CHAPTER III

RESULTS

3.1 System Analysis Result

3.1.1 System Analysis of Biochemical Requirements

Biochemical requirements were considered and then fields in the database were created. Considerations were as follow:

✦ A peptide is characterized by its amino acid sequence. All selected peptides have been named for easy references.

✦ Categorization of peptide is necessary in order to group the peptide, mostly by function.

✦ Source of peptide is also important since the peptide of the same name but existing in different organisms may posses significantly difference amino acid sequences potentially leading to different structures and functions.

✦ three-dimension structure of peptide is stabilized by countervailing forces. Disulfide bonds and data regarding hydrophobic interactions are included in this database.

✦ Product code number field was added so that users could place orders for the desired peptides after searching this database since all peptides in the database were collected from several chemical catalogues with product code numbers.

✦ The molecular weight of a peptide relates to several biochemical parameters. First, it relates to the size of the peptide. The total number of amino acid residues can be roughly predicted by the result of molecular weight divided by 100, an average molecular weight of one amino acid. The size of the peptide also relates to hydrodynamic properties, such as the diffusion constant (Bull, 1964). In addition, it may also relate to the immunogenicity of such peptide (Roitt, 1971).

✦ Net charge and pI have practical applications in the purification and identification of peptides.

From above reasons, the following fields were created for storing peptide data: Peptide Name, Sequence, Category, Source, Disulfide Bridges and Product Code Number field. The aforementioned data were collected from chemical catalogues, but MW, net charge at any pH, pI and hydropathic index were obtained from calculations, so, a computer script was created to calculate such properties.

3.1.2 System Analysis of CU Peptide Database Web Pages

Pages were designed for ease of use. First, an introductory page should be created to introduce background, definition and summary of the database. Second, a search form should be included. Users can search the desired peptide by filling in some data they know such as name, category or total amino acid residues of the peptide. Third, the results that match the user's search criteria should be presented in lists ready for the users to select. Finally, details of the peptide selected by the users are shown in the fourth page. Therefore, the CU Peptide Database should be composed of four web pages i.e. Introductory page, Search Form page, Results page and Details page.

In the Search Form page, in order to facilitate users to fill in the form, JavaScript language was used in controlling some interactive events such as

automatically filling some words in the blank form merely at the click of a button or allowing users to select some text to fill in the form without hitting the keyboard.

Some properties of peptide vary with environmental condition surrounding the peptide. For example, net charge of the peptide vary with surrounding pH. Thus, In the Details page, which shows information of the selected peptide, JavaScript was used to calculate immediately the peptide's net charge at any change of pH given by the user in case all residues of such peptide are of standard amino acids, with known pKs. In addition, in order to make the CU Peptide Database capable of calculating MW, net charge at any pH and pI of unknown peptide, JavaScript routines were used.

3.2 CU Peptide Database Web Pages

CU Peptide Database is designed to help users search the property information of peptide of interest. It contains about 1,000 peptides with their information such as name, category, sequence, hydropathy index, source, disulfide bridge, product code number and cross-reference. The web pages contains scripts which calculate the molecular weight (MW), the isoelectric point (pI) and the net charge at any pH. These calculations can be invoked not only on the 1,000 peptides in this database, but also on other peptides of known sequence. In addition, JavaScript is used to make the CU Peptide Database more convenient for users.

Internet users can access the CU Peptide Database via browsers such as Netscape Navigator or Microsoft Internet Explorer. Filling in the Search Form, which contains name, category, sequence, total residue, molecular weight and pI field, one can obtain the property information of the peptide.

3.2.1 Introductory Page

The Introductory page is the entry point or the "Home page" of the database. It gives a summary of the CU Peptide Database and presents hypertext links to other biochemical database (see figure 3-1).

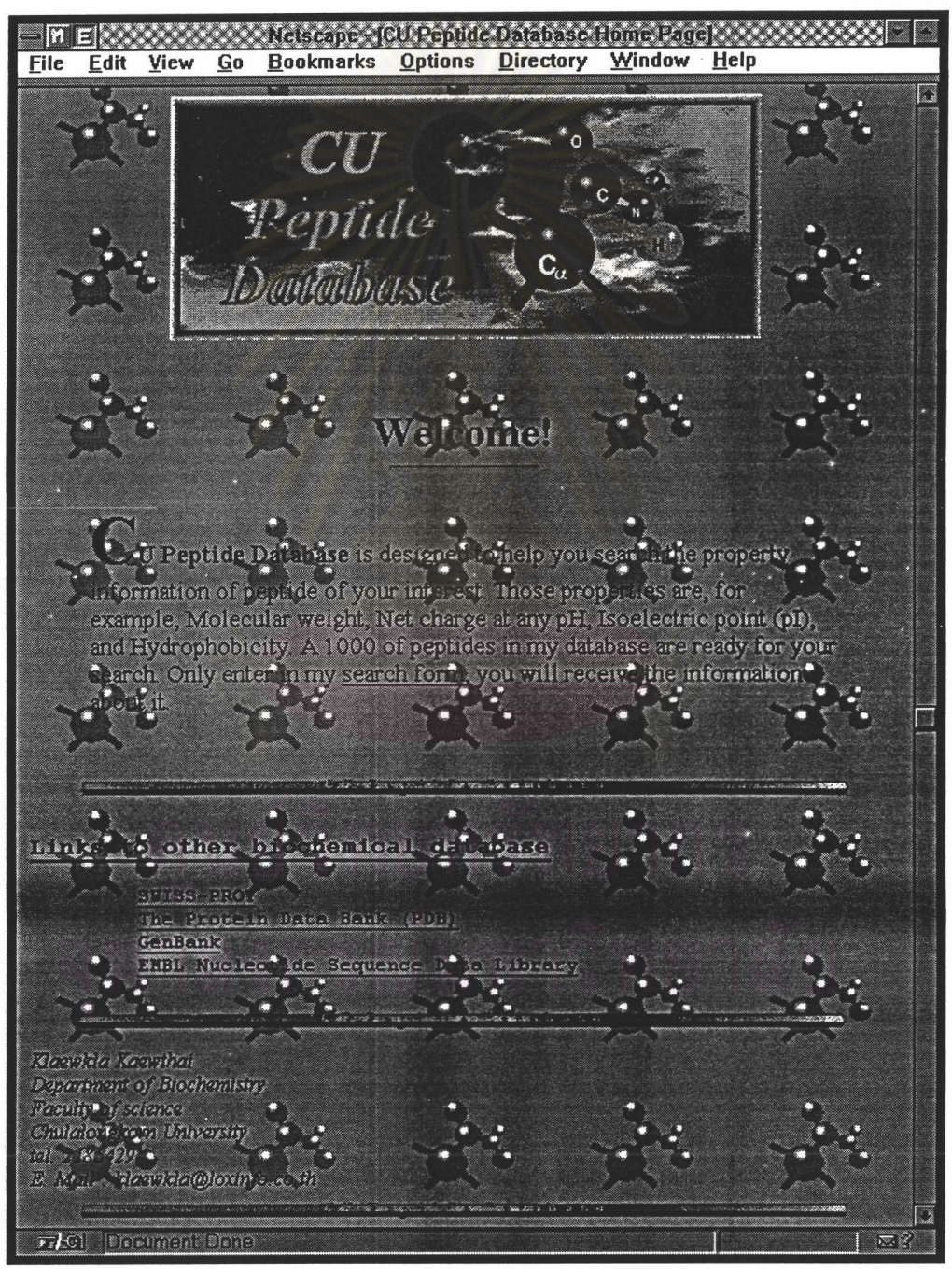

Figure 3-1 Introductory Page

3.2.2 Search Form Page

This web page is a search form for locating the desired peptide using several criteria. Search fields include Category, Peptide Name, Sequence, Source, Total residue, MW and pI (see figure 3-2). All fields except Category field are text field, that is, fields for inputting textual information. The Category field is a drop down list field with 23 choices mostly categorized by function. They are:

1. None (not specify category)
2. Adrenocorticotropic Hormone and Fragments
3. Angiotensin and related peptides
4. Atrial Natriuretic Peptides
5. Bradykinin and Related Peptides
6. Chemotactic Peptides
7. Dynorphin and Related Peptides
8. Endorphins and β-Lipotropin Fragments
9. Enkephalin and Related Peptides
10. Enzyme Inhibitors
11. Fibronectin Fragments and Related Peptides
12. Gastrointestinal Peptides
13. Growth Hormone Releasing Peptides
14. Luteinizing Hormone Releasing Hormone and Related Peptides
15. Melanocyte Stimulating Hormone and Related Peptides
16. Neurotensin and Related Peptides
17. Opioid Peptides
18. Oxytocin, Vasopressin,Vasotocin and Related Peptides
19. Parathyroid Hormone and Fragments
20. Protein Kinase Related Peptides

21. Somatostatin and Related Peptides
22. Substance P and Related Peptides
23. Miscellaneous Peptides

In addition, buttons written in JavaScript were added to facilitate users in filling the form. The buttons are as follow:

- *Select Operator button* is a button for selecting an operator in searching. The operator includes = (equal), < (less than), ≤ (less than or equal), > (greater than), ≥ (greater than or equal), ... (range), @ (single character), * (zero or more character)
- *Clear button* is for deleting all text in the field, every text field has its clear button.
- *Clear All button* is for deleting all text in all text fields.
- *Search Now! button* is a button for submitting search criteria.

The database allows users to search for the desired peptide in a variety of search criteria by using operators. For text fields i.e. Peptide Name, Sequence and Source, using '@' or '*' together with the searching word can yield different results as follow:

- Using the searching word without '@' or '*' means finding the peptide whose name, sequence or source begins with that searching word.
- Beginning the searching word with '*' means finding peptide whose name, sequence or source contains the searching word whether at the beginning, in the middle or at the end of the sequence.

- '@' is used instead of one character. For example, in searching the peptide by sequence, if the searching word is Gl@, it means finding the peptide whose sequence begins with Glu or Gln. Or If the searching word is Ala-@@@-Val, it means finding the peptide whose sequence begins with Ala, followed by any amino acid residue, then followed by Val.

For fields containing numeric value, i.e. total residue, molecular weight, and pI, operators that can use in searching are '=' (equal), '<' (less than), '≤' (less than or equal), '>' (greater than), '≥' (greater than or equal) and '...' (range).

In addition, the database allows users to search by multiple fields. Users can specify searching word, with or without operators, in two or more fields at the same time. the more fields used in searching, the more specific the results are.

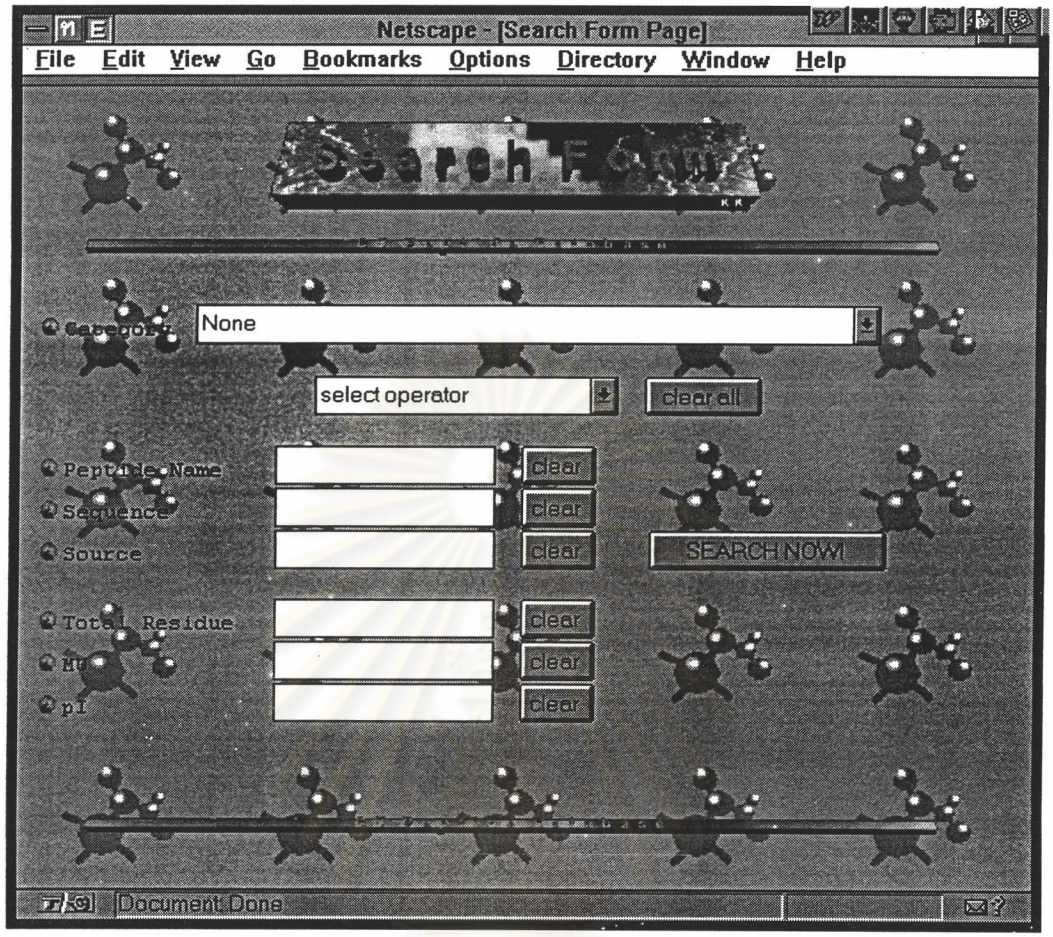

Figure 3-2 Search Form Page

3.2.3 Results Page

Result page shows list of peptide that match the user's search criteria (see figure 3-3). Clicking the name of the desired peptide leads to Details page.

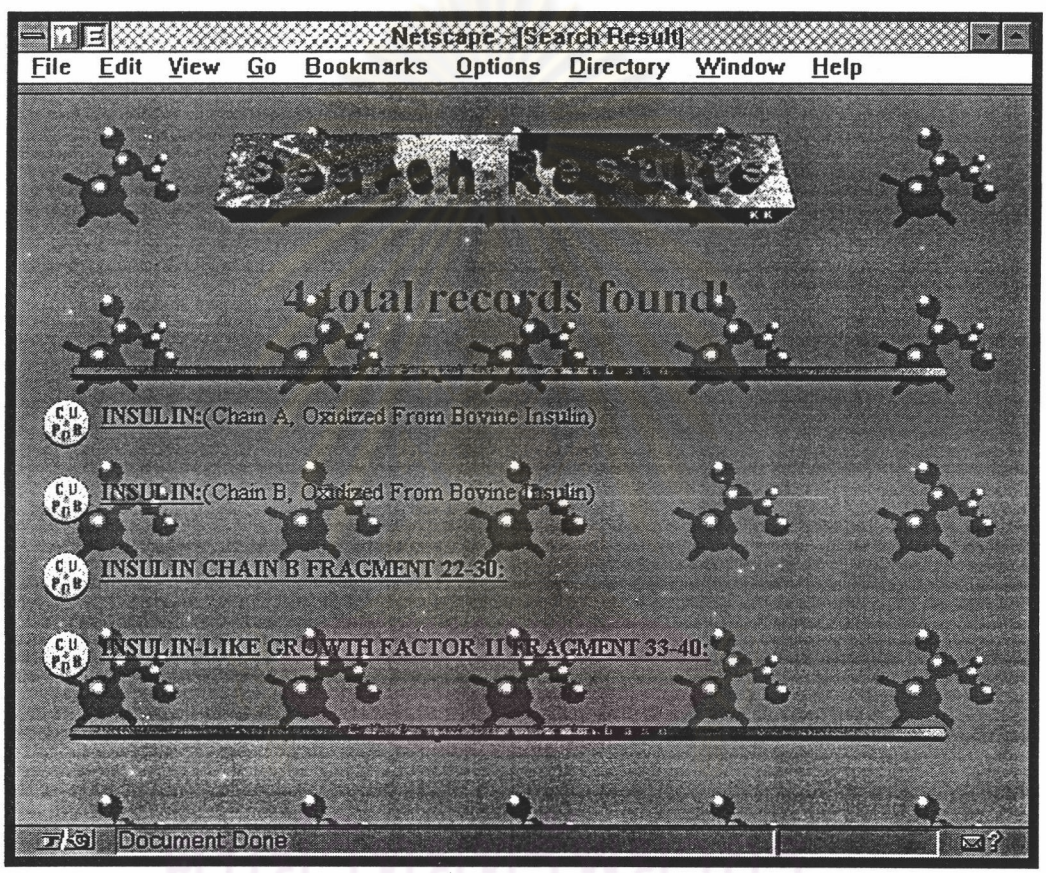

Figure 3-3 Results page

3.2.4 Details Page

Details page has two types. The first type is for a peptide that consists of one or more non-standard amino acids. It gives information about category, sequence, total residue, molecular weight, source, disulfide bridge, product code number and cross reference. (See figure 3-4.) In doing cross reference to SWISS-PROT, 509 related peptides and 7 peptides with the same name and amino acid sequence were found. In doing cross reference to PDB, 20 related peptides and 3 peptides with the same name and amino acid sequence were found.

The second type is for a peptide that consists of standard amino acids. The calculation part was added so that interactive calculations of molecular weight, net charge at any pH and pI can be performed. Calculation Area consists of Sequence, MW, Net Charge and pI fields. The Sequence field has 20 associated buttons for auto-fill amino acid residue and also has a "clear" button for deleting text in the field. The MW, net charge and pI fields each has a button for instructing JavaScript to calculate the MW, net charge and pI selectively when the appropriate button is clicked. The net charge also has special list field for selecting pH value since net charge depends on pH value. Selecting pH value and clicking calculate net charge button will return net charge at that pH (see figure 3-5).

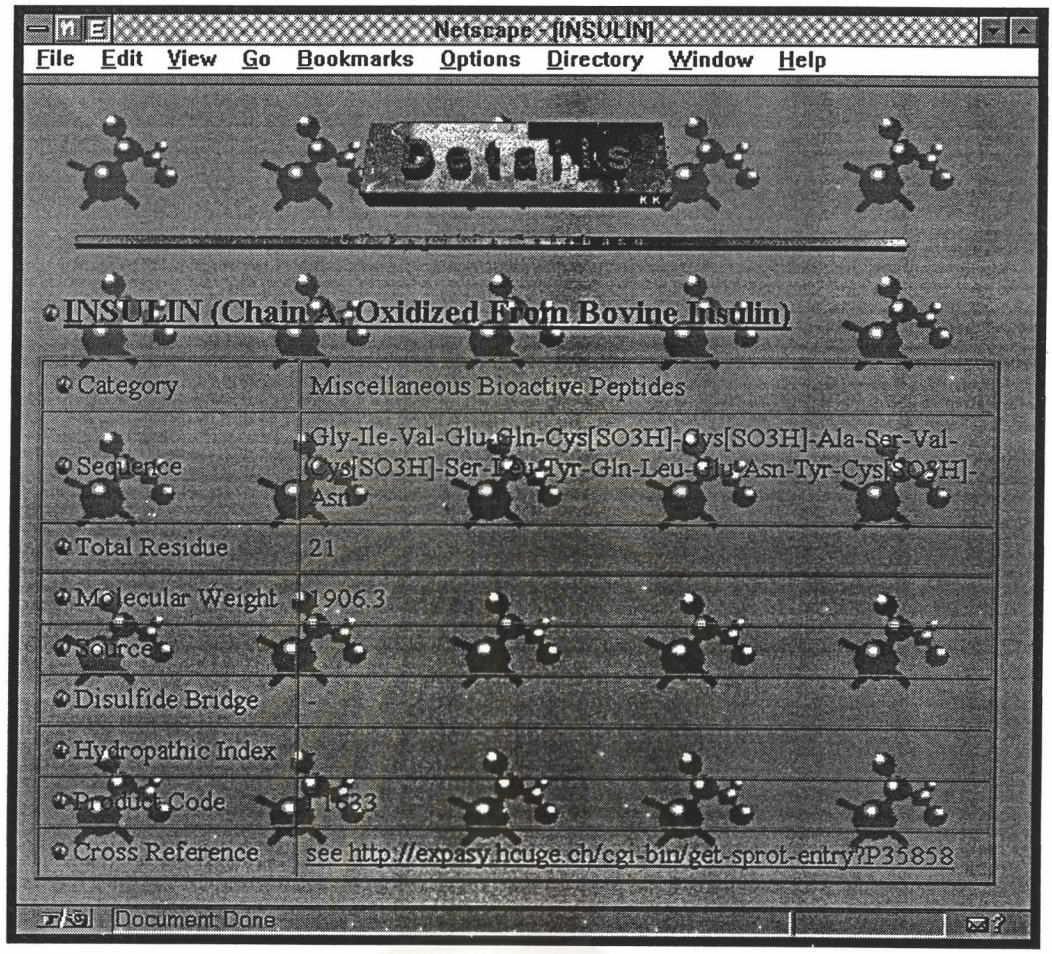

Figure 3-4 **Details page of a peptide that consists of non-standard amino acid residue**

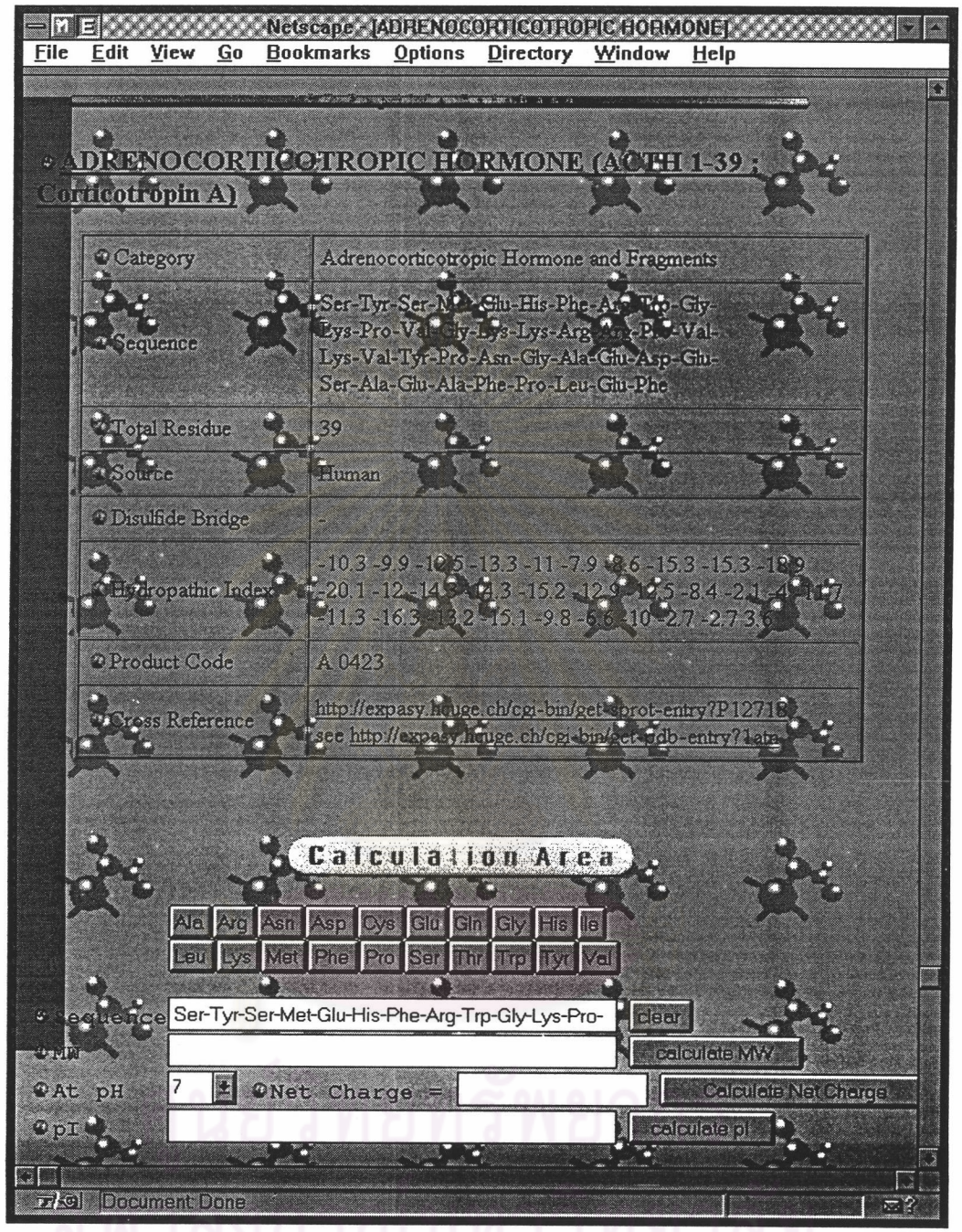

Figure 3-5 Details page of peptide with all standard amino acid residues

3.3 Primary Test of CU Peptide Database

CU Peptide Database web page was tested as follow:

✦ Linking between the four pages and linking to other biochemical database worked well.

✦ In the Search Form page, first, interactive list box and interactive buttons were tested. The interactive box called "select operator" was tested in auto-filling operator in the field that cursor exist and it work effectively. Buttons called "clear" and "clear all" which delete text in each field and all fields selectively also did their functions well. Second, each field was separately tested for searching and the results were satisfactory, that is, the database returned lists of peptide that match the search criteria. Finally, several searching two or more fields at the same time which database binded text in each field with the word "and" gave effectively result.

✦ In the Details page with Calculation Area, interactive buttons and calculation worked effectively. The calculation gave correct results, as predicted theoretically. (See section 2.2.2.2.)

✦ The database was tested on actual users on April 16, 1997. The test report is shown in Appendix A.

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

Only three peptides in this database were found to cross reference with PDB. The three peptides were Charybdotoxin, Echistatin and glucagon. Since hydrophobic forces are a major influence in causing protein to fold into their native structures (Kauzmann, 1950), hydropathic index should be a good predictor of which portions of a peptide are inside a protein and which portions are outside. Thus, the

plot of hydropathic index calculated in this database was compared with the three dimensional structure of peptide in PDB, which came from the experiment. Three dimensional structure of peptide was viewed by *Rasmol*, visualization program.

Hydropathy profile was ploted from hydropathic index calculated in this database (see figure 3-6, 3-8 and 3-10). Profile above zero value of hydropathy tends to be hydrophobic, that is, that range of residue tends to be inside of a peptide molecule. On the other hand, the profile under zero value of hydropathy tends to be hydrophilic and should be outside of a peptide, thus, expose to water. The bar above and under profile are residues inside and outside of a peptide, derived from observation of three dimensional structure of peptide viewing by Rasmol (see figure 3-7, 3-9 and 3-11 and table 3-1, 3-2 and 3-3).

3.4.1 CHARYBDOTOXIN

Category : Miscellaneous Bioactive Peptides

Sequence : Glu-Phe-Thr-Asn-Val-Ser-Cys-Thr-Thr-Ser-Lys-Glu-Cys-Trp-Ser-Val-Cys-Gln-Arg-Leu-His-Asn-Thr-Ser-Arg-Gly-Lys-Cys-Met-Asn-Lys-Lys-Cys-Arg-Cys-Tyr-Ser

Total Residue : 37

Hydropathic Index : -.4, 2.3, -4.4, -7.2, -1.2, -6.3
-6.3, -4.6, -1.4, -4.2, -7.9, -.2, .1, -5.9, -5.7, -5.7,
-14.4, -17.3, -17.7, -10.7, -12.6, -12.9, -13.3, -16.5, -13.2, -13.7,
-10.3, -7.7, -11

MW : 4305.5

pI : 8.8846

Cross Reference : PDB : http://expasy.hcuge.ch/cgi-bin/get-pdb-entry?2crd

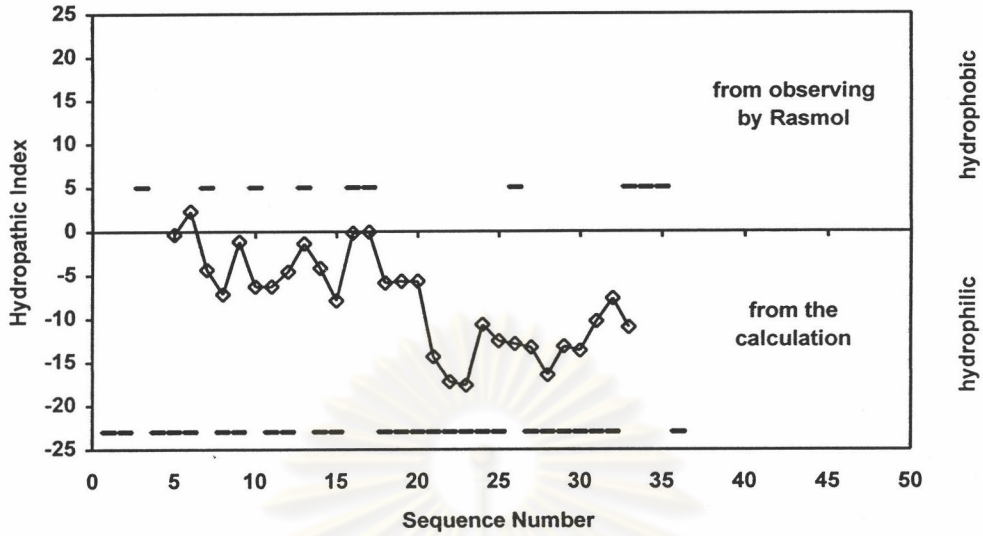

Figure 3-6 Hydropathy profile of Charybdotoxin

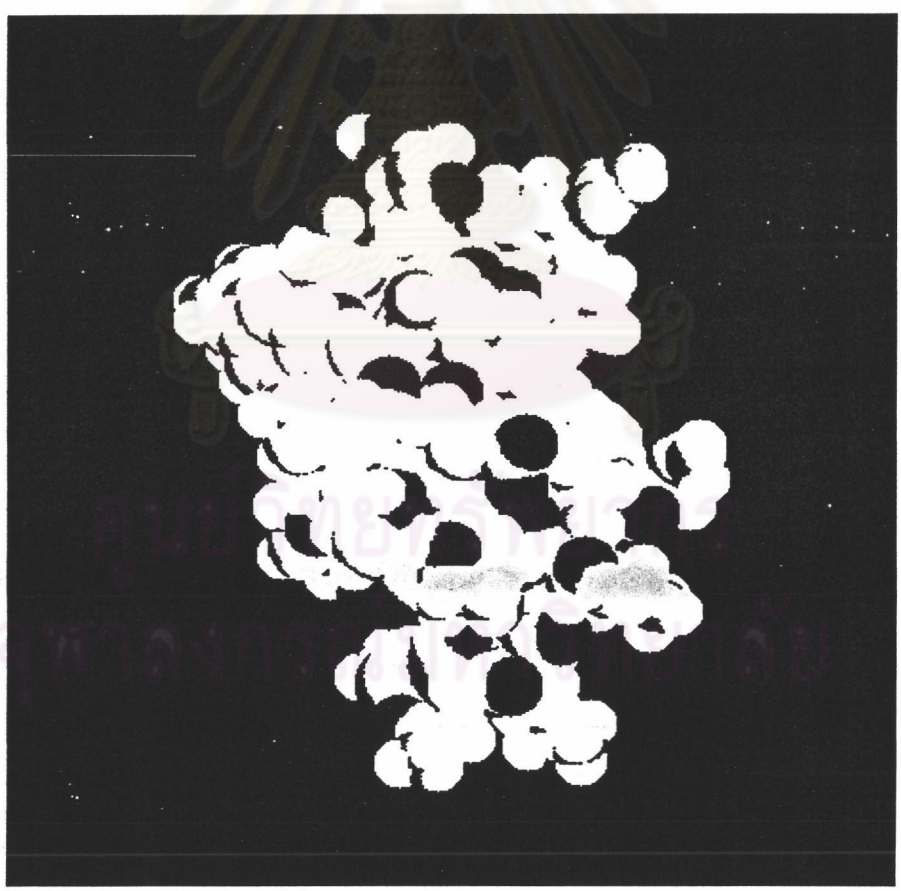

Figure 3-7 Three dimensional structure of Charybdotoxin used in the observation

Table 3-1 Result of observing inside and outside residue of Charybdotoxin

residue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
inside			•				•			•			•			•	•			
outside	♦	♦		♦	♦	♦		♦	♦		♦	♦		♦	♦			♦	♦	♦
residue	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37			
inside						•							•	•	•					
outside	♦	♦	♦	♦	♦		♦	♦	♦	♦	♦	♦				♦	♦			

3.4.2 ECHISTATIN

Category : Fibronectin Fragments and Related Peptides

Sequence : Glu-Cys-Glu-Ser-Gly-Pro-Cys-Cys-Arg-Asn-Cys-Lys-Phe-Leu-Lys-Glu-Gly-Thr-Ile-Cys-Lys-Arg-Ala-Arg-Gly-Asp-Asp-Met-Asp-Asp-Tyr-Cys-Asn-Gly-Lys-Thr-Cys-Asp-Cys-Pro-Arg-Asn-Pro-His-Lys-Gly-Pro-Ala-Thr

Total Residue : 49

Hydropathy Index : -6.8, -6.8, -6.8, -7.2, -3.6, .6,
-1.7, -7.7, -10.6, -6.8, 1.2, 1.2, 1.2, -6.1, -8.1, -8.7,
-5.6, -8.7, -11.5, -14.1, -20.1, -19.7, -16.5, -15.8, -14.8, -14.8,
-15.2, -12.4, -11.8, -11.8, -5.8, -6.1, -13.1, -13.1, -14.3, -13.6,
-16.8, -19.7, -17.8, -18.5, -17.6

MW : 5406.0

pI : 7.4952

Cross Reference : PDB : http://expasy.hcuge.ch/cgi-bin/get-pdb-entry?2ech

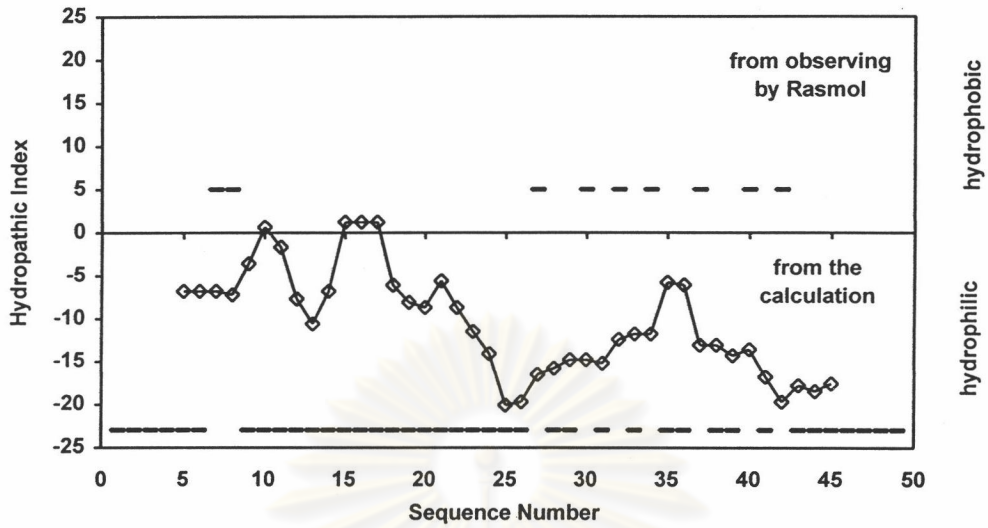

Figure 3-8 Hydropathy profile of Echistatin

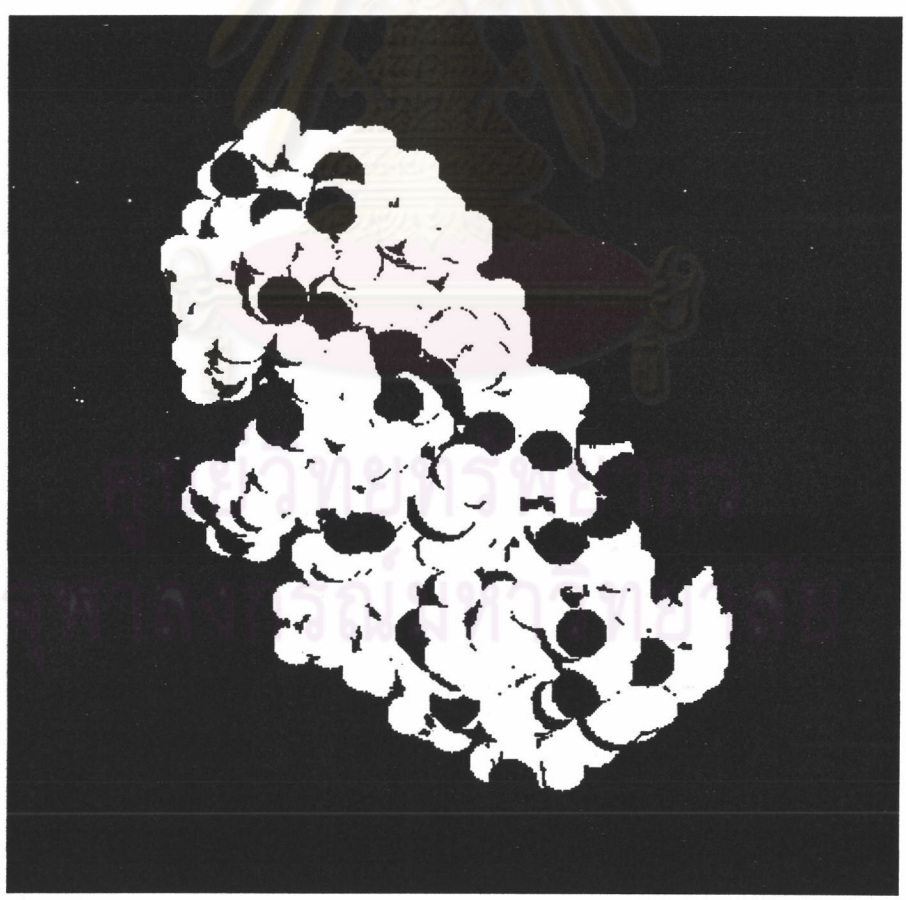

Figure 3-9 Three dimensional structure of Echistatin used in the observation

Table 3-2 Result of observing inside and outside residue of Echistatin

residue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
inside							•	•												
outside	◆	◆	◆	◆	◆	◆			◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
residue	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
inside							•			•		•		•			•			•
outside	◆	◆	◆	◆	◆	◆		◆	◆		◆		◆		◆	◆		◆	◆	
residue	41	42	43	44	45	46	47	48	49											
inside		•																		
outside	◆		◆	◆	◆	◆	◆	◆	◆											

3.4.3 GLUCAGON

Category : Gastrointestinal Peptides

Sequence : His-Ser-Gln-Gly-Thr-Phe-Thr-Ser-Asp-Tyr-Ser-Lys-Tyr-Leu-Asp-Ser-Arg-Arg-Ala-Gln-Asp-Phe-Val-Gln-Trp-Leu-Met-Asn-Thr

Total Residue : 29

Hydropathy Index : -10.8, -8.9, -8.9, -9.3, -10.2, -5.7, -12, -12.1, -15.8, -16.8, -13.7, -16.4, -16, -11.9, -11.5, -11.5, -11.6, -3.3, 3.1, -2.2, .6

MW : 3463.3

pI : 7.6220

PDB : see http://expasy.hcuge.ch/cgi-bin/get-pdb-entry?1gcn

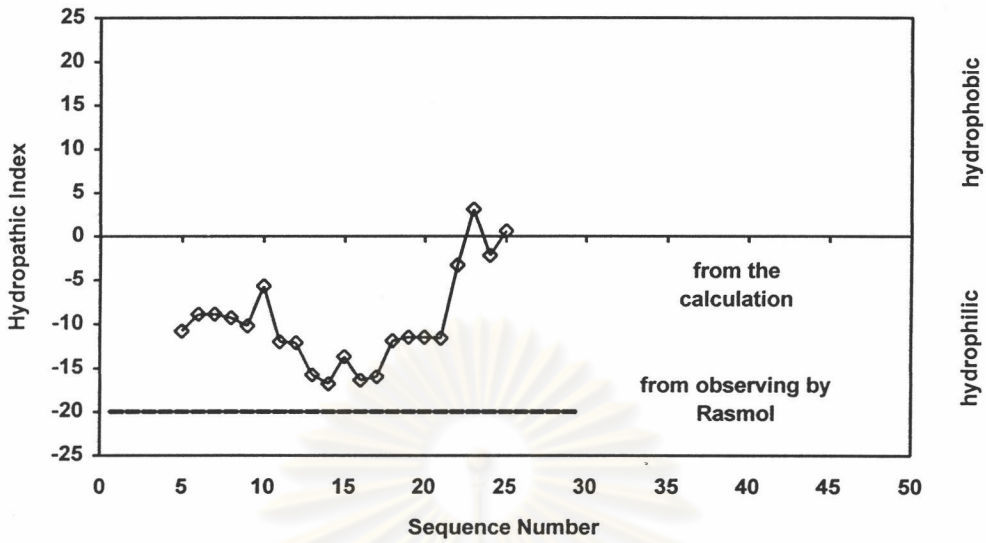

Figure 3-10 Hydropathy profile of Glucagon

Figure 3-11 Three dimensional structure of Glucagon used in the observation

Table 3-3 Result of observing inside and outside residue of Glucagon

residue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
inside																				
outside	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
residue	21	22	23	24	25	26	27	28	29											
inside																				
outside	◆	◆	◆	◆	◆	◆	◆	◆	◆											