CHAPTER IV

DISCUSSION

4.1 Technical Change in CU Peptide Database

CU Peptide Database originated from its original version that Tanasugarn, 1995 developed. The original database was created in HyperCard, one of the database management programs in Macintosh. It contained about 10 peptides. This research, started two years ago, continued developing the database in HyperCard. One year later, according to the great development of information technology, CU Peptide Database was designed to be accessed via the Internet. HyperCard has a CGI application support for Internet accessibility but the process is very slow. Therefore, FileMaker Pro, which also has a CGI application support for Internet accessibility and has rapid process, was used in place of HyperCard. Then, the research was started from zero again because of the great difference in features between HyperCard and FileMaker Pro. The WEBSTAR server and WEB FM, a CGI application were used for making CU Peptide Database can be accessed via the Internet. Information technology change rapidly, nowadays, JavaScript play an important role in the Internet. Thus, JavaScript was added into CU Peptide Database web page for facilitate internet users in searching peptide. In the future, CU Peptide Database may be apply to locate in UNIX platform which are common in worldwide server.

4.2 Advantages and Disadvantages of the CU Peptide Database

Biochemical researchers that study about peptide can take several advantages from the CU Peptide Database. The researchers can search the desired peptide and get its information. Results obtained by the calculations can offer preliminary predictions of peptide properties before doing experiments. The database management system can calculate the peptide properties not only existing in the database, but also of an unknown peptide if exact amino acid sequence is known. The calculation algorithm may be applied to protein, especially in case of molecular weight since it do not depend on the condition surrounded the protein.

In addition, in searching for the desired peptide, the database allows users to search in a variety of criteria by using search operators and searching by multiple fields. The more users search by mutiple fields, the more they gains the specific results.

However, the peptide information in this database was obtained by chemical catalogue, not by original papers, so, there may be some mistakes, especially the peptide sequence which give rise to molecular weight, net charge and pI value.

Besides, the net charge and pI value of peptide was obtained by calculating each residue seperately and summed to give total value, that is, the calculation is sequence independent. Therefore, these value may be different from the experiment's value.

4.3 Precision of Calculating Script

In this research, peptide information that come from calculation include molecular weight (MW), net charge of peptide at any pH, isoelectric point (pI) and hydropathic index. Molecular weight of a peptide was calculated as the sum of all atomic mass in all residue of a peptide. The molecular mass of standard amino acids were obtained from a popular biochemistry text book (Voet, 1990).

Equation in calculating net charge was derived from Henderson - Hasselbalsh equation. The calculation were on the basis that all residues of peptide completely expose to the water. Thus the theoretical result from this research may be different from experiment result, especially for larger peptide which tends to be fold into tertiary structure. Besides, this research assumed that peptides are in standard condition (solvent is water and temperature is equal to 25 °C), so it cannot be apply in other condition.

In the calculating of pI, the pI script started calculating the net charge of peptide at a very low pH value. If the net charge was still above zero, the script would run again at 0.0001 pH unit higher than last round, until the net charge approached zero (within -0.0001 to 0.0001 unit). The script would return the pH value, i.e. the calculated isoelectric point. The very small increase of pH give high precision but time consuming. The processing time also depends on size of the peptide and its own pI. If the peptide has low pI, the computer script will terminate faster than peptide that has high pI.

In the calculating of hydropathic index, the script assigns the appropriate hydropathy value to each residue of peptide and then sum the appropriate number of consecutive residue, called span, starting at the amino terminal, within overlapping

segments displaced from each other by one residue. A given sum was then plotted above the middle residue of the segment. Kyte and Doolittle, 1982 pointed out that the appropriate number of residues should be seven to eleven. The hydropathic profiles using span less than seven residues are noisy. On the other hand, using long span more than eleven tends to miss small consistent features. Thus, this research used nine consecutive residue each in calculating one point of hydropathic profiles.

4.4 Comparison the Result from the CU Peptide Database with Data Derived from Experimental Studies

For proteins, it is widely known that the hydropathic index is a good predictor of which portions are inside the protein molecule and which portions are outside (Kyte and Doolittle, 1982). For peptides, it was hypothesized that, owing to the small molecular size, the predictive value of hydropathic indexes may not be significant; the smaller the molecular size, the weaker the correlation. In order to demonstrate the utility of the CU Peptide Database, a case study was conducted in which the hydropathic index of a given peptide (a calculated parameter from the CU Peptide Database) was correlated with the relative location of amino acid residue side chain i.e. on the inside or on the outside of the peptide molecule (derived from 3-dimensional data in PDB). Since the three dimensional structure data of peptide can be obtained from PDB, cross-references of 971 peptide in this database to PDB were performed. Ideally, one would like to have as many sample peptides as possible to test the hypothesis. Unfortunately, only three peptide, Charybdotoxin, Echistatin and Glucagon, were found both in PDB and the CU Peptide Database.

The profiles of hydropathic index of the three peptides (figure 3-6, 3-8 and 3-10) did not seem to correlate with the observation result of three dimensional peptide structure through the visualization program. In spite of the small sample size (n=3), the poor correlation seems to support the hypothesis.

4.5 Maintenance of the Database

CU Peptide Database is secure in the sense that Internet users cannot add, delete or modify the databases. However, WEB FM, the CGI application used in this research has an option that would allow users to add new peptides into the databases. It may be a new feature in further studies of CU Peptide Database in the future.

4.6 Limitations of CU Peptide Database

✦ The computer script written in this database can calculate molecular weight of any size of peptide or protein. But the scripts for calculating net charge, isoelectric point and hydropathic index of peptide could not apply for peptide with non standard amino acid residues since pK of ionizable group and hydropathy of these residues were unknown.

✦ Since FileMaker Pro 3.0 has no features about creating graph, the result of hydropathic indexes were presented in numeric value.