Cellular and Molecular Biology
E-ISSN : 1165-158X / P-ISSN : 0145-5680
www.cellmolbiol.org
C M B Association Publisher

Original Research

Protein molecular defect detection method based on a neural network algorithm

Meiqing Zheng[1*], Somayeh Kahrizi[2]

[1] Digital Information Technology Department, Zhejiang Technical Institute of Economics, Hangzhou 310018, China
[2] Department of Computer Engineering, Zagros Institute of Higher Education, Kermanshah, Iran

[*]**Correspondence to:** zhengmeiqing@163.com

Received March 11, 2020; **Accepted** June 27, 2020; **Published** October 31, 2020

Doi: http://dx.doi.org/10.14715/cmb/2020.66.7.13

Abstract: Proteins, as the largest macromolecules in the body, are among the most important components of the body and play very vital and important roles. These substances are made up of a series of amino acid chains that, depending on the type of protein, the number of these amino acids can reach several thousand. Proteins function differently depending on the type and location of their presence, including enzymatic activity to catalyze the process, identify microbes and cancer cells, transport substances such as respiratory gases, and signalize. In the biochemical experiment, the problem of optimizing the detection of protein molecular defects, because of the randomness of the information, parameters, selection and setting, limits the detection accuracy of protein molecular defects. Based on the characteristics of fast learning speed and a robust network of neural network algorithm, a protein molecular defect detection method based on a neural network algorithm was proposed. Firstly, the protein secondary structure was predicted by the method of protein secondary structure prediction based on the generalized regression neural network to obtain the protein structural features; secondly, the protein defective molecular sequence classification model based on the neural network was used to classify the protein defective molecular sequence to achieve the protein molecular defect detection. The results showed that the detection accuracy of the proposed method was very high, which meets the needs of protein molecular defect detection, and has some application advantages compared with similar detection methods.

Key words: Neural network algorithm; Protein; Molecule; Defect; Detection; Secondary structure.

Introduction

Protein is an important component of all cells and tissues in the human body. All important parts of the body need to be involved in protein (1-4). Their structure, function, expression, value, shape, and presence are important (5-10). Generally speaking, protein accounts for about 18% of the total mass of the human body (11-12). The most important thing is that it is related to life phenomena. Protein is the material basis of life, is an organic macromolecule, is the basic organic substance of the cell, and is the main undertaker of life activities (13-14). There is no life without protein. Amino acids are the basic components of proteins (15-16). It is closely related to life and various forms of life activities. Protein is involved in every cell and all important components of the body (17). Protein accounts for 16%-20% of human body weight, that is, a 60 kg adult has about 9.6 to 12 kg protein in its body. There are many kinds of proteins in the human body, with different properties and functions, but they are all composed of more than 20 amino acids in different proportions, and they are constantly metabolized and renewed in the body. With the advent of the post-genomic era, life science has entered a new stage of vigorous development. The research on protein and its structure, function, activity, proteomics and so on have developed rapidly. As one of the main characteristic parameters of protein, molecular weight is an important prerequisite to determine a new protein and carry out the follow-up research activities of the protein. Therefore, it is very urgent to study the molecular weight test method of protein and accurately and effectively measure the molecular weight of protein (18).

It can be seen from the relevant literature that foreign Bock and Gough have carried out in-depth exploration experiments on proteins, which have verified that they can predict whether a pair of protein sequences have interaction only through amino acid sequence information of proteins without any biological information of protein sequences; Shen et al. Proposed a triplet coding algorithm to represent protein sequences, and trained them with SVM And prediction, with an accuracy of 83.9%; Guo et al. proposed to extract sequence features of proteins by self-covariance coding, and the classification algorithm was predicted by SVM based on radial basis kernel function, with an accuracy of 88.09%; the most important of these methods is to extract sequence features and efficient classification algorithm. However, the research on the detection of protein molecular defects has yet to be developed.

There are many methods to test the molecular weight of proteins. The most widely used and mature gel electrophoresis methods, gel chromatography, salting-out precipitation and laser light scattering, especially in recent years, are increasingly mature and rapidly developing biomass spectrometry technologies, including electrospray ionization mass spectrometry and matrix-

assisted laser desorption ionization mass spectrometry. The research of artificial neural networks originated in the 1940s. It was first proposed by neurobiologist Meculloch and mathematician Pitts, who jointly opened a road for researchers to study neural networks. Since then, the development of neural networks has gone through the initial stage, the depression stage to the prosperity stage in recent years. In recent years, with the rapid development of neural network technology, it has been widely used in the text, image, voice, recognition and other fields (19); In signal processing, welding, fault diagnosis and other industrial control systems, it has achieved great success and development; in agriculture, mining, geography, transportation, meteorology, power system and other practical fields, it has made significant improvement, and now it has become an important field of artificial intelligence research. After more than half a century of in-depth research, the neural network has become an increasingly mature and widely used subject. Therefore, this paper proposes a method of protein molecular defect detection based on a neural network algorithm to achieve high-precision protein molecular defect detection.

Prediction method of protein secondary structure based on generalized regression neural network

Protein structure prediction is an important task in the post-genomic era, and protein secondary structure prediction is the key step of protein feature extraction. At present, the speed of the protein secondary structure experiment is far behind that of the primary structure (amino acid sequence). Therefore, it is necessary to predict the secondary structure of protein theoretically based on the information provided by the primary structure of the protein. The prediction of protein secondary structure is to find out whether the corresponding structure of each amino acid in the protein sequence composed of 20 amino acids is spiral, β-folding or other types. The protein sequence determines its structure and structure determines its function. The prediction of protein secondary structure from amino acid sequence is of great practical significance in understanding the relationship between protein structure and function, as well as in molecular detection, biopharmaceutical and other fields (20).

GRNN system structure

GRNN is a kind of neural network proposed by Donald f. Specht in 1991. GRNN is a nonlinear regression radial basis function neural network based on nonparametric estimation. The network has good generalization performance and strong nonlinear mapping ability. The network training process is actually the process of determining the smooth parameters. When the sample data is scarce, the effect is also good, and finally converges to the optimal regression surface with more sample size (21).

GRNN structure is shown in Figure 1.

GRNN structure is divided into three layers: input layer, hidden layer (radial base layer) and output layer. The first layer of the network is the input layer, which is a column vector $P=(p_1,p_2,...,p_R)$ of $R\times1$ dimension; the input vector is transferred to the hidden layer, which has Q_1 neurons (22-23). $LW_{1,1}$ is the weight matrix of hidden-layer neurons, and the transfer function is the Gaussian function radbas, whose expression is as follows:

$$R_i(x)=\exp\left(-\|x-c\|^2\right)/2\sigma_i^2 \quad \{1\}$$

Among them, σ_i^2 determines the properties of the basic function in the position of the i-th hidden layer. The larger σ_i^2 is, the smoother the basis function is. It is also called the spreading factor. x and c represent the number of radial basis neurons, the number of linear neurons and the number of input vectors.

The input $n^1=\|LW_{1,1}-p\|b^1$ of the hidden layer node and the output $a_i^1=radbas\left(\|_iLW_{1,1}-p\|b_i^1\right)$ of the hidden layer node i, where a_i^1 is the i-th element of the hidden layer output vector a^1, b_i^1 is the i-th element of the hidden layer threshold vector b^1, and $_iLW_{1,1}$ is the i-th line of the weight vector $LW_{1,1}$ of the i-th hidden layer neuron.

The third layer of the GRNN network is the specific linear output layer (24). Firstly, the output of the hidden layer and the weight matrix $LW_{2,1}$ of the layer are normalized dot product, then they are used as weight input and then sent to the transfer function, which is a linear function (25). The output of the network can be expressed as follows:

$$y=a^2=pure\lim\left(LW_{2,1}\times a^1/sum\left(a^1\right)\right) \quad \{2\}$$

In GRNN training, the number of radial basis neurons and linear neurons is the same as the number of input vectors. The training of GRNN is divided into two steps: the first step is learning without teachers, determining the weight $LW_{1,1}$ between the input layer and the hidden layer, and the generated threshold is determined by the parameter spread, $b^1=0.8326/spread$; the second step is learning for teachers, training and generating the weight matrix $LW_{2,1}$ between the hidden layer and the output layer according to the provided target vector set.

Materials and Methods

Data set selection

In this research, 226 datasets were used, excluding the protein chains with an identity greater than 35%. Finally, 77 single chains of 46 protein complexes were selected as the data set of protein secondary structure prediction, 56 of which were randomly selected as the

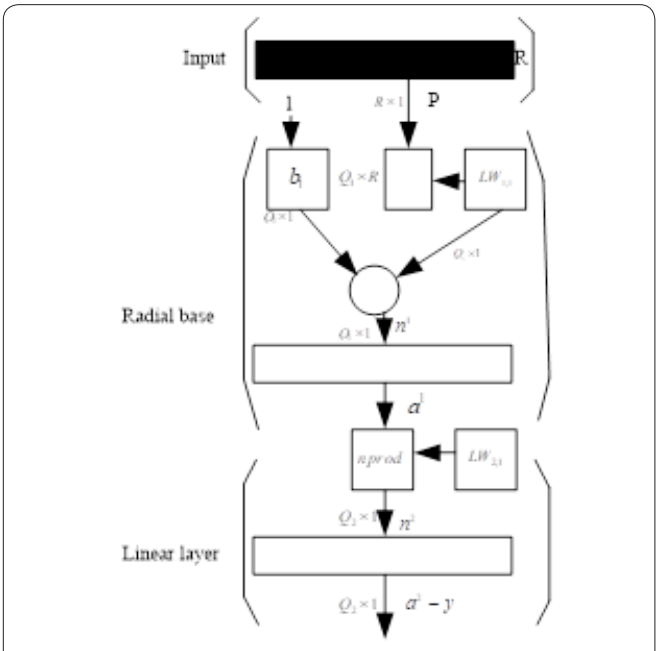

Figure 1. The GRNN model structure.

Table 1. Training data set.

Training data set	Training data set	Test data set
1ABY-A	1AOK-A	1RRL-M
1AGR-E	1AXI-B	1TCR-A
1ATN-A	1BRL-B	1TRN-B
1BFV-L	1EFV-A	1RLB-A
1EBV-L	1GUA-A	1TCR-B
1EBD-A	1LGB-A	1YUH-L
1FRV-A	1PHN-A	1RLB-E
1IBC-B	1AGR-A	1TMC-B
1MHL-C	1AQD-B	2BTF-P
1ABY-B	1BFV-H	1SCU-A
1AIS-A	1CAU-B	1TTR-B
1ATN-D	1FIN-B	2PCC-B
1BPL-A	1IBC-A	1SCU-B
1EBD-C	1MHL-A	1VOL-A
1FRV-B	1RBL-A	2REQ-A
1IGT-B	-	1SEB-D
1MIO-A	-	1YRN-A
1ADO-B	-	2REQ-B

Training concentration: H structure accounted for 40.81%; e structure for 21.49%; C structure for 37.70%.

Test focus: H structure accounts for 40.64%; e structure accounts for 20.58%; C structure accounts for 38.78%.

training datasets (Table 1), and the remaining 21 as the test datasets (26).

Construction of GRNN predictor

In this study, the input vector of the sliding window was constructed based on 5-bit encoding and profile encoding mode rich in biological evolution information, and different spread was set up to create multiple GRNN predictors to predict the secondary structure of the protein. All the networks used in the simulation were constructed with MATLAB neural network toolbox, and GRNN was created with newgrnn (P, T, spread) function, in which P is the input data (vector) and T is the target data (known protein secondary structure vector) (27). During GRNN training, the selection of smoothing parameters has a great impact on the prediction performance of the network, so the spread setting was the key. Experiments showed that the larger the spread, the stronger the network generalization ability. In this study, we set up a different spread to judge its impact on the prediction results. The experimental results showed that the larger the sliding window, the larger the corresponding spread value, because the larger the window, the larger the input of coverage and response input range (28-29).

The input layer of the simulation network encoded the amino acid sequence of the primary structure of the protein. The input layer was optimized and improved by constructing sliding windows of 9, 13 and 17 sizes respectively. The number of neurons in the input layer is 9×5, 13×5 and 17×5 respectively. The number of neurons in the input layer was 9×20, 13×20 and 17×20 respectively. The number of neurons in the hidden layer was equal to the number of vectors in the input layer. The training data set was used to build the GRNN predictor, and then the trained GRNN was used to predict the test data. The output layer vector encoded the secondary structure of the protein and represented H (α helix), E (β folding) and C (random curl) structure with [100], [010] and [001], respectively (30). The output of the network was optimized by the principle of "the winner is the king". Assuming that the output of three neurons in the output layer was (0.30.60.1), the current amino acid structure was determined as E-helix (010) according to the principle of "the winner is the king". By constantly moving the sliding window to the next amino acid position, we can predict the secondary structure of all amino acids one by one, so as to obtain all the characteristics of the protein under the secondary structure (31).

Classification model of protein defect molecular sequence-based on neural network

According to the requirements of protein defect molecular sequence data detection, this study designed a protein defect molecular sequence classification model based on the neural network. The neural network is a new theory and technology which has been widely used in artificial intelligence automatic control in recent decades. With continuous research and development in recent decades, the relevant theory and technology of neural networks have made a breakthrough. The classification technology based on neural network refers to the establishment of the neural network model, which is used to simulate the thinking mode of human thinking, and strive to use a neural network model composed of linear relations to simulate various nonlinear problems in the real world. Because there is no clear feature information of protein data to guide the classification of protein sequences, all kinds of feature-based protein sequence classification algorithms currently used are incomplete feature classification methods, that is, there is no strict current relationship between the determined features of protein data and the results of protein data classification. At present, all kinds of classification algorithms are based on a high probability classification strategy, that is, a protein sequence has a high probability of falling into a certain category of protein classification, so it is considered that the sequence belongs to this category of protein data, and it can be seen that all kinds of protein sequence classification problems studied and implemented at present are nonlinear problems. Therefore, the neural network can be used to classify protein defective molecular sequences, which can achieve a better classification effect in theory (32). Furthermore, support vector machine and others classification techniques are also verified for different purposes (33-35).

In the process of design and implementation, the model of protein defect molecular sequence classification based on neural network is mainly completed in three stages: first, a model of protein defect molecular sequence classification based on the neural network should be preliminarily estimated and established according to the characteristics of protein sequence, data scale and current classification situation. In general, the model adopts a classical neural network structure. By setting one or more hidden layers in the neural network, the linear function relation was used to approximate the nonlinear function problem. In the first design stage, a large number of representative data will be selected for

classification training, and the parameters of the neural network established will be calculated according to the training results, and the neural network classification model which can be used for training data classification relationship is consistent. In the second stage, the new protein molecular data are tested by the established data network classification model, and the rationality of data network design will be analyzed according to the classification results of these protein defect molecular sequences in the test process. If there is a big difference between the classification results of the designed neural network and the actual protein defect molecular sequence classification results, it is necessary to analyze the relevant parameters of the neural network, find the causes of errors and adjust the relevant parameters of the neural network and the threshold value of classification, so that the established neural network classification algorithm can be applied to the actual protein sequence classification. The third stage will be the application stage of the neural network classification algorithm. By classifying and testing the established neural network model and various protein sequence data generated in the actual sequencing process, the accuracy and accuracy of the defect molecule classification results obtained by the classification model are calculated (36). In addition, there are lots of classifications techniques using data mining, machine learning for various diseases detection (37, 38).

The structure of the neural network-based protein defect molecular sequence classification model is shown in Figure 2.

As can be seen in Figure 2, the neural network established in this paper was divided into three layers. The first layer was the input layer, which is mainly responsible for receiving the characteristic parameters of each protein sequence to be classified, and the second layer is the hidden layer. The hidden layer is an important layer used to describe and describe the linear relationship of neural networks. By setting the relevant condition parameters, a high-precision linear change function relationship from the input layer to the hidden layer can be realized (39). This kind of high precision linear variation function relation can realize the accurate simulation of nonlinear problems. The outermost layer is the output layer. The output layer will sum the results calculated by the hidden layer and make the final decision. The internal linear relation T of the neural network designed in this paper is shown in Formula 3, which describes the functional relationship between the input layer and the hidden layer. The functional relationship between the output layer and the hidden layer is described by Formula 4. Formula 4 indicates the influence of the output results of each hidden layer on the final neural network judgment results (40).

$$\begin{cases} T1 = \sum_{i-1}^{7} c_{1i} x_i - Q_1 \\ Tj = \sum_{i-1}^{7} c_{ji} x_i - Q_j \quad (j = 1,2,...,16) \\ T16 = \sum_{i-1}^{7} c_{16} x_i - Q_{16} \end{cases} \quad \{3\}$$

$$T = \sum_{j=1}^{16} T_j \quad \{4\}$$

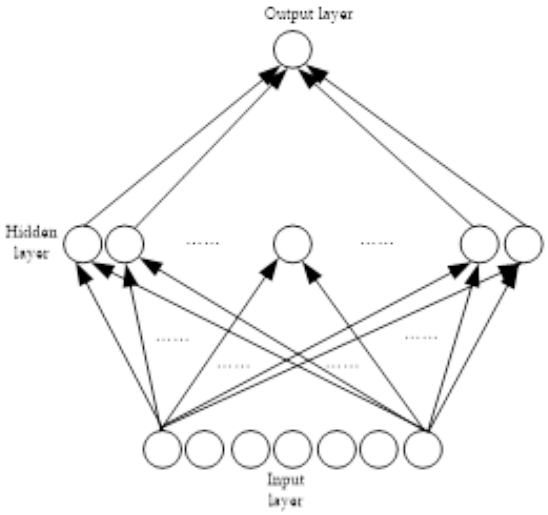

Figure 2. Structure of protein sequence classification model based on the neural network.

Among them, x_i represents the value of each input protein molecular characteristic parameter; c_{ji} represents the influence coefficient of each protein molecular characteristic parameter on the intermediate value of the neural network in this layer; Q_j represents the adjustment threshold of a neural network from the intermediate layer to the output layer, and i and j represent constants.

Results

Protein structure prediction results

The simulation selected the target amino acids and 8, 12 and 16 amino acids adjacent to them in the sequence to form a sliding window input vector with the size of 9, 13 and 17. The sliding window was set to be Q_H, Q_E, Q_C and Q_3 in turn, and a 5-bit coding and profile coding optimization GRNN was constructed to predict the secondary structure of the protein. The prediction results of this method are shown in Table 2.

When GRNN was constructed with 5-bit encoding, different spread values were used for different sliding windows. After repeated tests and experiments, changing the window size cannot continue to improve the prediction performance of the network. In order to further increase the generalization performance of the network and improve the prediction accuracy, the GRNN predictor was further improved by using the Profile coding rich in biological evolution information and setting multiple spreads.

When GRNN was optimized based on profile coding, three spread values were used for different sliding win-

Table 2. Prediction results of GRNN based on 5-bit coding.

Sliding window	Win9	Win13	Win17
Q_H	0.6757	0.6965	0.6009
Q_E	0.3026	0.3474	0.3158
Q_C	0.5733	0.5935	0.5256
Q_3	0.5593	0.5847	0.5131

Table 3. GRNN prediction results based on profile coding.

Sliding window	Win9	Win13	Win17	Sliding window	Win9	Win13	Win17	Sliding window	Win9
	5	6	7		5	6	7		5
Q_H	0.9358	0.9529	0.9529	Q_H	0.9358	0.9529	0.9529	Q_H	0.9358
$aveQ_H$	0.9472			$aveQ_H$	0.9472			$aveQ_H$	0.9472
Q_H	0.5192	0.5326	0.5326	Q_H	0.5192	0.5326	0.5326	Q_H	0.5192
$aveQ_E$	0.5281			$aveQ_E$	0.5281			$aveQ$	0.5281
Q_C	0.7256	0.7315	0.7315	Q_C	0.7256	0.7315	0.7315	Q_C	0.7256
$aveQ_C$	0.7295			$aveQ_C$	0.7295			$aveQ_C$	0.7295
Q_3	0.9873	0.9993	0.9993	Q_3	0.9873	0.9993	0.9993	Q_3	0.9873
$aveQ_3$	0.9953			$aveQ_3$	0.9953			$aveQ_3$	0.9953

Table 4. Test results of this method.

Types of protein sequences	Number of protein sequences to be tested	Number of correct detection sequences	Check the accuracy
Fer4	124	122	0.9839
Atp-synt-ab	149	148	0.9933
Gluts	88	88	1
Lipoclin	56	55	0.9821
Response-reg	95	94	0.9895
Hormone-rec	147	146	0.9932
Adh-short	50	50	1
Adh-zinc	134	132	0.9851
HLH	75	74	0.9867
Cytochrome-c	92	91	0.9891

dows, and the prediction results are shown in Table 3.

According to Table 3, the maximum Q_H, the maximum Q_3, the $aveQ_H$ and the $aveQ$ were 96.24%, 98.96%, 95.58%, and 99.53%, respectively.

The above results proved the effectiveness of this method and the feasibility of simulation.

Defect detection results

By using the neural network model trained in this research, 1000 groups of protein sequences in the actual application process were classified and tested, and the test results are shown in Table 4.

According to the data in Table 4, the method in this paper had a good detection effect, and the overall detection accuracy had reached more than 98%. Moreover, the model of protein defect molecular sequence classification based on a neural network designed in this paper was essentially a way of simulating human classification and thinking, and the whole model had good learning and expansibility. Therefore, in the actual test process, if more representative protein sequences can be selected for model training, the neural network classification model will have higher detection accuracy.

In order to further analyze the detection accuracy of this method, two protein secondary structures, α-helix and β-folding, were set up. By using this method, protein content detection method based on the near-infrared spectrum, protease and protein kinase detection method based on fluorescent protein, the defective protein molecules in the two protein structures were detected. The mean square error of the three methods was tested (Figures 3 and 4).

It can be seen from the analysis of Figures 3 and 4 that the mean square error of the detection of defective protein molecules by the three methods is quite different between the two protein secondary structures. The detection accuracy of this method is always lower than that of protein content detection method based on near-infrared spectrum, protease and protein kinase detection

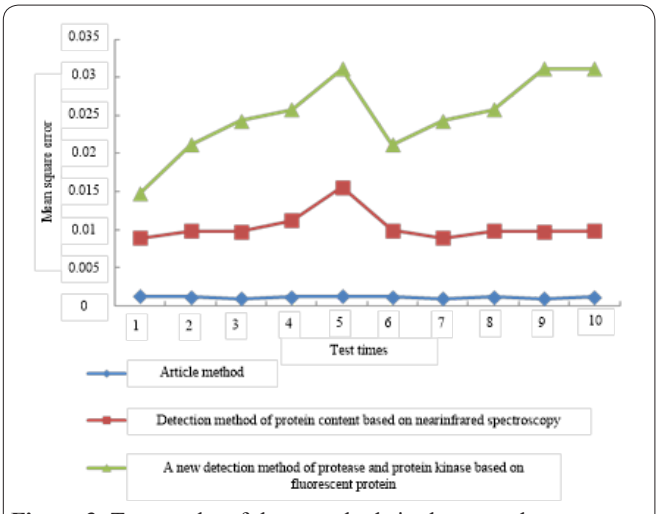

Figure 3. Test results of three methods in the secondary structure of α -helix protein.

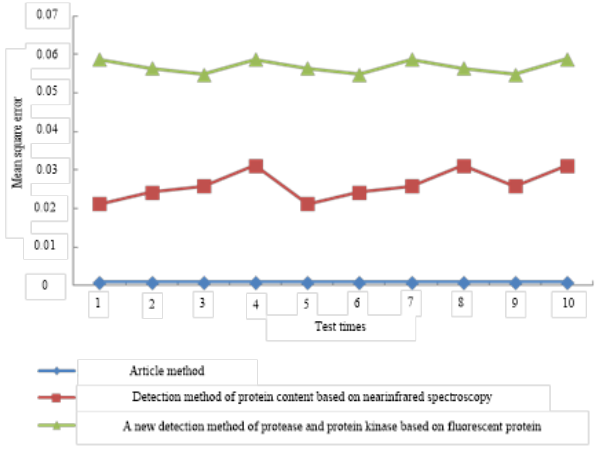

Figure 4. Test results of three methods in the secondary structure of β - helix protein.

method based on fluorescent protein, and the maximum error is no more than 0.0015. It showed that this method can detect protein molecular defects with high precision.

Discussion

In this study, the protein structure was analyzed as follows:

Primary structure

Protein is an organic macromolecule with an amino acid as its basic unit. At present, there are more than 300 kinds of natural amino acids known, but there are only 20 kinds of amino acids that make up the protein of the organism. At present, researchers pay more attention to the research of amino acids in the organism. The structures of these amino acids are all α - amino acids. A carboxyl group and an amino group, a hydrogen atom and an R group are connected to the α - carbon of each amino acid. The difference in the structure of 20 amino acids is that they contain different R groups.

The chemical bond formed in the dehydration condensation process of amino acids is named peptide bond. The process is completed by the amino group of one amino acid and the carboxyl group of another amino acid. The compound formed by the dehydration condensation of the peptide bond is called a peptide. The dehydrated condensation of two molecules of amino acids is called dipeptides. The dehydrated condensation of more amino acids is called polypeptides. The part without dehydration is called residues. Each protein molecule is composed of a peptide chain or multiple peptide chains, and each peptide chain usually contains dozens to hundreds of amino acids. The types of amino acids may vary greatly, and the number of amino acids is even more different. The different spatial structures of peptide chains formed by amino acid dehydration and condensation make the spatial structures of proteins complex and diverse. The primary structure of proteins refers to the arrangement order of amino acid residues that make up the polypeptide chains of proteins. It is the simplest and basic structure of proteins, which determines the structure of protein secondary structure, tertiary structure and advanced structure.

Secondary structure

The secondary structure of protein needs to rely on hydrogen bonds, which is a kind of local spatial structure model of atoms in the main chain. The main skeleton formed by it is a kind of spatial conformation formed in the auxiliary process of the hydrogen bond, but it does not include the interrelation of other segments (including the side chain conformation). Among them, α -helix and β -folding are two common secondary structures of the protein. In this study, we use these two structures as background to verify the detection accuracy of this method.

α-helix: α-helix is mainly right-handed helix. The specific parameters of - helix are given below: its pitch is 0.54 nm, each circle contains 3.6 amino acid residues, each amino acid residue accounts for 0.15 nm, and it rotates 100 degrees around the axis. α -helix is characterized by small steric hindrance, stable conception and easy formation during peptide folding.

β-folding: the formation of β-folding requires the orderly connection of two or more peptide chains with the help of hydrogen bond, and almost all of these peptide chains are extended. We can test the zigzag folding concept. In addition, in β-folding, Cα is always at the folding angle, and the R group of amino acids is at a special position. At the folding angle, the axial distance between each two amino acids has been measured to be 0.35 nm.

Three-level structure

The tertiary structure of the protein is formed by the interaction of hydrogen bond, van der Waals force, hydrophobic force, ionic bond and covalent disulfide bond. These chemical bonds provide necessary maintenance for the tertiary structure of the protein. Its spatial conformation is based on the secondary structure of the protein, and the more complex secondary structure forms a complex and diverse tertiary structure. The tertiary structure of the protein is complex and irregular, but the current research suggests that there are only hundreds to thousands of protein folding types.

In addition to the subject matter, neural networks have been widely used in many other subjects (41-44).

In this paper, a method of protein molecular defect detection based on a neural network algorithm was proposed. In the aspect of protein feature extraction, the prediction method of protein secondary structure based on GRNN was adopted. This method can construct multiple GRNN predictors based on 5-bit coding and different sliding windows, which can predict protein secondary structure with fast prediction efficiency. In the aspect of protein molecular defect detection, it is based on extracted secondary structure characteristics. In order to achieve high-precision detection of protein molecular defects, a model of protein defect molecular sequence classification based on the neural network was designed. After a large number of protein data training on neural network, the detection accuracy of this method is higher than 98%, especially for some new protein molecular defects detection effect is better.

References

1. Mondal S, Mukherjee S, Bagchi B. Origin of diverse time scales in the protein hydration layer solvation dynamics: A simulation study. J

Chem Physic 2017; 147(15): 154901.
2. Aghazadeh R, Zamani M, Motallebi M, Moradyar M. Agrobacterium-Mediated Transformation of the Oryza sativa Thaumatin-Like Protein to Canola (R Line Hyola308) for Enhancing Resistance to Sclerotinia sclerotiorum. Iran J Biotechnol. 2017; 15(3): 201-207.
3. Barkhordari F, Raigani M, Garoosi YT, Mahboudi F, Davami F. Optimization of EnBase Fed-Batch Cultivation to Improve Soluble Fraction Ratio of α-Luffin Ribosome Inactivating Protein. Iran J Biotechnol. 2018; 16(1): e1482.
4. Chen B, Jia Y, Xiong J, Chen L, Luo M, Cheng W, Wang Y, Liu F, Ma S. Effect of Pressure Stresses on Cell Viability and Protein Expression of Fascial Fibroblast. Iran J Biotechnol. 2019; 17(1): e1863.
5. Chen C, Kang C, Rong N, Wu N, Chen C, Wu S, Zhang X, Liu X. Evaluation of Immunogenicity, Protective Immunity on Aquaculture Pathogenic Vibrio and Fermentation of Vibrio alginolyticus Flagellin FlaC Protein. Iran J Biotechnol. 2019; 17(3): e2628.
6. Feizollahzadeh S, Kouhpayeh S, Rahimmansh I, Khanahmad H, Sabzehei F, Ganjalikhani-Hakemi M, Andalib A, Hejazi Z, Rezaei A. The Increase in Protein and Plasmid Yields of E. coli with Optimized Concentration of Ampicillin as Selection Marker. Iran J Biotechnol. 2017; 15(2): 128-134.
7. Hoseini ASH, Mirzarezaee M. Prediction of Protein Sub-Mitochondria Locations Using Protein Interaction Networks. Iran J Biotechnol. 2018; 16(3): e1933.
8. Jia H, Li H, Zhang L, Xu D. Development of a Novel Gene Expression System for Secretory Production of Heterologous Proteins via the General Secretory (Sec) Pathway in Corynebacterium Glutamicum. Iran J Biotechnol. 2018; 16(1): e1746.
9. Khoshtinat Nikkhoi S, Dorostkar R, Ranjbar S, Heydarzadeh H, Tat M, Ghalavand M, Farasat A, Hashemzadeh MS. Synergistic Effect of Expressed miR-128 and Puma Protein on Targeted Induction of Tumor Cell Apoptosis. Iran J Biotechnol. 2016; 14(3): 185-191.
10. Liao Z, Liu H, Liu J, Cai M, Chen T, Hong Q, Luo X, Li X, Ding X, Shen H, Chen D. Heat Shock Protein 60: Identification of an Undetected Allergen from Dermatophagoides farinae. Iran J Biotechnol. 2018; 16(3): e1697.
11. Mohammadi S, Nikkhah M. TiO2 Nanoparticles as Potential Promoting Agents of Fibrillation of α-Synuclein, a Parkinson's Disease-Related Protein. Iran J Biotechnol. 2017; 15(2): 87-94.
12. Mirnurollahi SM, Bolhassani A, Irani S, Davoudi N. Expression and Purification of HCV Core and Core-E1E2 Proteins in Different Bacterial Strains. Iran J Biotechnol. 2015; 13(3): 57-62.
13. Motamedi MJ, Shahsavandi S, Amani J, Kazemi R, Takrim S, Jafari M, Salmanian AH. Immunogenicity of the Multi-Epitopic Recombinant Glycoproteins of Newcastle Disease Virus: Implications for the Serodiagnosis Applications. Iran J Biotechnol. 2018; 16(4): e1749.
14. Sahmani M, Azari S, Tebianian M, Gheibi N, Pourasgari F. Higher Expression Level and Lower Toxicity of Genetically Spliced Rotavirus NSP4 in Comparison to the Full-Length Protein in E. coli. Iran J Biotechnol. 2016; 14(2): 50-57.
15. Savardashtaki A, Sharifi Z, Hamzehlou S, Farajollahi MM. Analysis of Immumoreactivity of Heterologously Expressed Non-structural Protein 4B (NS4B) from Hepatitis C Virus (HCV) Genotype 1a. Iran J Biotechnol. 2015; 13(4): 32-37.
16. Seyfi Abad Shapouri MR, Mahmoodi P, Gharibi D, Ghorbanpoor M, Yaghoubi S, Rezaei E, Rashno M, Mehravar N. Molecular Cloning, Expression and Peroxidase Conjugation of Staphylococcus aureus Protein A. Iran J Biotechnol. 2016; 14(4): 230-235.
17. Chen K., Song X, Liao, Z. A method for calculating the delay bound of 1553B bus by network calculus. J China Acad Electro Info Technol 2019; 2: 129-133.
18. Morinaka BI, Lakis E, Verest M, Helf MJ, Scalvenzi T, Vagstad AL, Sims J, Sunagawa S, Gugger M, Piel J. Natural noncanonical protein splicing yields products with diverse β-amino acid residues. Science 2018; 359(6377): 779-82.
19. Han J, Han J, Huang H. Research on hybrid cascaded multi-level converter with floating capacitor. J Power Suppl 2018; 16: 71-76.
20. Timr S, Kadlec J, Srb P, Ollila OS, Jungwirth P. Calcium sensing by recoverin: effect of protein conformation on ion affinity. J Phys Chem Let 2018; 9: 13-1619.
21. Zhao G, He M, Tang G. Research on re-usage evaluation method of retired lithium-ion power batteries. Chin J Power Sourc 2018; 11: 1632-1634.
22. Borodinova A, Zuzina P, Balaban M. Typical protein kinases in maintenance of long-term memory and synaptic plasticity. Biochem (Moscow), 2017; 82: 243-256.
23. Song L. Research on the development strategy of Intelligent Tourism Based on Internet. Automation Instrumentation 2018; 1: 37-39.
24. Ukuwela AA, Bush AI, Wedd AG, Xiao Z. Glutaredoxins employ parallel monothiol–dithiol mechanisms to catalyze thiol–disulfide exchanges with protein disulfides. Chem Sci 2018; 9(5): 1173-83.
25. Zhan D, Wang R, Min W. Characterization of enzymatic properties of mutant P184Q of aspartate kinase from corynebacterium pekinense. J Jilin University (Science Edition), 2017: 55: 1614-1620.
26. Lu X, Wang C, Li X, Zhao J. Temperature and pH influence adsorption of cellobiohydrolase onto lignin by changing the protein properties. Biores Technol 2017; 245: 819-825.
27. Zhu G, Zhu T, Hong X. Simulation of uniform mixing optimization for sludge agitator. Computer Simulation, 2017; 12: 411-414.
28. Anda A, De Vico L, Hansen T. Intermolecular modes between LH2 bacteriochlorophylls and protein residues: The effect on the excitation energies. J Physic Chem B. 2017; 121(22): 5499-4508.
29. Zhang R, Tao J. A nonlinear fuzzy neural network modeling approach using an improved genetic algorithm. IEEE Transact Indust Electron 2017; 65(7): 5882-5892.
30. Dora S, Sundaram S, Sundararajan N. An interclass margin maximization learning algorithm for evolving spiking neural network. IEEE Transact Cybernetic 2018; 49(3): 989-999.
31. Faris H, Aljarah I, Mirjalili S. Improved monarch butterfly optimization for unconstrained global search and neural network training. Appl Intelligence. 2018; 48(2): 445-64.
32. Vamvoudakis KG, Miranda MF, Hespanha JP. Asymptotically stable adaptive–optimal control algorithm with saturating actuators and relaxed persistence of excitation. IEEE Transact Neural Networks Learning Sys 2015; 27(11): 2386-98.
33. Rahman MZ, Ahmed MS. Flag identification using support vector machine. JU Journal of Information Technology (JIT). 2013 Jun; 2:11-6.
34. Ahmed K, Jesmin T. Comparative Analysis of Data Mining Classification Algorithms in Type-2 Diabetes Prediction Data Using WEKA Approach. International Journal of Science and Engineering. 2014 Oct;7(2):155-60.
35. Billah M, Waheed S, Ahmed K, Hanifa A. Real time traffic sign detection and recognition using adaptive neuro fuzzy inference system. work. 2015 Oct; 3(1):1-5.
36. Yu J, Wu Z, Wang M, Tan M. CPG network optimization for a biomimetic robotic fish via PSO. IEEE Transact Neural Networks Learning Sys 2015; 27(9):1962-1968.
37. Ahmed, K., Al-Emran, A., Jesmin, T., Mukti, R. F., Rahman, M. Z., & Ahmed, F. Early detection of lung cancer risk using data mining. Asian Pacific Journal of Cancer Prevention. 2013; 14(1): 595-598.
38. Ahmed K, Jesmin T, Rahman MZ. Early prevention and detection of skin cancer risk using data mining. International Journal of Computer Applications. 2013; 62(4): 1-6.
39. Fontenla-Romero O, Pérez-Sánchez B, Guijarro-Berdiñas B.

LANN-SVD: a non-iterative SVD-based learning algorithm for one-layer neural networks. IEEE Transact Neural Networks Learning Sys 2017; 29(8): 3900-5.
40. Jung HY, Heo YS. Fingerprint liveness map construction using convolutional neural network. Electron Lett 2018; 54(9): 564-566.
41. Akkurt I, Gunoglu K, Tekin H, Demirci Z, Yegin G, Demir D. Estimation of bremsstrahlung photon fluence from aluminum by artificial neural network. Int J Radiat Res. 2012; 10 (1) :63-65.
42. Abdolmaleki P, Abrishami-Moghddam H, Gity M, Mokhtari-Dizaji M, Mostafa A. Improving the performance of neural network in differentiation of breast tumors using wavelet transformation on dynamic MRI. Int J Radiat Res. 2005; 3 (3) :135-142.
43. Karbalaie K, Vallian S, Lachinani L, Tanhaei S, Baharvand H, Nasr-Esfahani MH. Analysis of Promyelocytic Leukemia in Human Embryonic Carcinoma Stem Cells During Retinoic Acid-Induced Neural Differentiation. Iran J Biotechnol 2016; 14(3): 169-176.
44. Abdolmaleki P, Yarmohammadi M, Gity M. Comparison of logistic regression and neural network models in predicting the outcome of biopsy in breast cancer from MRI findings. Int J Radiat Res. 2004; 1 (4) :217-228