Peptide Companion - User's Manual

Contents

Chapter 1 What to expect of Peptide Companion

This program enables one to evaluate elemental and amino acid composition of a peptide, to predict its MS fragmentation and to analyze problems when the molecular weight of your product does not match the theoretical value, and/or to estimate its HPLC behavior on a reversed phase (using eight different algorithms). You can add molecules of acids, salts or water to your peptide and/or try to fit the values found in the elemental analysis to the theoretical value by adding molecules of water and acetic acid (which are usually present in freeze-dried samples).

You may very quickly find the amino acid sequence of any of hundreds peptides (hormones, inhibitors, protein fragments, etc.) from the database and use these structures in calculations or HPLC predictions. You can also modify any peptide bond by any known replacement.

The program analyzes peptide/protein sequences for conformational profiles according to the Chou-Fasman algorithm and evaluates protein profiles and amphipathic profiles. Ten different scales are available (hydrophilicity, hydrophobicity, acrophilicity, accessibility, antigenicity, and hydropathy).

You can calculate the mass distribution of the library synthesized by randomizing an unlimited number of any building blocks in up to 30 positions. You can predict amino acid composition of a peptide based on found molecular weight.

Peptides or proteins can be submitted to simulated enzymatical or chemical digestion and fragments generated in this way can be analyzed in other parts of the program.

Any other type of molecule composed of repeating units, such as RNA, DNA, polysaccharides, can be analyzed using this program. However, the building blocks have to be defined and, understandably, the parts of this program dedicated to peptide and protein analyses will not be available or will provide unreliable results.

Chapter 2 Hardware requirements

To run Peptide Companion you need the following:
- o An IBM 386 , 486 or compatible computer
- o Windows 3.1 or higher

The faster the computer, the better. Evaluation of large proteins, libraries, or amino acid compositions may take considerable time on slow machines. The program needs at least 910KB of space on the hard disk.

Chapter 3 Installation

- o a) Check hardware requirements.
- o b) Make backup copy of your distribution disk:
  Use Copy command in the File menu or Copy Disk command in the disk menu in File Manager, or the Copy or Diskcopy commands in MS-DOS.
- o c) Run Setup:
  - 1. Insert distribution disk in drive A (or B).
  - 2. Choose Run command in the File menu of Program Manager.
  - 3. Type a:setup.
  - 4. Follow instructions on screen. (If you are upgrading from the older version (or from ELEM/PREDICT) and have previously defined your abbreviations or peptides, the computer will ask you if you want to upgrade the files, or leave them as they are.)

  Note: You can run Setup as many times as necessary, however, you will not be asked again for the name and password. As a result, you should be careful of the name under which you are going to register your software. Knowing your password, you can change it any time in the future. (If you ever forget your password, just give us a call.)

- o d) (Optional) Check all installed files:
  If you cannot find one of the following files in the directory which you defined for program installation, try to run the Setup program again or contact CSPS. If you have your data files in another directory, (from previous versions, etc...) you can change the default directory, by selecting "Change Default Directory..." item from the "File" menu. When the program does not find the files on the default directory in the beginning of the program it will ask for their location the same way.

PCOMP.EXE	(Application's .EXE file)
PCHELP.HLP	(Helpfile)
TUTOR.TXT	(Tutorial text)
PEPTIDY.DTB	(Peptide database)
PEPTIDE.DAT, ENZYM1.DAT, ENZYM2.DAT, HP.DTA, HP1.DTA, ZKRATKY.DTA, ZKRFAB.DTA, HESLO.ELM, MENO.ELM	(Data files)
Not required to run program:	
READ.IT	(Changes and additions not covered in the manual)
*.SEQ; *.PEP	(Sequence files that can be used in Predict window)
*.SCL	(Scales for protein predictions)
ITEMS.WRI	(Other items available from CSPS)

o e) (Optional) Run the Tutorial:
Start the program by double clicking on its icon or the computer will ask you if you want to run it at the end of installation. When you click on the "Tutorial" item in the Help menu, the program will restart itself. If you have written any peptide or changed the abbreviations and did not save it, the program will ask you

Figure 1

Tutorial

Welcome to Peptide Companion tutorial.

This tutorial will show step by step instructions to get you started in Peptide Companion.

It is recommended that you read the documentation for Microsoft WindowsTM.

To exit this tutorial, follow instructions on the bottom of this window.

To cancel tutorial and return to normal execution press the "Exit Tutorial" button or press Alt-X

Continue

Exit Tutorial

Screen Capture of Tutorial Pop-up Window

to confirm your action. A small window ^(See Figure 1) will appear, telling you what to do. Press the "Continue" button to continue the tutorial or the "Exit Tutorial" button to exit tutorial. When you exit tutorial you will be at the point in which the tutorial ended. (If you pressed for example the "Exit Tutorial" button when the tutorial was going over the helical wheel, you will be left in the "Helical Wheel" window.) The tutorial goes over only some basic functions. See the help function (press F1 at any moment) for more information.

o f) Take a moment to fill out and mail the Registration Card you received with the software. This will entitle you to (i) the technical support if you will have any difficulty using Peptide Companion; (ii) be automatically notified about upgrades or revisions; (iii) be notified about new programs released and be offered special introductory pricing.

Chapter 4 Nomenclature requirements

Chapter 4.1 Using amino acid and protecting group abbreviations

Chapter 4.1.1 General rules

In order to get satisfactory results you have to adhere to several rules of entering data into the computer. The nomenclature generally follows the published recommendations (e.g. Eur.J.Biochem. 138, 9-37 (1984)). Abbreviations used by the computer are defined below, but you may teach it new abbreviations (up to four letters long). In the definition of new abbreviations you may use the following elements: C, H, N, O, S, F, Cl, Br, I, P, Li, Na, K, Cs, Si. Both three letter and one letter codes can be used for peptide formula entries. Formulas used in the conformation and protein profile evaluation must be in one letter code. However, the program will make the translation for you.

The standard one-letter amino acid code:							
Ala	A	Arg	R	Asn	N	Asp	D	Cys	C
Gln	Q	Glu	E	Gly	G	His	H	Ile	I
Leu	L	Lys	K	Met	M	Phe	F	Pro	P
Ser	S	Thr	T	Trp	W	Tyr	Y	Val	V

Chapter 4.1.2 Defining new abbreviations

To teach the computer a new abbreviation, just enter it alone or in the peptide formula and then try to perform calculation of elemental analysis ("EA, AAA, M.W.") or mass spectrum ("MS Fragmentation"). The computer will respond by displaying the "New Abbreviation Editor" (See Figure 2) window. You will then have to type in the molecular formula (e.g. C5H10NOS) and the description of the abbreviation (e.g. the full name of the protecting group). This will simplify your work later when you may not remember the object for which you have defined the particular abbreviation. If you have used the molecular formula in the peptide formula (such as ...-Ile-Tyr(C3H7)-Gly-...), and you don't want to define C3H7 as the new abbreviation, click on the "Formula" button. In this case, the program will use it as a summary formula.
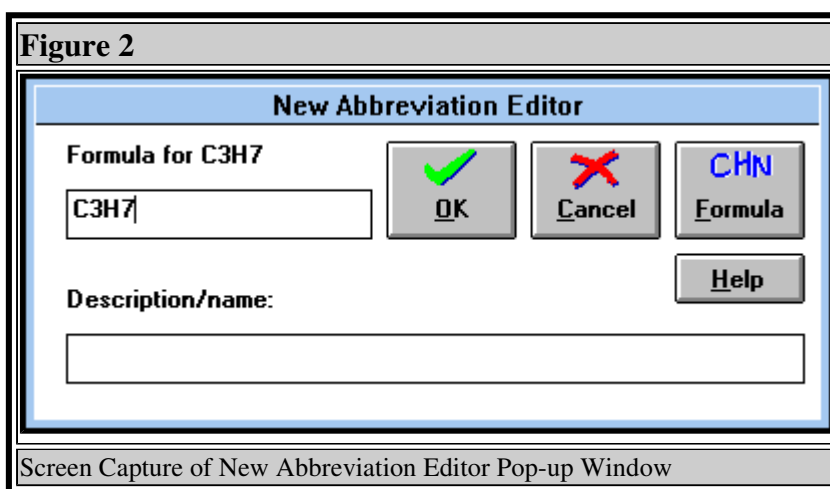
Figure 2

New Abbreviation Editor

Formula for C3H7

C3H7

OK Cancel CHN Formula

Help

Description/name:

Screen Capture of New Abbreviation Editor Pop-up Window

A more direct way to enter a new abbreviation is to press the "New Abbreviation" button. A very similar "Abbreviation Editor" [(See Figure 3)] window is displayed in which the abbreviation can then be defined. The same editor is displayed after selection of an abbreviation from the list and pressing "Edit Abbreviation" button. If you try to define an abbreviation which already exists, the computer will ask whether it should overwrite the old definition with the new one. Note that some "critical" (such as H, OH, NH2) abbreviations cannot be edited. The

Figure 3

Abbreviation Editor

Abbreviation: Asp

Has formula:

C 4 H 4 N1 O2 (default +OH)

Description/name:

residue of aspartic acid (default +H)

OK   Cancel   Help

Screen Capture of Abbreviation Editor Pop-up Window
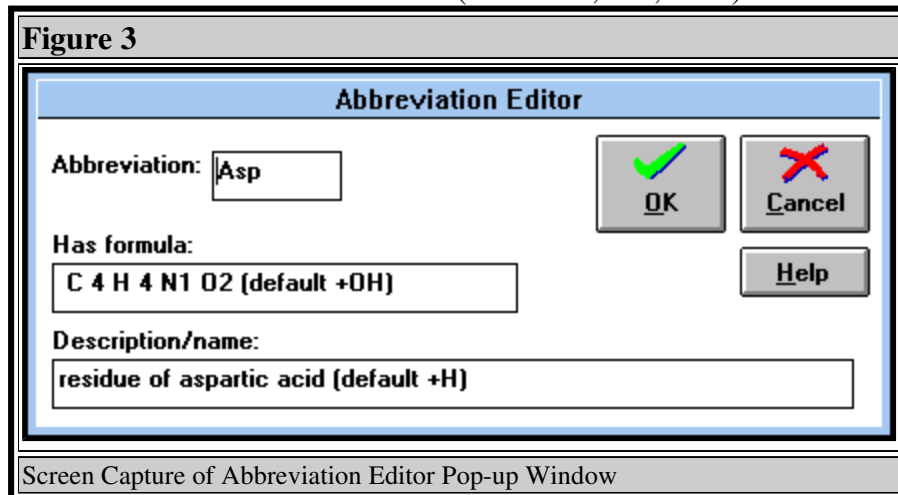
next question the computer asks is "Add default atoms?". If you answer "Yes", you can choose to automatically add an H or OH group to the amino acid residue formula in the cases when abbreviation is not followed by parentheses. This option will be explained in the following sections.

The next question is "Is this unnatural amino acid?". If you answer "Yes", this abbreviation will be interpreted in the amino acid composition calculation as an unnatural amino acid.

At the end of the session (when quitting Peptide Companion) you will be given a chance to save all of the abbreviation changes to the disk. You will have to enter your password (which you have defined during the installation). Any changes must be accompanied by the correct password. You can change the password by clicking at "File" item in the menu line and continuing by clicking at "Change Password". The window "Change Password" will appear and after correctly entering the assigned password, you will be able to define the password of your choice.

Chapter 4.1.3 Trifunctional amino acids

In the case of trifunctional amino acids it is better to define the molecular composition without the side chain H or OH, and then to specify the functionality of the side chain which will be added automatically, when there is no protecting group on the side chain. The program will ask you which default group you want to add in these cases (H, OH or nothing).

Sulfur (SH) containing amino acids (Cys, Pen, Hcy, Mpa, Cpp, Pmp, Dmp, Dep) written in the "-Cys-" form indicates the oxidized state (cyclic) peptide. If you want to specify free cysteine, you have to write "-Cys(H)-".

Chapter 4.1.4 Side chain substitution of amino acids

If you want to substitute the side chain of a normally unsubstituted amino acid, you must compensate for the loss of hydrogen in the side chain. The best solution is to define a new abbreviation for the modified amino acid. For example, "-Phe(Cl)-" will give you a summary formula one hydrogen too big. The correct result can be obtained by using "-PhCl-" as a new

abbreviation and entering C9H8NOCl as the molecular formula of PhCl. (If there is no editing window displayed after using the new abbreviation, then this abbreviation already exists in memory, but it can be defined for something else. In this case check the result carefully or check the composition of this abbreviation!)

Another method of amino acid formula modification is the alteration of an existing abbreviation. For example, to be able to write "-Phe(Cl)-" and obtain the correct result, you have to modify the definition of the "Phe" abbreviation. Find "Phe" in the list of abbreviations, highlight it by the mouse click, and click on the button "Edit Abbreviation". You will be prompted to type the new composition of "Phe". Type the composition one H smaller then that shown on the screen and Enter it by clicking on the "OK" button. Then answer "Overwrite existing abbreviation?" by clicking the "Yes" button, and "Add default atoms? (H,OH)" by again clicking the "Yes" button and then selecting the default atoms from the window "Default Atoms" (See Figure 4). You will now have the

Figure 4
Screen Capture of Default Atoms Pop-up Window
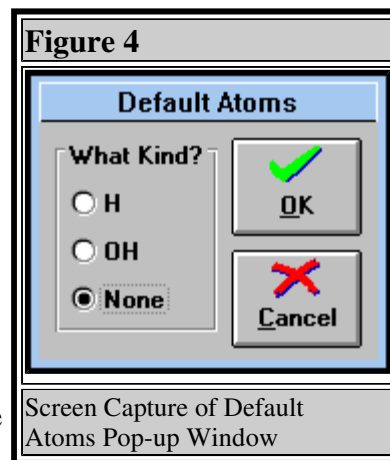
choice of using this newly edited abbreviation always (it can be stored in the datafile) or using it only in this session. For the permanent use you will have to save it to the disk at the end of the session. From now on, if you type "-Phe-" the program will automatically add C9H8NO plus H to the formula of the peptide, and if you type "-Phe(Cl)-", it will add C9H8NO plus Cl.

You must always think what is going on when you substitute the side chain - if you are replacing H (or OH, NH2, etc.) by any substituent (Cl, Br, I, NO2,...), or if you are just adding atoms (e.g. Met - Met(O) - Met(O2)). (Using the wrong nomenclature will produce the wrong result. As the program cannot guess what you want, it will do only what you ask it to do.)

The philosophy of data representation is best explained by the following examples:		
-Ala-	means	-NH-CH(CH3)-CO- (C3H5NO), no default atoms defined
-Asp- or -Asp (OH)-	means	-NH-CH(CH2COOH)-CO- (C4H4NO2 + default atoms O and H)
-Asp()- or -Asp(_)-	means	-NH-CH(CH2CO-)-CO- (C4H4NO2 + no default atoms)
-Asp(OMe)-	means	-NH-CH(CH2COOCH3)-CO- (C4H4NO2 + CH3O)

Chapter 4.2 List of defined abbreviations

Aaa	adamantaneacetyl
Aad	residue of 2-aminoadipic acid (default +OH)
Abu	residue of 2-aminobutyric acid
Ac	acetyl
Acm	acetamidomethyl
Ada	adamantyl
Adoc	adamantyloxycarbonyl

Adpo	1-(1-adamantyl)-1-methylethoxycarbonyl
Ahx	residue of 2-aminohexanoic acid
Ala	residue of alanine
bAla	residue of b-alanine
Ape	residue of 2-aminopentanoic acid
Apm	residue of 2-aminopimelic acid (default +OH)
Arg	residue of arginine (default +H)
Asn	residue of asparagine (default +H)
Asp	residue of (default +OH)
A2bu	residue of 2,4-diaminobutyric acid (default +H)
A2pr	residue of 2,3-diaminopropionic acid (default +H)
Boc	t-butyloxycarbonyl
Bom	benzyloxymethyl
Bpoc	2-(4-biphenylyl)-2-propyloxycarbonyl
Bum	t-butoxymethyl
But	t-butyl
ButS	t-butylthio
BzBr	bromobenzyl
Bzl	benzyl
Cha	residue of cyclohexylalanine
CHA	cyclohexylamine
cHx, Chx	cyclohexyl
Cit	residue of citrulline
Cpp	residue of b,b-cyclopentamethylene-b-mercaptopropionic acid (default -H)
Cyc	residue of S-carboxypropylcysteine (default -OH)
Cys	residue of cysteine (default -H)
Dab	residue of 2,4-diaminobutyric acid (default +H)
DCBz	dichlorobenzyl
DCHA	dicyclohexylamine
Dde	1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl
Ddz	2-(3,5-dimethoxyphenyl)-propyl-2-oxycarbonyl
Dns	dansyl (5-dimethylaminonaphthalene-1-sulfonyl)
Dpr	residue of 2,3-diaminopropionic acid (default +H)
Dep	residue of b,b-diethyl-b-mercaptopropionic acid (default -H)
Dmp	residue of b,b-dimethyl-b-mercaptopropionic acid (def. -H)
Dnp	dinitrophenyl
Et	ethyl
Fme	fluorenylmethyl
Fmoc	fluorenylmethyloxycarbonyl
For	formyl

Gln	residue of glutamine (default +H)
iGln	residue of isoglutamine
Glp	residue of pyroglutamic acid
Glu	residue of glutamic acid (default +OH)
Gly	residue of glycine
Har	residue of homoarginine (default +H)
Hcc	residue of S-carboxyethylhomocysteine (default -OH)
Hcy	residue of homocysteine (default -H)
His	residue of histidine
Hse	residue of homoserine (default +H)
Hsl	residue of homoserinelactone
Hyp	residue of hydroxyproline (default +H)
Ile	residue of isoleucine
Iva	isovaleroyl
Leu	residue of leucine
Lys	residue of lysine (default +H)
Mbh	dimethoxybenzhydryl
Mbs	methoxybenzensulfonyl
Me	methyl
Met	residue of methionine
Mmt	4-methoxytrityl
Mpa, Mpr	residue of b-mercaptopropionic acid (default -H)
Mtr	4-methoxy-2,3,6-benzenesulfonyl
Mts	mesitylene-2-sulfonyl
Mtt	4-methyltrityl
Mur	residue of muramic acid
Nle	residue of norleucine
Nps	o-nitrophenylsulfenyl
Npys	3-nitro-2-pyridinesulfenyl
Nva	residue of norvaline
OBut	t-butyl ester
OBz	benzyl ester
OChx	cyclohehyl ester
ODhb	3-hydroxy-4-oxo-3,4-dihydro-1,2,3-benzotriazinyl ester
OEt	ethyl ester
OMe	methyl ester
ONa	sodium salt
ONp	nitrophenyl ester
OPfp	pentafluorophenyl ester
OPcp	pentachlorophenyl ester

Orn	residue of ornithine (default +H)
OSu	N-hydroxysuccinimide ester
OtBu	t-butyl ester
OTcp	trichlorophenyl ester
Pen	residue of penicillamine (default -H)
pGlu	residue of pyroglytamic acid
Phe	residue of phenylalanine
Phg	residue of phenylglycine
Pms	pentamethylbenzenesulfonyl
Pmc	2,2,5,7,8-pentamethylchroman-6-sulfonyl
Pmp	residue of b,b-cyclopentamethylene-b-mercaptopropionic acid (default -H)
Pro	residue of proline
Sar	residue of sarcosine
Ser	residue of serine (default +H)
Sta	residue of statine
StBu	t-butylthio
Tacm	trimethylacetamidomethyl
TFA	trifluoroacetyl (trifluoroacetic acid is TFA-OH, CF3COOH)
Thi	residue of thienylalanine
Thr	residue of threonine (default +H)
Thx	residue of thyroxine (default +H)
Tic	residue of tetrahydroisoquinoline-3-carboxylic acid
Tos	tosyl
Tacm	trimethylacetamidomethyl
Trp	residue of tryptophan (default +H)
Trt	trityl
Tyr	residue of tyrosine (default +H)
TyrI	residue of iodotyrosine (default +H)
TyI2	residue of diiodotyrosine (default +H)
TyBr	residue of bromotyrosine (default +H)
TyB2	residue of dibromotyrosine (default +H)
Val	residue of valine
Xan	xanthyl
Z	benzyloxycarbonyl
ZBr	bromobenzyloxycarbonyl
ZCl2	dichlorobenzyloxycarbonyl
ZNO2	nitrobenzyloxycarbonyl
ZOMe	methoxybenzyloxycarbonyl

Chapter 4.3 Entering the peptide formula

Chapter 4.3.1 Entering linear, cyclic and branched peptides

The peptide formula is written in the usual way using the above given abbreviations and a dash ("-") to represent a peptide bond.

The amino acid or peptide chain must have specified N and C-terminal functionality. Any peptide without N-terminal H and C-terminal OH is interpreted as a cyclic structure. Additional cyclization (side chain to side chain) may be realized by subtraction of water (addition of -1 H2O) from the structure of a linear peptide, or adding empty parentheses to amino acids engaged in cyclization. Another possibility is to add parentheses containing underline ("_") to amino acid side chains forming the additional cycle. For the rules of disulfide cyclic structure representation see 4.1.3.

Repetition of a structural motif can be expressed by square brackets ("[...]n"). Everything inside of these brackets is counted n-times. This is a useful tool for writing the formulae of polymers, repeating sequences of amino acids, or the formulae of MAPs (in this case, do not forget that in repeating Lys, you must show that the side chain is substituted - add parentheses with an underline "(_)", or empty parentheses "()").

Chapter 4.3.2 Entering peptides with modified peptide bonds

You can modify the peptide bond by replacing the "-" sign in the peptide formula by "<...>" and writing either an abbreviation or a summary formula of the peptide bond replacement within these brackets (e.g. ...TyrIle... or ...GlyGly...).

Chapter 4.3.3 Examples of peptide formula representation acceptable to Peptide Companion

H-Gly-Gly-Phe-Leu-OH-DCHA	
Boc-Arg(Tos)-OPfp	
Ac-Leu-Leu-Arg-H-[H2SO4]0.5	(Leupeptin hemisulfate)
H-Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly-NH2	(oxytocin - disulfide)
H-Cys(H)-Tyr-Ile-Gln-Asn-Cys(H)-Pro-Leu-Gly-NH2	(oxytocein - linear peptide)
Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly	([9-1 cyclo]oxytocin - bicyclic analog)
H-Cys-Tyr-Ile-Glu-Asn-Cys-Pro-Lys-Gly-NH2	([Glu4, Lys8]oxytocin - cyclic disulfide)
H-Cys-Tyr-Ile-Glu(_)-Asn-Cys-Pro-Lys(_)-Gly-NH2	([Glu4, Lys8, 4-8 cyclo]oxytocin - bicyclic analog)
H-Ala-Lys(Ac-Leu-Leu)-Asp(Leu-Leu-OH)-Gly-OH	(branched peptide)
[H-Gly-Pro-Phe]4-[H-Tyr-Ile-Val-Trp]4-[Lys(_)]7-Ala-OH	(example of MAP)

# Chapter 5 Functions of Peptide Companion

Software Peptide Companion is driven by standard Windows tools. You can use either the mouse with its left button to point and click the chosen buttons, or you can use the Tab key to move on the screen and Alt + underlined character keys to perform the chosen task. In the following section, all functions of the software are described.

After you start the program, the "Main" ^(See Figure 5) window ("Peptide Companion") is displayed. It contains the following buttons: "List Peptides", "Write from Scratch", "Three Letter Code", "EA, AAA, M.W.", "HPLC Prediction", "MS Fragmentation", "Difficult Sequences", "Conformation /

Figure 5

Screen Capture of Peptide Companion
Main Window

Profile", "Library", "Digest", "AA Composition from M.W." and "Exit".
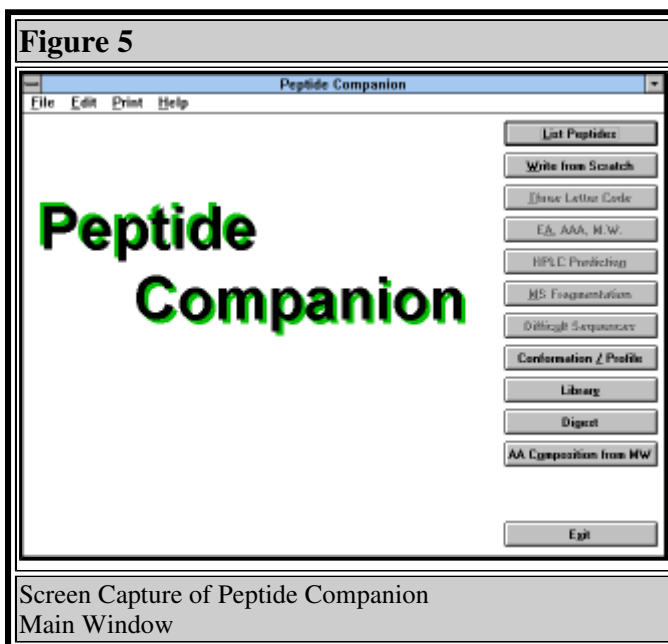
Chapter 5.1 Use or modify peptide formulae (Database of sequences, "List Peptides" ^(See Figure 6) button)

Pressing this button gives you the opportunity to choose the amino acid sequence of the peptides stored in the database. First of all, you can choose the database from which you will select the formula ("Default" or "Choose Dbase" buttons, if you are using a different database than the default database, the button "Current Dbase" will be displayed). You can also change your mind and go back and type the formula by yourself ("Write from Scratch" button). After selection of the database, the list of peptides in the database is displayed. You can highlight any peptide by the mouse pointing or by using Up and Down arrows, PgUp, PgDn, Home and End.

Figure 6

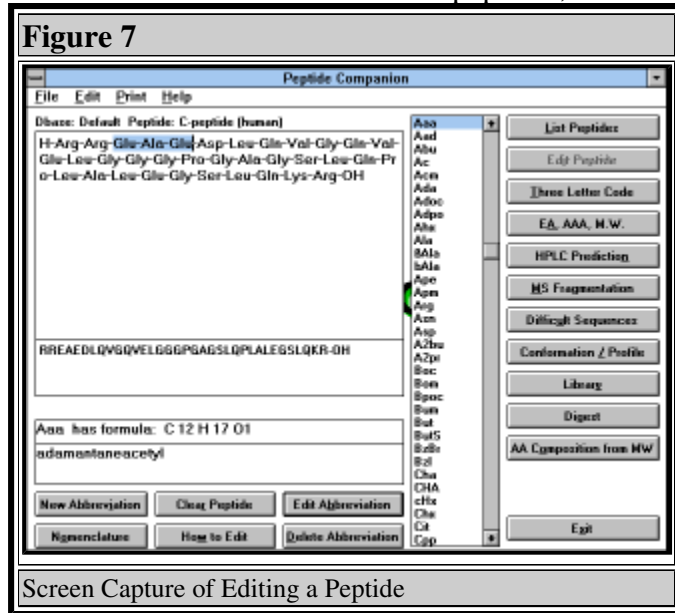Screen Capture of List Peptides Function
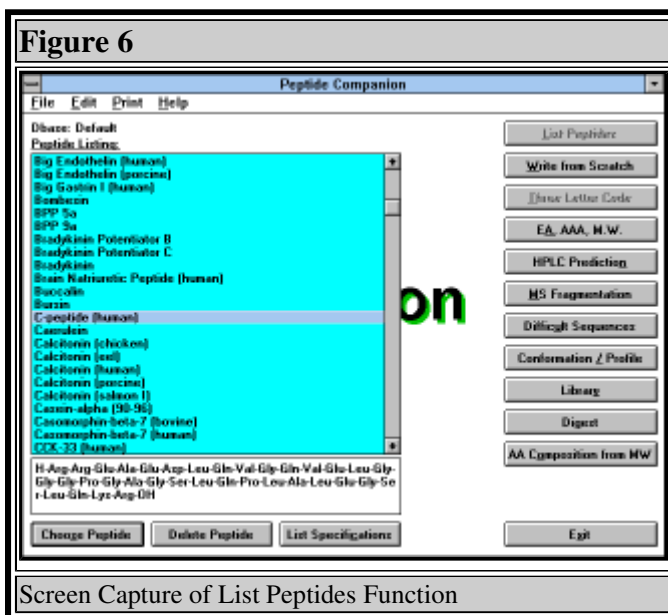
The formula of a selected peptide appears in the lower window. By pressing Enter or by double clicking the mouse, or by clicking on the "Choose Peptide" ^(See Figure 7) button you load the chosen peptide for the editing. If you do not intend to edit the selected peptide, you can start the calculation using this peptide by clicking on buttons "EA, AAA, M.W.", "HPLC Prediction", "MS Fragmentation", "Difficult Sequences", "Conformation / Profile". The
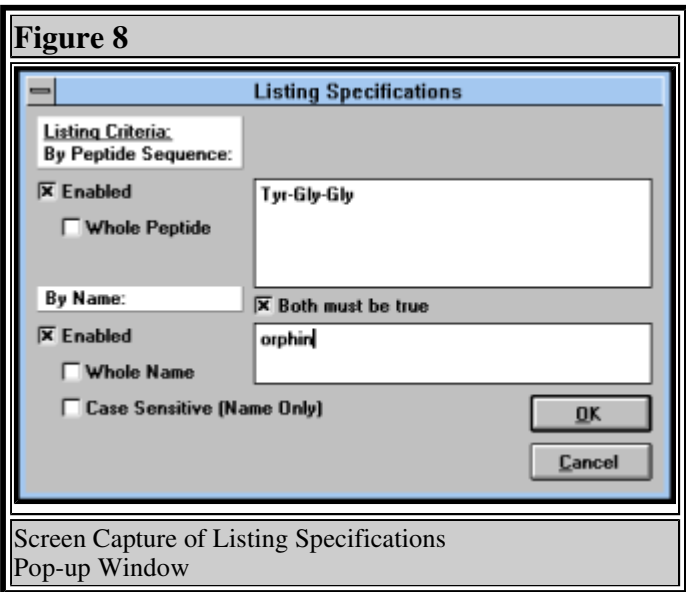
database contains several hundred peptides, and other peptides can be added. Any peptide can be deleted by highlighting it and clicking on the "Delete Peptide" button. To modify the database, you must enter your password. The reason for this is that the peptides in the database were very thoroughly checked and any unauthorized modification might result in sequence information not confirmed by CSPS or by the software licensee.
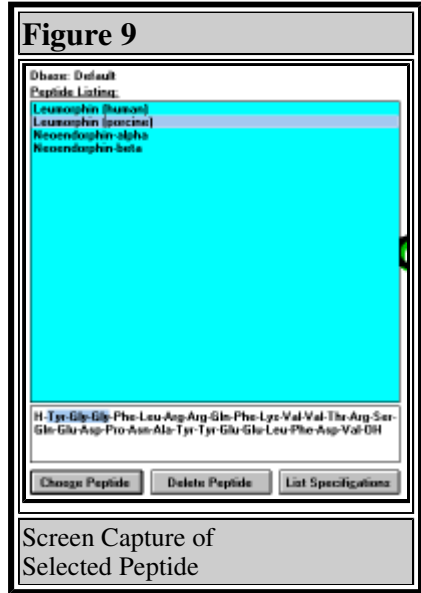
Figure 7

Screen Capture of Editing a Peptide

You can select all peptides from the given database fulfilling the selection criteria. Click on "List Specifications" [(See Figure 8)] and a window which allows you to define a character string which will be sought in the peptide formulae or in peptide names. A search in peptide names can be case sensitive or case insensitive, i.e. "Vaso" and "vaso" can be different or the same -- search in peptide formulae is always case sensitive. You can specify the search for the whole formula or name; when this option is enabled, only fully defined items will be listed. You can combine both conditions in an "OR" or "AND" manner. If "Both must be true" is checked, both conditions have to be fulfilled to give you the positive answer. If this box is not checked, both peptides containing the particular sequence or defined string in their names will be listed.

Figure 8

Screen Capture of Listing Specifications Pop-up Window

Figure 9

Screen Capture of Selected Peptide

Clicking on a peptide name [(See Figure 9)] selected by the use of selection criteria will display its structure in the lower field. If a structural element (partial sequence) was used as the selection criteria, the first occurrence of the defined sequence will be highlighted in the selected peptide formula. If no peptide meets the selection criteria, the original database will be displayed.

Chapter 5.2 Input formula ("Write from Scratch" ^(See Figure 10) button)

If there is no peptide sequence loaded, this button will let you type in the new peptide formula. An alternative way of building the formula is double clicking on the abbreviations displayed in the window with the list of defined abbreviations on the right. You can save the typed peptide formula by clicking on "File" menu and selecting the "Save Peptide" submenu. You will be given a chance to save the peptide formula in either the database which is in use at the moment ("Current Dbase"), in any other database ("Different Dbase"), or you can generate the new database ("New Dbase"). The computer will then ask you for the name of the

Figure 10

Screen Capture of Write Peptide from Scratch

peptide. If a name you specify already exists in the database the program will ask you to confirm overwriting it. A formula not added to the database can be used for the calculations, or it can be modified, but after loading another sequence the prior one is lost. The "Edit Peptide" button is activated when a peptide formula is loaded into the editing window. It uses the same editing tools as defined above.
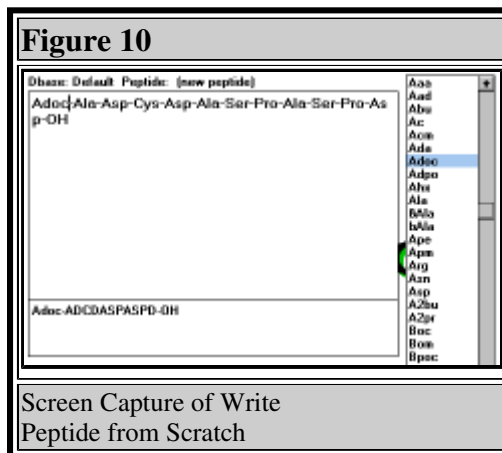
# Chapter 5.3 Translation of codes ("Three Letter Code" button)

You can enter a peptide formula either in the one-letter code or in the three-letter code. One-letter code can be entered in both the upper (bigger) and lower (smaller) windows. Three-letter code typed in the upper window is being translated into one-letter code in the lower window automatically. If the formula is being typed in a one-letter code (in any window), the button "Three Letter Code" must be clicked before any calculation is attempted. The code is then translated into the three letter code in the upper window. You will be warned to check the termini of the peptide since the program is automatically assuming that you want to calculate for an unprotected linear peptide.

Chapter 5.4 Calculation of elemental analysis, amino acid composition and molecular weight ("EA, AAA, M.W." ^(See Figure 11) button)

The results of the calculations are shown on the screen and can also be sent to the printer (see section 6). The information on the screen shows the name of the compound, its formula, summary formula, molecular weight, molecular weight for mass spectroscopy (M.W. (monoisotopic)), elemental composition, and amino acid composition. You can modify the formula in the formula window and recalculate the result ("Recalculate" button). You can also save the peptide formula ("Save Formula" button). The "Other Elements" button will display composition of all elements in the sample.

Figure 11

Screen Capture of Elemental Analysis Window
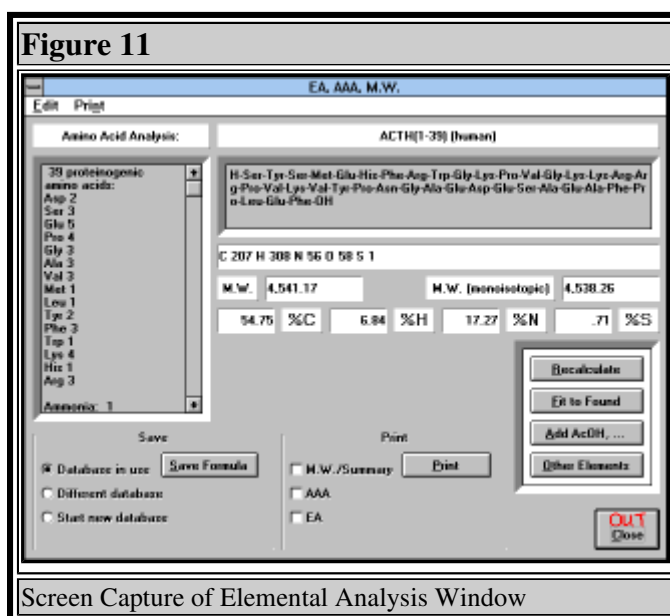
Chapter 5.4.1 The addition of various molecules ("Add AcOH ..." ^(See Figure 12) button) and calculation of probable composition of the lyophilizate ("Fit to Found" button)

Figure 12

Screen Capture of Add AcOH Function
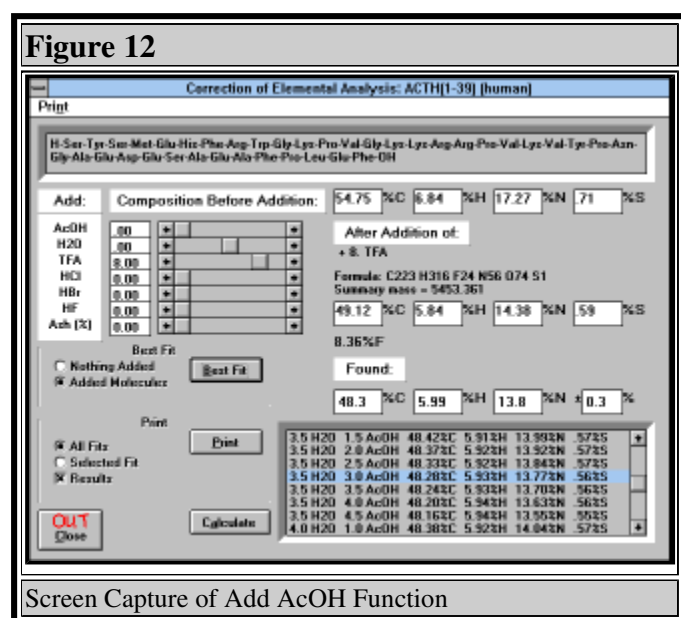
Workup or purification procedures of peptides result in the formation of various salts (trifluoroacetates, hydrofluorides, hydrochlorides, acetates, etc.). The calculation of elemental analysis of peptides with added different molecules may show the most probable composition of the peptide sample.

Clicking buttons "Add AcOH..." or "Fit to Found" transfers you to the window allowing the addition of molecules of AcOH, H2O (also subtraction), TFA, HCl, HBr, HF, and/or inorganic residues (% of Ash). Clicking on the appropriate scroll bar can add molecules in 0.25 or 1.00 increments. Pulling the scroll bar button or entering the value in the value window can enter any value.

Freeze-dried samples of peptides usually contain various amounts of water and acetic acid (if

this was the last solvent system from which the peptide was lyophilized). You can try to fit the found values from the result of elemental analysis to the theoretical values obtained by the addition of molecules of water and acetic acid to the peptide (either with or without added salts). You are prompted to input values found by elemental analysis and a range in % under which you will consider the result satisfactory. Elemental compositions satisfying the condition are calculated. All satisfactory results are displayed in the lower right hand corner with the best fitting analysis highlighted.

Chapter 5.5 Prediction of RP HPLC retention times ("HPLC Prediction" ^(See Figure 13) button)

You may compare the behavior of your peptide with any standard peptide available in your laboratory. This comparison does not give you the absolute value of k', which depends largely on the column, instrument, and other conditions of your experiment, but it gives you the relative elution order of two peptides at two pH values of mobile phase. According to our experience, large deviations from the predicted elution order usually mean that the analyzed sample has other than expected structure (missing or redundant amino acids, uncleaved protecting groups, modified amino acid residues). For the best results always compare peptides of reasonably similar

Figure 13

Screen Capture of HPLC Prediction Window

lengths. Comparison of a peptide with a free amino acid is not sensible.
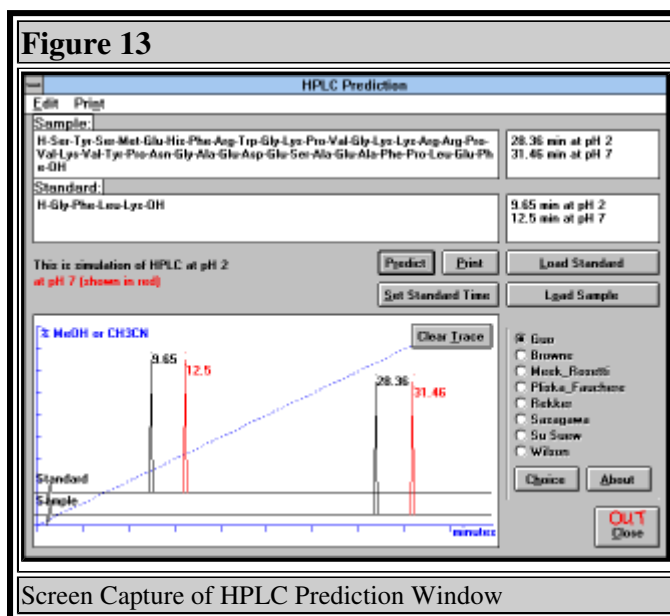
After pressing the "HPLC Prediction" button you will be able to choose the prediction algorithm. You can use one of the following set of parameters:
- Guo D., Mant C.T., Taneja A.K., Parker J.M.R. and Hodges R.S.: J. Chromatogr. 359, 499 (1986) (This is used as the default algorithm)
- Browne C.A., Bennett H.P.J. and Solomon S.: Anal. Biochem. 124, 201 (1982)
- Meek J.L.: Proc. Natl. Acad. Sci. U.S.A. 77, 1632 (1980) and Meek J.L. and Rossetti Z.L.: J. Chromatogr. 211, 15 (1981)
- Pliska V. and Fauchere J.L.: Proc.6th Amer.Pept.Symp., p.249. Pierce 1979
- Rekker R.F.: The Hydrophobic Fragmental Constants, p.301. Elsevier 1977
- Sasagawa T. and Teller D.C.: Handbook of HPLC, vol.II, p.53. CRC Press 1984
- Su Suew J., Grego B., Niven B. and Hearn M.T.W.: J. Liq. Chromatogr. 4, 1745 (1981)
- Wilson K.J., Van Wieringen E., Klauser S., et al: J. Chromatogr. 237, 407 (1982)

If you want to use a generic set, based on the data of Guo D., et al. (J. Chromatogr. 359, 499 (1986)) with the correction for molecular weight (Mant C.T., Burke T.W.L., Black J.A. and Hodges R.S.: J. Chromatogr. 458, 193 (1988)) just press the "Predict" button. This set is the most complete since it was completed by the calculated values for various protecting groups.

If you want to use a different set of parameters, select the desired option and press the "Predict" button again. The new prediction will be overlaid on the trace. If you don't want to overlay the new prediction, press "Clear Trace". The buttons "Load Standard" and "Load Sample" will let you select the appropriate peptide from any database of peptides available to you. Formulas are loaded either by pressing the "OK" button or by double clicking on the chosen name. Peptide formulas can be edited within their windows, or they can be written from scratch. The peptide formula last edited in the main window will be used as the sample in the transfer to the "HPLC Prediction" window. An arbitrary peptide is used as default standard in the prediction, unless otherwise specified.

For better comparison of calculated and experimentally determined values, you can enter retention time of the standard peptide observed in your experiment - - press "Set Standard Time". You will be prompted to enter the "dead time" of your column (which depends on the column dimensions, stationary phase, and flow rate) and retention time of your standard (or retention times of your standard under different conditions). Prediction will then be corrected for the values observed experimentally. Newly entered column parameter (dead time) will be used in all subsequent calculations during the present session.

If you use an abbreviation, for which the retention characteristics are not defined, you will be informed about this fact. However, the prediction will be performed and in the case that both the standard and your peptide contain the same unknown abbreviation, their relative positions will be predicted correctly.

Results of predictions can be printed either as the values or also as the traces. For the details see section 6.
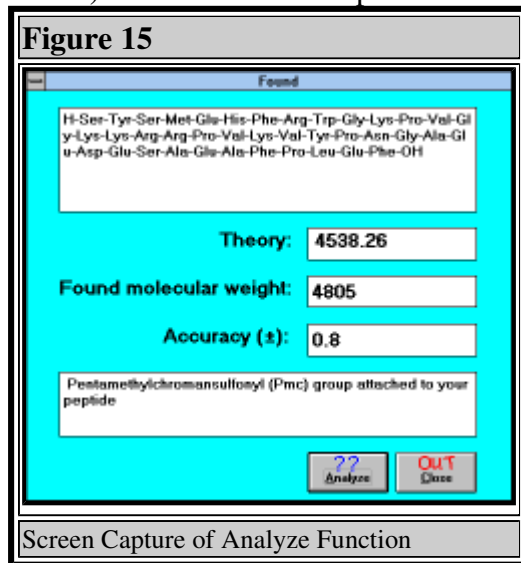
Chapter 5.6 Fragmentation for Mass spectroscopy ("MS Fragmentation" (See Figure 14) button)

This option enables you to generate probable fragments of mass spectra of a peptide. The peptide is cleaved in all peptide bonds, and the list of peptide fragments and a sorted list of all probable ions is produced. However, only unbranched linear peptides give the correct results.

Clicking on any value in the sorted list will highlight the appropriate fragment. Double clicking on the value or fragment structure will display its full structure and values for the appropriate possible types of fragmentation. For more information about fragmentation of peptides see e.g. Stults, J.T. in "Biomedical Applications of Mass

Figure 14

Screen Capture of MS Fragmentation Window

Spectrometry", Vol. 34, p. 145-201, Wiley, New York, 1990, or Bieman, K. in "Methods in Enzymology", Vol. 193, p. 455-479, Academic Press, New York, 1990. The formula of the
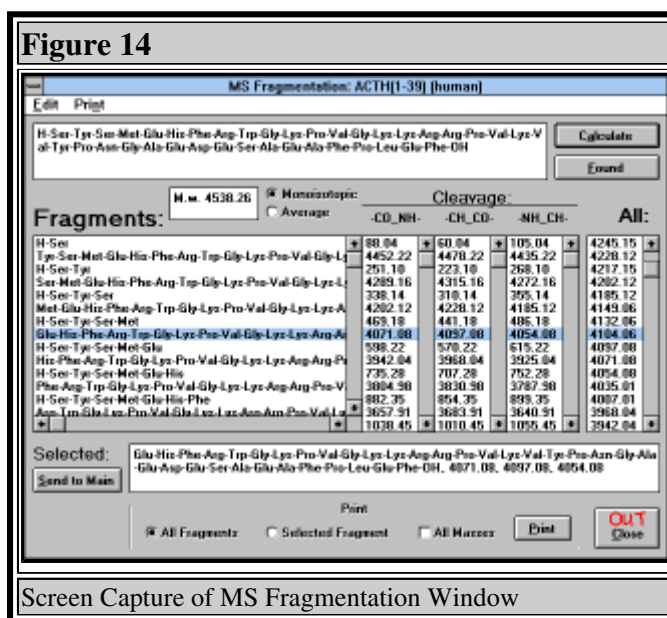
peptide can be edited and a new fragmentation recalculated by clicking on the "Calculate" button). The results can be presented in masses calculated using monoisotopic values (click "Monoisotopic") for all atoms, or using average values (click "Average").

Figure 15

H-Ser-Tyr-Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Val-Gly-Lys-Lys-Arg-Arg-Pro-Val-Lys-Val-Tyr-Pro-Asn-Gly-Ala-Glu-Asp-Glu-Ser-Ala-Glu-Ala-Phe-Pro-Leu-Glu-Phe-OH

Theory: 4538.26

Found molecular weight: 4805

Accuracy (±): 0.8

Pentamethylchromansulfonyl (Pmc) group attached to your peptide

?? Analyze   OUT Close

Screen Capture of Analyze Function

To compare your results from the spectral measurement with the theoretical data, click the button "Found". The displayed window will prompt you to enter the value which you consider to be the molecular peak of your compound. The program will analyze possible reasons for a difference between your value and theory (by pressing the "Analyze" [(See Figure 15)] button or Enter). It will tell you whether you did not completely remove a protecting group, modified the peptide during work up, forgot to couple some amino acid, and so on. Remember that this analysis depends on the precision with which you have determined your experimental data (and which you can change by editing the "Accuracy" window.

Chapter 5.7 Prediction of synthetically difficult sequences ("Difficult Sequences" [(See Figure 16)] button)

The following rules are employed for the prediction of difficult sequences:

Figure 16

Screen Capture of Prediction of Difficult Sequences Window

- ○ (i) Coupling difficulties begin at residue number five from the carboxy terminus. A pendent peptide chain starts to be prone to aggregation after acquiring its fifth residue and its susceptibility to aggregation declines with the increasing peptide length.
- ○ (ii) The carboxy terminal six to nine residues dominantly influence the character of the whole sequence.
- ○ (iii) The sequence starts being potentially difficult after two or three consecutive potentials have exceeded the value 1.1 and ends with the same number of consecutive potentials below 0.9. With the higher aggregation potential comes the higher proportion of pendent peptide chains tending to aggregate. The result is the slower aminoacylation.
- ○ (iv) In most cases aggregation profiles with seven or more consecutive average potentials greater than 1.1 will cause difficult couplings for the entire sequence (even if the potential drops below 0.9 for the rest of the sequence).
- ○ (v) Residues more than ten amino acids distant from the carboxy terminus can be coupled without significant problems even if the aggregation profiles show four consecutive values over 1.1 (but do not show strong aggregation in the carboxy terminal decapeptide).
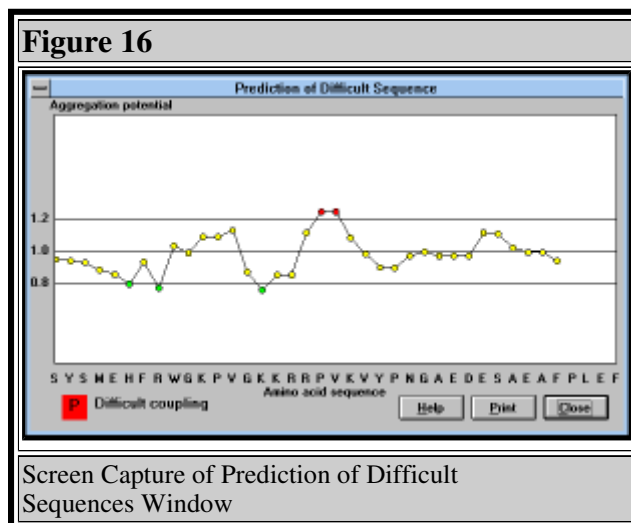
- (vi) The type of the coupled amino acid contributes significantly to the overall aminoacylation rate. Since prediction of difficult sequences by using the aggregation profile eliminates the effect of random coupling problems, the relative reactivity of a coupled amino acid has to be considered. This effect is particularly pronounced when slow reacting amino acids are coupled to a difficult sequence.

Clicking on the "Difficult Sequences" button will show graphical representation of aggregation potential for amino acids starting from the fifth position in the peptide chain. If it is difficult to associate the point on the graph with the amino acid residue on the x axis, point to the particular position with the mouse pointer, and the amino acid label will be displayed in the lower left corner, together with the predicted coupling difficulty. Only sequences containing natural unprotected amino acids can be evaluated since the aggregation parameters are not defined for unnatural building blocks. If you want to use this program to evaluate unnatural sequences, replace unnatural building blocks with natural amino acids having similar coupling properties (make a qualified guess). This program also will not suggest the optimal combination of protecting groups to overcome formation of difficult sequences -- there is not enough experience to write a successful algorithm. However, once the sequence is predicted as difficult, one should consider various synthetic alternatives for the given sequence (various protecting groups, protected asparagine and glutamine, protected backbone, cyclic alternatives of Cys and/or Ser, double couplings, chaotropic salts, high temperature coupling, coupling in ultrasonic bath, etc.).

Chapter 5.8 Evaluation of conformational parameters or protein profile ("Conformation / Profile" button)

Chapter 5.8.1 General rules

This part of the program analyzes peptide/protein sequences. The protein analysis includes calculation of conformational profiles according to the Chou-Fasman algorithm and evaluation of protein profiles and amphipathic profiles. Ten different scales are available (hydrophilicity, hydrophobicity, acrophilicity, accessibility, antigenicity, and hydropathy). The program uses peptide/protein sequences stored in files in standard NIH format (see below); the sequence can be also typed from the keyboard, edited, and stored on disk. Although this option is intended for evaluation of longer peptides and proteins, peptide formulas typed in the main window, or retrieved from the database of peptides can be used. However, only peptides composed of natural amino acids and without any side chain modifications can be evaluated. The program will recognize any unnatural parts of the peptide and truncate the formula.

Amino acid sequences of proteins are stored in files in standard NIH format. Each file can contain any number of comment lines; each comment line must start with a semi-colon (;), and all of these lines are ignored by the program. The line immediately preceding an amino acid sequence contains the name of the protein and it will appear on every output as the protein name. The program uses amino acid sequences written in the standard one-letter amino acid code using upper- or lower-case letters. Peptide/protein sequences written in the three-letter code will be converted to one-letter code (when entered in the main window, or when generated by Digest option).

Standard termination of a sequence is indicated by the digit 1 or 2 or an asterisk (1, 2, or *).

Any further lines are ignored. The present version of the program can run proteins up to 3000 amino acid residues long.

```
; An example of an acceptable file. Semi-colon indicates line with comments
; First line without semi-colon contains the name of a protein
Name of protein
ACSFDGERTHNVMNSDGFARSFEGRILKSLDPLKASKNDKIEKRLNAKSLKDLKAMNVSTAFSREFDGNAMMMSNDNDKLITKLSHFGAKAMS*
```

Chapter 5.8.2 Functions of the "Predict" ^(See Figure 17) window

Clicking on the "Conformation / Profile" button brings you to the "Predict" window. There are three categories of commands:

- (1) Basic commands: "Load/Save", "Setup" and "Close" buttons
- (2) Evaluate commands: "Chou-Fasman", "Protein profile" and "Amphipathic" profile buttons
- (3) View commands:
  - (3.1) View graph containing "Chou-Fasman" and "Profile" buttons
  - (3.2) View consisting of "Chou-Fasman", "Turn", "Profile", and "Helical wheel" buttons

Figure 17

Screen Capture of Predict Window
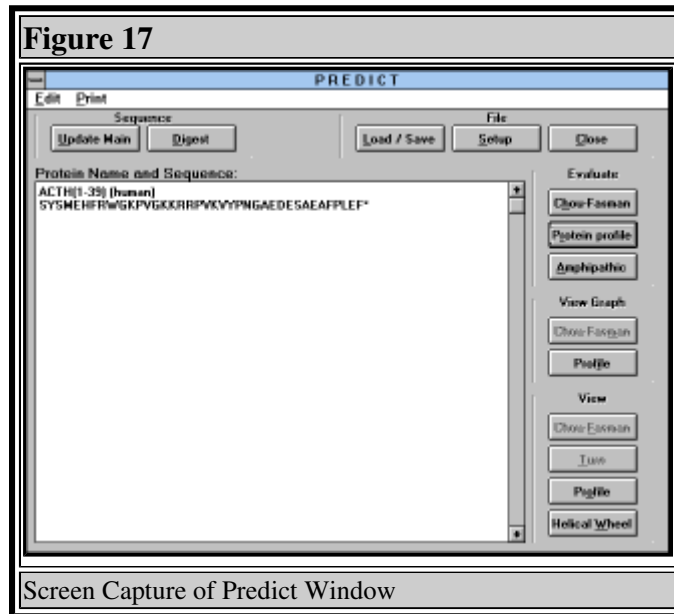
Chapter 5.8.2.1 Basic (File) commands

"Load/Save" button: This option allows you to load (open) a file containing a protein sequence, save the protein sequence in a file, and delete any file (not only those containing protein sequences).

To load a protein sequence, click on the selected file in the current path and directory that contains the amino acid sequence of the peptide/protein, e.g. TEST.SEQ and click on "Load" button (or double click on selected file). The name of the protein (but not the name of the file) and its amino acid sequence appears on the screen. The default directory is the working directory (e.g. C:\PCOM), it can be changed by double clicking on the new path from the "Load/Save" window.

To save the protein sequence in a file (e.g. after editing the sequence or writing a new sequence), type the name of the file and click on the "Save" button. When using a name that already exists on the disk, the old file will be overwritten by the contents of the file currently residing in the memory. However, you will be warned about the possibility of loosing previously stored information.

To delete a file, click on the selected file (or type its name in the File name window), and

click on the "Delete" button. You will be asked to confirm the operation.

"Setup" [See Figure 18] button: This command enables one to change the length of the window in the protein profile, the length of the amphiphilic helix, the probability limit for beta-turn formation, and the path from which data are read. You can use the new values temporarily ("Close and Use New Setup" button), save them to disk ("Save Setup to Disk and Close") or cancel the setup ("Cancel" button).

"Close" button: This command returns you to the Peptide Companion main menu. The currently loaded protein sequence and results of calculation will be lost.

Editing a sequence: The sequence residing currently in the memory can be edited. Peptide Companion uses the standard Windows editor including features Copy, Delete, Paste, and Undo. For details see your Windows manual.

Figure 18

Screen Capture of Setup Pop-up Window
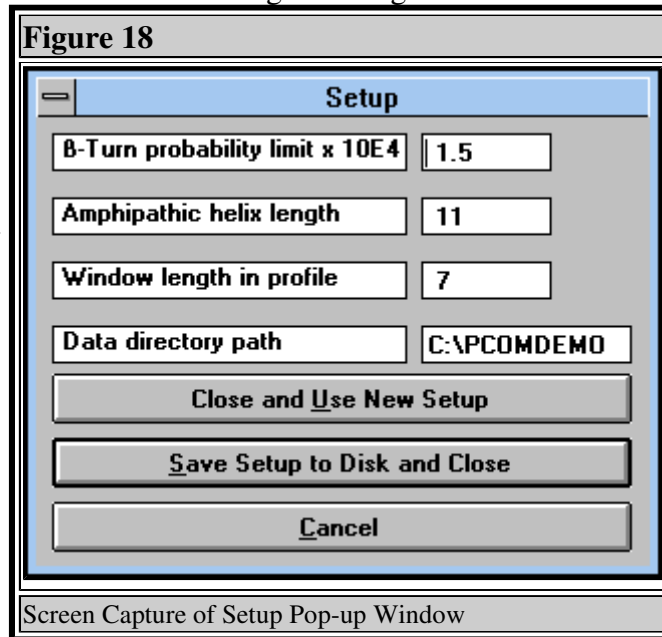
Chapter 5.8.2.2 Evaluate commands

In the "Predict" window you can evaluate protein profiles, amphiphilic profiles, and conformational parameters according to the Chou-Fasman method.

Chou-Fasman parameters ("Chou-Fasman" [See Figure 19] button):

Conformational parameters are evaluated by the Chou-Fasman algorithm, one of the most common secondary-structure prediction methods (Chou, P.Y., and Fasman, G.D. (1978) Adv.Enzymol. 47, 46-148). In this method, each amino acid has been assigned a conformational potential for a-helix, b-sheet, and b-turn formation, derived from the frequency of its occurrence in the particular secondary structure. Then, the mean potential of all consecutive tetrapeptides for all three conformational states (Ph, Ps, and Pt) is calculated. As a measure of propensity to form a b-turn, an extra parameter, the positional probability of b-turn occurrence, p, is calculated by multiplying empirical constants for four consecutive amino acids. According to Chou and Fasman, the tetrapeptide tends to form a b-turn when p 0.75 - 1x10-4.

Figure 19

Screen Capture of Chou-Fasman Method
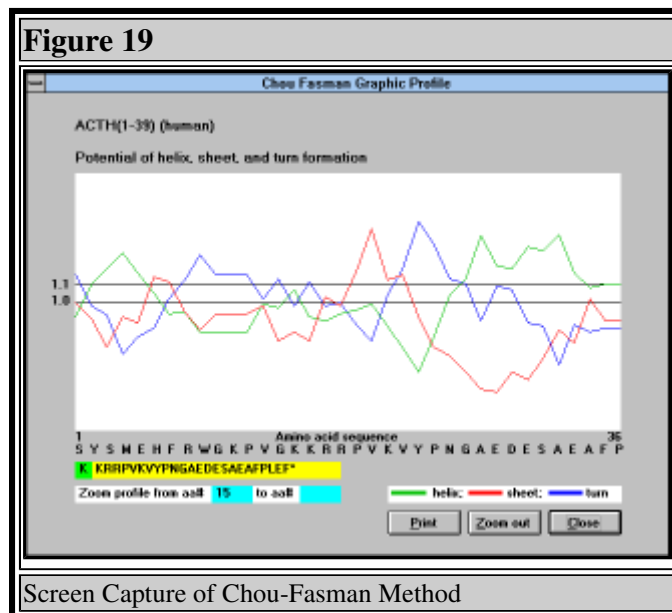
Protein profile ("Protein Profile" button):

Nine different protein profiles can be evaluated. Each amino acid is assigned a numerical parameter value according to the selected scale. The program determines the mean value of a seven-peptide window (default value) moving along the protein sequence. The profiles are normalized for the purpose of comparing different sets of parameters. The mean hydrophilicity (or accessibility, etc.) over the entire protein is calculated and a zero value is arbitrarily set at this average. The values +1 and -1 are set for the maximum and minimum local hydrophilicity, respectively.

The following scales are available:

Figure 20

Screen Capture of Scales Selection Window

- Parker et al. hydrophilicity scale J.M.R.Parker, D.Guo, and R.S.Hodges, Biochemistry 25, 5425 (1986).
- Janin accessibility scale J.Janin, Nature (London) 277, 491 (1979).
- Bulk hydrophobic scale P.Manavalan and P.K.Ponnuswamy, Nature (London) 275, 673 (1978). P.A.Karplus and G.E.Schulz, Naturwissenschaften 72, 212 (1985).
- Hopp and Woods hydrophilicity scale T.P.Hopp and K.R.Woods, Proc.Natl.Acad.Sci. U.S.A. 78, 3824 (1981).
- Fraga global scale S.Fraga, Can.J.Chem. 60, 2606 (1982).
- Welling et al. antigenicity scale G.W.Welling, W.J.Weijer, R.van der Zee, and S.Welling-Wester, FEBS Lett. 188, 215 (1985).
- Hopp acrophilicity scale T.P.Hopp, in Synthetic Peptides in Biology and Medicine (K.Alitalo, P.Partanen, and A.Waheri, eds.), p.3. Elsevier, 1985.
- Kyte and Doolittle hydropathy scale J.Kyte and R.F.Doolittle, J.Mol.Biol. 157, 105 (1982).
- Novotny large sphere accessibility scale J.Novotny, M.Handschumacher, and R.E.Bruccoleri, Immunol. Today 8, 26 (1987).
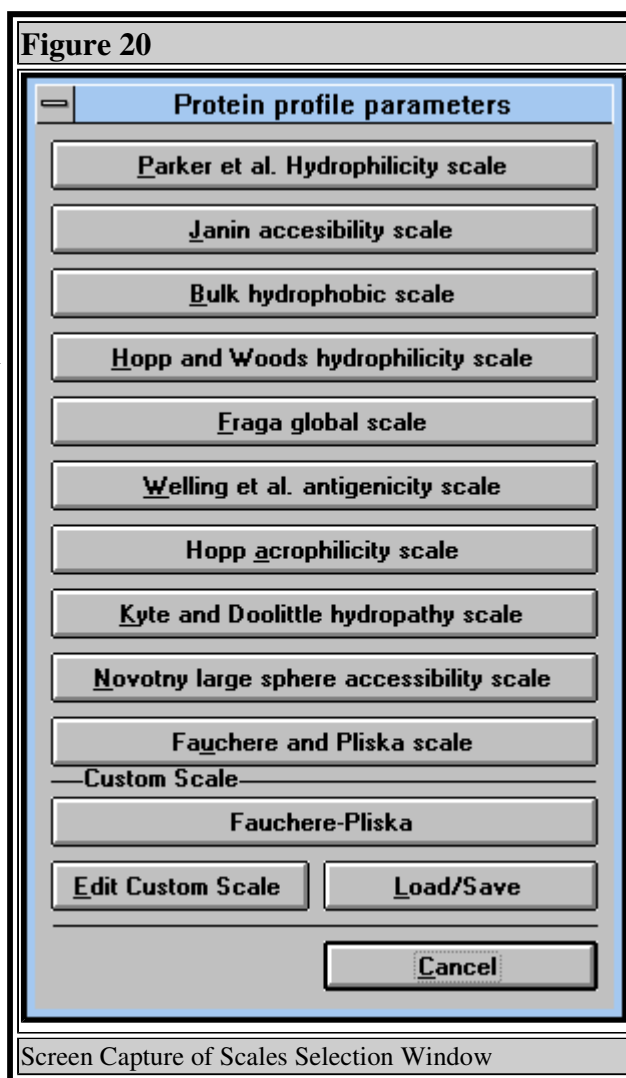
You can also define your own scale. Pressing "Edit Custom Scale" (See Figure 20) will let you introduce new parameters, which can be saved for later use ("Load/Save" button). Above defined scales are stored independently on the disk (extension *.scl) and can be loaded and modified. However, we strongly recommend saving newly defined scales under different names to avoid possible confusion.
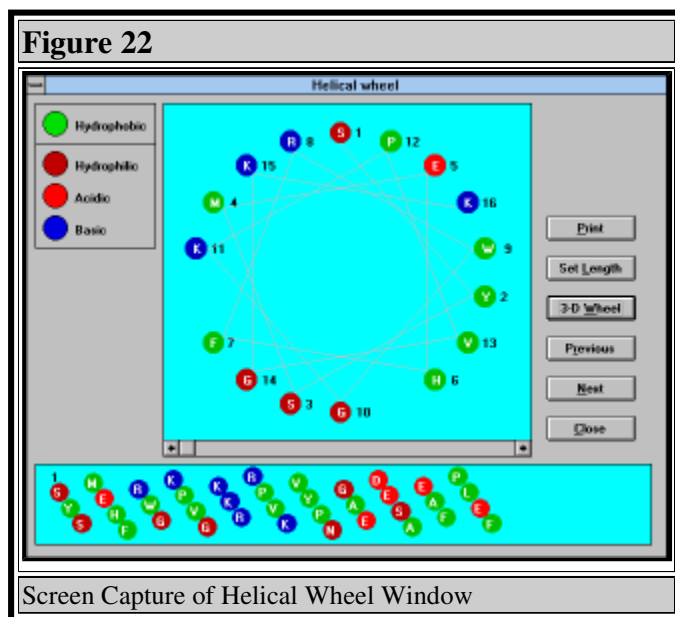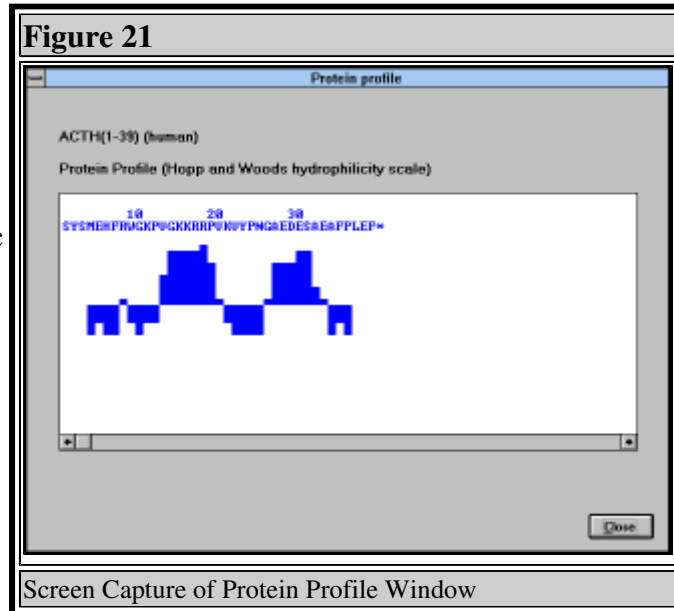
Amphipathic profile ("Amphipathic" button):

This subroutine evaluates the amphipathicity of all consecutive protein fragments of a selected length (default value 11). Using the scale selected (ten scales available), the hydrophilicity vector is calculated and normalized in the same way as described for the protein profiles. The highest value of local amphipathicity indicates the most hydropathic segment of the protein. The helical amphipathicity has been shown to correlate with the localization of T-cell determinants (Margalit H. et al. J.Immunol. 138, 2213 (1987)).

Chapter 5.8.2.3 View commands

To inspect results obtained in the Evaluate procedure, click on the appropriate button and profiles or parameters will appear on the screen.

View graph: Clicking on the "Profile" (protein profile or amphipathic profile) or "Chou-Fasman" button will show you results in a graphical representation. You can zoom in on any portion of the protein: select the first amino acid of the zoomed window by mouse (the aa numbering appears in the window below the profile), click and hold the left mouse button, move the mouse to the last amino acid of the zoomed window, and release the mouse button. The "Zoom out" button will show you the whole protein.

Figure 21

Screen Capture of Protein Profile Window

Figure 22

Screen Capture of Helical Wheel Window

View: Click on the "Chou-Fasman" button and a table with Chou-Fasman parameters will appear on the screen. Use the scroll bar to see any part of the protein. Since B-cell determinants very often reside in those parts of a protein molecule that have a high tendency to form a b-turn, the Chou and Fasman algorithm can be used for the prediction of potential B-cell determinants. Click on the "Turn" button and the program output lists only tetrapeptide sequences having a probability of b-turn occurrence greater than a preselected limit, usually 1.5x10-4 (default value).

The "Profile" button will show you a segment of protein or amphipathic profile in the bar graph form. This is useful if you want to see the full profile in the same
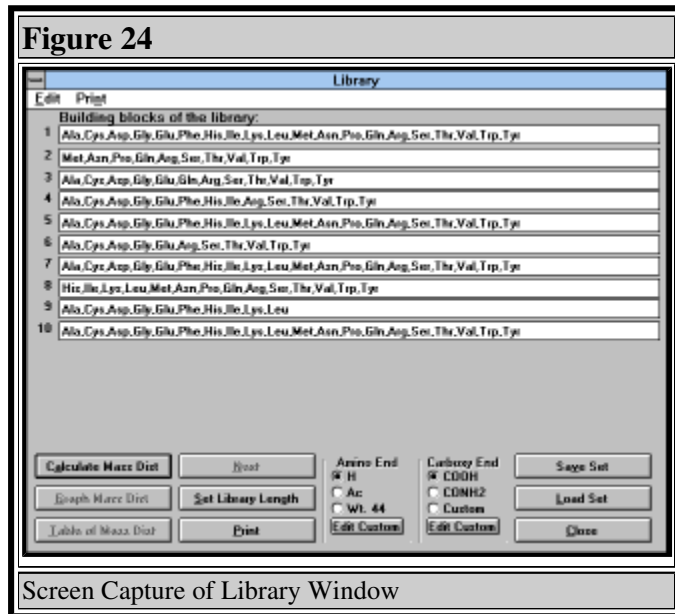
scale together with the amino acid sequence. Use the scroll bar to move along the protein sequence. This option can display only 1000 amino acids.

An example of an output: [See Figure 21] Normalized profile: each bar represents the mean value of a heptapeptide (default value) window centered at the 4-th amino acid. The bars extend upwards when the local hydrophilicity is greater than average, and vice versa.

The "Helical wheel" [See Figure 22] button will show you a standard representation of the helix in an axial view (default length 18 can be changed permanently by pressing the "Setup" button and saving the change after editing the "Setup" window, or temporarily by pressing "Set Length" button) and in an unwrapped tubular view (standard length 50 amino acid residues). Hydrophobic amino acids are in green circles, polar ones are in red (acidic), blue (basic), and/or brown (other) circles. Use the "Next" or "Previous" buttons to walk along the protein. For fast changes use the horizontal scroll bar. The "3D Wheel" [See Figure 23] shows helical arrangement in three dimensional space. The view can be turned by defined (editable) increments by clicking on the arrows below the picture box.

Figure 23

Screen Capture of 3D Helical Wheel Window
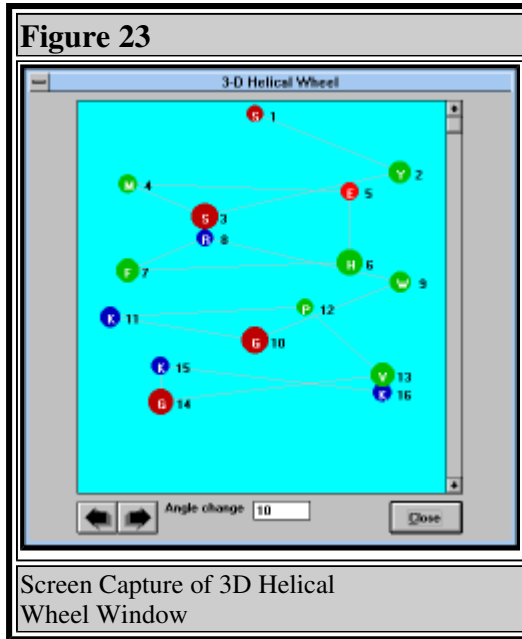
Chapter 5.9. Library ("Library" [See Figure 24] Button)

Figure 24
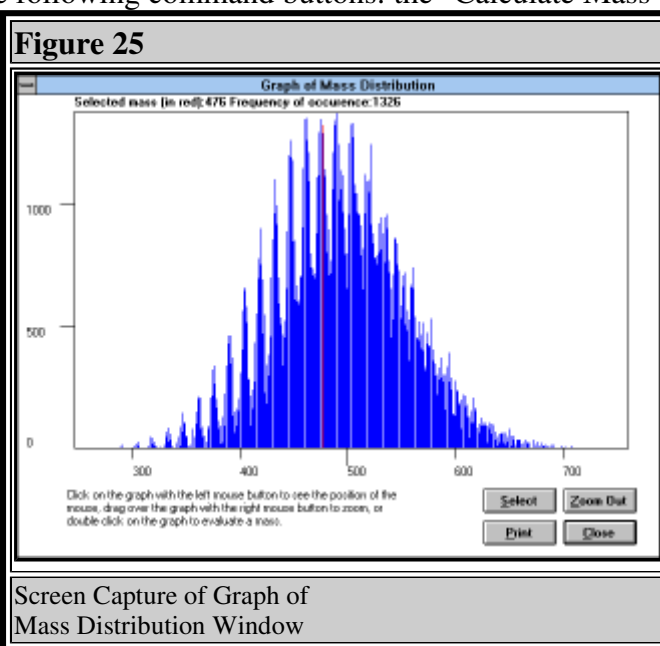

Screen Capture of Library Window

Calculates and graphs the mass distribution in a peptide (or non-peptide) library.

Distribution of the masses is calculated using integer values of building block masses and it does not take into account the isotopic distributions. It can be used for rough evaluation of the quality of prepared library (Andrews et al., in Techniques in Protein Chemistry, Vol. V, p. 485, Academic Press, Orlando 1994) or for library design. The proper selection of library building blocks can simplify mass spectroscopic evaluation of the results.

New building blocks can be defined for library construction. Care must be taken not to use an abbreviation already defined for protecting groups or other structural features. If the newly defined abbreviation should be made permanent, it has to be saved when you exit the program. Clicking on the "Library" button will bring up a dialog box asking you for the length of the library you wish to

randomize. The default is 5 (minimum is 2 and maximum is 30). After you click "OK", the "Library" window will appear. It uses the following command buttons: the "Calculate Mass

Dist" button (calculates the mass distribution), the "Graph Mass Dist" button (shows the graph of mass distribution), the "Table of Mass Dist" button (shows the table of mass distribution), the "Save Set" button (saves the current set of amino acids (building blocks) used for the prediction of mass distributions), the "Load Set" button (loads a set of amino acids (building blocks) used for the prediction of mass distributions), the "Next / Previous" buttons (shows next or previous screens of blocks used for randomization if you set the length of library to more then 15), the "Print" button (prints the current set of building blocks), the "Set Library Length" button (sets the length of the library and

Figure 25

Screen Capture of Graph of Mass Distribution Window

it also puts the default set of amino acids in each randomization), and the "Close" button (returns to the "Main" window).
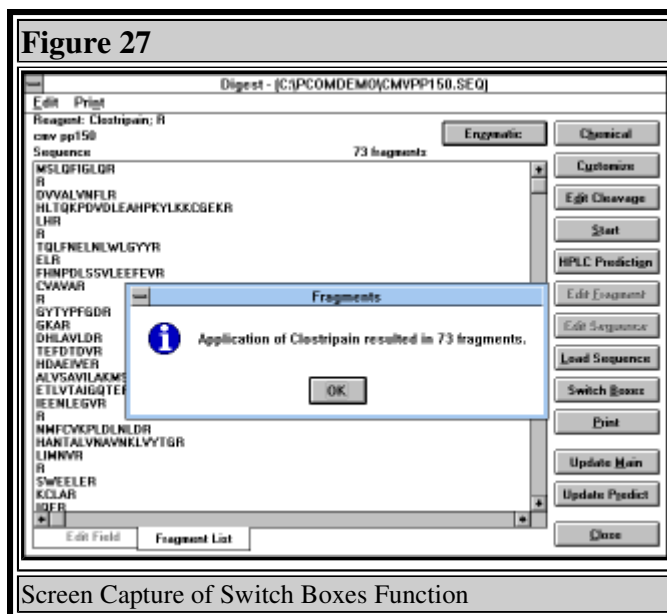

Figure 26
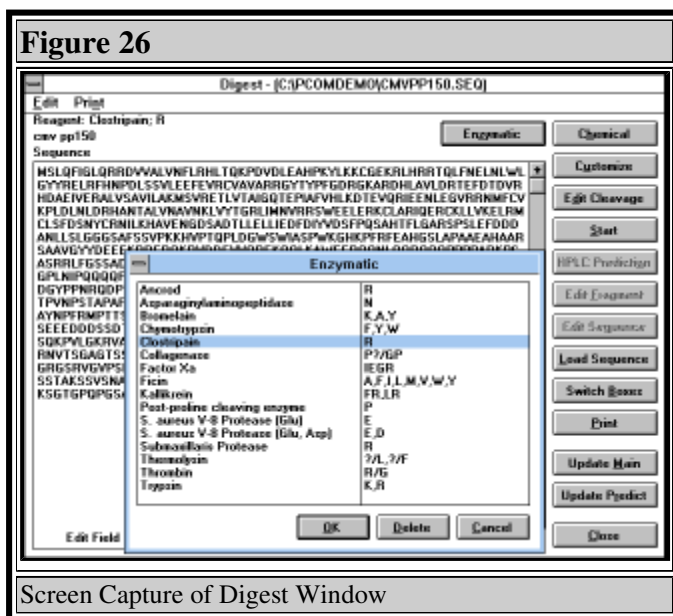
Screen Capture of Digest Window

Clicking on the "Table of Mass Dist" button will show the "Table of Mass Distribution" window. In this window all possible molecular weights generated in the library, together with the frequency of their occurence are displayed. In the lower part of the table the highest, lowest and average values are compiled. You can also save the table ("Save" button) for use in other spreadsheet programs (e.g. Excel, Quattro Pro) as a text (.TXT) file. This feature can be used for that purpose only as the table cannot be loaded in Peptide Companion. Use the "Save Set" button in the "Library" window instead. For printing the table see 6.3.7.

When you press the "Graph Mass Dist" (See Figure 25) button, the "Graph of Mass Distribution" window appears. In this window you can press the right mouse button to see the coordinates of the cursor, double click on the graph and select a mass to show its value and frequency of its occurrence in the library (this can also be done with the "Select" button), or drag the left mouse button over the graph to zoom (press "Zoom Out" button to zoom out).
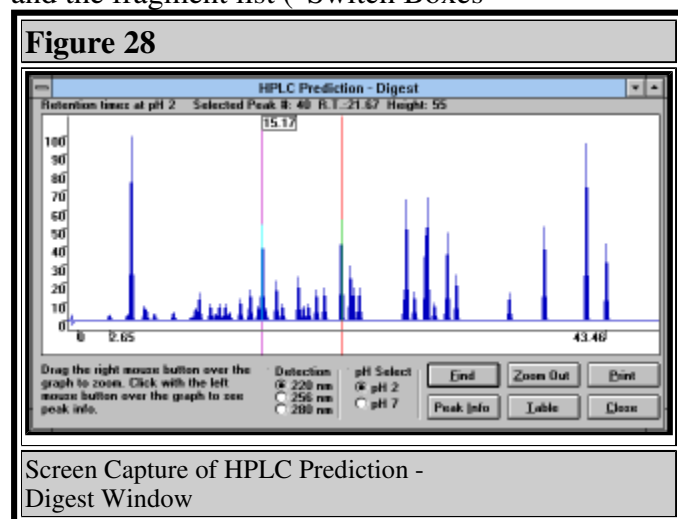
Chapter 5.10 Enzymatic or chemical degradation of peptides and proteins ("Digest" ^(See Figure 26) button)

Proteins or peptides can be submitted to simulated enzymatic or chemical (or combined) degradation. Loading this window brings selected peptides to the sequence edit field. When no peptide has been selected, an empty field will be displayed. A new sequence can be loaded by pressing the "Load Sequence" button. Pressing "Enzymatic", "Chemical", "Customize", or "Edit Cleavage" will let you select the reagent or customize the cleavage specificity. The number of expected fragments and their list would be shown in the fragment list.

You can switch between the edit field and the fragment list ("Switch Boxes"

Figure 27

Screen Capture of Switch Boxes Function

Figure 28
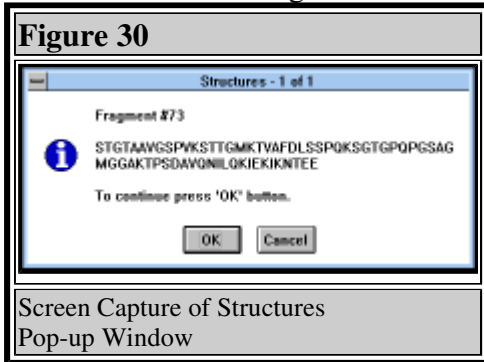

Screen Capture of HPLC Prediction - Digest Window

button or click on the tab on the bottom of the screen). After the fragment is selected (by clicking on it), it can be loaded into the edit field ("Edit Fragment" button) and can be edited for further cleavage (by different reagent) or it can be loaded into the main window or predict window ("Update Main" or "Update Predict" buttons) and to any application from there. The "Customize" button allows addition of any cleavage reagent and saving it in the database. The "Digest" window uses the following command buttons. "Enzymatic" / "Chemical" buttons load

reagents from disk. The "Customize" button enables you to write your own reagent and save it to the disk as enzymatic, or chemical reagent (requires your password). The "Edit Cleavage" button displays a dialog box in which you type the cleavage specificity which you want to use for the degradation. This definition cannot be saved. If you want to use a defined specificity in the future, use the "Customize" button for saving. The "Start" button starts the cleavage. Choosing any cleavage specificity using the buttons above will start the cleavage automatically. The "Edit Fragment" button
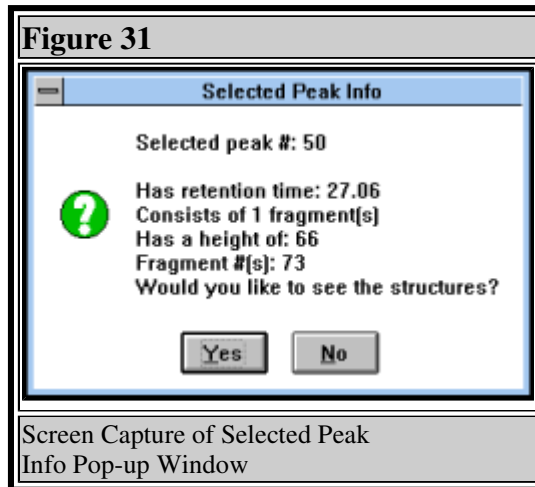
Figure 29

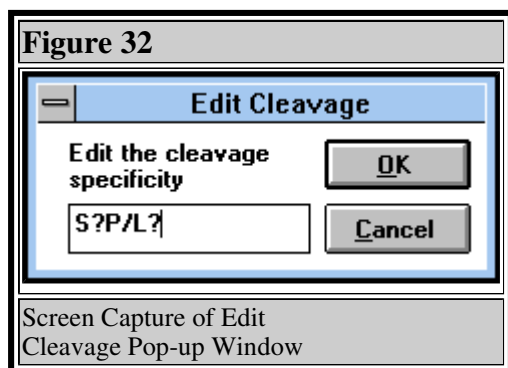Screen Capture of Find Pop-up Window

moves the selected fragment from the fragment list to the edit field. By starting cleavage, the originally loaded sequence will be replaced by this fragment and the old sequence will be cleared. The "Edit Sequence" button puts the sequence last used for cleavage or last loaded into the edit field. The "Load Sequence" button loads the sequence in NIH format. For example see 5.8.1. The "Switch Boxes" [(See Figure 27)] button switches between the edit field and the list of fragments. The "Print" button prints fragments and sequences. The "Update Main" / "Update Predict" buttons place the selected fragment or the fragment in the edit field (if the edit field is on top) into the "Main" or "Predict" window. The "Close" button returns to the "Main" window.

Figure 30

Structures - 1 of 1
Fragment #73
STGTAAWGSPWKSTTGMKTVAFDLSSPQKSGTGPQPGSAG MGGAKTPSDAVQNILQKIEKIKNTEE
To continue press 'OK' button.
OK Cancel

Screen Capture of Structures
Pop-up Window

"HPLC Prediction" [(See Figure 28)] button predicts the elution order of fragments on the reversed phase column. This prediction uses algorithm of Guo et al. (see section 5.5) and all limitations discussed earlier apply here as well. You can choose HPLC with detection at different wavelengths, using different pH of mobile phase, you can zoom into the crowded area of the trace, or you can select [(See Figure 29)] the information [(See Figure 30)] about any peak [(See Figure 31)].

Figure 31

Selected Peak Info
Selected peak #: 50
Has retention time: 27.06
Consists of 1 fragment[s]
Has a height of: 66
Fragment #(s): 73
Would you like to see the structures?
Yes No

Screen Capture of Selected Peak
Info Pop-up Window

Chapter 5.10.1 Cleavage specificity nomenclature [(See Figure 32)]

Figure 32

Edit Cleavage
Edit the cleavage specificity OK
S?P/L? Cancel

Screen Capture of Edit
Cleavage Pop-up Window

The one-letter code is used in the definition of specificity. The single letter indicates that the peptide will be cleaved always after this residue. The specified sequence has the same meaning. Cleavage between defined residues is specified by the "/" symbol between the specified residues. The unspecified amino acid is symbolized by "?". Multiple specificity is defined by putting a comma (",") between the specificities.
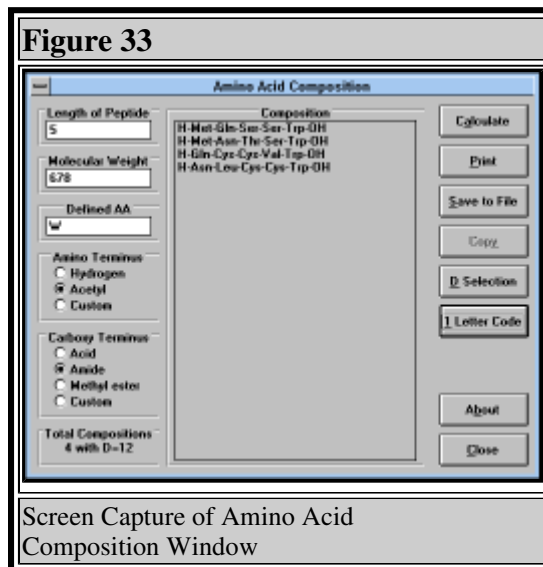
Examples:

- S,P Cleavage after serine or proline
- S/P Cleavage of all -Ser-Pro- sequences
- S?P/L? Cleavage between sequences Ser-Xxx-Pro and Leu-Xxx

Chapter 5.11 Prediction of amino acid composition "AA Composition from M.W." ^(See Figure 33) button

You can calculate probable amino acid composition of peptides composed of natural amino acids based on their molecular weight. To make the calculation feasible with present computer technology, molecular weight is calculated using integer values of monoisotopic compositions. The calculation is performed for the defined length of the peptide (Length of Peptide window). If you want to evaluate all possible peptide lengths, you have to calculate the compositions for each length separately. Long peptides have usually unreasonably high number of possible compositions and it is a good idea to narrow the search down by defining some of the amino acids in the peptide (type their list in Defined AA window (you can use single letter or three letter code)). Selection of possible

Figure 33

Screen Capture of Amino Acid Composition Window

compositions can be further simplified by measuring number of exchangeable protons in the peptide molecule (enter the "D Value" -- the number of exchangeable protons -- ("D Selection" ^(See Figure 34) button) which can be obtained by measuring mass spectrum of the sample dissolved in D2O -- see Sepetov et al. Rapid Comm. Mass Spectrom. 7, 58 (1993)).

Figure 34
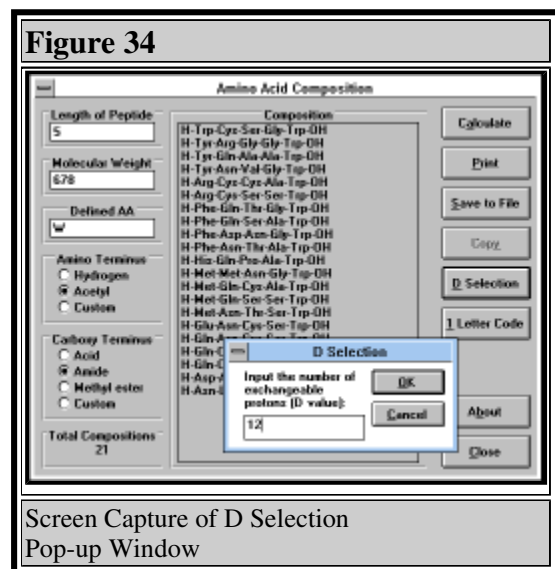
Screen Capture of D Selection Pop-up Window

It is not possible to distinguish between Ile (I) and Leu (L) , and Lys (K) and Gln (Q) and therefore only L and Q are used in the calculations. If you want to calculate composition of cyclic peptide, you must linearize it, i.e. add the molecular weights of molecules subtracted during cyclization (e.g. +2 for disulfides, + 18 for lactams).

Fields with blue background signalize that recalculation using newly defined parameters was not performed. Parameters which can be changed are: (i) Length of peptide; (ii) Molecular weight; (iii) Defined amino acids; (iv) Amino terminal group (default H, choices Ac and Custom); (v) Carboxy terminal group (default OH, choices NH2, OMe, and Custom). After selection of Custom option you are prompted to define molecular weight of terminal group and number of exchangeable protons in this group. (If you will not be using D Selection, you don't have to bother defining the last value.)

Any calculated composition can be transferred in the form of sequence (generated by pressing "3 Letter Code" button) into any other window ("Copy" button). Calculated amino acid compositions can be saved in the file ("Save to File" button), or printed ("Print" button). If the number of possible compositions exceed the available memory space (usually above

4500 compositions), only fraction of compositions is reported. (However, we believe that information of this size is not valuable anyway.)

Chapter 5.12 Quit ("Exit" button)

Leaves the program. Program will confirm your intention to quit and it will ask you whether you want to save any changes to abbreviation definitions you may have made during the session. You can save the changes only if you know your password.

Chapter 6 Printing the results

Chapter 6.1 Two ways to print in Peptide Companion

- 1. Regular printing (Default): Every new print job is done separately (one print job can have several elements). If you are using Regular printing, the information will print immediately on the whole page, even if it doesn't take as much space. Graphs (and any other information that can be printed from the "Predict" window), however, are printed on a separate sheet of paper every time.
- 2. Consecutive printing: If you are using Consecutive printing, the computer doesn't print anything until you choose "Print Memory" in the "Print" menu. In this way you save paper, because you can use one page to print several outputs. The "Predict" window output is always printed on a separate sheet of paper, but not printed until you choose "Print Memory". If you switch off Consecutive printing, the computer will automatically print everything in the memory.

You can switch between Consecutive and Regular printing on any window that has the "Print" menu. There are three items in the "Print" menu, "Consecutive Printing", "Print Memory", and "Paper Orientation". If "Consecutive Printing" is unchecked "Regular printing" is on and "Print Memory" will be disabled. When you check "Consecutive Printing", Consecutive printing will be on and "Print Memory" will be enabled. When you switch the Consecutive printing on or off the change will be done in "Print" menus in all windows.

Chapter 6.2 Changing the paper orientation

You can change the paper orientation by selecting the "Print" menu, choosing "Paper Orientation" submenu and choosing the orientation. When "Consecutive Printing" is enabled the program will print the memory when the orientation is changed.

Chapter 6.3.1 EA, AAA, M.W. window

First select what part you want to print. There are three checkboxes in the "Print" frame: "M.W./Summary" - this option prints molecular weight and summary formula; "AAA" prints the amino acid analysis; "EA" prints the elemental analysis. Then press the "Print" button.

Chapter 6.3.2 Correction of Elemental Analysis window

The printing procedure is very similar to the previous one. There are two option buttons: "All Fits" prints all fits if you calculated the best fit; "Selected Fit" prints just the selected fit. There is also a "Results" checkbox; if it is checked, it will print the results obtained by calculation. Press the "Print" button.

Chapter 6.3.3 HPLC Prediction window

When you press the "Print" button in the "HPLC Prediction" window the computer will ask you if you want to "Print All", which means that the values plus the graph are going to be

printed (this option cannot be printed into memory and will be printed immediately), or "Only Values" which means that only the peptide formulas plus the values are going to be printed.

Chapter 6.3.4 MS Fragmentation window

You can print "All Fragments" or just the "Selected Fragment" by choosing an option in the "Print" frame. You can choose to print All Masses by checking the "All Masses" check box; they will be printed as the last information.

Chapter 6.3.5 Prediction of Difficult Sequence window

Refer to printing Graphs in the "Predict" window.

Chapter 6.3.6 Predict window

All graphs in the "Predict" window that have a "Print" button can be printed; you can also print the Chou-Fasman parameters. If you zoom the graph, it will be printed zoomed.

Chapter 6.3.7 Library window

In the "Library" window you can print the set of building blocks by pressing the "Print" button. In the "Table of Mass Distribution" window you can print the table by pressing the "Print" button. In the "Graph of Mass Distribution" window you can print the graph (as seen on screen) also by clicking the "Print" button (this is the same as in the "Predict" window).

Chapter 6.3.8 Digest window

You can print the fragments of the peptide by pressing the "Print" button. The computer will ask you whether you also want to print the whole sequence.

Chapter 7 Troubleshooting

1	**Problem:**
	This manual is not clear enough about a certain point or does not contain the information you want.
	Solution:
	See Help for more detailed information.
2	**Problem:**
	Setup program stops, saying it cannot copy VBRUN300.DLL.
	Solution:
	If you already have a copy of VBRUN300.DLL in your "\windows\system" directory try renaming it, to VBRUN300.000 for example, (do not delete it) and run the setup program again. To do this start File Manager and go to your windows system directory, usually "C:\WINDOWS\SYSTEM", then locate the file "VBRUN300.DLL" and select it. From the "File" menu choose the "Rename..." item. Type in the new name and press the "OK" button.
3	**Problem:**
	Message "Critical Datafile Missing."
	Solution:
	Run setup program again without upgrading any datafiles.
4	**Problem:**
	When you scroll on a scrollbar to see data not currently displayed and they appear distorted, or when a graphical representation, like helical wheel, is cut into many small fragments.
	Solution:
	Try using standard VGA or SuperVGA display driver. Some drivers can cause problems.
5	**Problem:**
	Some text results do not fit the boxes or are illegible.
	Solution:
	Maybe your display driver does not use standard Windows fonts, use another font size or a different driver.
6	**Problem:**
	"Out of memory" message when starting.
	Solution:
	• I. Quit all unnecessary applications. • II. Start "Peptide Companion Startup" and turn off the "Load all windows" checkbox, then check the "Unload windows not in use" checkbox and press the "OK" button. Restart Peptide Companion.
7	**Problem:**
	"Start" button in Digest window is not enabled.
	Solution:
	No sequence or reagent was specified. When you specify a reagent start button is pressed automatically.

Problem:
I have further problems, questions or suggestions.
Solution:
If you have any further problems, questions or suggestions contact: **CSPS Pharmaceuticals, Inc.** P.O.Box 22567 San Diego, CA 92192-2567 U.S.A. **FAX:** (858) 550-9666 **Email:** csps@5z.com

8

Disclaimer:

CSPS SHALL HAVE NO LIABILITY WITH RESPECT TO ANY LOSS OR DAMAGE DIRECTLY OR INDIRECTLY ARISING OUT OF THE USE OF THE DISK, THE PROGRAMS OR THE DOCUMENTATION. WITHOUT LIMITING THE FOREGOING, CSPS SHALL NOT BE LIABLE FOR ANY LOSS OF PROFIT, INTERRUPTION OF BUSINESS, DAMAGE TO EQUIPMENT OR DATA, INTERRUPTION OF OPERATIONS OR ANY OTHER DAMAGE, INCLUDING BUT NOT LIMITED TO DIRECT, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES.

Appendix A Command Buttons

- Main window
    - List Peptides Button - By clicking on it you can choose a Peptide Dbase.
    - Write From Scratch Button / Edit Peptide Button - Shows the Peptide Edit Tools.
    - Three Letter Code Button - Translates one letter code to three letter code.
    - EA, AAA, M.W. Button - Calculates elemental analysis, amino acid analysis and the molecular weight of a peptide.
    - HPLC Prediction Button - Calculates RP HPLC retention times.
    - MS Fragmentation Button - Fragmentation for mass spectrometry.
    - Difficult Sequences Button - Prediction of synthetically difficult sequences.
    - Conformation / Profile Button - Displays the Predict window.
    - Library Button - Displays the Library window.
    - Digest Button - Displays the Digest window.
    - Delete Peptide Button - Deletes Peptide from the database.
    - List Specifications - Lists Peptides according to custom specifications.
    - New Abbreviation Button - Adds custom abbreviation to the list.
    - Edit Abbreviation Button - Edits selected abbreviation.
    - Exit Button - Exits Peptide Calculator.
- Predict window
    - Load / Save Button - Loads and saves proteins.
    - Setup Button - Sets defaults for Predict window.
    - Close Button - Returns to Main window. All calculations are lost.
    - Evaluate Commands - Lets you choose the following:
        - Chou-Fasman Button - Performs the predictions according to Chou and Fasman.
        - Protein Profile Button - Lets you choose the appropriate scale, or define a new one.
        - Amphipathic Button - Lets you choose the appropriate scale, or define a new one.
    - View Commands
    - View Graph
    - Chou-Fasman Button - graph representation of the calculated data is shown.
    - Profile Button - shows calculated data in linear graph form.
    - View
    - Chou-Fasman Button - tabular representation of the calculated data is shown.
    - Turn Button - shows table of probable turns in the protein.
    - Profile Button - shows calculated data in bar graph form.
    - Helical Wheel Button - shows helical wheel representation of the protein or peptide.
- EA, AAA, M.W. window
    - Recalculate Button - Recalculates all calculations with the edited peptide.
    - Add AcOH ... / Fit to Found Buttons - Addition of various molecules (Add AcOH ... button) and calculation of probable composition of the lyophilizate (Fit to Found button).
    - Other Elements Button - Displays composition of all elements in the sample.
    - Save Formula Button - Saves Formula of the peptide.
    - Print Button - Prints results.
    - Close Button - Returns to Main window.

- Correction of Elemental Analysis window
    - Best Fit Button - Calculates the best fit and all satisfactory results.
    - Print Button - Prints results.
    - Calculate Button - Calculates elemental analysis with added molecules.
    - Close Button - Returns to EA, AAA, M.W. window.
- HPLC Prediction window
    - Predict Button - Calculates HPLC Prediction.
    - Print Button - Prints results.
    - Set Standard Time Button - Allows you to input the experimentally observed retention time of the standard peptide.
    - Clear Trace Button - Same as Predict Button, but clears the graph first.
    - Load Standard Button - Loads the formula of a standard peptide.
    - Load Sample Button - Loads the formula of a sample peptide.
    - Close Button - Returns to Main window.
- MS Fragmentation window
    - Calculate Button - Recalculates MS fragmentation after editing the sample.
    - Found Button - Compares results from spectral measurement with theoretical data.
    - Print Button - Prints results.
    - Close Button - Returns to Main window.
- Library window
    - Calculate Mass Dist Button - Calculates the mass distribution.
    - Graph Mass Dist Button - Shows the graph of mass distribution.
    - Table of Mass Dist Button - Shows the table of mass distribution.
    - Save Set Button - Saves the current set of amino acids (building blocks) used for the prediction of mass distributions.
    - Load Set Button - Loads a set of amino acids (building blocks) used for the prediction of mass distributions.
    - Next / Previous Buttons - Shows next or previous screens of blocks used for randomization if you set the length of library to more then 15.
    - Print Button - Prints the current set of building blocks.
    - Set Library Length Button - Sets the length of the library. It puts the default set of amino acids in each randomization.
    - Close Button - Returns to Main window.
- Digest window
    - Enzymatic / Chemical Buttons - Loads reagent from disk.
    - Customize Button - Enables you to write your own reagent and save it to the disk as enzymatic, or chemical (requires your password).
    - Edit Cleavage Button - Displays a dialog box in which you type custom cleavage specificity. Cannot be saved, use Customize button for saving.
    - Start Button - Starts the cleavage. Choosing any cleavage specificity using the buttons above will start the cleavage automatically.
    - HPLC Prediction Button - Calculates retention times of fragments on reversed phase column.
    - Edit Fragment Button - Puts the selected fragment from the fragment list to the edit field. By starting cleavage the sequence will be replaced by this fragment and the old sequence will be cleared.
    - Edit Sequence Button - Puts the sequence last used for cleavage or last loaded into the edit field.
    - Load Sequence Button - Loads Sequence in NIH format.
    - Switch Boxes Button - Switches between the edit field and the list of fragments.

- Print Button - Prints fragments and sequence.
- Update Main / Update Predict Buttons - Put the selected fragment or the fragment in the edit field (if the edit field is on top) in the Main or Predict window.
- Close Button - Returns to Main window.
  - Amino Acid Composition window
    - Calculate Button - Performs the calculation.
    - Print Button - Prints the list of possible amino acid compositions.
    - Save to File Button - Saves the results in either one- or three-letter code (in the form of sequence).
    - Copy Button - Copies the selected composition from the list to clipboard.
    - D Selection - Selects the compositions from the list which fulfil the condition of number of exchangeable protons.
    - 3 Letter Code Button - Translates one letter code list of amino acids to peptide sequence written in the three letter code.
    - 1 Letter Code Button - Performs just the opposite.
    - About Button - Describes function of this window.
    - Close Button - Returns to Main window.

In all windows Close or Cancel button will return to the previous window. Cancel button cancels all operations on the windows closed, Close button does not.

Appendix B Menus

- File Menu (Main window)
  - Change Password ... Item - Changes the user password.
  - Change Default Directory... Item - Changes where datafiles are loaded from.
  - Save Peptide ... Submenu - Saves the peptide to a dbase.
  - Sort Current Dbase! Item - Sorts peptides in the dbase currently in use by name.
  - Save Abbreviations! Item - Saves the Changes to the abbreviations list as you would if you'd leave the program.
  - Exit Item - Exits the program.
- Edit Menu (most major windows)
  - Undo Item - Undoes the last change in text involving edit menu.
  - Cut Item - Cuts selected text and puts it to clipboard.
  - Copy Item - Copies selected text to clipboard.
  - Paste Item - Pastes contents of the clipboard to the cursor position or instead of the selected text.
  - Delete Item - Deletes a character or selected text.
- Print Menu (most major windows)
  - Consecutive Printing Item - If checked, consecutive printing is enabled.
  - Print Memory! Item - Prints memory content if consecutive printing is on.
  - Clear Memory Item - Clears the printing memory.
  - Print Setup... Item - Displays a dialog with the printer setup (orientation etc...).
- Help Menu (Main window)
  - Contents Item - Displays the contents of this help.
  - Search Help On ... Item - Displays the Help Search dialog.
  - How to Use Help Item - Displays the How to Use Help helpfile.
  - Tutorial Item - Starts the Tutorial. Follow instructions on screen.
  - Other Items Available Item - Shows a write file with other items from CSPS.
  - Manual Changes - New Stuff Item - Displays a file with updates to this manual.
  - About Peptide Companion... Item - Displays the About Peptide Companion... dialog box..

Appendix C Index of Tables

Appendix D Index of Figures