International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:4, No:1, 2010

On the Prediction of Transmembrane Helical Segments in Membrane Proteins Based on Wavelet Transform

Yu Bin, Zhang Yan

***Abstract*—**The prediction of transmembrane helical segments (TMHs) in membrane proteins is an important field in the bioinformatics research. In this paper, a new method based on discrete wavelet transform (DWT) has been developed to predict the number and location of TMHs in membrane proteins. PDB coded as 1KQG was chosen as an example to describe the prediction of the number and location of TMHs in membrane proteins by using this method. To access the effect of the method, 80 proteins with known 3D-structure from Mptopo database are chosen at random as the test objects (including 325 TMHs), 308 of which can be predicted accurately, the average predicted accuracy is 96.3%. In addition, the above 80 membrane proteins are divided into 13 groups according to their function and type. In particular, the results of the prediction of TMHs of the 13 groups are satisfying.

***Keywords*—**discrete wavelet transform, hydrophobicity, membrane protein, transmembrane helical segments

I. Introduction

STUDY of membrane protein structures and functions is an important job in present bioinformatics. About 20-30% of genome products have been predicted as membrane proteins, which have significant biological functions in the life activity of the cells, such as composing nerve signal molecules and drug receptors, transporting ion and alimentation, conducting immunoreactions, etc [1]. As the stable natural conformations of membrane proteins need the assistant of the biology membrane, it is not easy to measure their three-dimensional structures with X-ray diffraction or nuclear magnetic resonance. The structure data of membrane proteins only occupy 0.6% in the latest PDB database. Thus there is a huge gap between known membrane protein sequences and their unknown structures that greatly restrict the further research of the functions. Therefore, it is very necessary to develop an accurate and efficient approach to predicting the structure of the membrane protein.

So far many transmembrane helical segments (TMHs) predicting algorithms for membrane proteins have been proposed. In 1982 Kyte and Doolittle firstly suggested a hydrophobicity analysis method of membrane protein sequences [2]. Thereafter von Heijne [3] put forward the well-known positive-inside rule to guide prediction in 1986. SOSUI [4] and PRED-TMR [5] were based on the foregoing two methods. In recent years, some statistical methods have been developed that like DAS [6], TMAP [7], ANN-based PHDhtm [8], HMM-based TMHMM [1] and HMMTOP [9]. Wavelet transform was first introduced into bioinformatics research in 1996 and raised extensive attention immediately [10-14]. In this paper, we make full use of the hydrophobicity of amino acids and multiresolution feature of DWT to decompose the amino acids of TM proteins into a series of structures in different layers, then predicting the location of TMHs according to the information of the amino acids sequence in different scales. 80 proteins with known 3D-structure from Mptopo database are chosen at random as the test objects (including 325 TMHs), 308 of which can be predicted accurately, the average predicted accuracy is 96.3%. In addition, the above 80 membrane proteins are divided into 13 groups according to their function and type. In particular, the results of the prediction of TMHs of the 13 groups are satisfying.

YU Bin is with the College of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao, 266061, China (phone: 86-0532-88956917; e-mail: yubin@qust.edu.cn).

ZHANG Yan is with Qingdao University of Science and Technology, Qingdao, 266061, China (e-mail: zy@qust.edu.cn).

II. Materials and Methods

A. Materials

The test dataset is retrieved from the latest MPtopo database (http://blanco.biomol.uci.edu/mptopo)[15], which collects a set of membrane protein structure data identified by crystallography or other experimental technologies such that they can be treated as reliable samples.

B. Methods

Proteins are biomacromolecules that are consisted of twenty different amino acids joined with peptide bonds. Different amino acids have different side-chains that define diverse physico-chemical characteristics of different types of amino acids. Hydrophobic effects are of the most importance among the features because the hydrophobic effects determine to a great degree the stability of protein structures [16]. So considering the critical importance of hydrophobicity in holding the secondary and tertiary structures of proteins, we should map the amino acid sequence of protein onto a sequence of hydrophobicity values that are regarded as raw signals for the wavelet analysis. The hydrophobicity values of 20 amino acids are given in Table I.

TABLE I
UNITS FOR MAGNETIC PROPERTIES

Amino Acids	A	C	D	E	F	G	H	I	K	L
H-Values	1.8	2.5	-3.5	-3.5	2.8	-0.4	-3.2	4.5	-3.9	3.8
Amino Acids	M	N	P	Q	R	S	T	V	W	Y
H-Values	1.9	-3.5	-1.6	-3.5	-4.5	-0.8	-0.7	4.2	-0.9	-1.3

The wavelet transform (WT) is relatively new mathematical technique and has some similarities with the Fourier Transform (FT). Wavelets differ from Fourier methods in that they allow the localization of a signal in both time and frequency. A WT of a signal typically outperforms an FT when the signal under consideration contains discontinuities and sharp spikes. The Discrete Wavelet Transform (DWT) decomposes a function into its wavelet coefficients. From a computational point of view, it proceeds by recursively applying two convolution functions, known as quadrature mirror filters, each producing an output stream that is half length of the original input, until the resolution level zero is reached. Mallat brought out the most important concept multiresolution analysis (MRA) in a discrete wavelet theory as well as fast algorithm of orthonormal wavelet tranform-Mallat algorithm [17].

Let $\phi(x)$ is a scaling function which satisfies the following two-scale equation:

$$\phi(x) = \sqrt{2}\sum_{n\in Z} h_n \phi(2x-n) \tag{1}$$

where Z be a set of integers, and the coefficients $\{h_n, n \in Z\}$ denote a low-pass filter (H). The wavelet function $\psi(x)$ can be constructed using the scaling function $\phi(x)$ as

$$\psi(x) = \sqrt{2}\sum_{n\in Z} g_n \phi(2x-n) \tag{2}$$

where the coefficients $\{g_n, n \in Z\}$ denote a high-pass filter (G).

Assume that the shifted scaling function $\{\phi(x-k), k \in Z\}$ and the shifted wavelet functions $\{\psi(x-k), k \in Z\}$ are orthonormal, respectively. Let $\{c_l^0\}$ denote a sequence of hydrophobicity values, and we define a linear combination $f(x)$ of the sequence with scaling functions $\{\phi(x-k), k \in Z\}$:

$$f(x) = \sum_{k\in Z} c_k^0 \phi(x-k) \tag{3}$$

According to a wavelet theory, we have another expansion of $f(x)$:

$$f(x) = \frac{1}{\sqrt{2}}\left(\sum_{k\in Z} c_k^1 \phi(2^{-1}x-k) + \sum_{k\in Z} d_k^1 \psi(2^{-1}x-k)\right) \tag{4}$$

From (3) and (4) and using orthonormality of the scaling and wavelet functions, we can decompose the sequence $\{c_l^0\}$ into low frequency and high frequency components.

$$c_k^1 = \sum_{l\in Z} c_l^0 \overline{h}_{l-2k} \tag{5}$$

and

$$d_k^1 = \sum_{l\in Z} c_l^0 \overline{g}_{l-2k} \tag{6}$$

Repeatedly application of this decomposition, we can deduce

$$c_k^{j+1} = \sum_{l\in Z} c_l^j \overline{h}_{l-2k}, \qquad j = 0,1,2,\cdots, \tag{7}$$

and

$$d_k^{j+1} = \sum_{l\in Z} c_l^j \overline{g}_{l-2k}, \qquad j = 0,1,2,\cdots, \tag{8}$$

Conversely, we can derive a reconstruction formula form (3) and (4):

$$c_k^j = \sum_{l\in Z} c_l^{j+1} h_{k-2l} + \sum_{l\in Z} d_l^{j+1} g_{k-2l}, \qquad j = 0,1,2,\cdots, \tag{9}$$

Above-mentioned formulas can refer to the literature of Mallat [17].

In (7) and (8), the sequences $\{c_k^{j+1}\}$ and $\{d_k^{j+1}\}$ mean low and high frequencies. In this paper, only the first formula (9) is used because as far as most of the protein hydrophobicity signals are concerned, low frequency domain is especially important and it can reflect the general characteristics of signals. However the high frequency domain is always connected with noise and disturbance, so the basic features of signals will be reserved when the high frequency domain is discarded by putting $d_k^{j+1} = 0$. Using (9), we reconstruct a new sequence $\{\tilde{c}_k^j\}$ only from $\{c_k^{j+1}\}$, that is, we utilize low-pass filtering of wavelet transform, study the general trend and set an optimal threshold to locate TMHs. The threshold here is determined by the biggest average prediction accuracy among a set of protein sequences.

In this paper, we adopted the important Daubechies (dbN) wavelet series as mother wavelet and selected db10 as the optimum wavelet base after analyzing the all data of the test dataset as well as reconstruct wavelet from five different scale levels. To reach a high accuracy in the detection of TMHs, our method is dependent upon the post-treatment of the signals obtained after wavelet reconstruction. The post-treatment can be generalized in the following three steps.

Step1 Discard those predicted TMHs that have less than 7 amino acid residues.

Step2 If the predicted TMHs is between 30 and 50 residues, which means the TMHs is too long and is not factual, then the TMHs is expanded 10 amino acid residues from the two sides respectively and further we cut this TMHs into two equal parts

to seek for potential TMHs.

Step3 If the length of the predicted TMHs is greater than 50 residues, then the TMHs is cut into three equal parts using the same above-mentioned method.

For convenience, our prediction method is called WavePrd that is coded in MATLAB programming language, and the wavelet process can also be directly executed with the Wavelet Toolbox of MATLAB software that is developed by the MathWorks Company.

In order to test the accuracy of prediction methods, we study TM proteins from two aspects—TMHs and amino acid residues.

There are three important evaluation indexes: (1) FP (false-positive): the number of wrongly predicted TMHs; (2) FN (false-negative): the number of not-predicted TMHs; (3) Prediction accuracy of TMHs [9]: $Q_p=\sqrt{M*C}\times 100\%$, here M=Ncor/Nobs (Ncor stands for the number of correctly predicted TMHs, Nobs stands for the number of observed TMHs), M can be regard as a measure index of sensitivity; C=Ncor/Nprd (Nprd stands for the total number of predicted TMHs), C is regarded as a measure index of specificity.

Prediction accuracy of residues is another evaluation index. The calculation fomula is FAAcor=(NAAcor/NAAall)100%, where NAAcor is the number of correctly predicted TMHs residues and NAAall is the total residues.

III. Results and Discussion

We pick PDB ID 1KQG from MPtopo database as an example to illustrate the validity of WavePrd. This protein sequence has 217 amino acid residues with 4-TMHs shown in Figure 1 [18]. The original hydrophobicity plots and the wavelet neural network signal graphs at each scale level are shown in Figure 2. It is known that the peaks of WavePrd filtering are possibly corresponding with the real TMHs and each peak matches one TMHs core. Thus a group of TMHs can be predicted with our method.

MSKSKMIVRTKFID**RACHWTVVICFFLVALSGISFFF**P
TLQWLTQTFGTP**QMGRILHPFFGIAIFVALMFMFVR**F
VHHNIPDKKDIPWLLNIVEVLKGNEHKVADVGKYNA**G**
QKMMFWSIMSMIFVLLVTGVIIWRPYFAQYFPM**QVV**
RYSLLIHAAAGIILIHAILIHMYMAFWVKGSIKGMIEG
KVSRRWAKKHHPRWYREIEKAEAKKESEEGI

Fig. 1 Linear sequence of the 1KQG protein and the parts of bold-face denote the real TMHs

It can be seen that at the scale level 4, according to the wavelet filtering graph for the hydrophobicity sequence of 1KQG protein the predicted TMHs are correspondent well with the real TMHs. The selection of level 4 is based on our comparisons of wavelet filtering at each scale level. The effects of filtering at scale level 2 and 3 are not apparent but the hydrophobicity signals are excessively filtered at scale level 5, which further results in the lose of much information about the original sequences. The TMHs prediction accuracy reaches 100% and the amino acid residues prediction accuracy reaches 98.8% at the scale level 4 with optimal threshold 0.836. The contrast data in TableII show above result more clearly. And amino acid sequence of membrane proteins 1KQG was prediced by the method DAS[6], HMMTOP2.0 [9], PHDhtm [8], PRED-TMR2 [5], SOSUI [4], TMAP [7], MHMM2.0 [1]. The result is shown in Table 2. From the table, we can see the four TMHs of membrane proteins 1KQG were correspondence of TMHs we have predicted, i.e. All TMHs of membrane proteins have been predicted by WavePrd. One more TMHs was predicted by DAS; Three TMHs were predicted by PHDhtm and the third TMHs has 64 residues, i.e. The third and the fourth TMH were predicted together and result in big error. Better result has been achieved by other methods, yet TMHs and residues prediction accuracy are the highest by WavePrd.

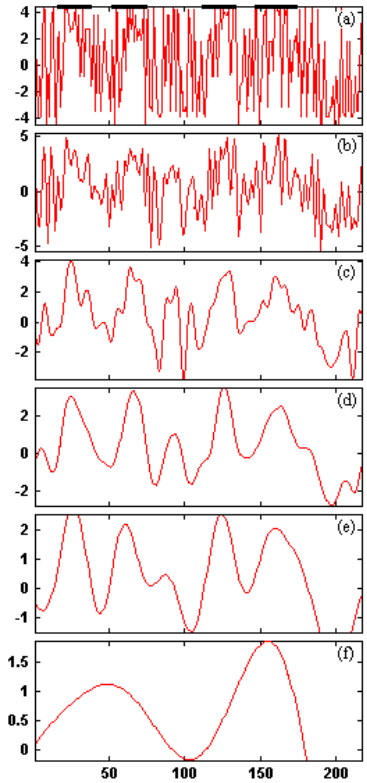

Fig. 2 The hydrophobicity signal plot and low frequencies at five different scale levels for 1KQG protein (a) j=0; (b) j=1; (c) j=2; (d) j=3; (e) j=4; (f) j=5

To access the effect of the method, 80 membrane proteins are randomly selected as test sets to be predicted by WavePrd, including 325 TMHs,19396 amino acid residues altogether. Through analysis, we choose db10 as the optimal wavelet base. The total number of predicted TMHs is 321 at the scale level 4 with optimal threshold 0.836, among which 308 TMHs are identical to real TMHs. The average prediction accuracy of TMHs is 96.3% and that of residues is 83.6%. The total residues of TMHs is 6580, among which 5501 are predicted rightly. The number of false positive segments is 7 and the number of false negative segments is 17.

TABLE II
LOCATION OF TMHS OF THE SEQUENCE OF 1KQG (TOP ROW), WAVEPRD PREDICTION AND RESULTS FROM OTHER CURRENTLY USED PREDICTION METHODS

	TM1	TM2		TM3	TM4
Observed	15-37	51-74		112-134	146-175
WavePrd	17-36	53-70		116-134	149-176
DAS	18-39	57-75	90-92	118-136	149-175
HMMTOP2.0	20-38	55-73		116-135	152-176
PHDhtm	18-45	55-76		117-180	
PRED-TMR2	19-37	55-73		115-135	156-176
SOSUI	18-40	55-77		115-137	150-172
TMAP	14-42	51-78		112-134	148-172
TMHMM2.0	21-40	55-77		117-139	154-176

These results are better than that obtained by using other wavelets or levels. Table III shows the detailed results. According to the function and classification of membrane proteins, we divided 80 membrane proteins sequences into 13 groups, which were shown in Table Ⅳ. Mentioned above, the choosing of threshold was ascertained by the maximal and average prediction accuracy given by test data. For every group of membrane proteins, at the scale level 4, we tried to choose a proper threshold to raise prediction accuracy. Membrane proteins of the same family are homologous, so the optimal threshold of each group of membrane proteins may be different. The values of parenthesis in Table 5 are the optimal threshold used in prediction.

TABLE III
PREDICTION RESULTS OF WAVEPRD METHOD

Nobs	Nprd	Ncor	Q_p	M	C	FP	FN	FAAcor
325	315	308	96.3%	94.8%	97.8%	7	17	83.6%

TABLE IV
MEMBRANE PROTEIN FAMILIES USED IN OUR PREDICTIONS

Family name	**PDB code**					
ABC transporters	1jsq	1l7vA	1pf4			
Bacteriorhodopsin	1ap9					
Channel proteins	1fqyA	1fx8A	1msl	1mxm	1oedA	1oedB
	1oedC	1oedE	1p7b	1rc2A	1rhzA	1rhzB
Cytochrome bc_1 complexes	1bgyE	1bgyJ	1bgyK			
Cytochrome b6f complexes	1um3A	1um3B	1um3D	1um3F	1um3G	1um3H
Cytochrome c oxidases	1ehkA	1ehkB	1ehkC	1occA	1occB	1occC
	1occD	1occG	1occI	1occJ	1occK	1occL
	1occM	1qleA	1qleB	1qleC	1qleD	
Glycophorin	1afoA					
Light-harvesting complexes	1kzuA	1lghA				
Photosynthetic reaction centers	1eysH	1eysL	1eysM	1prcH	1prcL	1prcM
	2rcrL	2rcrM				
Photosystems	1jboA	1jboB	1jboF	1jboI	1jboJ	1jboK
	1jboL	1jboM				
Respiratory proteins	1a91C	1fftA	1fftB	1fftC	1fumC	1kqgB
	1kqgC	1lovD	1nekC	1nekD	1okcA	1q16C
	1qlaC					
Rhodopsins	1f88	1h2sB	1h68A			
Translocation proteins	1pw4A	1s7b	2cpb			

From Table V, we can see that to choose different threshold for every group of membrane proteins can raise the prediction accuracy of TMHs and residues. With the threshold 0.566, the prediction accuracy of TMHs of ABC transporters is from 95.3% to 100%; With the threshold 0.915, the prediction accuracy of residues of light-harvesting complexes is from 93.9% to 97.9%. From another angle, we can see theTMHs and residues average prediction accuracy of the five groups of

membrane proteins are the best, which are cytochrome c oxidases, glycophorin, light-harvesting complexes, respiratory proteins and photosynthetic reaction centers. For the thirteen groups of membrane proteins data, with the threshold 0.836, the maximal and average prediction accuracy of membrane proteins TMHs is 96.5% and that of residues is 83.6%. But by choosing different threshold for every group of membrane proteins of data base, the average prediction accuracy of TMHs and residues is the highest, which are 96.8% and 84.4%. These comparisons indicate that our method is more accurate and effective in predicting the TMHs number and location of membrane proteins, which provide important information for research of membrane protein structure and function.

TABLE V
PREDICTION ACCURACY AND OPTIMAL THRESHOD IN EACH PROTEIN FAMILY

Family name	Prediction accuracy %			
	Qp [a]	FAAcor [a]	Qp [b]	FAAcor [b]
ABC transporters	95.3 (0.836)	74.8	100 (0.566)	75.6
Bacteriorhodopsin	100 (0.836)	70.7	100 (0.885)	71.3
Channel proteins	91.4 (0.836)	81.1	91.4 (0.847)	81.3
Cytochrome bc_1 complexes	86.6 (0.836)	66.7	86.6 (0.765)	68.9
Cytochrome b6f complexes	95.7 (0.836)	82.5	95.7 (0.891)	82.6
Cytochrome c oxidases	99.2 (0.836)	93.7	99.2 (0.836)	93.7
Glycophorin	100 (0.836)	91.3	100 (0.668)	92.0
Light-harvesting complexes	100 (0.836)	93.9	100 (0.915)	97.9
Photosynthetic reaction centers	98.4 (0.836)	90.6	98.4 (0.866)	91.2
Photosystems	97.0 (0.836)	82.6	97.0 (0.836)	82.6
Respiratory proteins	93.7 (0.836)	91.6	93.7 (0.836)	91.6
Rhodopsins	100 (0.836)	79.5	100 (0.885)	79.9
Translocation proteins	97.0 (0.836)	88.1	97.0 (0.868)	88.5
Average	**96.5**	**83.6**	**96.8**	**84.4**

[a] The average prediction accuracy of every group of membrane proteins with the threshold 0.836.

[b] With the different threshold for every group of membrane proteins, the prediction accuracy will be raised.

IV. CONCLUSION

Some achievements have been gained when wavelets are introduced into bioinformatics. We make use of multiresolution analysis theory to decompose the original signals into low frequency and high frequency domains in amino acid hydrophobicity scale format. In wavelet analysis, the low frequency can be easily obtained from a raw function by the decomposition and reconstruction formula. So high frequency domain is deleted and low frequency region is left for reconstructing wavelet because we only study the general features of protein sequences. In the results, we obtain precise filtering signals that can help us to find the actual location of TMHs in the protein sequences.

With the advancement of high-throughput sequencing technology and the practice of sequencing model organisms' genomes, more and more DNA and protein sequences are swarming into biological sequence databases with an unprecedented rate. How to mine valuable information efficiently from mass biological sequences is crucial to the research of bioinformatics as well as to many significant fields of systems biology. The study of the structure and function of TM proteins is increasingly emphasized since TM proteins play an extraordinarily important role in the life activity of the cells, such as signal transduction, immune response and membrane transport. However, the structural determination of proteins needs a considerable number of purified proteins and it is a hard task because the peptide chains of the TM proteins span a lipid bilayer and sometimes transverse membrane many times. Because transmembrane helix combines closely with membrane, first of all, the membrane must be disintegrated by eradicator to separate TM proteins. Then, the TM protein can be purified and made crystal. This is not only difficult in technique but also is expensive. Thus, the high-resolution 3D structural determination and analysis of TM proteins cannot answer the need of the research for TM protein functions. The computer prediction and analysis of the TMHs is able to provide much important information to disclose the relationship between the structure and function of TM proteins. At the same time wavelet transform is taking effects on dealing

with biological sequence information and is frequently applied to many other fields of bioinformatics. We believe that wavelet methods will exert great action in bioinformatics-related fields.

ACKNOWLEDGMENT

This work was supported by grants from the National Natural Science Foundation of China (No. 3057l059).

REFERENCES

[1] A. Krogh, B. Larsson, G. von Heijne, E. Sonnhamme, "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes,"*J. Mol. Biol.*, vol. 305, 2001, pp. 567-580.
[2] J. Kyte, R. F. Doolittle, "A simple method for displaying the hydrophathic character of a protein," *J. Mol. Biol.*, 1982, 157: 105-132.
[3] G. Heijne, "The distribution of positively charged residues in bacterial inner membrane proteins correlates with the transmembrane topology," *EMBO J*, vol. 5, 1986, pp. 3021-3027.
[4] T. Hirokawa, S. Boon-Chieng, S. Mitaku, "SOSUI: classification and secondary structure prediction system for membrane proteins," *Bioinformatics*, vol. 14, 1998, pp. 378-379.
[5] C. Pasquier, V. J. Promponas, G. A. Palaios, J. S. Hamodrakas, S. J. Hamodrakas, "A novel method for predicting trsnsmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm," *Protein Eng*., vol. 12, 1999, pp. 381-385.
[6] M. Cserzö, E. Wallin, I. Simon, G. von Heijne, A. Elofsson, "Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method," *Protein Eng*., vol. 10, 1997, pp. 673-676.
[7] B. Persson, P. Argos, "Prediction of transmembrane segments in proteins utilizing multiple sequence alignments," *J. Mol. Biol.*, vol. 237, 1994, pp. 182-192.
[8] B. Rost, R. Casadio, P. Fariselli, "Topology prediction for helical transmembrane segments at 86% accuracy," *Protein Sci*., vol. 5, 1996, pp. 1704-1718.
[9] G. E. Tusnady, I. Simon, "Principles governing amino acid composition of integral membrane proteins: application to topology prediction," *J. Mol. Biol*., vol. 283, 1998, pp. 489-506.
[10] Altaiski, M. Mornev, O. Polozov, "Wavelet analysis of DNA sequence," *Genet. Anal.* , vol. 12, 1996, pp. 165-168.
[11] B. Yu, X. H. Meng, H. J. Liu, et al, "Prediction of transmembrane helical segments in transmembrane proteins based on wavelet transform," *Journal of Shanghai University (English Edition)*, vol. 10, 2006, pp. 308-318.
[12] K. B. Li, P. Issac, A. Krishnan, "Predicting allergenic proteins using wavelet transform," *Bioinformatics*, vol. 20, 2004, pp. 2572-2578.
[13] P. Liò, "Wavelets in bioinformatics and computational biology: state of art and perspectives," *Bioinformatics*, vol. 19(1) 2003, pp. 2-9.
[14] J. P. Mena-Chalco, Y. Zana, and R. M. Cesar, "Identification of protein coding regions using the modified Gabor-wavelet transform," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 5, 2008, pp. 198-207.
[15] S. Jayasinghe, K. Hristova, S. H. White, "MPtopo: A database of membrane protein topology," *Protein Sci.*, vol. 10, 2001, pp. 455-458.
[16] D. Eisenberg, A. D. Mclachlan, "Solvation energy in protein folding and binding," *Nature*, vol. 319, 1986, pp. 199-203.
[17] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Math.Intell*, vol. 11, 1989, pp. 674-693.
[18] M. Jormakka, S. Tornroth, B. Byrne, S. Iwata, "Molecular basis of proton motive force generation: Structure of formate dehydrogenase-N," *Science*, vol. 295, 2002, pp. 1863-1868.

YU Bin received the ME degree in computational mathematics from Shanghai University, Shanghai, China, in 2005. His research interests include bioinformatics, systems biology and matrix theory.

ZHANG Yan received the BE degree in computer science and technology from Northwestern Polytechnic University, Xi'an Shanxi, China, in 2004, the ME degree in control theory and control engineering from the Qingdao University of Science and Technology, Qingdao, Shandong, China, in 2009. His research interests include wavelet theory, pattern recognition and image processing.