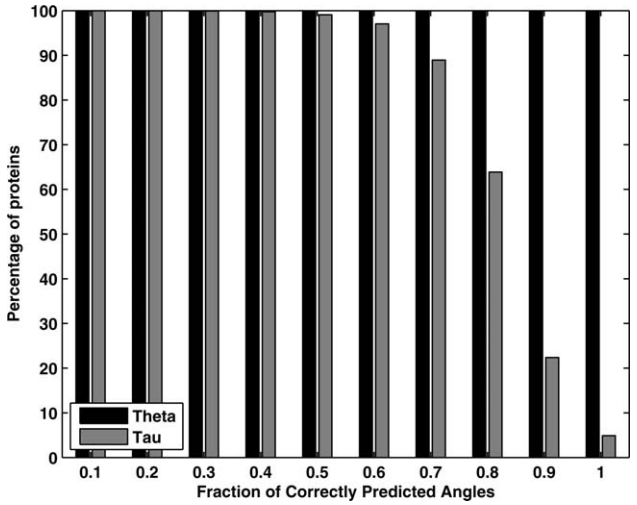

Figure 6. Percentage of proteins with more than a fraction of correctly predicted angles (θ and τ angles are less than 36° from native values, respectively) for the TS1199 dataset.

structure constructed from predicted θ and τ angles (*X*-axis) is plotted against the RMSD between a native local structure and its corresponding structure constructed from predicted ϕ and ψ angles (*Y*-axis) in a density plot. The majority of RMSD values are less than 6 Å. The average RMSD values of local structures from predicted θ and τ are 1.9 Å for 10 mer, 3.1 Å for 15 mer, 4.3 Å for 20 mer, and 7.0 Å for 30 mer. By comparison, the average RMSD values from predicted ϕ and ψ are 2.2 Å for 10 mer, 3.4 Å for 15 mer, 4.8 Å for 20 mer, and 7.7 Å for 30 mer. The improvement of θ/τ derived structures over ϕ/ψ derived structures is more than 10%. More local structures from predicted θ and τ angles are more accurately predicted than those from predicted ϕ and ψ angles as demonstrated by the size of the triangle at the bottom-right corner. The spread from the diagonal line confirms the complementary role of these four predicted angles.

The difference (RMSD) between local structures generated by predicted θ and τ angles and those by predicted ϕ and ψ angles can serve as an effective measure of how accurate a predicted local structure is. Figure 7b shows the density plot of the RMSD from the native (*Y*-axis) versus the RMSD from the ϕ and ψ-derived structure (*X*-axis) for 15-residue fragments. There is a trend that the larger the structural difference from different types of angles is, the less accurate the predicted local structure (larger RMSD) will be. For example, if the RMSD between θ/τ derived and ϕ/ψ-derived local structures is less than 2 Å, the RMSD of a θ/τ-derived structure from its native structure is most likely less than 4 Å based on the most populated region in red.

Discussion

This study developed the first machine-learning technique for prediction of the angle between $C\alpha_{i-1}$—$C\alpha_i$—$C\alpha_{i+1}$ (θ) and a dihedral angle rotated about the $C\alpha_i$—$C\alpha_{i+1}$ bond (τ). These angles reflect a local structure of three to four amino acid residues. By comparison, ϕ and ψ angles are the property of a single residue while secondary helical and sheet structures involve more than three residues. Thus, direct prediction of θ and τ angles is complementary to sequence-based prediction of ϕ and ψ angles and secondary structures. Predicting θ and τ angles also has one advantage over ϕ and ψ angles because θ has a narrow range of 0 to 180° while ϕ and ψ, similar to τ are both dihedral angles ranging from −180° to +180°. Indeed, by using the stacked sparse auto-encoder deep neural network, we achieved MSE values of 9° for θ and 34° for τ. By comparison, MAE is 22° for ϕ and 33° for ψ by SPINE-X. As a result, θ and τ calculated from predicted ϕ and ψ angles are less accurate with an MAE of 10° for θ and 38° for τ, 10% higher than direct prediction of θ and τ.

Complementarity between predicted θ/τ angles and predicted ϕ/ψ angles is demonstrated from the accuracy of

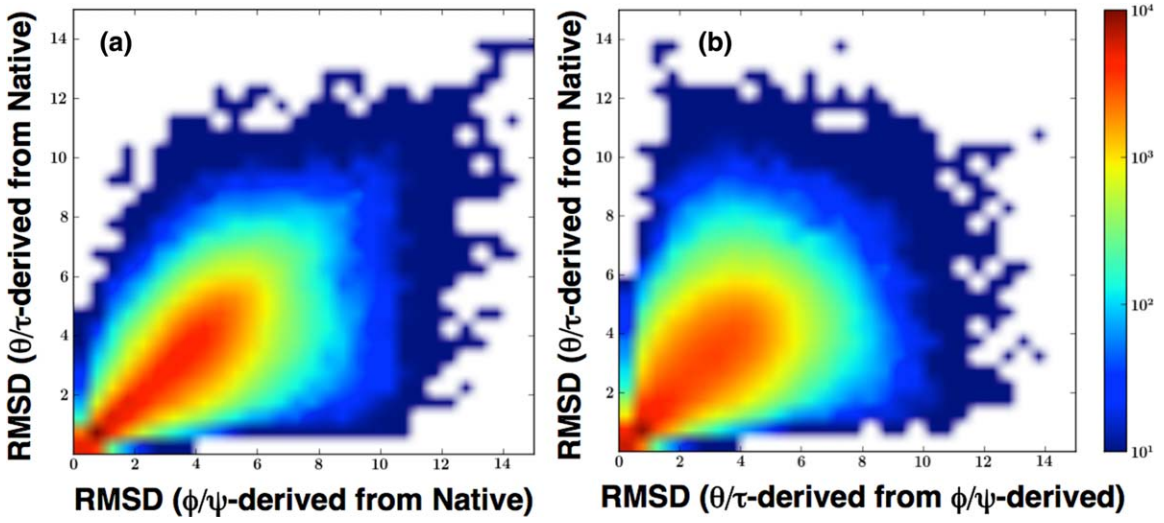

Figure 7. a) Consistency between 15-residue local fragment structures derived from predicted ϕ/ψ (*X*-axis) and those from predicted θ/τ angles (*Y*-axis) in term of their RMSD (in Å) from the native structure for the TS1199 dataset. b) RMSD values between two angle-derived local structures (*X*-axis) are compared to RMSD of a θ/τ-derived structure from its native structure.

local structures constructed based on these predicted angles. As shown in Figure 7a, some local structures are more accurately constructed by θ and τ angles while others are more accurately constructed by ϕ and ψ angles. Moreover, RMSD values between θ/τ-derived and ϕ/ψ-derived structures can be utilized as a measure for the accuracy of a predicted local structure (Fig. 7b). Usefulness of predicted angles for fragment structure prediction is illustrated by the fact that the average RMSD of 15-residue fragments is only 3Å from the corresponding native fragment structures. Currently, the most successful techniques in structure prediction (e.g. ROSETTA[31] and TASSER[32]) are based on mixing and matching of known native structures either in whole (template-based modelling) or in part (fragment assembly).[33,34] Fragment structures based on predicted θ and τ angles provide an alternative but complementary approach to the homolog-based approach for generating fragment structures. In addition to fragment-based structure prediction, predicted θ and τ angles can also be used directly as a constraint for fragment-free or *ab initio* structure prediction[1,2] as predicted ϕ and ψ angles did.[9]

How to handle the periodicity of torsion angles is an issue facing angle prediction (-180° is same as 180°). In our previous work for predicting ϕ and ψ angles, we used a simple angle shift,[7] and prediction of peaks (two-state classification), followed by prediction of deviation from the peaks.[9] Here, we introduced a sine and cosine transformation of θ and τ angles, a technique commonly used in signal processing and speech recognition.[30] We have compared the sine and cosine transformation with angle shifting and its combination of two-state classification because the distributions of θ and τ angles also have two peaks (Fig. 3). We found that the MAE of τ is 54° by direct prediction, 41° by angle shifting, and 36° by combining two-peak prediction with angle shifting. Thus, a MAE of 34° by sine and cosine transformation has the highest accuracy. We also examined the use of arcsine or arccosine, rather than arctangent. We found that using arccosine (with sine for phase determination) yields similar prediction accuracy as using arctangent but using arcsine leads to significantly worse prediction. We expect that such sine and cosine transformation of ϕ and ψ angles will also likely improve over existing SPINE-X prediction. For SPINE-X, MAE values are 33° for ψ angles and 22° for ϕ angles, respectively.

We also examined how much improvement in angle prediction is due to the use of deep learning neural networks. We found that when only one hidden layer (150 nodes) is utilised, MAE values are 8.8° for θ angles and 34.1° for τ angles, respectively. Thus, using deep 3-layer neural networks yields minor but statistically significant improvement over simple neural networks.

The most difficult angles to predict are the angles of coil residues (Table 1). This is true for θ and τ angles as well as for ϕ and ψ angles. Angles in coil regions have a MAE of 11° for θ and 50° for τ, compared to 32° for ϕ and 56° for ψ. This is likely because coil regions are structurally least defined. Despite of large errors, predicted ϕ and ψ angles in coil regions have been proved to significantly improve the accuracy of predicted structures.[9] Thus, we expect that predicted θ and τ angles in coil regions will also be useful as restraints for *ab initio* structure prediction[9] or template-based structure prediction.[14]

Acknowledgments

We also gratefully acknowledge the support of the Griffith University eResearch Services Team and the use of the High Performance Computing Cluster "Gowonda" to complete this research. This research/project has also been undertaken with the aid of the research cloud resources provided by the Queensland Cyber Infrastructure Foundation (QCIF).

Keywords: local structure prediction · protein structure prediction · secondary structure prediction · fragment structure prediction · fold recognition · deep learning · neural network

How to cite this article: J. Lyons, A. Dehzangi, R. Heffernan, A. Sharma, K. Paliwal, A. Sattar, Y. Zhou, Y. Yang *J. Comput. Chem.* **2014**, *35*, 2040–2046. DOI: 10.1002/jcc.23718

[1] Y. Q. Zhou, Y. Duan, Y. D. Yang, E. Faraggi, H. X. Lei, *Theor. Chem. Acc.* **2011**, *128*, 3.
[2] J. T. Guo, K. Ellrott, Y. Xu, *Methods Mol. Biol.* **2008**, *413*, 3.
[3] E. Faraggi, T. Zhang, Y. Yang, L. Kurgan, Y. Zhou, *J. Comput. Chem.* **2011**, *33*, 259.
[4] B. Rost, *J. Struct. Biol.* **2001**, *134*, 204.
[5] D. Kihara, *Protein Sci.* **2005**, *14*, 1955.
[6] O. Dor, Y. Zhou, *Proteins: Struct. Funct. Bioinf.* **2007**, *68*, 76.
[7] B. Xue, O. Dor, E. Faraggi, Y. Zhou, *Proteins: Struct. Funct. Bioinf.* **2008**, *72*, 427.
[8] E. Faraggi, B. Xue, Y. Zhou, *Proteins: Struct. Funct. Bioinf.* **2009**, *74*, 847.
[9] E. Faraggi, Y. D. Yang, S. S. Zhang, Y. Zhou, *Structure* **2009**, *17*, 1515.
[10] K. T. Simons, R. Bonneau, I. Ruczinski, D. Baker, *Proteins: Struct. Funct. Genet.* **1999**, Suppl 3, 171.
[11] Y. M. Huang, C. Bystroff, *Bioinformatics* **2006**, *22*, 413.
[12] C. Mooney, A. Vullo, G. Pollastri, *J. Comput. Biol.* **2006**, *13*, 1489.
[13] M. J. Wood, J. D. Hirst, *Proteins: Struct. Funct. Bioinf.* **2005**, *59*, 476.
[14] Y. Yang, E. Faraggi, H. Zhao, Y. Zhou, *Bioinformatics* **2011**, *27*, 2076.
[15] R. Karchin, M. Cline, Y. Mandel-Gutfreund, K. Karplus, *Proteins: Struct Funct Bioinf* **2003**, *51*, 504.
[16] W. Zhang, S. Liu, Y. Zhou, *PLoS ONE* **2008**, *6*, e2325.
[17] A. Korkut, W. A. Hendrickson, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15667.
[18] Y. Zhou, M. Karplus, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 14429.
[19] D. Kihara, H. Lu, A. Kolinski, J. Skolnick, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10125.
[20] A. Liwo, Y. He, H. A. Scheraga, *Phys. Chem. Chem. Phys.* **2011**, *13*, 16890.
[21] M. M. Flocco, S. L. Mowbray, *Protein Sci.* **1995**, *4*, 2118.
[22] G. J. Kleywegt, *J. Mol. Biol.* **1997**, *273*, 371.
[23] G. Wang, R. L. Dunbrack, Jr., *Nucleic Acids Res.* **2005**, *33*, W94.
[24] G. E. Hinton, *Trends Cogn. Sci.* **2007**, *11*, 428.
[25] Y. Bengio, *Foundations and trends® in Machine Learning* **2009**, *2*, 1.
[26] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, *Adv. Neural Inf. Process. Syst.* **2007**, *19*, 153.
[27] R. B. Palm, Prediction as a Candidate for Learning Deep Hierarchical Models of Data; Technical University of Denmark, Kongens Lyngby, **2012**.
[28] S. F. Altschul, T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res.* **1997**, *25*, 3389.

[29] J. Meiler, M. Müller, A. Zeidler, F. Schmäschke, *J. Mol. Model.* **2001**, *7*, 360.
[30] B. Bozkurt, L. Couvreur, T. Dutoit, *Speech Commun.* **2007**, *49*, 159.
[31] K. T. Simons, C. Kooperberg, E. Huang, D. Baker, *J. Mol. Biol.* **1997**, *268*, 209.
[32] Y. Zhang, J. Skolnick, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7594.
[33] J. M. Bujnicki, *Chembiochem* **2006**, *7*, 19.
[34] Y. Zhang, *Curr. Opin. Struct. Biol.* **2009**, *19*, 145.

Received: 5 June 2014
Revised: 12 July 2014
Accepted: 9 August 2014
Published online on 12 September 2014