THE CHEMICAL SYNTHESIS OF PEPTIDES

Peptides are the long molecular chains that make up proteins. Synthetic peptides are used either as drugs (as they are biologically active) or in the diagnosis of disease. Peptides are difficult to make as the synthetic chemist must ensure that the amino acids that make up the chain are added in the correct order and that they don't undergo any other reactions. This involves adding one amino acid, washing away any unreacted acid then adding the next and so on. As can be imagined, this is very time consuming and only gives very low yields.

A technique that has been relatively recently developed involves attaching one end of the peptide to a solid polymer, meaning that the peptide cannot get washed away along with the excess acid. This is much quicker than classical synthesis, and leads to dramatically improved yields. The process consists of five steps carried out in a cyclic fashion.

Step 1 - Attaching an amino acid to the polymer
The amino acid is reacted with a molecule known as a "linkage agent" that enables it to attach to a solid polymer, and the other end of the linkage agent is reacted with the polymer support.

Step 2 - Protection
An amino acid is an acid with a basic group at one end and an acid group at the other. To prevent an amino acid from reacting with itself, one of these groups is reacted with something else to make it unreactive.

Step 3 - Coupling
The protected amino acid is then reacted with the amino acid attached to the polymer to begin building the peptide chain.

Step 4 - Deprotection
The protection group is now removed from the acid at the end of the chain so it can react with the next acid to be added on. The new acid is then protected (**Step 2**) and the cycle continues until a chain of the required length has been synthesised.

Step 5 - Polymer removal
Once the desired peptide has been made the bond between the first amino acid and the linkage agent is broken to give the free peptide.

Peptides synthesised at Massey University are widely used in medical research, as they can be synthesised quickly and accurately. When larger amounts of peptides are required they are synthesised in the same manner, although the equipment to produce peptides on this scale is not available in New Zealand.

INTRODUCTION

A peptide is a chain of special acids called amino acids linked together by bonds known as amide bonds. A protein consists of one peptide folded in a particular way, or several peptides folded together. Such peptides are synthesised very rapidly within living cells, but until recently could only be artificially synthesised in very long, slow processes that had poor yields and gave impure products. Recently a new technique known as solid phase peptide synthesis (SPPS) has been developed. SPPS results in high yields of pure products and works more quickly than classical synthesis, although still much more slowly in than living cells. This technique is discussed below.

Uses of synthetic peptides

Synthetic peptides have two main uses: as peptide drugs and as peptides for diagnostic purposes.

Peptide drugs

Peptide drugs are either naturally-occuring peptides or altered natural peptides. There are many naturally-occuring peptides that are biologically active. If a patient does not naturally produce a peptide that they need, this peptide can be synthesised and given to them. In addition, the amino acids in an active peptide can be altered to make *analogues* of the original peptide. If the analogue is more biochemically active than the original peptide it is known as an *agonist* and if it has the reverse effect is known as an *antagonist*. Contraceptives have been made by synthesising the antagonists of fertility peptides.

Diagnostic peptides

Peptides can be designed that change colour under certain conditions, and these can be used for diagnostic purposes. For example, a chromogenic peptide substrate can readily detect the presence, absence and varying blood levels of enzymes that control blood pressure and blood clotting ability.

The SPPS laboratory at Massey University has supplied peptides for research purposes to universities, CRIs, research institutes and private industry since 1973. These peptides have been used for medical research into areas such as heart disease, leprosy and tuberculosis. The laboratory itself is involved in research into and development of synthetic methods and peptide production.

Introduction to protein chemistry

Peptides are polymers of amino acids made using anything from two to hundreds of amino acids. They are all based on the α-amino acid structure

$$H_2N-\underset{\displaystyle R}{\underset{|}{CH}}-\overset{\displaystyle O}{\overset{||}{C}}-OH$$

There are twenty amino acids that commonly occur in nature (**Table 1**) and many others have been synthesised.

Table 1 - Side chains of the common naturally occurring amino acids

Name	Side chain	Protection
glycine (gly)	$-H$	never necessary
alanine (ala)	$-CH_3$	never necessary
valine (val)	$-CH(CH_3)_2$	never necessary
leucine (leu)	$-CH_2-CH(CH_3)_2$	never necessary
isoleucine (ile)	$-CH(CH_3)-CH_2CH_3$	never necessary
lysine (lys)	$-CH_2CH_2CH_2CH_2NH_2$	protection essential agent = tBOC
arginine (arg)	$-CH_2CH_2CH_2NHC(=NH)NH_2$	protection often useful agent = PMC
aspartic acid (asp)	$-CH_2-C(=O)-OH$	protection essential agent = tBu
asparagine (asn)	$-CH_2-C(=O)-NH_2$	protection sometimes necessary agent = tBu
glutamic acid (glu)	$-CH_2-CH_2-C(=O)-OH$	protection essential agent = tBu
glutamine (gln)	$-CH_2-CH_2-C(=O)-NH_2$	protection sometimes necessary agent = Trt
threonine (thr)	$-CH(OH)CH_3$	protection often useful agent = tBu
methionine (met)	$-CH_2-CH_2-S-CH_3$	protection sometimes necessary oxidise to S=O

cysteine (cys)	$-CH_2-SH$	protection essential agent = Trt or tBu
serine (ser)	$-CH_2-OH$	protection often useful agent = tBu
phenylalanine (phe)	$-CH_2-$	never necessary
tyrosine (tyr)	$-CH_2-$ $-OH$	protection often useful agent = tBu
histidine (his)	$-CH_2-$ N N	protection often useful agent = Trt
tryptophan (trp)	$-CH_2$	protection sometimes necessary agent = tBOC
proline (pro)	CH-COOH N H [1]	never necessary

These peptides (or combinations of them) fold in characteristic ways to give proteins. In mammals all optically active[2] amino acids are in the L form, so one change that can be made to peptides is to substitute a D amino acid for one of the amino acids in the chain.

---

[1]The full structure of proline (rather than just its side chain) is given because proline is cyclic and so doesn't have a side chain as such.

[2]When a molecule is exposed to plane polarised light (i.e. light in which all the light waves are moving up and down in the same plane), some molecules rotate the light. Such molecules are spoken of as being *optically active*. Each optically active molecule has a mirror image which cannot be super-imposed on the orginal and which rotates light the same amount in the opposite direction. In glyceraldehyde (an optically active molecule) the form which rotates light to the left is called L-glyceraldehyde (from the Greek *laevo*, meaning left) and the other form is called D-glyceraldehyde (from the Latin *dexter*, meaning right). For amino acids, the form that looks like L-glyceraldehyde is the 'L' form of the amino acid and the form that looks like D-glyceraldehyde is known as the 'D' form, regardless of which way they rotate plane polarised light. All naturally occuring amino acids except glycine are optically active.

OH, H, CHO, CH_2OH | OH, H, CH_2OH, CHO | NH_2, H, COOH, CH_3 | NH_2, H, H_3C, COOH

L-glyceraldehyde D-glyceraldehyde L-alanine D-alanine

The bonds between amino acids are called *amide* bonds. These are strong bonds but they can be hydrolysed by either strong acid or strong base, meaning that only neutral substances and weak acids and bases can be used in peptide synthesis.

SOLID PHASE PEPTIDE SYNTHESIS

Peptide synthesis is much more complicated than simply forming amide bonds by mixing the desired amino acids together in a test tube. With twenty natural amino acids and a number of unnatural ones as well the possible combinations formed with this technique are numerous. This complexity makes the synthesis of peptides both fascinating and challenging.

If solutions containing two amino acids are mixed together, four different dipeptides (as well as other longer peptides) will be formed. (e.g. for a mixture of glycine and alanine the four dipeptides would be glygly, glyala, alagly, alaala. In this representation of peptides the free amino group or N-terminus is on the lefthand amino acid and the free carboxylic acid group, the C-terminus is at the righthand end.). To ensure that only the desired dipeptide is formed the basic group of one amino acid and the acidic group of the other must both be made unable to react. This 'deactivation' is known as the *protection* of reactive groups, and a group that is unable to react is spoken of as a protected group. In classical organic sythesis the acids are protected, allowed to react and deprotected, then one end of the dipeptide is protected and reacted with a new protected acid and so on. In SPPS the amino acid that will be at one end of the peptide is attached to a water-insoluble polymer and remains protected throughout the formation of the peptide, meaning both that fewer protection / deprotection steps are necessary and that the reagents can easily be rinsed away without losing any of the peptide.

Step 1 - Attaching an amino acid to the polymer

Peptide chains have two ends, known respectively as the N-terminus and the C-terminus[3], and which end is attached to the polymer depends on the polymer used. This article assumes that polyamide beads are used and hence that the the C-terminus of the peptide is attached to the polymer. The attachment is done by reacting the amino acid with a linkage agent and then reacting the other end of the linkage agent with the polymer. This means that a peptide-polyamide link can be formed that will not be hydrolysed during the subsequent peptide-forming reactions. Common linkage agents are di- and tri-substituted benzenes such as those shown below:

MeO

$HOCH_2$—C_6H_4—CO_2H $HOCH_2$—C_6H_3—OCH_2CO_2H

These then join the C-terminus amino acid and resin together as follows:

[3]The N-terminus is the end that has a free NH_2 group and the C-terminus is the end that has a free carboxyl group.

MeO

R

NH_2—CH—C(=O)—OCH_2—[benzene ring]—OCH_2C(=O)—$NHCH_2CH_2NH$—

amino acid   linkage agent   resin

Step 2 - Protection
The next amino acid also needs to have its amino group protected to prevent the acids reacting with each other. This is done by protecting it with FMOC (9-fluorenylmethoxy-carbonyl). In addition, any amino acid side chains that are aromatic, acid, basic or highly polar are likely to be reactive (see **Table 1**). These must also be protected to prevent unwanted branched chains from forming. There are four main groups used in this way: tBu (a tertiary butyl group), Trt (a triphenylmethyl group), tBOC (a tertiary butyloxycarbonyl group) and PMC (a 2,2,5,7,8-pentamethylchroman-6-sufonyl group). Examples of a carboxyl group protected with FMOC and examples of the different types of side chain protection are given below.

H CH_2–O—C(=O)—Cl

FMOC

H CH_2–O—C(=O)—NH—CH(R)—C(=O)—OH

an amino acidprotected wuth FMOC

HO—C(=O)—CH(NH_2)—CH_2–OH

serine

HO—C(=O)—CH(NH_2)—CH_2–O—C(CH_3)(CH_3)—CH_3

serine protected with tBu

HO—C(=O)—CH(NH_2)—CH_2–SH

cysteine

HO—C(=O)—CH(NH_2)—CH_2—S—C(phenyl)₃

cysteine protected with Trt

$HO-C(=O)-CH(NH_2)-(CH_2)_4-NH_2$

lysine

$HO-C(=O)-CH(NH_2)-(CH_2)_4-NH-C(=O)-O-C(CH_3)_3$

lysine protected with tBOC

$HO-C(=O)-CH(NH_2)-(CH_2)_3-NH-C(=NH)-NH_2$

arginine

$HO-C(=O)-CH(NH_2)-(CH_2)_3-NH-C(=NH)-NH-SO_2-$ (ring with H_3C, CH_3, H_3C, CH_3, O, CH_3, CH_3)

arginine protected with PMC

Step 3 - Coupling
The FMOC protected amino acid is then reacted with the last amino acid attached to the polyamide. The reaction is catalysed by DCC (1,3 dicyclohexylcarbodiimide), which is itself reduced to DCU (1,3-dicyclohexylurea). The reaction is shown in **Figure 1** (RCOOH represents the FMOC protected acid and H_2NR represents the reactive end of the growing peptide chain).

Step 4 - Deprotection
Excess DCC is washed off the insoluble polymer with water, then the FMOC group removed with piperidine (a cyclic secondary amine). This is a trans amidification reaction.

N—H

piperidine

Steps 2 to **4** are repeated as each new amino acid is added onto the chain until the desired peptide has been formed.

Step 5 - Polymer removal
Once the peptide is complete it must be removed from the polyamide. This is done by cleaving the polyamide - peptide bond with a 95% solution of trifluoro acetic acid (TFA). The side-chain protecting groups are also removed at this stage.

Figure 1 - The amino acid coupling reaction

PRACTICAL AND FINANCIAL ADVANTAGES OF SPPS

The primary advantage of SPPS is its high yield. As peptides consists of many amino acids, if the yield for each amino acid addition is much less than 100%, overall peptide yields are negligible. For example, if each amino acid addition has a 90% yield then the overall yield of a 50 amino acid peptide is only 0.5%. Modern SPPS instrumentation pushes coupling and deprotection yields to greater than 99.99%, giving an overall yield of greater than 99% for a 50 amino acid peptide.

SPPS is also much quicker than conventional step-by-step solution synthesis. With SPPS, a 20 amino acid peptide can be synthesised in a 24 hour period and longer ones in less than a week. With the advent of automated synthesizers and sophisticated analytical and purification equipment the peptide chemist can now make peptides in the range of 20-50 amino acids in length and in amounts from 20-100 milligrams. This is often more than enough for biochemists and biologists to carry out extensive pilot studies, and as they often only look at a particular peptide once this speed is particularly useful. If very large amounts of peptide are required (e.g. for the industrial production of peptide drugs) then this speed is

sacrificed for purity. However, production rates are still high, and hundreds of grams of peptide can be produced on kilograms of polymer every year. As often only milligrams of polymer is needed per dose, this represents hundreds of thousands of doses.

UTILITIES

The Massey facility has one 430A and one 431A Applied Biosystem fully automated peptide synthesizers. The latter is controlled by a Macintosh computer. For developmental work the Massey group also has a semi manual LKB 4174 Biolynz apparatus. In addition to purification equipment (HPLC and FPLC facilities), the group has access to the Department of Biochemistry for amino acid analysis (AAA) and protein sequencing to aid in the analysis of the peptide products.

ENVIRONMENTAL IMPLICATIONS

SPPS, like much of organic chemistry, makes use of organic solvents which are hazardous to the environment. The SPPS group at Massey University is currently researching ways of producing peptides in aqueous or partially aqueous (e.g. water/ethanol mixtures) solutions to avoid the use of organic solvents.

Written by David Harding (Massey University) and edited by Heather Wansbrough with reference to:

- Beavis, Ronald C.; *Protecting groups used in peptide synthesis*; [Online] http://128.122.10.5/aainfo/deriv.htm; 26 February 1997

- Atherton, E. and Sheppard, R. C.; *Solid Phase Peptide Synthesis: a practical approach*; Oxford University Press; 1989

- Bodanszky, M.; *Principles of Peptide Synthesis (2nd. ed.)*; Springer-Verlag; 1993