

BACHEM



WELCOME TO THE WORLD OF PEPTIDES

INSIGHTS INTO THE PEPTIDE CHEMISTRY OF BACHEM

AN INTRODUCTION INTO THE WORLD OF AMINO ACIDS AND PEPTIDES

presented by Monika Mergler, Ph.D., Bachem AG

Contents

1.	The business areas of Bachem	3
2.	What are peptides?	6
3.	What are amino acids?	7
4.	Manufacture of peptides	14
5.	Analysis of amino acids and peptides	24
6.	Modification of peptides	26

Abbreviations

ADS	Analytical Datasheet
API	Active Pharmaceutical Ingredient (drug substance in medicines)
CCD	Counter Current Distribution (purification method)
CofA	Certificate of Analysis
DMF	Dimethylformamide (solvent used in SPPS)
GMP	Good Manufacturing Practice*
HPLC	High Performance Liquid Chromatography (method of analysis and purification)
MS	Mass Spectrometry (method of analysis)
NCE	New Chemical Entity (drug substance not yet licensed)
NMP	N-Methylpyrrolidone (solvent used in SPPS)
RSLC	Rapid Separation Liquid Chromatography (method of analysis)
SPPS	Solid-phase Peptide Synthesis
TLC	Thin Layer Chromatography (method of analysis)
Xaa, Yaa	any amino acid (3-letter code)

Symbols for chemical elements

C	Carbon
H	Hydrogen
N	Nitrogen
O	Oxygen
S	Sulfur

* GMP = Good Manufacturing Practice, production guidelines that assist quality assurance. These guidelines must be followed when substances to be used in humans are manufactured. Their adherence is regularly checked by specialist authorities (such as Swiss-medic in Switzerland, FDA in the USA). Drug substances in medicines, ingredients in cosmetics and additives in foods are some of the items that sell under these guidelines.

© Copyright by Bachem AG, CH-4416 Bubendorf, Switzerland. Reproduction only with permission.

The business areas of Bachem

The core activities of Bachem are the contract manufacture of pharmaceutical active substances and the synthesis and sale of amino acid derivatives, peptides and fine chemicals. Bachem produces peptides as well as small molecule APIs as drug substances for the pharmaceutical industry under GMP conditions*.

The business areas of Bachem are:

Research Chemicals

Catalog products, available from stock

These amino acid derivatives, peptides and other chemicals are not manufactured under GMP conditions and may only be used for laboratory and research purposes. They may not be used in humans.

Custom Synthesis

Contract synthesis, exclusively for a customer

Mainly peptides, that are manufactured in quantities ranging from a few milligrams up to several grams. On request, the products can be synthesized under GMP conditions.

New Chemical Entities (NCEs)

Contract synthesis

NCEs are patented peptides or organic compounds that Bachem manufactures exclusively for the patent holder or a licensee. Even during the first, small-scale manufacture, the key point to remember with NCEs is that the product may prove to be pharmaceutically active and therefore must be capable of manufacture in larger quantities and potentially under GMP conditions.

Bachem strives to develop a long-term partnership with such customers. This begins with the initial synthesis of an NCE on a laboratory scale, moves on to the manufacture of larger quantities under GMP conditions for the clinical phases and continues right up to large-scale manufacture of the drug substance for the approved medicine.

Generic Active Pharmaceutical Ingredients (Generic APIs)

Pharmaceutically active compounds

Substances used as active ingredients in medicines have to be manufactured under GMP conditions. They are sold in large quantities.

GENERIC APIs

- peptide and small molecule generics
- supply of commercial quantities
- reputation for sustained quality & supply
- high-demand APIs available from stock
- close and long term partnerships

CATALOG PRODUCTS

- more than 7500 products available online
- peptides, amino acids, inhibitors, substrates
- fast & easy ordering @ shop.bachem.com
- bulk quantities according to your needs
- excellent customer and technical support

NCEs

- process development, optimization and validation
- comprehensive analytical services
- supply of cGMP material for all clinical phases
- cGMP-manufacturing of APIs for commercial supply

CUSTOM SYNTHESIS

- from discovery stages to early clinical candidates
- strong focus on quality and timely execution
- development of synthetic routes for scale up
- best industry practice
- highly motivated and experienced team

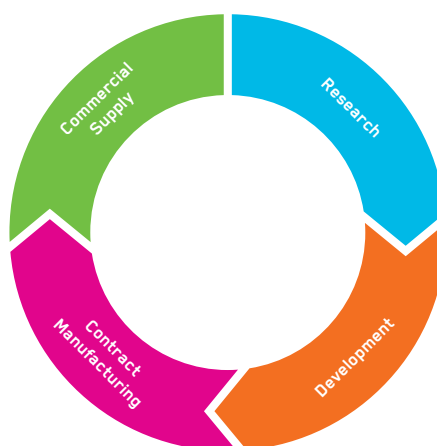


Fig. 1: The Bachem 360° business model.

Generic Active Pharmaceutical Ingredients (Generic APIs)

Generics are drug substances or medicines, whose patent has expired.

Medicine = “formulated” drug substance:

Depending on how medicines are administered (as tablet, as injection etc.), various non-active additives or “excipients” are mixed with the actual drug substance. This process is called formulation and it can be patented independent of the drug substance patent. Suitable formulation can, for example, enable the drug substance to be released only slowly (depot formulations), to be delivered as a nasal spray or absorbed through the skin. The development and patenting of new formulations of generic drug substances offers the drug substance manufacturer the possibility of good, long-term sales.

Table 1: Example of some peptide APIs manufactured by Bachem and their indications

Peptide	Indication
Calcitonin	Osteoporosis
Deslorelin	Fertility control (veterinary medicine)
Desmopressin	Diabetes insipidus
Glucagon	Hypoglycemia (low blood sugar) in diabetes
Goserelin	Cancer
Gonadorelin	Reproduction medicine
Leuprolide	Cancer
Octreotide	Acromegaly
Secretin	Diagnosis (pancreas)
Triptorelin	Cancer

The discovery of other pharmacological activities can lead to additional indications for a generic and thereby increase the demand for the drug substance. This is another case when patents (usage patents) can be granted. Generics are also used to diagnose diseases (i.e. diagnostic agents). A further market with great potential is the use of generics in veterinary medicine. The patent protection of a drug substance generally expires after 18 years. Small quantities of the peptide can be sold for research and development purposes even before the patent has expired (Bolar Exemption). Some peptide APIs manufactured by Bachem and their indications are listed in Table 1.

The structures of calcitonin, glucagon, gonadorelin (GnRH or LHRH) and secretin correspond to those of the body's own “endogenous” natural peptide hormones; the other peptides listed in the table are modified natural peptides. Through changes made to their structure, they are more slowly broken down in the body than endogenous peptides.

Triptorelin is available in the form of two different salts, acetate and pamoate. The pamoate is released more slowly and is therefore used in depot formulations.

The small molecule APIs sold by Bachem are manufactured by our subsidiary in Vionnaz.

Examples are propofol - used to induce general anesthesia, or for sedation in artificially ventilated adults during intensive care - and zolpidem, a most frequently prescribed drug against insomnia.

If you are interested, more about our peptide and non-peptide generic APIs can be found on our homepage www.bachem.com

You can download our generics catalog and request a quote.

New Chemical Entities

During preclinical development, lead finding and lead optimization require large panels of peptides. These are generated as custom synthesized molecules for customers around the world. Frequent consultation with Bachem experts allows further refining of target compounds. As such, a clear partnering aspect is required to come up with pioneering concepts and molecules to bring into clinical development.

There is a large pipeline of peptide drug candidates in clinical development. Also, over 50 peptides have been approved to treat various diseases. Bachem has the largest portfolio of projects for which we supply the peptide.

When clients have selected their lead compound they commence clinical trials. It is a decade-long process to approval of the drug. During this time, there is a close collaboration to learn more about the product. Each production step is scrutinized and manufacturing reproducibility strived for. Scale-up and full control of the process is targeted. Validation and control of the process is the end result of an intense partnership.

Custom Synthesis

In addition to our research grade products listed in our catalog and available from stock, Bachem offers an efficient and high quality custom synthesis service for peptides and amino acids derivatives. Our peptide chemists can prepare compounds according to your specifications in any scale you need - from milligram to kilogram and ton quantities. We are experts in producing fluorescently labeled peptides, peptides with multiple disulfide bridges, peptides containing stable isotopes, and long and difficult peptides. Other specialties include multi-step organic transformations and peptidomimetic molecules.

Research Chemicals

„Catalog products“

These products are sold from stock in various pack sizes or in larger quantities (minibulk and bulk). Most of them are manufactured by Bachem UK or Bachem AG at Bubendorf (CH).

The catalog products cover a broad customer base. The products in the main catalog are divided into two categories. The first section of the main catalog contains amino acid derivatives and products for the synthesis of peptides, i.e. products mainly ordered by chemists. It also contains biochemicals that are not amino acid derivatives or peptides. So this section also serves customers involved in biological and medical research.

The second section is larger than the first one. It lists the available peptides and is addressed mainly to life science research. This division into two product categories (amino acids, products for peptide synthesis, and non-peptide biochemicals/peptides) is mirrored in the Online Shop (s. Fig. 3).

The catalog business is also important for the acquisition of customers for our other services, such as contract synthesis. It is a simple way by which we can show potential customers the high product quality and excellent service offered by Bachem.

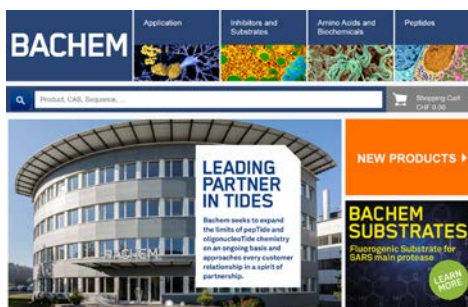


Fig. 3: Website entry page of the Bachem Online Shop.

What are peptides ?

Peptides are small proteins.
 Peptides and proteins differ only in their length:
 Proteins >100 amino acids
 Peptides 2-100 amino acids

Peptides and proteins are macromolecules, i.e. long molecules made from small subunits. They can be imagined as chains of pearls, composed of differently shaped beads. The “pearls” are the 20 proteinogenic amino acids (amino acids that are used in the cell to build peptides and proteins). They are linked together in any combination, i.e. the number of possible combinations is virtually infinite. The amino acids may be repeated in any frequency, so for example, glycine (Gly) in the protein collagen is found in every third position in the chain:
 ...-Xaa (mostly Pro (L-proline)-Yaa-Gly-...

Apart from water, our body consists mainly of proteins. Although sharing the same structural principle, they have a huge variety of functions (see Table 2). Proteins are also a vitally important constituent of our diet.

What makes this variety possible?
 Through variation in the amino acid building blocks and in length, macromolecules can be “constructed” with the most diverse properties imaginable!

Biologically active peptides are, for instance, hormones or toxins (see p. 23). The majority of our hormones are peptides of widely varying length:
TRH (thyrotrophin-releasing hormone) is a tripeptide (it consists of 3 amino acids).
LHRH (also known as GnRH (gonadotrophin-releasing hormone) is a peptide made from 10 amino acids (decapeptide).
Calcitonin consists of 32 amino acids.
pTH (parathormone), with its 84 amino acids, is almost a protein.
Insulin consists of 2 chains of peptides, one composed of 30 amino acids and one of 21 that are linked to each other via disulfide bridges (see p. 23).

In general, peptide hormones are produced by specialized cells, released into the blood and transported to the target organ. The cells to be stimulated have receptors that recognize the hormones and bind them specifically. These receptors are specialized proteins that are embedded in the cell membrane. The binding of the hormone acts as a signal for the cell and the desired effect is triggered.

- Short peptides
- 2 amino acids = dipeptide
 - 3 amino acids = tripeptide
 - 4 amino acids = tetrapeptide
 - 5 amino acids = pentapeptide etc.
 - a few amino acids (2-20) = oligopeptide

Table 2: Endogenous proteins and their functions

Examples	Occurrence of the protein	Requirements/function
Myosin, actin	Muscle tissue	Flexible molecule, contractible
Collagen (the most common protein in the body)	Connective tissue, tendons, skin	Molecule stable in shape, stretch-resistant
Hemoglobin, albumines	Blood	Transporter molecule, soluble
Enzymes such as the digestive enzyme trypsin in the small intestine (trypsin is a “protease”, which splits peptides and proteins)	Throughout the body	Enzymes = catalysts
Hormones such as thyrotrophin hormone (TSH) (stimulates the thyroid glands)	Throughout the body	Hormones = messengers
Antibodies (immunoglobulins)	Blood	Immune defense

With the 20 amino acids that can occur several times in the peptide, even with short peptides many combinations are possible, e.g. approx. 3.2 million different pentapeptides!

Even dipeptides can be biologically active: Leu-Trp* and other dipeptides lower blood pressure.

Acetyl-Asp-Glu (NAAG) is an important neurotransmitter, a substance that mediates the transmission of signals between nerve cells.

The peptides found in the body are usually obtained by enzymatic splitting or "cleavage" of proteins.

What are amino acids?

Structure of an α -amino acid

Four different substituents** are attached to a central carbon atom (the α -C atom):

- the acid group
(carboxylic acid group = carboxyl group, COOH, which Bachem abbreviates to -OH in the three-letter code*)
- a basic group
(amino group, NH₂, abbreviated H- in the three-letter code)
- the side chain
(R can vary greatly and thereby determines the properties of the peptide)
- a hydrogen atom
(H)

Accordingly, the α -C atom carries 4 groups with different chemical properties that play a very large role in what follows (Fig. 4).

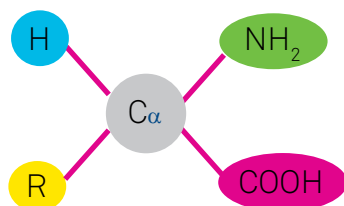


Fig. 4: α -amino acid, the α -C atom with the four different substituents.

As already mentioned, there are 20 different "proteinogenic" amino acids from which

proteins of the cell are made. They differ solely in the side chain R. In addition many other α -amino acids are found in nature, either free, as products of metabolism or incorporated into peptides/proteins. Examples are L-hydroxyproline (Hyp) in collagen, or L-ornithine (Orn) free in the urine. Other α -amino acids, e.g. L-norleucine (Nle) have so far only been obtained through chemical synthesis. Bachem calls non-proteinogenic amino acids such as Hyp and Nle "unusual amino acids". Often, they are designated as "unnatural amino acids" irrespective of their occurrence in nature.

Amino acids (whether proteinogenic or not) can be biologically active, e.g. L-tryptophan, L-glutamic acid.

The side chain R

- can be a hydrogen atom:
glycine, the simplest α -amino acid
- can carry an additional acid group (COOH):
aspartic acid, glutamic acid or a modified acid group (an amide, CONH₂): asparagine, glutamine
- can carry an additional basic group:
arginine (strong base), lysine, histidine (weak base)
- can carry a polar group:
serine, threonine
- can be a hydrocarbon (non-polar):
alanine (R = methyl), phenylalanine (R = benzyl), valine (R = isopropyl)
- can contain sulfur:
cysteine, methionine

In addition to the α -amino acids from which peptides and proteins are built, Bachem has many other amino acids on offer, including those in which the amino group is bound to another carbon atom.

L- und D-amino acids

The four different substituents of the α -carbon atom are not arranged in one plane, but lie at four corners of a tetrahedron with the α -C atom in the center (Fig. 5).

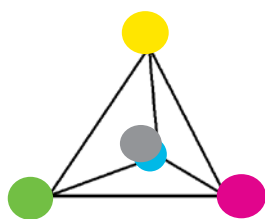
Therefore two forms of the amino acid molecule are possible that behave like mirror

*See Tab. 4 on p.9 for the meaning of the abbreviations and codes.

**Substituents are atoms or groups of atoms, that are bound to a central carbon atom (which is, in case of α -amino acids, the α -C atom) instead of a hydrogen atom.

Table 3: The functional groups important for amino acids

Name	Structure	Name	Structure
Amino group	NH ₂	Carboxyl group	COOH
Hydroxyl group	OH	Amide group (Carboxamide group)	CONH ₂
Thiol or mercapto group (in cysteine)	SH	Guanidino group (in arginine)	NH-C(=NH)-NH ₂


Fig. 5: Arrangement of the substituents around the α -carbon atom (gray).

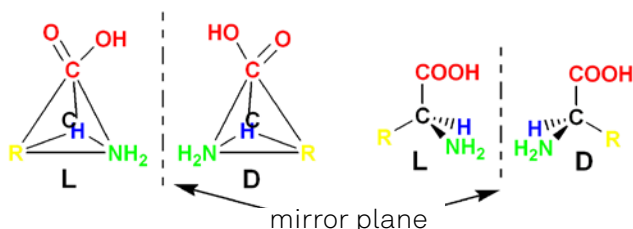
images of each other and are known as stereoisomers or enantiomers:

The two enantiomers do not differ in their physical properties, except that in solution, they rotate the plane of polarized light.

Compounds that do this are called "optically active". Optical activity is found very widely

images"; the D-amino acids, occur much less often in nature. D stands for dexter (Latin for "right"). To reiterate this point, L or D denote the arrangement of the four different substituents on the α -C atom, i.e. the respective enantiomer (see Fig. 6). The direction of rotation cannot be deduced from the designation L or D. In the case of L-amino acids, it can be positive or negative (see below), with the corresponding D-amino acids, it is always in the opposite direction.

The values of rotation are often stated as $[\alpha]$ on the ADS (see p. 24) of amino acids and their derivatives, since they are characteristic for the particular compound. Values of rotation are often also measured for peptides.


Fig. 6: Absolute configuration of the L- and D-isomers of amino acids.

in nature. Apart from glycine ($R = H$, image and mirror image identical), all proteinogenic amino acids show this phenomenon. Glucose (sugar) also rotates the light and this property is used to determine the concentration of glucose solutions, as does DNA and its building blocks.

All proteinogenic amino acids (apart from Gly) are L-enantiomers (L stands for laevus, Latin for "left"), e.g. L-alanine. Their "mirror

Some examples:

L-alanine: + 14.3°

D-alanine: -13.9°

L-tryptophan: - 31.8°

D-tryptophan: + 30.7°

(the discrepancies between the absolute values for the L- and corresponding D-enantiomer lie within the range of accuracy of the method)

Table 4: Abbreviations of the proteinogenic amino acids (L-form, except Gly)

Name	in peptide*	3-letter	1-letter	Name	in peptide	3-letter	1-letter
Alanine	alanyl	Ala	A	Leucine	leucyl	Leu	L
Arginine	arginyl	Arg	R	Lysine	lysyl	Lys	K
Asparagine	asparaginyl	Asn	N	Methionine	methionyl	Met	M
Aspartic acid	aspartyl	Asp	D	Phenylalanine	phenylalanyl	Phe	F
Cysteine	cysteyl	Cys	C	Proline	prolyl	Pro	P
Glutamine	glutamyl	Gln	Q	Serine	seryl	Ser	S
Glutamic acid	glutamyl	Glu	E	Threonine	threonyl	Thr	T
Glycine	glycyl	Gly	G	Tryptophan	tryptophyl	Trp	W
Histidine	histidyl	His	H	Tyrosine	tyrosinyl	Tyr	Y
Isoleucine	isoleucyl	Ile	I	Valine	valyl	Val	V

*s.Table 5 on p. 15

A 1:1 mixture of the two enantiomers is called a racemate. In this case, the rotations of the L- and D- form compensate each other.

In the Bachem notation, the L-form is not explicitly shown because of its ubiquity (e.g. L-alanine = H-Ala-OH). Only the less common D-form (D-alanine = H-D-Ala-OH) and the racemate, which is shown as DL, are designated as such.

Notation for amino acids

If the name of an amino acid is written out in full, the enantiomeric form is shown as well: L-alanine and D-alanine.

In the case of a racemate - as is common in organic chemistry - the DL is routinely omitted.

Amino acids and peptides are represented by their three-letter codes (usually the first three letters of the name).

In addition, there is also a one-letter code that is preferred, especially for longer peptides and proteins

Examples:

L-alanine → Ala, L-Ala, H-Ala-OH (3-letter);
A (1-letter)

L-arginine → Arg, L-Arg, H-Arg-OH (3-letter);
R (1-letter)

Bachem notation:

H-Ala-OH, H-Arg-OH

The notation signifies that it is the L-form and that the amino and carboxyl groups are free.

The abbreviations of the 20 proteinogenic amino acids are given in Table 4.

The one-letter code is only used for the 20 proteinogenic amino acids, for D-amino acids, the usual practice is to use small letters, e.g. f = D-Phe.

On the other hand, for many of the non-proteinogenic amino acids there is often also an obligatory three-letter code, e.g. Hyp (L-trans-hydroxyproline), Nle (L-norleucine), Orn (L-ornithine). Not all abbreviations usually found in the literature are also used by Bachem. For some amino acids, the non-abbreviated name is used, e.g. L-thiazolidine-4-carboxylic acid (otherwise Thz).

Brief portraits of the 20 proteinogenic amino acids

An organism such as the human body must have all 20 proteinogenic amino acids available in order to synthesize proteins. They all need to be in the L-form (except Gly).

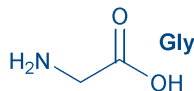
A distinction is drawn between essential and non-essential amino acids.

Non-essential amino acids can be synthesized by the body itself; essential cannot, i.e. they must be supplied in the diet.

Amino acids can also be subdivided according to the side chain (see p. 7) that determines their properties. There are acidic, basic, polar and non-polar amino acids. They are distinguished by different colors in the "Periodic Chart of Amino Acids" (see p. 13).

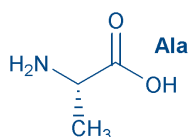
The properties of peptides and proteins are determined by the side chains of the amino acids from which they are constructed.

Glycine (Gly, G)



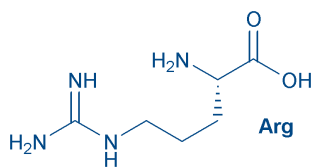
is the simplest amino acid and one of the most frequently encountered in peptides/proteins. Gly is non-essential. Glycine owes its name to its sweet taste.

Alanine (Ala, A)



is the simplest optically active amino acid and the second most common in proteins. The L-form shown here of this non-essential amino acid is incorporated into proteins by the body, but D-alanine is also found quite frequently in nature in peptides. Alanine is one of the non-polar amino acids.

Arginine (Arg, R)

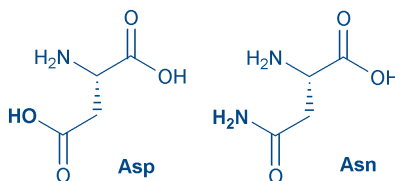


is a strongly basic amino acid. The nitrogen-containing guanidino group in the side chain is such a strong base that it forms a salt with virtually any acid. It owes its name to the fact that it was isolated as the silver salt. Arg is semi-essential (the body can produce it, but often in insufficient amounts). Arg-containing peptides are highly polar and readily water-soluble.

Asparagine (Asn, N)

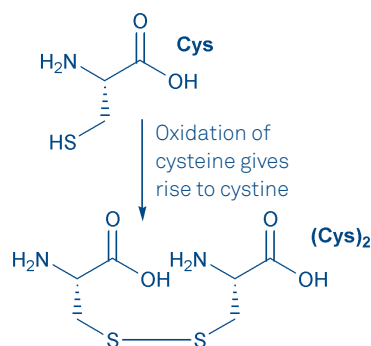
is actually a derivative of aspartic acid (the amide) and like the latter, was discovered in asparagus. Asn is a polar amino acid and is relatively labile in peptides and proteins.

Aspartic acid (Asp, D)



is an acidic amino acid because it contains a further acid group in the side chain. This persists when Asp is incorporated into a peptide. Aspartic acid was discovered in asparagus, which helped give it its name. Asp is non-essential for mammals.

Cysteine (Cys, C)

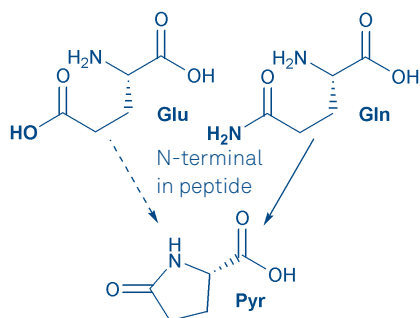


is a rare amino acid, although it is extremely important for the structure of peptides and proteins. The non-essential Cys contains sulfur in the side chain, as a slightly acidic thiol group. Thiol groups are readily oxidized, with the formation of a disulfide bond. One molecule of cystine (Cyt or (Cys)₂) is formed from two molecules of cysteine. In peptides that contain a Cys, two chains are linked together. If there are two Cys in the peptide, then a ring can be produced, as we will see again later.

Glutamine (Gln, Q)

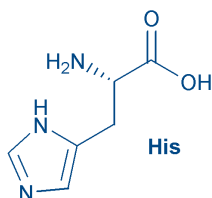
is actually a derivative (the amide) of glutamic acid. Glutamine is a polar amino acid that occurs as the free form in large quantities in the human body. Glutamine at the start of a peptide chain forms pyroglutamic acid spontaneously, or with the help of an enzyme.

Glutamic acid (Glu, E)



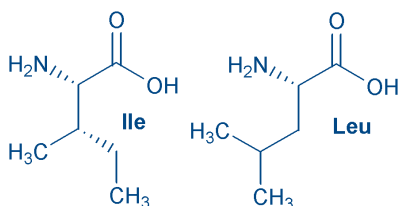
is an acidic amino acid. Glu is non-essential. The salts of glutamic acid are called glutamates and the sodium salt (monosodium glutamate, MSG) is known as a taste-enhancer. Glu has great physiological importance in our nervous system as a neurotransmitter. Glu forms pyroglutamic acid much less readily than glutamine.

Histidine (His, H)



is a weakly basic, polar amino acid that is essential for humans. The name comes from "histos" (Greek for tissue). The imidazole ring that it carries in the side chain, contains two nitrogen atoms. Imidazoles catalyze many reactions and therefore His (in combination with Cys, Ser or Thr) is often to be found in the active center of enzymes.

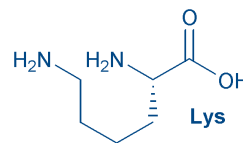
Leucine (Leu, L) and Isoleucine (Ile, I)



are isomeric α -amino acids. Isomeric compounds have the same molecular formula and the same molecular weight, but differ in their structure. Both are non-polar molecules. Leu and Ile are essential for humans.

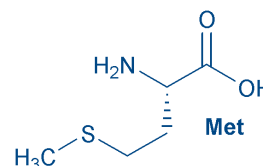
The name leucine comes from the white (Greek: leucos) leaves in which it crystallizes. Further stereoisomers are possible with isoleucine (allo-Ile).

Lysine (Lys, K)



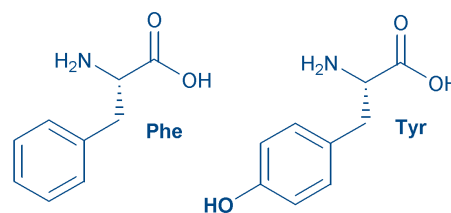
is a basic amino acid, that carries a second amino group in its side chain. It is an α,ϵ -diamino acid. Lysine is essential.

Methionine (Met, M)



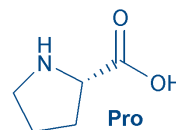
is, like cysteine, a sulfur-containing amino acid. The sulfur is present as methyl thioether, as reflected in the name "Me-thio". Thioethers are sensitive to oxidation, which should be heeded when handling Met-containing peptides. This essential amino acid is one of the non-polar amino acids.

Phenylalanine (Phe, F)



is an essential, non-polar amino acid. Together with His, Tyr (that can be considered as a phenylalanine derivative) and Trp, it is one of the aromatic amino acids.

Proline (Pro, P)



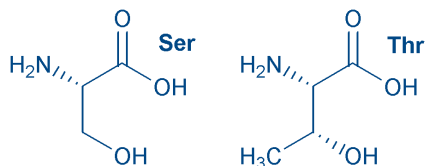
is the only cyclic proteinogenic amino acid, a ring of 5 atoms that contains an α -amino acid group and an α -C. Pro is non-polar and

non-essential. Due to its special structure, Pro has a huge influence on the spatial structure of peptides and proteins.

Serine (Ser, S)

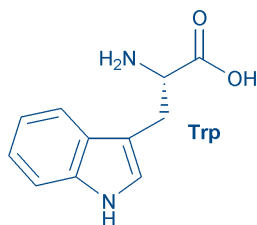
is a polar amino acid due to the hydroxyl group in the side chain. The name of this non-essential amino acid is derived from the Latin (sericum) for silk.

Threonine (Thr, T)



like serine, contains a hydroxyl group in the side chain, but is an essential amino acid. Threonine, in common with isoleucine, contains a second asymmetric carbon atom with four different substituents, so other stereoisomers are possible (allo-Thr).

Tryptophan (Trp, W)

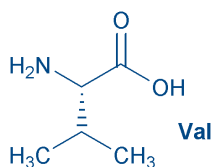


is a non-polar essential amino acid, an indole derivative. The free amino acid has an antidepressant effect and is a precursor of serotonin. Tryptophan fluoresces in the UV range (308-350 nm).

Tyrosine (Tyr, Y)

since it can be formed in the body from phenylalanine, tyrosine is a non-essential amino acid. The relatively non-polar amino acid was first isolated from cheese (Greek: tyros).

Valine (Val, V)





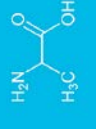

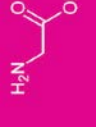
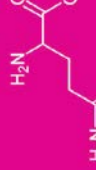
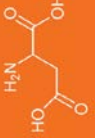
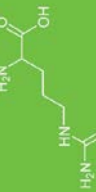
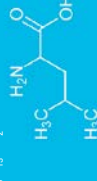
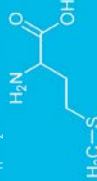
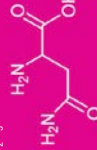
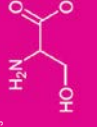
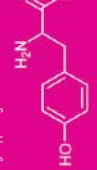

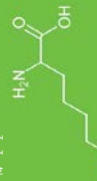
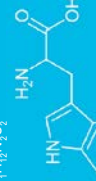
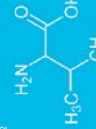
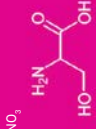
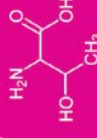

is a non-polar essential amino acid.

Amino acid derivatives

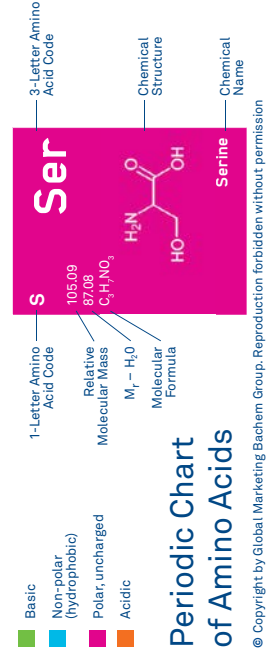
Derivatives in organic chemistry are understood to mean descendants of a compound. A derivative of an amino acid is obtained by modification (chemical alteration) of the amino group and/or the carboxyl group and/or the side chain. If this modification can be reversed under certain conditions that do not alter the amino acid, it can serve as a protecting group. Such derivatives are suitable building blocks for peptide synthesis.

Parent Compound	Derivative
H-Ala-OH	→ Ac -Ala-OH (amino group blocked)
	→ H-Ala- NH₂ (carboxyl group blocked)
	→ Fmoc -Ala-OH (amino group blocked)
	However Fmoc can be selectively split off ("cleaved") under mild conditions, i.e. is a suitable protecting group for peptide synthesis.

BACHEM

<p>H</p> <p>His</p> <p>155.16 137.14 C₆H₉N₃O₂</p>  <p>Histidine</p>	<p>F</p> <p>Phe</p> <p>165.19 147.18 C₉H₉N₃O₂</p>  <p>Phenylalanine</p>	<p>A</p> <p>Ala</p> <p>89.09 71.08 C₃H₇NO₂</p>  <p>Alanine</p>	<p>C</p> <p>Cys</p> <p>121.16 103.14 C₃H₇NO₂S</p>  <p>Cysteine</p>	<p>G</p> <p>Gly</p> <p>75.07 57.05 C₂H₅NO₂</p>  <p>Glycine</p>	<p>Q</p> <p>Gln</p> <p>146.15 128.13 C₅H₉N₃O₃</p>  <p>Glutamine</p>	<p>D</p> <p>Asp</p> <p>133.10 115.09 C₄H₇NO₄</p>  <p>Aspartic Acid</p>
<p>R</p> <p>Arg</p> <p>174.20 156.19 C₆H₁₄N₄O₂</p>  <p>Arginine</p>	<p>L</p> <p>Leu</p> <p>131.18 113.16 C₆H₁₃NO₂</p>  <p>Leucine</p>	<p>M</p> <p>Met</p> <p>149.21 131.20 C₅H₁₁NO₂S</p>  <p>Methionine</p>	<p>N</p> <p>Asn</p> <p>132.12 114.10 C₄H₈N₂O₃</p>  <p>Asparagine</p>	<p>S</p> <p>Ser</p> <p>105.09 87.08 C₃H₇NO₃</p>  <p>Serine</p>	<p>Y</p> <p>Tyr</p> <p>181.19 163.17 C₉H₉NO₃</p>  <p>Tyrosine</p>	<p>E</p> <p>Glu</p> <p>147.13 129.11 C₅H₉NO₄</p>  <p>Glutamic Acid</p>
<p>K</p> <p>Lys</p> <p>146.19 128.17 C₆H₁₄N₂O₂</p>  <p>Lysine</p>	<p>W</p> <p>Trp</p> <p>204.23 186.21 C₁₁H₁₇N₃O₂</p>  <p>Tryptophan</p>	<p>V</p> <p>Val</p> <p>117.15 99.13 C₅H₁₁NO₂</p>  <p>Valine</p>	<p>1-Letter Amino Acid Code</p> <p>S</p> <p>105.09</p> <p>Relative Molecular Mass</p> <p>C₃H₇NO₃</p> <p>M_r - H₂O</p> <p>Molecular Formula</p>	<p>3-Letter Amino Acid Code</p> <p>Ser</p>  <p>Serine</p> <p>Chemical Structure</p> <p>Chemical Name</p>	<p>T</p> <p>Thr</p> <p>119.12 101.10 C₄H₉NO₃</p>  <p>Threonine</p>	<p>I</p> <p>Ile</p> <p>131.18 113.16 C₆H₁₃NO₂</p>  <p>Isoleucine</p>

Periodic Chart of Amino Acids



© Copyright by Global Marketing Bachem Group. Reproduction forbidden without permission

Chemical Manufacture of Peptides

How is a peptide synthesized?
 Why synthesize peptides at all?
 Can't nature do it better?

Chemical Synthesis

has advantages as:

- Any scale is possible
 - It is simpler than biological methods for shorter peptides
 - The product is BSE/TSE-free
- and additional possibilities:
- A change of the amino acid sequence of known peptides is feasible
 - Modifications of peptides can be achieved (see p. 26)

The aim is to improve on nature, i.e. to modify the biological activity of the peptide:

- Increase the desired activity (biologically active peptides usually show several different activities)
- Optimize stability
 "not better, but longer"
 (peptides cannot be taken orally and are rapidly broken down in the body, which is not favorable for a drug)
- Minimize side effects
- Elucidate structure-activity relationships
- Synthesize peptides that do not occur naturally, and many more.

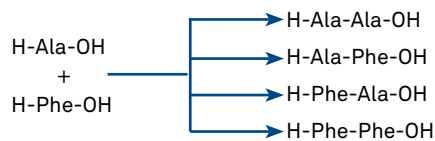
Synthesis of a dipeptide from two different amino acids

Example:

L-Alanyl-L-phenylalanine (H-Ala-Phe-OH)

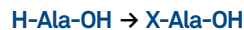


Two amino acids react with each other to form the dipeptide, with separation of water. This does not happen spontaneously, so they have to be activated for the reaction to occur. For this purpose "coupling reagents" have been developed. They generate more reactive derivatives of the amino acids and remove the water which is formed concomitantly from the system. However, if a coupling reagent is simply added to a solution of alanine and phenylalanine, it results in a mixture of the 4 possible dipeptides; tri- and longer peptides can be produced as well:

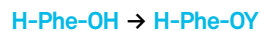


In order to actually obtain the desired dipeptide H-Ala-Phe-OH and not a mixture, two protecting groups are needed. They must not be split off under the conditions of the coupling reaction, but be readily cleaved in a separate step after coupling has taken place.

The amino group of alanine must be blocked:



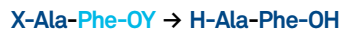
and the acid group of phenylalanine:



Now only the acid group of alanine can react with the amino group of the phenylalanine:

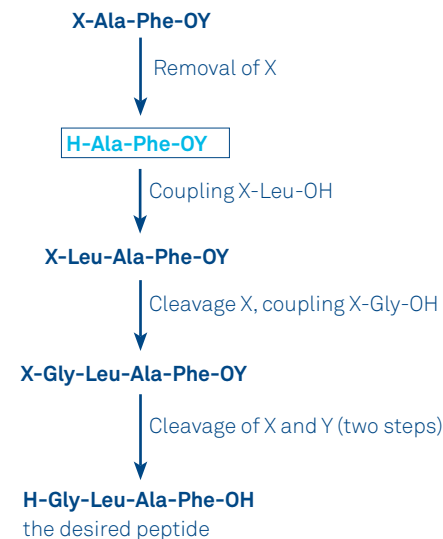


Finally, the protecting groups are removed:



i.e. X and Y are removed under the same conditions.

Now let us choose X and Y so that X is removed under conditions in which OY is retained. This allows us to synthesize longer peptides:



In this way, a peptide can be elongated to any desired length. X is called a temporary protecting group, because it is only used for the coupling step. Y is a permanent protecting group. It must be stable enough to endure all the coupling and X-cleavage steps. However, it must be removable in the final step, without damaging the peptide. A peptide always has two different groups at each end, and thereby a "direction". The amino group at the end of a peptide is known as the N-terminus, the carboxyl group at the other end, the C-terminus. In chemical synthesis, the peptide is constructed from the C-terminus in the direction of the N-terminus and therefore the terminal carboxyl group must be protected throughout the entire synthesis. Irrespective of the method of notation, when representing a peptide sequence, one starts at the N-terminus, see Table 5.

Table 5: Notation of peptides

Notation	Explanation
H-Gly-Leu-Ala-Phe-OH	Bachem notation, here the N- and C-terminals are easily recognized, both are "free", not modified
Gly-Leu-Ala-Phe	Terminal groups are only explicitly stated if they are not free
GLAF	1-letter code. Terminal groups are only explicitly stated if they are not free
Glycyl-L-leucyl-L-allyl-L-phenylalanine	"written out in full", especially common for shorter peptides

Amino protecting groups

Table 6 lists the protecting groups most commonly used by Bachem for the α -amino group. They are cleaved with a variety of methods. Examples of N^{α} -protected amino acid derivatives:

Z-Leu-OH
 Boc-Ala-OH
 Fmoc-Phe-OH

Many Boc or Z-amino acid derivatives are available as DCHA or CHA salts ((Di)cyclohexylammonium salts). Salt formation improves the stability on storage of acid-sensitive derivatives, because amino acids are quite strong acids, stronger than acetic acid. Some Boc or Z-derivatives are oils; they can only be obtained in solid, crystalline form as salts. Solid products are much easier to handle than oils, e.g. when weighing.

Benzylloxycarbonyl (Z or Cbz) is the oldest usable N^{α} -protecting group for amino acids (M.Bergmann & L.Zervas 1932). The development of Z brought about the start of modern peptide synthesis and hence also the abbreviation Z in honor of Leonidas Zervas. The Z group is often today still used to protect the amino group, also in organic synthesis.

Protecting groups for carboxyl moieties

In contrast to the α -amino group, the terminal acid group must be protected throughout the entire synthesis. The most commonly used protecting groups are listed in Table 7.

Examples of esters of amino acids:

H-Ala-OtBu · HCl
 H-Val-OMe · HCl
 H-Glu(OBzl)-OBzl · p-tosylate

Esters of amino acids are only stable on storage in the form of their salts with strong acids such as hydrochloric acid (hydrochloride) or p-toluene sulfonic acid (p-tosylate). These salts are also more readily available in crystalline form than the free amino acid esters.

The choice of amino and carboxyl protecting group also depends on the method of synthesis. The two most important methods are described in the next section. In the case of solid-phase synthesis (SPPS), the C-terminal protecting group is an insoluble polymer.

Side chain protecting groups

The almost infinite variety of the side chains of amino acids and their great significance for the properties of the peptide has already been mentioned. During peptide synthesis, some side chain groups must be permanently protected.

Table 6: Common temporary amino protecting groups

Abbr.	Chemical name	Cleavage reagent
Fmoc	9-Fluorenylmethoxycarbonyl	Piperidine
Boc	t-Butoxycarbonyl	Trifluoroacetic acid
Z	Benzyloxycarbonyl	Catalytic hydrogenation Hydrogen / palladium

Table 7: Common permanent acid protecting groups for the C-terminus and the side chain of Asp and Glu

Abbr.	Chemical name	Cleavage reagent
OtBu	t-Butylester	Trifluoroacetic acid
OBzl	Benzylester	Catalytic hydrogenation Hydrogen / palladium
OMe	Methylester	Bases (OMe used only for C-terminus)

Table 8: Examples of amino acid derivatives from the Bachem catalog

Derivative	Use
Fmoc-Arg(Pbf)-OH	Standard derivative of arginine for Fmoc-SPPS, must be activated
Fmoc-Asn(Trt)-OPfp	Standard derivative for Fmoc-SPPS, already activated ("active ester"), can be used directly
Z-Glu(OtBu)-OSu	Derivative for solution synthesis, already activated
Boc-Ser(tBu)-OH	ONLY for N-terminal serine! N ^α and side chain protection are cleaved under the same conditions

Table 9: Common side chain protecting groups

Amino acid	Protecting group	Main use
Arg	Pbf, Pmc	Fmoc-SPPS
Arg	(salt formation)	Solution synthesis
Asp, Glu	OtBu	Fmoc-SPPS, solution synthesis
Asp, Glu	OBzl	Solution synthesis
Asn, Gln	Trt, Mtt	Fmoc-SPPS
Cys	Trt	Fmoc-SPPS, solution synthesis
Cys	Acm	SPPS, solution synthesis
His	Trt	Fmoc-SPPS
Lys	Boc	Fmoc-SPPS, solution synthesis
Lys	Z	Solution synthesis
Ser	tBu	Fmoc-SPPS, solution synthesis
Ser	Bzl	Solution synthesis
Thr, Tyr	tBu	Fmoc-SPPS
Trp	Boc	Fmoc-SPPS

Table 10: Meaning of the abbreviations for the side chain protecting groups

Abbr.	Meaning	Main use
Acm	Acetamidomethyl	Cys, Fmoc-SPPS, solution synthesis
Boc	t-Butyloxycarbonyl	Lys & Trp, Fmoc-SPPS
tBu	t-Butyl	Ser, Thr & Tyr, Fmoc-SPPS
Bzl	Benzyl	Ser (Thr, Tyr) solution synthesis
OtBu	t-Butyl ester	Asp & Glu, Fmoc-SPPS
OBzl	Benzyl ester	Asp & Glu, solution synthesis
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl	Arg, Fmoc-SPPS
Pmc	2,2,5,7,8-Pentamethylchroman-6-sulfonyl	Arg, Fmoc-SPPS
Trt	Trityl (Triphenylmethyl)	Cys, His, Asn & Gln, Fmoc-SPPS
Z	Benzyloxycarbonyl	Lys, solution synthesis

Table 11: Comparison of methods – solution (or solution-phase) synthesis versus SPPS

	Solution synthesis	Solid-phase synthesis (SPPS)
Reaction medium	Solution	Gel (swollen insoluble polymer)
Batch size	Any (especially large)	Any (also very small)
Rapidity of synthesis	Slow	Fast
Automation	Difficult to achieve	Semi-automated or commercially available fully automated
Synthesis strategy	Generally convergent*	Generally stepwise
Temporary protecting group	Generally Boc or Z	Fmoc (Fmoc-SPPS), Boc (Boc-SPPS)
Side chain protection	Minimum	Maximum
Consumption of materials (amino acid derivatives)	Moderate	High
Optimization	Possible	Possible
Purification + analysis of intermediates	Usual	Not possible
Final purification	Relatively simple	Laborious

*convergent synthesis of longer peptides: first of all, several fragments are synthesized that are combined at the end to the final product. Convergent synthesis proceeds faster than a stepwise approach, because several fragments can be constructed at the same time. The combination “SPPS of protected fragments followed by coupling in solution” is also possible (e.g. synthesis of enfuvirtide).

Let us now look at the 20 proteinogenic amino acid in more detail:

Ala, Leu and Phe have side chains that do not react under the conditions of peptide synthesis. Likewise **Ile, Pro** and **Val**.

Other side chain functions can be modified during peptide synthesis, if they are not blocked. The carboxyl groups of **Asp** and **Glu** and the amino group of **Lys** can take part in the coupling. This produces a peptide with an incorrect link or a branched peptide.

Other side reactions occur with the side chain functions of

Cys
Ser, Thr, Tyr
Arg
His
Asn, Gln
Trp
Met (protection only in special cases)

The side chain functions of lysine and cysteine must be permanently blocked during peptide assembly. For the rest of the amino acids carrying a third functionality the need for permanent side-chain protection depends on the method of synthesis and the peptide sequence.

The most common side chain protecting groups are listed in Table 9 and the abbreviations are explained in Table 10. During solution synthesis, the side chains of Asn, Gln, His, Thr, Trp and Tyr do not necessarily have to be protected. In the case of Arg, salt formation with a strong acid such as HCl can suffice. Solution synthesis also does not make such high demands on the stability of protecting groups as SPPS.

As a general rule:

Amino acid derivatives for peptide synthesis should be available in microcrystalline form (no amorphous mass or oil) and thus readily soluble. The Fmoc-derivatives should dissolve readily and rapidly in DMF (or NMP); this is important for fully automated Fmoc-SPPS.

Methods of peptide synthesis

As already mentioned, there are two standard methods of peptide synthesis:

- Solid-phase synthesis (SPPS)

- Solution (or solution-phase) synthesis (“classic peptide synthesis”)

Table 11 shows the similarities and differences of the two methods.

Solid-phase synthesis (SPPS)

Most peptides sold by Bachem or manufactured as contract synthesis, are produced by SPPS. Solution synthesis is still chosen for the synthesis of some of our peptide generics, for very short peptides (e.g. dipeptides) and for peptides modified at the C-terminus.

The essential advantages of the solid-phase synthesis over the solution synthesis are its rapidity and ease of automation.

However, it is often said that the larger the batch, the slower the SPPS proceeds, because the “manual work” increases.

Small and medium quantities of peptides can be manufactured in fully automated “synthesizers” (Fig. 7). Very large quantities (several kg of raw peptide) are produced in semi-automated machines (only the washing program still runs automatically) or manually.

Automation is possible, because the peptide is constructed in steps from the C-terminus to the N-terminus. One amino acid is coupled after another.



Fig. 7: Synthesizer, fully automated machine for solid phase peptide synthesis. The left picture shows reactors in which synthesis is already proceeding.

Table 12: Protecting groups for Fmoc-SPPS

Group	Protecting group	Cleavage reagent	Stable against
N ^α	Fmoc	Piperidine	Trifluoroacetic acid
Side chains	Boc, tBu, OtBu, Trt, Pbf	Trifluoroacetic acid (also released from the resin)	Piperidine

The process always involves four repeated steps

1. Cleavage of the temporary protecting group
2. Washing
3. Coupling of the next protected amino acid
4. Washing

One disadvantage of SPPS is that purification can only take place at the end of synthesis, after release of the peptide from the resin. In solution synthesis, partial sequences are first constructed and put together during the final steps (see p. 17 and 21).

Cross-linked polystyrene is usually employed as support material (resin) in SPPS. The polystyrene is obtained in the form of beads. Bachem only uses polystyrene cross-linked with 1% divinylbenzene (polystyrene-co-1% divinylbenzene) as support. Two particle sizes are available for selection:

200 – 400 mesh = 38 – 75 μm

100 – 200 mesh = 75 – 150 μm

(mesh: the beads resulting from polymerization are separated by sieving into fractions of a certain particle size).

Due to the cross-linking, the polystyrene beads are not soluble in any solvent, but in some, they swell to form a gel (see Fig. 8). What is wanted is good swelling in dimethylformamide (DMF)

or N-methylpyrrolidone (NMP), the standard solvents of SPPS.

The better the swelling, the more gel-like the beads and the faster the reaction occurs in the resin beads.

Depending on the nature of the temporary N^α protecting group, a distinction is drawn between Fmoc- and Boc-SPPS. A switch Fmoc <-> Boc during synthesis is not possible. Boc- and Fmoc-SPPS also differ in the choice of the side chain protecting group.

Due to its great importance, only Fmoc-SPPS will be discussed in detail here. Because of the milder reaction conditions and hence better quality of the raw peptide, Fmoc-SPPS has prevailed over the cheaper Boc-variant.

Wang resin (4-alkoxybenzyl alcohol resin) and 2-chlorotrityl resin (2-chlorotrityl chloride resin) are the resins most commonly used for the Fmoc-SPPS of peptides with a free C-terminus. Resin loading with the carboxy-terminal Fmoc amino acid is not a simple matter and therefore Bachem offers a wide choice of loaded resins. The Fmoc group is split off (when using loaded Wang resin) and the next Fmoc amino acid coupled. Completion of the coupling is



Fig. 8: Swollen resin beads under the microscope.

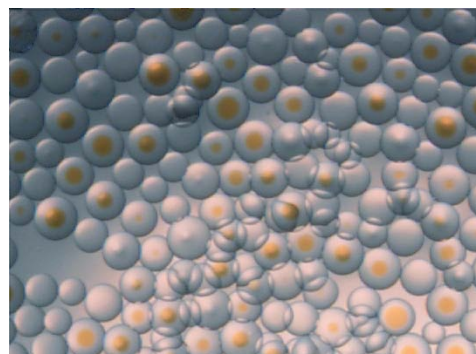


Fig. 9: Incomplete coupling. The color shows unconverted amino groups.

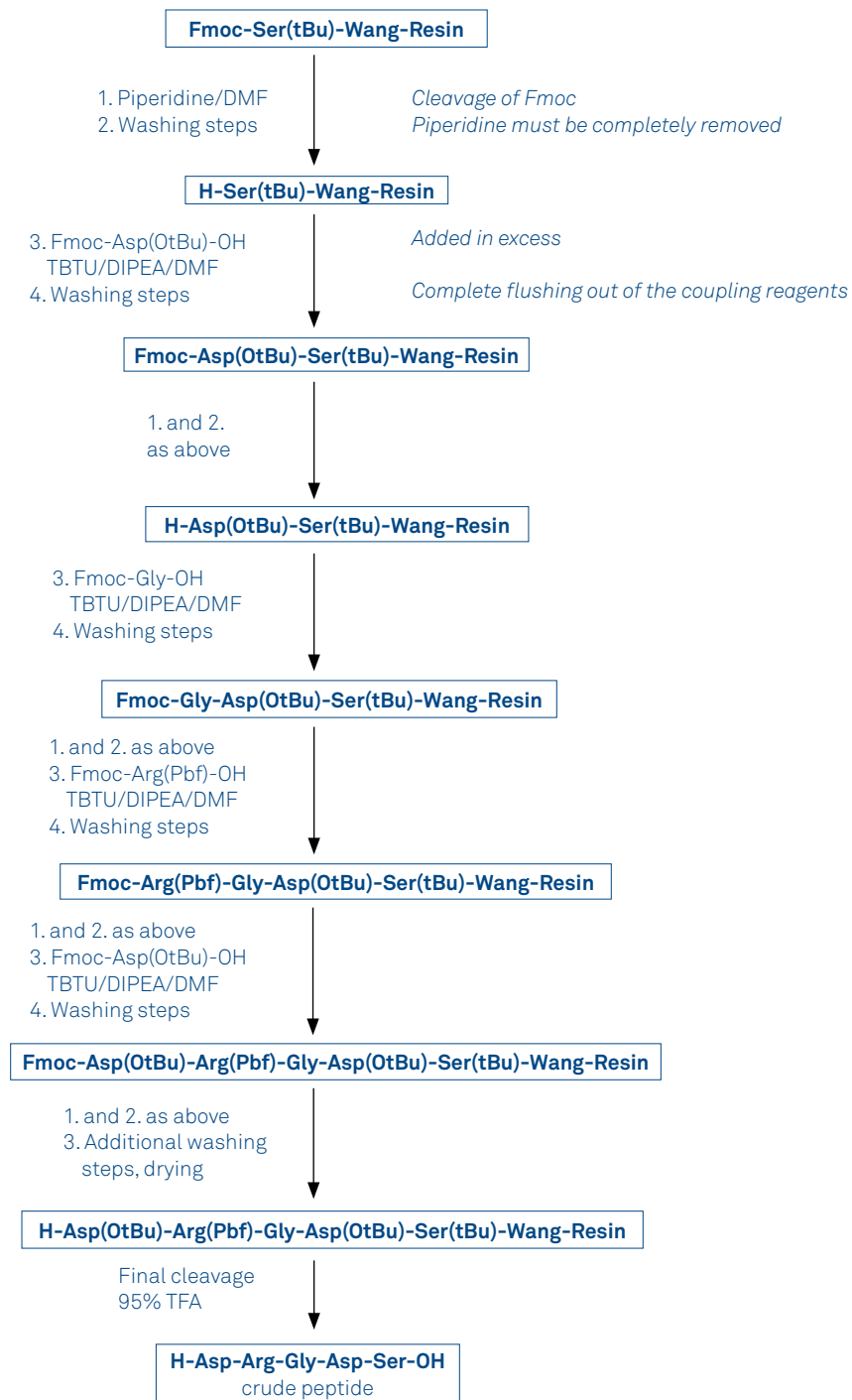


Fig. 10: Synthesis of **H-Asp-Arg-Gly-Asp-Ser-OH** (DRGDS).

TBTU (Product No. 4013268) is a common coupling reagent. To activate the Fmoc amino acid with TBTU, an equivalent amount of base (DIPEA, diisopropylethylamine) must be added.

determined with color tests (positive: free amino groups are “dyed”; negative: colorless, because there are no more free amino groups).

Fig. 9 shows how the test appears under the microscope in the case of incomplete coupling. Since the peptide is cleaved from the resin with trifluoroacetic acid at the end of synthesis, the side chain protecting groups must be labile against this acid. The protecting groups used are listed in Table 12.

Fig. 10 gives an example of an SPPS of an “RGD” peptide.

A different resin such as Ramage resin 4027095 or Rink amide AM-resin 4014571 is needed if a C-terminal peptidamide is to be synthesized. After cleavage of Fmoc from the resin, the C-terminal Fmoc amino acid is coupled to it and then the procedure is continued as described above. Resins such as 2-chlorotrityl resin or SASRIN enable even fully protected peptide fragments to be synthesized.

Solution synthesis

A solution synthesis requires careful planning. There is a standard protocol for Fmoc-SPPS that works for the synthesis of most peptides (although the quality of the raw peptide can be poor), but no such protocol exists for solution synthesis. There is a far greater variety with solution synthesis - not only for the possible protecting groups - but also of coupling methods, the solvents used and the preparative methods.

Not every conceivable protecting group combination works.

If an α -amino protecting group has been specified (Z, Boc, Fmoc...), then this also restricts the choice of side chain protecting groups, especially if one wants to remove all of them in one step at the end of synthesis. The protecting group combination Fmoc/tBu is particularly efficient and favored in SPPS, because the groups can be selectively cleaved and in any order. A comparable combination in solution synthesis is Boc/Bzl or Z/tBu, as shown in Table 13.

An example of a solution synthesis of a peptide modified at the C-terminus (a “peptide

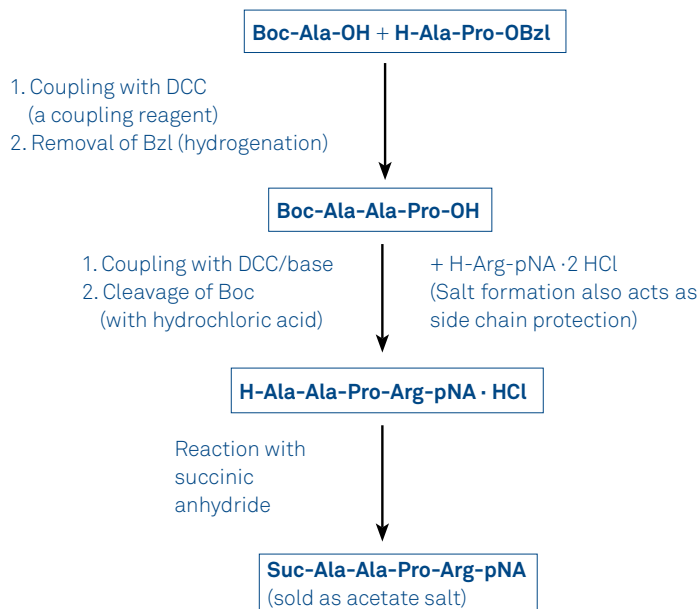


Fig. 11: Synthesis of **4017320 Suc-Ala-Ala-Pro-Arg-pNA** in solution.

Table 13: Protecting groups for solution synthesis

Group	Protecting group	Cleavage reagent	Stable against
N ^α	Boc	Trifluoroacetic acid	catalytic hydrogenation
Side chains, possibly C-terminal	Bzl, OBzl, Z	catalytic hydrogenation	Trifluoroacetic acid*
N ^α	Z	catalytic hydrogenation	Trifluoroacetic acid*
Side chains, possibly C-terminal	tBu, OtBu, Boc	Trifluoroacetic acid	catalytic hydrogenation
C-terminal	OMe**	Bases	Trifluoroacetic acid, catalytic hydrogenation

*Not completely stable against trifluoroacetic acid, but sufficiently stable for use in solution synthesis.

**The C-terminal methyl ester is often used for moderate-length fragments. On treatment with the base hydrazine hydrate, the hydrazide is obtained from the methyl ester (...Xaa-OMe →...Xaa-NHNH₂). This derivative can be “activated” (as azide) and coupled to another fragment with free N-terminus (“azide coupling”).

substrate”) is illustrated in Fig. 11.

The example also shows the benefits of the modular principle:

The fragment Boc-Ala-Ala-Pro-OH is also used in the synthesis of other substrates such as 4015678 (Boc-Ala-Ala-Pro-Ala-pNA) and inhibitors such as 4012071 (MeO-Suc-Ala-Ala-Pro-Ala-chloromethyl ketone). All intermediate steps are isolated, characterized and purified if necessary. This naturally takes time.

However, frequently-needed fragments can be synthesized in large quantities and held in stock, so saving time again.

Peptides that have been synthesized in solution are often purified by countercurrent distribution (CCD). This method allows to purify very large quantities, but it is

mostly used only for pre-purification.

In the example in Fig. 11, Boc was used as temporary N^α protecting group. Boc or Z are generally used in solution synthesis for temporary protection.

Reactive “active esters” of protected amino acids are also often used instead of a coupling reagent. Hydroxysuccinimide esters are popularly used in solution synthesis (Bachem: OSu, Fig. 12).

Purification of the raw product

If the quality is poor, purification of the raw peptide by preparative HPLC can be very time-consuming. Particularly with large quantities of peptide, the purification can prove a bottleneck.

With CCD and also with preparative HPLC,

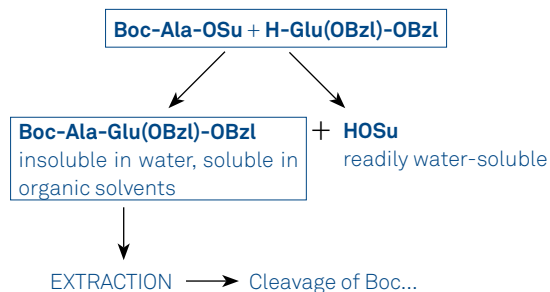


Fig. 12: Solution synthesis, coupling with active ester.

the purified peptide is obtained as a solution. The solvent is removed without stressing the product by freeze-drying (lyophilization) and a fine white powder is usually obtained. Water cannot ever be completely removed, but this is neither necessary nor worthwhile.

If a peptide contains basic groups (N-terminus, Arg, Lys, His) it binds trifluoroacetic acid (peptides obtained by Fmoc-SPPS) or acetic acid as salt. This cannot be removed. With few exceptions, our catalog peptides are available as trifluoroacetate or acetate.

Disulfide bridges

Many peptides in the catalog have the addition "Disulfide bond" in the name. Such a bond is formed by the oxidation of the thiol groups of two cysteines in the peptide chain. In the case of Fmoc-SPPS, this oxidation is carried out after cleavage from the resin, during which Trt is split off from Cys. Acm groups are cleaved with iodine and simultaneously oxidized yielding disulfide bonds.

A peptide can contain several disulfide bridges. Disulfide bridging and the correct linkage of the cysteine pairs (if there are several bridges) are extremely important for the biological activity of a peptide. Peptides do not exist as elongated chains, instead they form three-dimensional structures that are stabilized by disulfide bridges. Examples are shown in Fig. 13 and Fig. 14.

- One disulfide bridge can be found in e.g. calcitonin, somatostatin, vasopressin, oxytocin (natural peptide hormones); octreotide, desmopressin (hormone analogs, APIs)

- Several disulfide bridges are present in e.g. endothelin-1 (2), conotoxin (3), GaTx1, hepcidin-25 (4) (peptide hormones and toxins)

- Two peptide chains, linked by disulfide bridges: insulin, relaxins (peptide hormones)

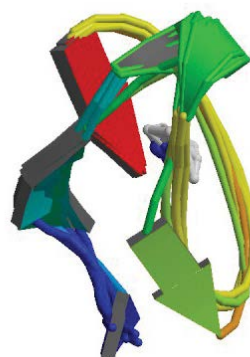


Fig. 13: omega-Conotoxin MVIIA.

Source: PDB **10MG**

Kohno, T., Kim, J. I., Kobayashi, K., Koder, Y., Maeda, T. and Sato, K.

Three-dimensional structure in solution of the calcium channel blocker omega-conotoxin MVIIA.

Biochemistry 34, 10256-10265 (1995)

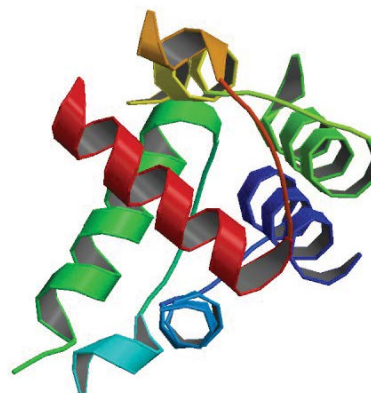


Fig. 14: Human relaxin (a hormone, has a similar structure to insulin).

Source: PDB **6RLX**

Eigenbrot, C., Randal, M., Quan, C., Burnier, J., O'Connell, L., Rinderknecht, E. and Kossiakoff, A. A.

X-ray structure of human relaxin at 1.5 Å. Comparison to insulin and implications for receptor binding determinants.

J. Mol. Biol. **221**, 15-21 (1991)

BACHEM
LEADING METHODS IN TESTS

Analytical Data Sheet	
Lot number	1029062 not for drug use
Catalog Number	H-9972
Product number	4042822
Product	Octreotide acetate salt H-D-Phe-Cys-Phe-D-Trp-Lys-Thr-Cys-L-threonine acetate salt (Disulfide bond)
Molecular formula	$C_{249}H_{344}N_{10}O_{52}$
Relative molecular mass	1019.26
Tests	Results
Appearance	white lyophilisate
Appearance of solution	clear, colorless solution (1 mg/mL in 90% acetic acid)
Identification (ESI-MS)	$m = 1019.2$ (average mass)
Purity (HPLC)	99.1% (TFA) 98.7% (TEAP)
Assay (elemental analysis)	83.8% (NH: 13.74%, NIG 11.52%)
Latest update: February 17, 2010	

Fig. 15: ADS of octreotide, a peptide also manufactured by Bachem as an API.

BACHEM
LEADING METHODS IN TESTS

Analytical Data Sheet	
Lot number	1041659 not for drug use
Retest date	April 2017
Catalog Number	B-1415
Product number	4003170
Product	Fmoc-Phe-OSu
Molecular formula	$C_{22}H_{24}N_2O_4$
Relative molecular mass	484.51
Tests	Results
Appearance	white powder
Appearance of solution	clear, colorless solution (50 mg/mL in dioxane)
Identification (TLC)	complexes with authentic material
Melting point	154 °C
Specific optical rotation	$[\alpha]_D^{25} = -25.1$ (2% in dioxane)
Purity (HPLC)	98.6% (TFA)
Enantiomer content (chiral chromatography)	< 0.1% D-enantiomer
Assay (elemental analysis)	99.8% (CB: 69.41%, CN: 69.24%)
Latest update: April 16, 2012	

Analytical Data Sheet
Lot number 1041659

Fig. 16: ADS of Fmoc-Phe-OSu, an amino acid derivative.

Analysis of amino acids and peptides

SPPS, purification and lyophilization produce a white powder. This must now undergo a detailed analysis (Quality Control, QC) to find answers to the following questions:

- Is it, in fact, the desired product?
- How pure is the product?
- What by-products does it contain?
- How high is the product content?

Our QC departments have a huge arsenal of equipment and methods available to analyze our products. The choice of analytical method also depends on the nature of

the product: an amino acid, a very short peptide or a peptide.

Identity

Is it the desired product?

If so, a measured value corresponds with that in the literature, or the value obtained with a reference.

The following parameters are used to demonstrate identity:

- Melting point
- Optical rotation $[\alpha]$ (see p.8)
- The R_f value is determined by thin layer chromatography (TLC):

R_f value = ratio of the distance travelled by the substance spot/solvent front. It depends on the solvent (mobile phase), R_f always < 1.

- Co-elution with reference (TLC):

Mixture of product with reference gives a single spot (and not two).

- The molecular weight is determined by mass spectrometry.
- The chemical composition of the compound is determined by elemental analysis: The content of carbon, hydrogen and nitrogen is measured and compared with the values that can be calculated from the molecular formula $C_xH_yN_zO_wS_v...$ Oxygen is only determined in exceptional cases. Sulfur can also be determined in-house.

- Spectroscopic methods:

IR, NMR, UV spectroscopy, for amino acid derivatives.

- Co-elution with reference (analytical HPLC):

Mainly with peptides.

- Amino acid composition:

Amino acid analysis (AAA), the peptide is cleaved into the individual amino acids with strong acid, the resulting mixture is separated by chromatography and the amino acids quantified. The correctness of the amino acid sequence cannot be checked with this method.

Purity

Purity is determined by:

- Thin layer chromatography (TLC) for amino acid derivatives, very short peptides, and biochemicals:

TLC is often a valuable method, because reference material of potential impurities e.g. the educts (reactants) can also be run at the same time and a large choice of detection methods is available.

- Analytical high performance liquid chromatography (analytical HPLC or the faster variant RSLC):

standard method for peptides.

- "Optical purity": content of incorrect enantiomers (see p. 7).

An important value for amino acid derivatives and loaded resins.

Product content

In addition to impurities arising during the synthesis, products may contain trace amounts of residual solvents and/or water. Peptides may also contain trifluoroacetic acid or acetic acid that is firmly bound as salt.

These compounds are not measured with the methods used to determine the purity of the desired product. In order to calculate a reaction mixture for example, parameters such as the water content must be known.

- Water content: determined with Karl Fischer titration (KF-titration).
- Residual solvents: determined by gas chromatography.
- Acetic acid content: by HPLC or ion chromatography.
- Residual trifluoroacetic acid, for our API peptides that are sold as acetate: by ion chromatography.
- Titration with solutions of acids or bases: for products that are bases or acids.
- Halide determination: the chloride or bromide content can be determined by titration with silver nitrate solution for products that are sold as the corresponding salts.
- Nitrogen content (from elemental analysis): for peptides, this is a measure of the peptide content.

A selection of the results obtained can be found on the ADS (Analytical Data Sheet) or CofA (Certificate of Analysis), that is made available to the customer who is buying the product (see Figs. 15 and 16).

In the case of GMP products, microbiological analyses are added that are carried out in a specialized laboratory at Bachem.

For many of our generics, the analyses to be undertaken are specified by the European Pharmacopoeia*. "Ph. Eur." in the name signifies that the drug substances conform to the specifications of the pharmacopoeia, e.g. Desmopressin Acetate Ph. Eur.

*A pharmacopoeia is a compilation of recognized pharmaceutical rules concerning the quality, purity, storage and naming of pharmaceuticals and the substances, materials and methods used in their manufacture and testing. The European Pharmacopoeia applies in Switzerland.

Modification of peptides

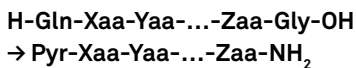
The modification of a peptide means a permanent chemical alteration of the molecule, in contrast to the protecting groups that are removed after synthesis.

Peptides can be modified in a great variety of ways.

In nature too, one finds a large number of peptide/protein modifications. Some are important for the bioactivity or function of the molecule (e.g. Pro in collagen is oxidized by enzymes to hydroxyproline; side chain phosphorylation of Ser, Thr or Tyr) others are “undesired”, or even indicators of pathological processes.

Modifications such as N-terminal acetylation or pyroglutamine formation and C-terminal amidation also serve in nature to stabilize peptides. “Blocked” peptides are less rapidly broken down by enzymes. Natural peptide hormones are often modified at the end groups, e.g. gonadorelin, TRH.

The C-terminal amide group is generated by enzymatic degradation of a glycine residue, another enzyme catalyzes the cyclization of Gln to pyroglutamate (Pyr):



Since the oxidation of methionine (see p. 11) in peptides often leads to deactivation, biologically active peptides are stabilized by the replacement of Met with the analogously structured, but non-oxidizable norleucine (Nle).

Chemical modifications also allow to fix or change the spatial structure of a peptide. The chemical synthesis of peptides offers many possibilities for modification. The N-terminus can be altered most easily: it only requires one additional step in SPPS. Important N-terminal modifications

- Acetylation
 - Biotinylation – biotin (= Vitamin B7)
- (many examples of acetylated and biotinylated peptides can be found in the cata-

log). Biotinylated peptides bind specifically to the protein avidin.

- “Fluorophore” – Mca, Abz, FITC
- (The modified peptides fluoresce and are detectable in the minutest amounts). The side chains of Cys and Lys can also be easily modified and examples can be found in the catalog.

The peptide “backbone” can be modified by the incorporation of special amino acids:

- D-amino acids
- N-methyl amino acids
- Non-proteinogenic amino acids (unusual amino acids)
- Replacement of the peptide bond by other bonds (e.g. reduced peptides)
- Ring closure, e.g. disulfide bridges (see p. 23)

Peptidamides, the most important modification of the C-terminal carboxyl group are readily accessible by SPPS. They are synthesized on specially developed resins such as Ramage resin (4027095). Much effort is required for other C-terminal modifications.

An example of modifications:

natural peptide

Leu-Enkephalin (Product No. 4000332)
 H-Tyr-Gly-Gly-Phe-Leu-OH

slightly modified

(3,5-Dibromo-Tyr¹)-Leu-Enkephalin
 H-**3,5-Dibromo-Tyr**-Gly-Gly-Phe-Leu-OH

Leu-Enkephalin amide

H-Tyr-Gly-Gly-Phe-Leu-**NH₂**

strongly modified

DAMGO (Product No. 4007829)

H-Tyr-**D-Ala**-Gly-**N-Me-Phe-glycinol**

(D-Pen²,p-chloro-Phe⁴,D-Pen⁵)-Enkephalin
 (Product No. 4017158)

H-Tyr-**D-Pen**-Gly-**p-chloro-Phe-D-Pen**-OH
 (disulfide bond)

BACHEM

BACHEM

Marketing & Sales Contact

Americas

Bachem Americas, Inc.

Tel. +1 888 422 2436 (toll free in USA & Canada)

+1 310 539 4171

sales.us@bachem.com

Asia Pacific

Bachem Japan K.K.

Tel. +81 3 6661 0774

sales.jp@bachem.com

Europe, Africa, Middle East and India

Bachem AG

Tel. +41 58 595 2020

sales.ch@bachem.com

Visit our website

www.bachem.com

or shop online

shop.bachem.com

All information is compiled to the best of our knowledge.
We cannot be made liable for any possible errors or misprints.
Some products may be restricted in certain countries.



www.bachem.com



shop.bachem.com