Chem Soc Rev

The Road to the Synthesis of "Difficult Peptides"

Journal:	*Chemical Society Reviews*
Manuscript ID	CS-SYN-08-2015-000680.R3
Article Type:	Review Article
Date Submitted by the Author:	13-Nov-2015
Complete List of Authors:	Albericio, Fernando; IRB Barcelona, Barcelona Science Park Paradis-Bas, Marta; IRB-Barcelona, Tulla-Puche, Judit; IRB Barcelona,

The Road to the Synthesis of "Difficult Peptides"

Marta Paradís-Bas,[*ab] *Judit Tulla-Puche,*[*abc] *and Fernando Albericio* [*abcd]

[a] *Institute for Research in Biomedicine (IRB Barcelona), Baldiri Reixac 10, 08028 Barcelona, Spain. E-mail: marta.paradis@irbbarcelona.org, judit.tulla@ub.edu, albericio@irbbarcelona.org; Fax: +34934037126; Tel: +34934037088*

[b] *CIBER-BBN, Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Barcelona Science Park, Baldiri Reixac 10, 08028 Barcelona, Spain*

[c] *Department of Organic Chemistry, University of Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain*

[d] *School of Chemistry & Physics, University of Kwazulu-Natal, Durban 4001, South Africa*

ABSTRACT

The last decade has witnessed a renaissance of peptides as drugs. This progress, together with advances in the structural behavior of peptides, has attracted the interest of the pharmaceutical industry in these molecules as potential APIs. In the past, major peptide-based drugs were inspired on sequences extracted from natural structures of low molecular weight. In contrast, nowadays, the peptides being studied by academic and industrial groups comprise more sophisticated sequences. For instance, they consist of long amino acid chains and show a high tendency to form aggregates. Some researchers have claimed that preparing medium-sized proteins is now feasible with chemical ligation techniques, in contrast to medium-sized peptide syntheses. The complexity associated with the synthesis of certain peptides is exemplified by the so-called

"difficult peptides", a concept introduced in the '80s. This refers to sequences that show inter- or intra-molecular β-sheet interactions significant enough to form aggregates during peptide synthesis. These structural associations are stabilized and mediated by non-covalent hydrogen bonds that arise on the backbone of the peptide and—depending on the sequence—are favored. The tendency of peptide chains to aggregate is translated into a list of common behavioral features attributed to "difficult peptides" which hinder their synthesis. In this regard, this manuscript summarizes the strategies used to overcome the inherent difficulties associated with the synthesis of known "difficult peptides". Here we evaluate several external factors, as well as methods to incorporate chemical modifications on sequences, in order to describe the strategies that are effective for the synthesis of "difficult peptides". These approaches have been classified and ordered to provide an extensive guide for achieving the synthesis of peptides with the aforementioned features.

1. INTRODUCTION

Peptide synthesis has become an attainable process since 1963,[1] when Professor Merrifield developed the revolutionary concept of constructing peptides using the solid-phase strategy. This breakthrough earned him the Nobel Prize in Chemistry while simultaneously providing researchers with an extraordinary tool to facilitate the synthesis of these molecules. Since the discovery of solid-phase peptide synthesis (SPPS), various synthetic improvements have emerged regarding the design of orthogonal protecting groups,[2] new functionalized solid supports,[3,4] efficient coupling reagents,[5] and even methodologies to overcome difficulties with sequence elongation. The synthesis of large peptides is now achievable thanks to relatively recent advances regarding the development of new synthetic tools. However, in spite of all the efforts made in these fields, certain sequences still require laborious syntheses.

In this regard, the so-called "difficult peptides" are an evident example of the need to further design strategies to enable their synthesis. Established in the '80s,[6] this peptide concept refers to sequences that show inter- or intra-molecular β-sheet interactions

significant enough to form aggregates. These structural associations occur during peptide synthesis using both solid-phase and solution-phase methodologies (Fig. 1)[7] and they are stabilized and mediated by non-covalent hydrogen bonds, which—depending on the sequence—are favored. These interactions arise on the backbone of the peptide, in particular between the hydrogen amides and the carbonyls.[8] The tendency of peptide chains to aggregate is translated into a list of common behavioral features attributed to "difficult sequences", detailed by Kent[9] and later by Milton.[10] The main relevant synthetic evidence provided by these authors are the following: repetitive incomplete aminoacylations (<15%) despite re-couplings; accentuated difficulties when resin loading is high[11] or when sterically hindered amino acids (AAs) are present in the sequence; and more importantly, slow or incomplete 9-fluorenylmethoxycarbonyl (Fmoc) removal.[12]

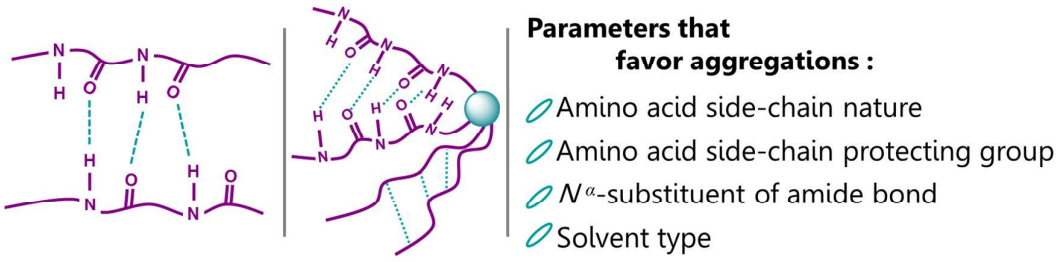


Fig. 1 Hydrogen bond inter-chain interactions during peptide synthesis (in solution or on solid-phase) and the factors that contribute to their formation.

Studies analyzing the influence of each AA on β-sheet interactions have confirmed that the capacity to induce intermolecular associations is sequence-dependent.[13] Nevertheless, the prediction of a "difficult peptide" by the AAs present in its sequence, *a priori*, is not evident. Several factors contributes to increasing the propensity to form inter-chain associations during the peptide synthesis which are associated with the AA

side-chains, with the protecting group of the AA, with the nature of *N*-substitution of the backbone amide, and even with the characteristics of the solvent used to synthesize them (Fig. 1).[10,14] Although the literature does not provide a comprehensive list of "difficult peptides", below we refer to some that fit the description and that have been synthesized or whose β-sheet folding mechanisms have been evaluated.

Table 1 Common examples of sequences that aggregate on solid-phase: "difficult peptides"

Sequences	Peptide Type	Ref.
RADA16-I [Ac-(RADARADA)$_4$-NH$_2$] **RADA16-II** [Ac-(RARADADA)$_4$-NH$_2$]	ionic self-assembling (Type I and Type II)	46 56
oligo-Ala, **oligo-Val** **oligo-Gln, oligo-Leu**	homooligo-pept.	17-23
Aβ (1-40) or **Aβ (1-42)**	amyloidogenic	29-33
Amylin	amyloidogenic	34, 35
ACP (65-74) (H-VQAAIDYING-NH$_2$)	-	13, 20 57-60
HIV-1 PR (81-99) (H-PVNIIGRNLLTQIGCTLNF-NH$_2$)	-	20 57-59
PnIA (A10) (H-GCCSLPPCALNNPDYC-NH$_2$)	-	20, 59
Thymosin α1	-	61-64

On the other hand, some peptides form β-sheet associations in solution only when their sequences are fully unprotected, not during the elongation. These peptides assemble in liquid media to establish stable interactions that evolve to aggregates. Although this behavior may be considered an appreciated property in certain fields, in terms of peptide characterization and/or purification it poses a significant drawback. These two post-

synthetic steps demand perfectly dissolved sequences in order to provide an accurate peptide analysis. Again, this phenomenon is not easily predictable by the AAs present in the sequence, as occurs for "difficult peptides". However, studies based on structural interaction experiments[15] have led to the development of computational methods to establish similarities between sequences that anticipate the aggregation of peptides.[16]

Table 2 Common examples of sequences that aggregate in solution

Sequences	Peptide Type	Ref.
Ac-**V_6D**-NH_2, Ac-**A_6D**-NH_2, Ac-**V_6D_2**-NH_2, Ac-**L_6D_2**-NH_2 Ac-**G_nD**-OH (n=4,8,10)	negatively charged surfactant-like	40-42
Ac-**V_6K_2**-NH_2, Ac-**L_6K_2**-NH_2, Ac-**A_6K**-NH_2, Ac-**KA_6**-NH_2, Ac-**V_6H**-NH_2, Ac-**L_6K**-NH_2, **H_2V_6**-NH_2, **KV_6**-NH_2	positively charged surfactant-like	41, 43
KFE-8 [Ac-$(FKFE)_2$-NH_2] **KLD-12** [Ac-$(KLDL)_3$-NH_2]	ionic self-assembling (Type I)	54 46, 47
DAR16-IV [Ac-(A-$(DA)_4$-$(RA)_3$-R)-NH_2]	ionic self-assembling (Type IV)	56
ELK8-II (Ac-LELELKLK-NH_2)	ionic self-assembling (Type II)	56
EAH16-II [Ac-$(AE)_2$-$(AH)_2$-$(EA)_2$-$(AH)_2$-NH_2] **EAK16-II** [Ac-$(EA)_2$-$(AK)_2$-$(AE)_2$-$(AK)_2$-NH_2]	ionic self-assembling (Type II)	56
RADSC-14 [Ac-$(RADSC)_2$-A_5-C-NH_2] **RADSC-16** [Ac-$(RADSC)_3$-A_3-C-NH_2]	ionic self-assembling (Type I)	55
Q7 (Ac-KFQFQFE-NH_2) **Q11** (Ac-QQKFQFQFEQQ-NH_2)	ionic self-assembling	50-52
DN1 (Ac-QQRFQWQFEQQ-NH_2) **P11-4** (Ac-QQRFEWEFEQQ-NH_2)	ionic self-assembling	26 43, 44
Sup35 (7-13) (H-GNNQQNY-NH_2)	amyloidogenic	36, 37

To sum up, peptide aggregation can arise during the synthesis or after the global deprotection. In this section, the peptides have been organized into families that present

aggregation in one or both stages. Therefore, the peptides described below share a β-sheet association tendency, and we highlight whether these interactions occur after the elongation (in solution) or during the synthesis (solid-phase).

1.1. Homooligo-Peptides

Homooligo-peptides are composed by only one type of AA. They are *de novo* synthetic sequences that show a high tendency to aggregate and they comprise hydrophobic AAs such as Ala, Gly, Ile, Leu or Val (Table 1). The aggregation is manifested not only during the synthesis of peptides on solid-phase, but also when the sequence is unprotected in solution. Those AAs most commonly analyzed are poly-alanine and poly-valine,[17] which show a notable capacity to assemble—in the case of oligo-alanine starting from the fifth residue. Moreover, the aggregating mechanisms of other homooligo sequences, such as poly-glutamine, have also been explored,[18,19] while poly-leucine has been studied to validate new synthetic strategies for "difficult peptides".[20] Hydrogen bond formation between the backbone amides of the peptide chains drives the assembly. The authors who initially studied the behavior of these sequences reported that, due to their characteristics, AA deletion occurs during the synthesis, thus demonstrating the structural conformation that explains this phenomenon.[21–23] In particular, the oligo-alanine has been widely used as a model to study conformational changes after chemical modifications.

1.2. Self-Assembling Peptides

Self-assembling peptides (SAPs), as their name denotes, are another class of sequences characterized by their capacity to assemble spontaneously in solution. This behavior appears when the peptide is completely unprotected after the cleavage from the

resin, thus precluding its purification once in solution phase. These dissolved molecules assemble into energetically favored β-structures that may subsequently form fibrils[24], which may evolve to hydrogelation. SAPs are not necessarily "difficult peptides", although some of them may also form β-sheet interactions during solid-phase synthesis and would also fit the description. Given their capacity to aggregate in solution, SAPs have a broad number of applications in biomedicine.[25–27] Despite the fact that the self-assembly is a desired property from a biomaterial perspective, this phenomenon underlies a number of human diseases.[28] In this regard, probably the most known SAP is amyloid-β peptide (Aβ), the primary peptide responsible for Alzheimer's disease.[29,30] Specifically, the two amyloid sequences involved in this neurodegenerative illness are the 40-mer Aβ (1-40) and the 42-mer Aβ (1-42) (Table 1), which misfold and cause insoluble non-native disordered aggregates, forming oligomers and fibrils.[31–33] Among the amyloidogenic peptides described, the hormone amylin (Table 1), another relevant sequence also present in humans, is highlighted because of its involvement in type 2 diabetes mellitus.[34,35] This hormone is a 37-mer peptide secreted in the pancreas, together with insulin. In a pathological condition, the excess of amylin self-assembles to produce deposits in the pancreas, thus causing non-insulin-dependent diabetes. One of the few amyloid-like peptides characterized by x-ray crystallography has been the 7-mer fragment of yeast prion protein Sup35, specifically the region (7-13)[36] (Table 2). The Sup35 protein exists in a non-infectious form but when folded as an infectious variant it is called a prion, the 7-mer region playing a key role in the latter. Moreover, studies on Sup35 heptapeptide fibril formation have been performed to understand the aggregation mechanisms.[37] Apart from the aforementioned peptides found in biological systems, there are other *de novo* designed sequences, classified as SAPs, which are not present in nature. These are commonly sub-divided into two families, which are described below.

1.2.1. Amphiphilic Self-Assembling Peptides

The first family comprises amphiphilic or amphipathic peptides (also named surfactant-like peptides),[38,39] whose sequences have two differentiated parts, one hydrophobic (more than three non-polar AAs) and one hydrophilic (one or two polar or charged AAs). Depending on the length of the hydrophobic core, SPPS may be hindered as occurs in "difficult peptides", the longest chain being the most difficult to achieve. Nevertheless, in solution, all these amphiphilic self-assembling peptides tend to arrange by forming nanotubes or nanovesicles[40] (see Fig. 2).

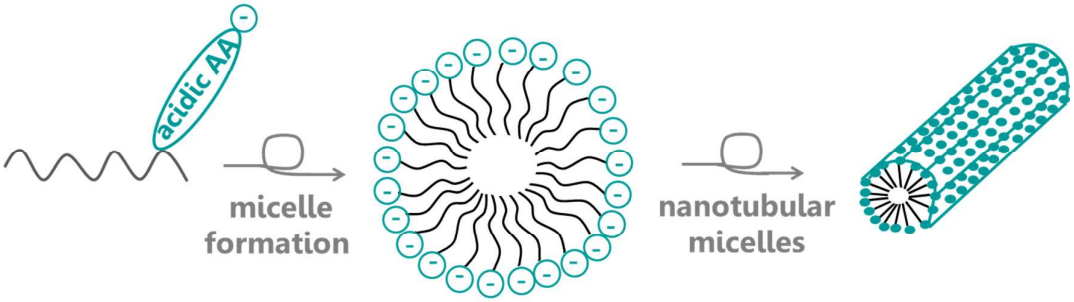


Fig. 2 Schematic representation of the structural organization adopted by a negatively charged amphiphilic SAP model.

Regarding their applications, it is worth highlighting the membrane protein stabilization conferred by some of these surfactant-like peptides.[41] Sequences such as V_6D, A_6D, and V_6K_2 are examples of amphiphilic peptides with a demonstrated capacity to assemble in water solutions when at a certain concentration[40–43] (Table 2). When the hydrophilic part of these peptides holds one kind of ionic charge, they are simply amphiphilic, whereas when they hold the two charges, they belong to the ionic peptide group.

1.2.2. Ionic Self-Assembling Peptides

The sub-type of ionic self-assembling peptides is characterized by the stabilization of aggregates as a result of ionic interactions, effect caused by positively and negatively charged AA side-chains.[44,45] The ion-ion associations occur once sequences are unprotected, thus the stabilization of aggregates in these peptides in solution is strong. Although most sequences belonging to this classification can be produced without difficulties on SPPS, some also present β-sheet interactions during the synthesis and can therefore also be classified as "difficult peptides". The *de novo* designs composed by a rational arrangement of oppositely charged AAs have served as self-assembling sequences (Table 1 and Table 2). Some of these structures, which are often referred to as "peptide lego", exhibit supra-structures that have reached the market for regenerative medicine.[46–48] Ionic self-assembling peptides are characterized mainly by the formation of complementary ionic interactions that are electrostatically orientated between two peptide chains facing opposite charges, thus favoring the self-assembling design when this behavior is desired. Peptide designs with alternating hydrophobic and hydrophilic AAs were among the initial methods proposed to achieve sequences with β-sheet interactions of interest[49–53] (see Fig. 3). Later, Zhang, a relevant author in the field of well-defined self-assembling ionic sequences, proposed the division of these ionic self-assembling peptides into modules (I, II, III,...) (Table 1 and Table 2) on the basis of their positive and negative charge arrangement (type I: + − + − + −; type II: + + − − + +; type III: + + + − − − ; ...).[46,47,54–56] One known example of this kind of peptide is RADA-16, also known as RADA-16-I because it belongs to type I. This peptide is part of a group of sequences in which the first described member was EAK-16,[57,58] which is a segment of a yeast protein with Z-DNA binding properties. In contrast, RADA-16 is

an artificial *de novo* designed structure.[59,60] The primary structure of RADA-16 is composed by a total of 16 residues with the peculiarity of comprising four repeats of the AA sequence Arg-Ala-Asp-Ala (R-A-D-A). It is precisely the alternating and complementary charges characteristic of this kind of sequences under an appropriate pH that lead them to adopt a β-sheet conformation. Various structural analyses of RADA-16 have been performed to understand nanofiber formation and its tridimensional organization.[55,61–63] In 2002, the north-American company BD Bioscience launched RADA-16 as a bioactive hydrogel named PuraMatrix®,[64] which was later commercialized under a different name by other companies and with other degrees of purity. This peptide has been extensively described for several medical applications and a few years ago we undertook the synthesis of RADA-16 using standard SPPS strategies.[65]

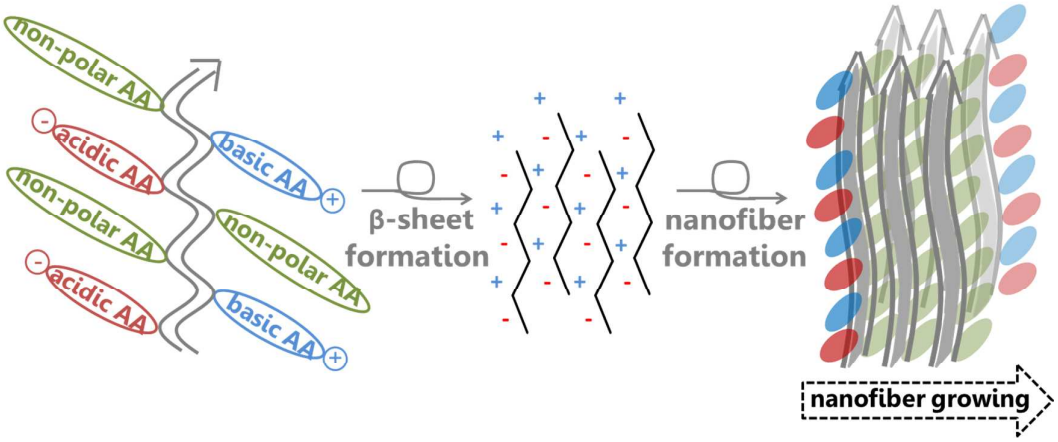


Fig. 3 Peptide model constructed with complementary charged AAs, its subsequent bi-dimensional organization, and finally a schematic representation of a tridimensional disposition through nanofiber formation.

1.3. Other "Difficult Peptides"

Moreover, some "difficult peptides" are not classified in a specific sub-type described in the literature. In this sense is included the extensively used protein fragment named Acyl Carrier Protein [ACP (65-74)][13,20,66–69] (Table 1) and the fragment of human immunodeficiency virus protease [HIV-1 PR (81-99)][20,66–68] (Table 1), two models used to evaluate new synthetic methodologies. Researchers specialized in conotoxin peptide synthesis have used the Ala10→Leu mutant of PnIA [PnIA(A10L)] (Table 1) as a "difficult peptide" to validate synthetic methodologies.[68] This sequence has also been used for the same purpose by other authors.[20] Another example is the commercially available Thymalfasin (also known as Thymosin α1)[70] (Table 1) described for hepatitis B and C.[71] It is a 28-mer peptide produced by the thymus gland that participates in the T cell maturation and has a wide range of reported medical applications,[72] which have promoted the development of strategies to obtain it.[73]

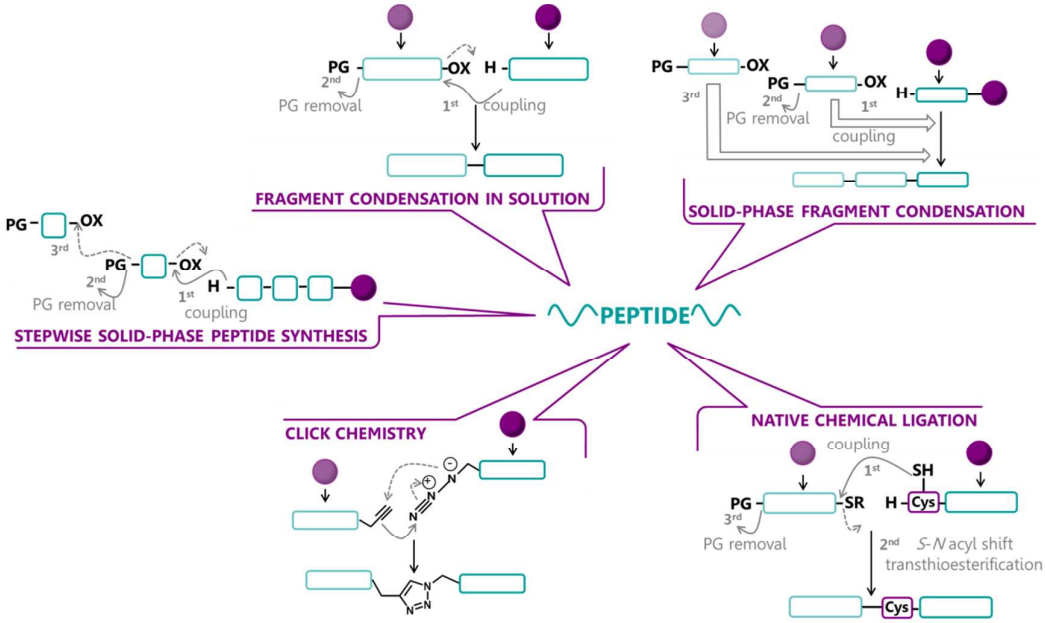


Fig. 4 The most commonly used solid-phase strategies to synthesize large peptides.

1.4. Scope of This Review

The use of peptides as therapeutic drugs has increased in recent years because of their biocompatibility and their application for the treatment of some diseases. Efforts channeled into the development of new synthetic tools to achieve peptide sequences based on natural products and into *de novo* synthesis have made it feasible to produce a large number of peptides. Several strategies based on solid-phase peptide synthesis have been proposed for large peptide sequences that cannot be achieved by the classical stepwise solid-phase synthesis (see Fig. 4). Initially, the fragment condensations, such as the solid-phase- or solution-phase based protocols (see Fig. 4), were described to favor the construction of complex peptides. Both methods consist of the rational and retro-synthetic detachment of segments in the native sequence, which in a synthetic flow would be connected in the appropriate manner on solid- or in solution-phase. Later, other more sophisticated methods, for example native chemical ligation or the

application of click chemistry (see Fig. 4), have facilitated the synthesis of certain peptides in which the coupling of two fragments occurs by a non-conventional amide bond formation. However, some sequences, due to their AA composition or their intrinsic structural arrangement cause not only difficulties regarding synthesis but also troubles related to their purification and characterization. In this regard, these obstacles are encountered in peptide sequences with self-assembling capacity. This assembly can occur in two stages, during the synthesis of the peptide or after its elongation when the sequences are fully unprotected. The sequences that show assembly during the synthesis are classified as "difficult peptides" while when these interactions are observed after the synthesis they are referred to as simply self-assembling peptides. On the basis of the origin of self-assembly, the literature describes a range of tools to minimize these interactions by favoring peptide handling. Regarding this concept, in the present review we have focused on existing methods to facilitate the synthesis and manipulation of these kinds of peptide. The reported methods described herein are all based on the enhancement of peptide solubility. The strategies have been classified in such a way as to facilitate the choice of the appropriate method for a given peptide sequence. Moreover, the most recent developments in this field have been covered in this review in order to expand knowledge regarding such self-assembling peptides.

In addition, published strategies addressed to facilitate the manipulation of peptides with self-assembling behavior in solution have also been evaluated in the present review. Peptides with this capacity allow the obtaining in aqueous solution initially of aggregates, later fibers, and subsequently hydrogel formations. Hydrogels are supramolecular organizations that show sufficient stability for use in biology, for example as regenerative tissues, drug delivery systems, and nanotubular materials,

among others. The wide range of applications in medicinal engineering of amphiphilic peptides not necessarily extracted from natural structures has attracted interest in designing new peptide sequences mainly addressed to force self-assembly.

Interestingly, here we present various strategies that allow both the synthesis of "difficult peptides" and the improvement of solubility, thus facilitating the purification or characterization of these molecules. These temporary chemical modifications in the peptide sequences allow the native sequence to be obtained after a specific final treatment that removes the modification introduced.

2. Strategies to Solubilize Peptides

The main obstacle common to all "difficult peptides", and also to those sequences that aggregate in solution, is the insolubility. The non-soluble nature of these molecules precludes their synthesis and/or hinders their characterization, and even prevents their purification. The peptide sequence and, most importantly, the AA composition, play a key role in terms of secondary structure, thereby directly affecting solubility. Both SPPS and synthesis of peptides in solution require protected sequences to allow the two functional groups to react properly to produce the desired amide bond. The nature of these protecting groups contributes to enhancing the hydrophobicity of the sequences and may induce interactions, which favor the insolubility of the peptides. In general, peptides that adopt α-helix secondary structures in solution are generally soluble in water, whereas β-sheet conformations are insoluble because of their capacity to interact by hydrophobic interactions, which evolve to aggregates.

In order to overcome this drawback, efforts have been made to develop strategies to enhance the solubility of peptides by disrupting β-sheet interactions.[74,75] Thus, the low

solubility of some peptides is considered not only a barrier but also a challenge in terms of the development of new solubilizing strategies. These methods can be classified into two main groups, namely those that modulate an *external factor* and those that introduce a *chemical modification* into the peptide sequence (see Fig. 5).

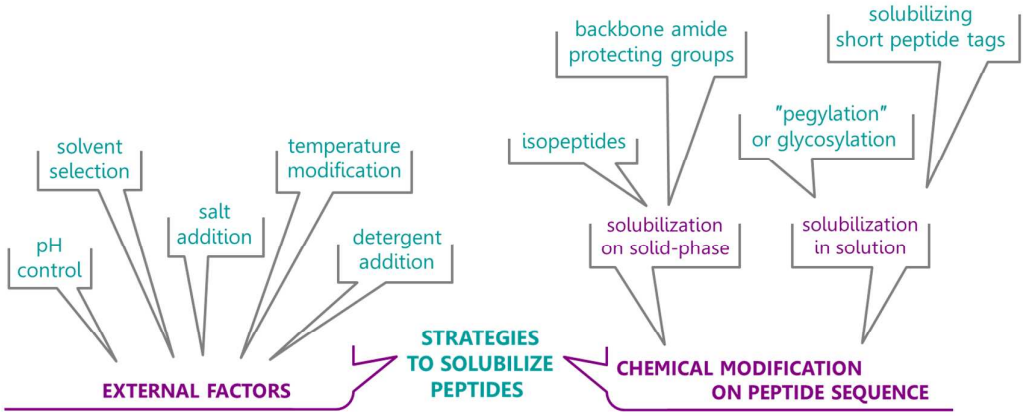


Fig. 5 Scheme of the principal solubilizing peptide strategies described in the literature.

2.1. External Factors

One of the first aspects to be considered when addressing peptide insolubility is the modification of external parameters. In some cases, such factors are introduced on solid-phase to favor peptide elongation by improving AA coupling or the removal of protecting groups. These parameters can be introduced not only during SPPS, but also to solubilize sequences in solution. Specifically, the main external parameters evaluated in the literature are the solvent, temperature, pH, the detergents and salts added to the peptide solution.

2.1.1. Solvent Selection

Although *N,N*-dimethylformamide (DMF), a dipolar aprotic solvent, is commonly used in solid- and solution-phase peptide synthesis, several studies have explored the capacity of other solvents to disrupt the secondary interactions that make peptides insoluble. In SPPS, the resin plays a key role in preventing intermolecular peptide associations. Consequently, an appropriate solvent to swell the resin may circumvent these limitations and the low loading may minimize the intermolecular connection.[76] In some cases, dimethyl sulfoxide (DMSO) achieves the solubilization of hydrophobic sequences because, as a polar aprotic hydrogen bond-accepting solvent, it increases peptide mobility and consequently favors its solvation.[77] A number of studies have demonstrated the use of DMSO as a solvent for SPPS to perform AA couplings and Fmoc removal treatments, thus enhancing the solubility and purity of peptides.[7] However, the oxidative properties of DMSO poses limitations for its use in SPPS and in solution when sequences contain AAs susceptible to oxidization, such as Met and Cys.

Another group of solvents that have been studied extensively and used to solubilize peptides in solution are fluorinated alcohols, of which the most widely used are 2,2,2-trifluoroethanol (TFE), 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP), and 1-phenyl-2,2,2-trifluoroethanol (PhTFE)[78] (Fig. 6), among others. In contrast to non-fluorinated alcohols, the main characteristic of these fluorine-containing ones is their strong hydrogen-bond donor capacity, a property that makes them suitable for interfering with the secondary structure of peptides.[78,79] Mixtures of H_2O/TFE in peptide copolymers with alternating neutral and charged AAs favor the α-helix as the proportion of TFE increases,[80] thereby supporting the general idea that TFE stabilizes the α-helix and only in some cases stabilizes the β-helix.[81] Furthermore, for certain

peptide sequences there is a reversible α-β transition, which modulates the conformation of the peptide and consequently its solubility.[82,83] Thanks to the compatibility of fluorinated alcohols with solid-support protocols, HFIP has also been used to enhance solubility during the SPPS of "difficult peptides", in those cases in which DMF is not as effective as a solvent. HFIP is a suitable alternative to perform acylations, Fmoc removal,[84] and even for peptide cyclization[85]. Other research groups have also explored the use of diverse solvents during SPPS to improve certain couplings.[86,87] These solvents include tetrahydrofuran (THF), acetonitrile, and *N*-methyl-2-pirrolidone (NMP), which affect the swelling of the resin.[76,88]

2.1.2. Temperature Modification

Peptide elongation on solid-phase, commonly performed at room temperature, can also be carried out at a higher one. In particular, synthesis performed at high temperatures is known as SPPS-elevated temperature (SPPS-ET). The most crucial aspects in SPPS-ET include side-reactions, such as AA racemization, and the alteration of solvent-swelling on resin, among others. In the '90s, studies based on SPPS-ET developed protocols at 60 ºC with several coupling reagents, solvents, and resins, which showed efficiency in peptide yields and purities.[89] Furthermore, other reported peptide synthesis that followed this methodology demonstrated minimized racemization at temperature ranges of 55–75 ºC[90] and 30–80 ºC[91] for short peptides even using triethylamine. Moreover, similar protocols at 50 ºC can be used for large sequences for certain AA couplings, when their introductions are hindered under common conditions,[76] or to facilitate "difficult peptide" synthesis.[92] In contrast, in solution peptide manipulation, the increase in temperature disrupts the secondary structure of these molecules, which may enhance solubility. The effect on unfolding the peptide by

temperature depends on the secondary structure transition of each type of self-assembling sequence. Thus, the modification of the characteristic β-sheet structure to a random coil by temperature has been proposed as a means to mediate the transition between insoluble to soluble peptide.[93]

Another way to enhance the temperature in order to solubilize peptides is via microwave (MW)-assisted synthesis. Studies on combinatorial chemistry under MW conditions have reported accelerated reactions for a large variety of reactions.[94] Several reports have validated this strategy to enhance solubility and to allow the synthesis of a wide range of peptides.[95–97]

2.1.3. pH Control

When AAs in the peptide sequence are ionizable—positively or negatively—another external parameter to take into account is the pH, which allows modulation of the neat charge of the sequence in solution, thereby modifying its interactions with the solvent and also its solubility.[98,99] The pH value at which a sequence exhibits a conformational α-β transition depends on the pKa of the AAs or on their isoelectric point. It is believed that the conformation of an ionizable peptide sequence depends more on external factors than on its intrinsic tendency to form a specific secondary structure. Several authors have taken advantage of ionizable peptides to control peptide interactions. In this regard, they have achieved interesting materials by modifying the pH.[100–102] On SPPS, interfering in the pH of the system to enhance solubility is not common practice because this parameter directly affects the integrity of the protecting groups used in such synthesis and could lead to their removal or even cleavage of the peptide.

2.1.4. Salt Addition

Other strategies to enhance the solubility of peptide sequences consist of adding chaotropic salts or denaturant agents (guanidinium chloride, urea or detergents) to the aqueous peptide solution. Regarding chaotropic salts, as their definition indicates, they are added to destabilize hydrophobic interactions and consequently the aggregates. Lithium salts, such as LiCl, LiBr or $LiClO_4$[103,104], are the compounds most used for this purpose, together with other salts such as KSCN and $NaClO_4$ (Fig. 6). These salts act by disrupting intermolecular hydrogen-bonding, first by increasing the polarity of the solvent (non-polar, aprotic or non-nucleophilic) and then by displacing it for the most favorable ion-dipole interaction between the salt and the peptide. The arrangement of peptide sequences in solution by ionic groups have been described[105,106] and have stimulated research about the addition of these salts to solubilize peptides, including NMR and circular dichroism as techniques to evaluate the salt role on the structure. Those studies demonstrated that chaotropic salts can be used during the manipulation of a peptide in solution and, based on the same concept, during solid-phase synthesis to improve peptide synthesis, mainly when THF is used as solvent.[107]

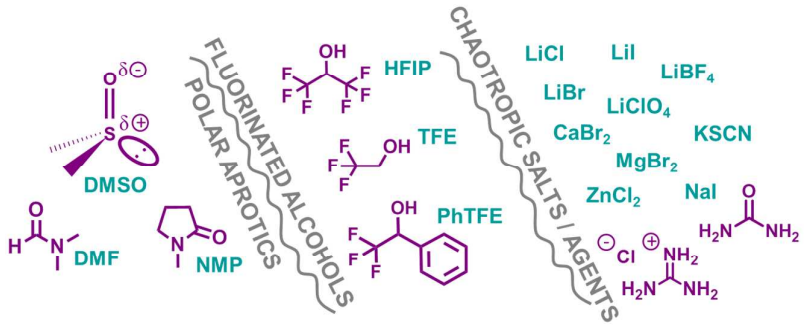


Fig. 6 Common solvents and chaotropic agents used to enhance the solubility of peptides in solution and on solid-phase.

The addition of certain concentrations of guanidinium hydrochloride (Gnd·HCl or Gdm·HCl) or urea to a peptide has resulted in increased solubility of the sequence, especially those containing non-polar AAs. The mechanism of action of these two agents suggests the formation of a complex between the peptide backbone and the denaturant agent through hydrogen-bonds (2 hydrogen donors per agent; and 2 acceptors from carbonyl and amine).[108] The stability of this interaction is stronger than the β-sheet association between peptide chains which form aggregation and this is translated into effective denaturant agents. In peptide fragment ligations, guanidinium hydrochloride is used mainly to dissolve peptides in solution or to prevent long peptide folding, by precluding the reaction.[109,110] NMR studies have confirmed these denaturant phenomena on the basis of the changes in chemical shift observed for structured conformations.[111–113] Furthermore, given the effect of guanidinium hydrochloride on solubility, some studies have proposed the introduction of L-Arg in the aqueous solution of proteins to improve their solubility.[114,115] The extension of the arginine application to some non-soluble peptides has been used to improve peptide detection by HPLC. However, this is not an appropriate solubilizing method for peptide synthesis on solid-phase, in solution-phase, or for purification purposes as arginine addition may compete with AA coupling and it implies the presence of an extra impurity, which may hinder purification.

2.1.5. Detergent Addition

Detergents are considered another kind of denaturing agent and are also a useful additive for hydrophobic sequences. Although these agents are extensively used for proteins, the addition of some kinds has also been reported for peptides. The most commonly used detergents are the anionic and non-ionic sub-types, specifically sodium

dodecyl sulphate (SDS) as an example of the former, and Triton X100 or Tween 20 as example of the latter (Table 3). They are used to solubilize peptides both in solution and during solid-phase synthesis, thus favoring yields and even peptide folding mechanisms.[116] Detergents favor micelle formation, thus enhancing the aqueous solubility of peptides; these structural changes have been analyzed to further understand the solubilizing power of detergents.[117,118] Non-ionic detergents also preserve certain structured sequences, protecting the peptides in well-organized micelles.[119,120] Although detergents have been commonly used for protein solubilization for many years,[121,122] some lines of research indicate that they should be avoided because they may interfere with HPLC chromatography and decrease the signal intensity in mass spectrometry analysis.[123] Recent publications have proposed protocols to improve HPLC compatibility with detergents.[124–126] The most effective way to solubilize hydrophobic peptides or facilitate peptide reactions in solution is probably by combining appropriate solvents with certain detergents. In this regard, the best known blend is the "magic mixture", which comprises dichloromethane (DCM)/DMF/NMP (1:1:1) containing 1% of Triton X-100 and 2 N of ethylene carbonate.[127–131]

Table 3 Some detergents used in peptide synthesis to increase sequence solubility

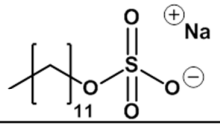
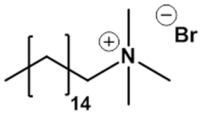
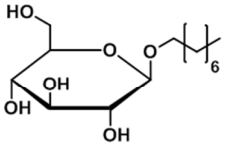
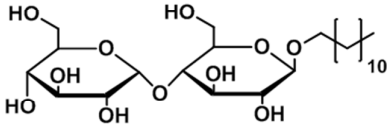
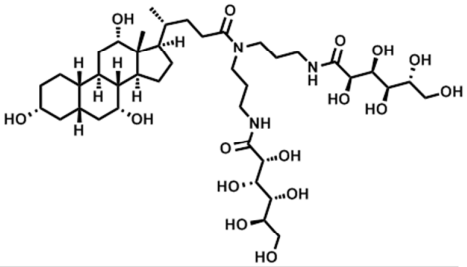
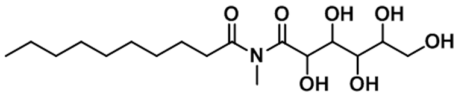
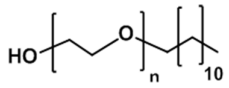
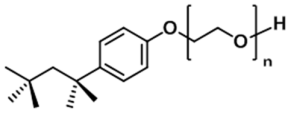
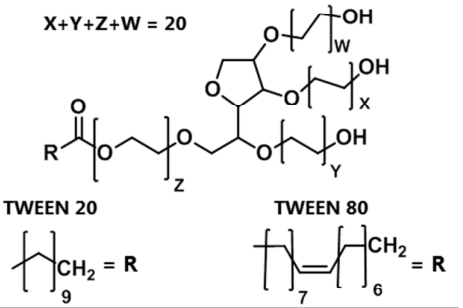
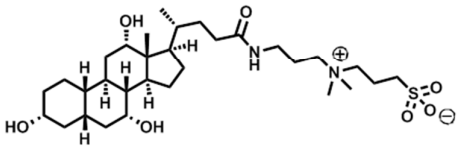
Detergent Name	Detergent Type	Detergent Structure
Sodium dodecyl sulfate (SDS)	anionic	
Hexadecyltrimethyl ammonium bromide (CTAB)	cationic	
Octyl glucopyranoside	non-ionic	
Dodecyl maltopyranoside	non-ionic	
Big CHAP	non-ionic	
N-Decanoyl-*N*-methylglucamine (MEGA-10)	non-ionic	

Table 3 (Continued)

Detergent Name	Detergent Type	Detergent Structure
Genapol X-080	non-ionic	
Triton X-100	non-ionic	
TWEEN 20 TWEEN 80	zwitter-ionic	
CHAPS	zwitter-ionic	

2.2. Chemical Modifications on the Peptide Sequence

In order to circumvent the limitations associated with some of the external factors, other kinds of strategies have been developed in recent years. Specifically, more sophisticated tools to improve the solubility of peptides, focused on chemical modifications in the peptide sequences, have emerged. These structural changes modulate the physico-chemical properties of peptides. Chemical modifications introduced into the peptide sequence can be *permanent* as long as the newly introduced moieties do not affect the objective of study of the desired peptide. On the other hand, the chemical changes introduced can be *temporarily* coupled to a certain functional

group. In the latter case, peptides with enhanced solubility can be momentarily manipulated, and later the solubilizing tool can be removed to afford the native peptide.

These strategies based on synthetic peptide modifications comprise two differentiated groups on the basis of the stage at which insolubility originates: during the synthesis on solid-phase, or after the peptide cleavage in solution. All these approaches are detailed below. Specifically, those tools designed to overcome the insolubility issue during SPPS synthesis are addressed to synthesize "difficult peptides". On the other hand, those methods developed to dissolve sequences in solution are focused on facilitating peptide characterization and purification.

2.2.1. Solubilization on Solid-Phase

2.2.1.1. Backbone Amide Protecting Groups

The development of new protecting groups for peptide chemistry has focused mainly on α-amino, α-carboxylic acid, and the side-chain of functionalized AAs. However, orthogonal protection of backbone amides may be necessary in specific cases.[2] Regardless of whether the amides are functional groups commonly non-protected in peptide synthesis, a small number of relevant side-reactions involve amides. Under certain conditions, both the formation of diketopiperazine (DKP)[132,133] (Fig. 7b) and aspartimide[134–136] (Fig. 7a) are initiated by the nitrogen from the amide bond.

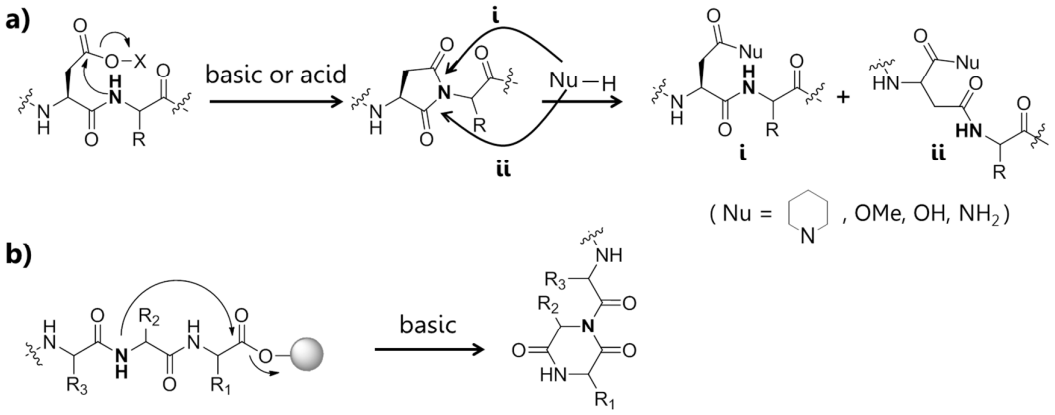


Fig. 7 Mechanisms involved in the formation of two side-products: (a) aspartimide; and (b) DKP.

The introduction of a substituent in the amide position significantly reduces the nucleophilicity of the nitrogen and subsequently precludes these internal cyclizations. Moreover, in addition to contributing to the minimization of these two by-products, this substitution on the amides is mandatory when the aggregation in SPPS needs to be avoided.

The backbone amide functions in peptides can be protected permanently or temporarily. In the first type, the group remains on the sequence as a non-cleavable substituent (such as methyl groups), while in the second case the group is removable under certain conditions. Given that these kinds of protecting group facilitate the synthesis of "difficult peptides", numerous efforts have been devoted to their design.

2.2.1.1.1. Pseudoprolines

Pseudoprolines (**ψpro**), developed by Mutter in 1992,[137,138] were the first backbone amide protecting groups described and have been highly relevant for the synthesis of "difficult peptides".[139–141] These building blocks are based on the proline structure, as the name suggests. Previous reports, also by Mutter, demonstrated the improvement of peptide synthesis when some AAs are substituted by prolines.[142–144] These benefits result from the structural peculiarity of prolines, which do not have hydrogen on their α-amino group, thus disrupting the continuity of hydrogen bonding on the backbone, a process responsible for the formation of insoluble aggregates. Furthermore, the *cis*-amide conformation is induced by the proline-like moiety, which destabilizes the β-sheet folding of peptides.[144] Following these concepts and in order to mimic the proline structure, Mutter built these dipeptide derivatives based on oxa- and thiazolidine moieties. The dipeptides are composed by any AA at the *N*-terminal position which form a cycle through the hydroxyl or thiol groups from the side-chains of Ser, Thr, or Cys, which are placed in the *C*-terminal position (Fig. 8).

The most commonly used pseudoprolines are those that contained two methyl moieties in the R_1 position (Fig. 8), the dimethyloxazolidines and the dimethylthiazolidines $AA_x(\psi^{Me,Me}pro)$. It has been described that the 2,2-dimethylated derivatives induce mainly the *cis* amide conformation, whereas the non-substituted pseudoprolines adopt an equilibrium *cis*-*trans* similar to the proline.[141,145] However, some authors, such as Brigaud and collaborators, proposed other substituents in the R_1 position to favor the *cis* amide conformation, specifically the introduction of a trifluoromethyl group to reach the $AA_x(\psi^{CF3,H}pro)$ for synthesizing fluorinated peptides.[146,147] In addition to the capacity of these derivatives to enhance the solubility of peptides by adding pseudoprolines,

another appreciated property is their ability to induce reversible turns. This property, which is attributed to the *cis* amide conformation preferred for pseudoprolines, allows macrocyclization,[148] as described by Jolliffe and co-workers[149] for Ser/Thr-based pseudoprolines or as described by Albericio and collaborators[150] for Cys-based ones. The use of Cys($\psi^{Me,Me}$pro) has allowed chemists to open the range of AAs involved in the pseudoproline structure (AA_x), thus extending options beyond Ser/Thr/Cys. As an example, the recent synthesis of the penicillamine-derived pseudoproline [Pen($\psi^{Me,Me}$pro)] and its introduction into a sequence has permitted the synthesis of a cyclic peptide where in the last step the pseudoproline is converted to Val by desulfurization.[151]

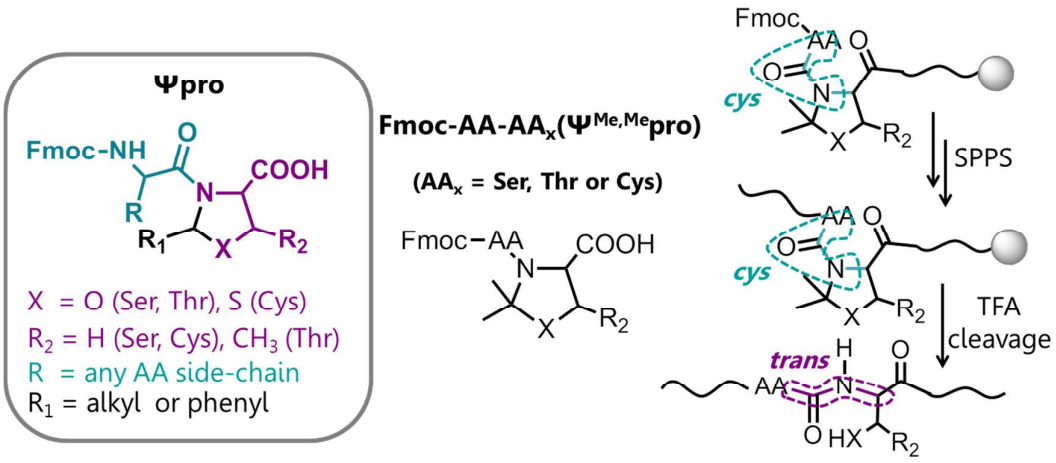


Fig. 8 Pseudoproline structure and its introduction on solid-phase peptide synthesis.

In fact, pseudoprolines composed by almost all the AAs on its *N*-terminal are commercially available. The introduction of pseudoprolines into a peptide sequence can be performed on solid-phase under standard coupling conditions. Once the peptide is cleaved from the resin by acydolysis, the pseudoproline is hydrolyzed, providing the two corresponding native AAs.

Furthermore, many publications have demonstrated that this strategy greatly contributes to the synthesis of "difficult peptides",[152] improves the cyclization of peptides,[153] facilitates the synthesis of peptides of biological interest,[154] achieves large glycopeptides,[155,156] and even facilitates protein synthesis.[154,157] Despite the successful results attributed to the use of pseudoprolines in peptide sequences, their limitation resides in the need to expand the type of AAs involved in the pseudoproline units. This drawback has led to an increased demand to explore new backbone protecting groups.

2.2.1.1.2. ortho-Hydroxybenzyl structure-based

Another research line focused on the design of backbone amide protectors was initiated by Sheppard and collaborators, also in the beginning of the '90s (1993), when they proposed the use of the *N*-(2-hydroxy-4-methoxybenzyl) (**Hmb**)[14,158] moiety for Fmoc/*tert*-butyl (*t*Bu) SPPS (see Fig. 9a and Table 4). The *N*-Fmoc-*N*(Hmb)-AA is previously prepared in solution and then coupled on solid-phase to synthesize the desired peptide (Fig. 9a). After the Hmb-derived AA has been introduced into the sequence, the Fmoc group is removed, and the incoming AA is coupled through a non-standard mechanism (Fig. 9c). Thus, during the introduction of the incoming AA, the proton from the *ortho*-hydroxyl in one of the tautomers forms a hydrogen bond with the nitrogen of the secondary amine (α-amino), thus favoring the initial acylation at the hydroxyl position. Once this ester bond is formed, an intra-molecular *O*→*N*-acyl transfer occurs, affording the desired amide bond and leaving the *ortho*-hydroxyl free. The peptide is elongated by standard SPPS conditions, and Hmb is removed (see Table 4) from the sequence, together with other acid-labile protecting groups, during the final peptide acidic cleavage from the resin to afford the desired peptide target.

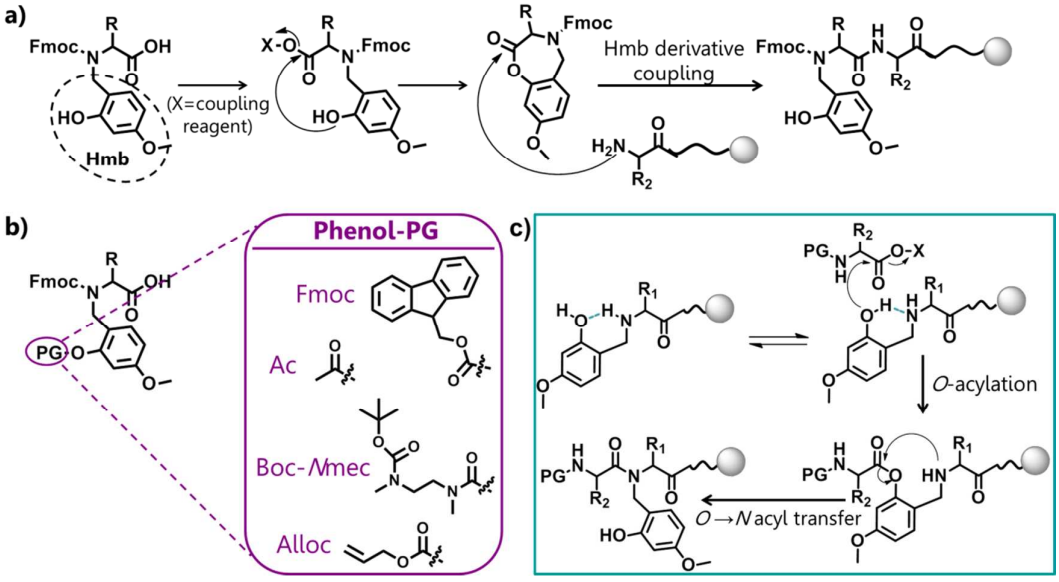


Fig. 9 Hmb backbone protecting group: (a) mechanism during the coupling on peptidyl-resin with mono-*N*-Fmoc-*N*(Hmb)-AA derivative; (b) phenol protecting groups; and (c) mechanism of incoming amino acid introduction on *N*(Hmb)-peptidyl-resin mediated by *O*→*N* acyl transfer.

Although some researchers have reported the syntheses of peptides incorporating the building block as mono-*N*-Fmoc-*N*(Hmb) derivative,[159,160] the major syntheses described, including Sheppard's proposal, have put forward the use of *O,N*-bisFmoc-*N*(Hmb) derivative (Fig. 9b).[14,161,162] The main reason for synthesizing and incorporating Hmb-derivatives bearing extra Fmoc protection on the phenolic hydroxyl group is addressed to prevent the formation of the intermediate isolated and described by Nicolás and collaborators (see Fig. 9a).[160] Thus, the mechanism that takes place during the introduction of the mono-*N*-Fmoc-*N*(Hmb)AA occurs through an internal cyclization mediated by the nucleophilic attack of the hydroxyl on the activated carboxyl group. The product afforded is the lactone 4,5-dihydro-8-methoxy-1,4-benzoxazepin-2(3H)-one, which undergoes subsequent amino attack on the carbonyl to

render the expected amide. This intra-molecular reaction is faster than the inter-molecular amino attack and has been demonstrated to be the main cause of the slow introduction of Hmb-derivatives. In particular, the R substitution in lactone intermediate (see Fig. 9a) becomes a steric hindrance when the amino group has to reach the carbonyl moiety. This steric obstacle is the main cause of the poor reactivity of lactone and consequently of the slow introduction of Hmb-derivative. This drawback can be solved with the protection of the phenolic hydroxyl group (Fig. 9b).

When Fmoc is used to protect the *ortho*-hydroxyl from Hmb, the protection is temporary. Commonly, once the Fmoc group is removed in a standard peptide synthesis, the hydroxyl group is left free, which does not affect subsequent AA introductions. However, when certain molecules susceptible to reacting with this hydroxyl group are introduced into the sequence, the semi-permanent protection of the phenol group becomes mandatory, being possible to protect it directly on resin (Fig. 9b). One strategy commonly chosen is acetylation (AcHmb), which is performed on solid-phase by acetic anhydride in the presence of DIEA. AcHmb is removed in two steps, first, by using hydrazine treatment to afford the Hmb and second by applying TFA cleavage conditions to release the unprotected peptide.[163] Some authors have adopted the option of acetylation[163–166] while others have selected alternatives to protect the phenol from Hmb, such as the use of the allyloxycarbonyl group (Alloc)[163] or the *tert*-butoxycarbonyl-*N*′-methyl-*N*-[2-(methylamino)ethyl]carbamoyl (BocNmec)[167] (Fig. 9b). In a similar process as acetylation, Alloc is introduced on solid-phase by diallyl pyrocarbonate ($Alloc_2O$) in the presence of DIEA, and BocNmecHmb is incorporated by activating the phenol with *p*-nitro-phenylchloroformate followed by treatment with mono-Boc-*N*,*N*-dimethylethylenamine/*N*,*N*′-ethyldiisopropylamine (DIEA) in DCM.

Removal of AllocHmb and NmecHmb also require two steps, namely $Pd(PPh_3)_4$ treatment (for Alloc)[168] or *N*-methylmorpholine treatment (for Nmec),[169] and the standard TFA Hmb removal mentioned before.

On the other hand, ester bond formation when the AA is being coupled onto the *N*(Hmb)-peptidyl-resin occurs kinetically slower than a standard amide bond and this is translated into a prolonged final *O*→*N*-acyl shift. In order to avoid possible side-reactions derived from this slow acylation, it was proposed that the dipeptide Fmoc-AA_2-*N*(Hmb)-AA_1-OH be pre-synthesized in solution and further introduced into the sequence. Although aspartimide prevention has been reported not only for synthesis using mono-*N*-Fmoc-*N*(Hmb)-AA,[170] but also using the dipeptide Fmoc-AA_2-*N*(Hmb)-AA_1-OH,[171] slow introductions of the bulky dipeptide leading to racemization[171] have been observed.

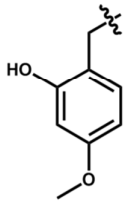
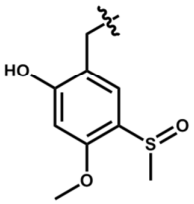
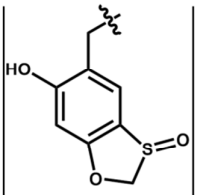
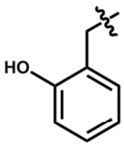
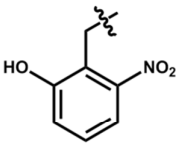
In spite of the limitations attributed to Hmb backbone amide protector, some dipeptides containing this protecting group are commercially available. Over the years, a number of peptide syntheses[92,170,172,173] have described that Hmb inhibits aspartimide formation and enhances the solubility of peptides, thus facilitating the synthesis of "difficult peptides". Specifically, the introduction of Hmb after the fifth or sixth residue of a sequence abolishes aggregations.

Parallel to the use of Hmb for the Fmoc/*t*Bu strategy, an equivalent backbone amide protecting group suitable for the Boc/benzyl (Bzl) SPPS strategy was required to cover the principal peptide strategies on solid-phase. In this regard, in 1994, two of Sheppard's collaborators for the Hmb proposal, Johnson and Quibell, defined the *N*-(2-hydroxybenzyl) group (**Hbz**)[174] (Table 4). This group shows enhanced acid stability compared to Hmb and is therefore resistant to the continuous TFA treatments

required to remove the temporary Boc groups common in Boc/Bzl SPPS strategies. The authors described the preparation of a modified AA protected by Hbz in solution and analogous to the Hmb derivatives, obtaining the *O,N*-bisFmoc-*N*(Hbz)-AA derivative. Although the synthesis was performed using the Fmoc/*t*Bu strategy, they demonstrated the stability of the protecting group to TFA and its final removal by a mixture of trifluoromethanesulfonic acid (TFMSA) and TFA (Table 4). The coupling of the incoming AA onto the *N*(Hbz)-peptidyl-resin occurs more slowly than in Hmb-containing peptidyl-resin, probably because of the absence of the electron-withdrawing effect of the methoxy in meta position with respect to the hydroxyl group.

Other approaches to prevent the poor *O*→*N*-acyl transfer associated with Hmb or Hbz have relied on the development of new backbone protecting groups that modify the benzyl substituent to afford a more efficient transacylation. In 1999, Meutermans and Smythe described a new nitro-activated group, namely the 6-nitro-2-hydroxybenzyl (**Hnb**)[175–177] (Table 4). The electron-withdrawing nitro group decreases the pK_a of the *ortho* ionizable group, thus favoring *O*-acylation and enhancing the reactivity of the *O*→*N*-acyl transfer. This group was introduced in a different manner to that used for the *ortho*-hydroxybenzyl analogs described previously. In this case, 6-nitro-2-hydroxybenzaldehyde is reacted with the α-amino from the peptidyl-resin by a reductive amination. The peculiarity of *o*-nitro derivatives is the process by which they are removed, which is mediated by photolysis at a certain wavelength which afford the corresponding nitrosobenzaldehyde.[178]

Table 4 Backbone amide protecting groups found in the literature based on the *o*-hydroxybenzyl structure

					
	Hmb	**Hmsb**	**1,3-Benzoxathiole-3-oxide derivative**	**Hbz**	**Hnb**
REMOVAL CONDITIONS	TFA/phenol/EDT/TES (92:3:3:2 v/v), 2 h, rt **or** TFA/phenol/TIS (94:3:3 v/v), 2 h, rt	TFA/TMSBr/EDT/thioanisole (89:4.5:2:4.5 v/v), 2 h, rt **or** $SiCl_4$/TFA/EDT/anisole (5.:90:2.5:2.5 v/v), 2 h, rt	NH_4I/Me_2S in TFA (20:20 eq) at 1 mg/mL, 2 h, 0 ºC	TFA/EDT/thioanisole/TFMSA (78:4:8:10 v/v), 2 h, rt	λ= 366 nm UV light (1 mM in 1% AcOH/MeOH or 1% AcOH/DMSO) 2-3 h, rt
REF.	14, 92, 158-167, 170-173	20, 180, 181	179	174	175-177

Subsequent to the Hmb group and based on the *ortho*-hydroxy moiety that enables the *O*-*N* rearrangement, in 1997, Offer, together with Quibell and Johnson, proposed the introduction of an electron-withdrawing group in *para* position to the 2-hydroxyl function, specifically the benzoaxazepsin-2(3H)-one moiety. This group, the (6-hydroxy-3-oxido-1,3-benz[d]oxathiol-5-yl)methyl (**1,3-Benzoxathiole-3-oxide derivative**),[179] (Table 4) can be used in both Boc and Fmoc strategies and the advantages associated with the modification on the Hmb protector are similar to those offered by nitro derivatives, since both groups share the same chemical electron properties. These authors proved that acylation occurs at a higher rate in less potent coupling conditions, thus considerably suppressing the epimerization side-reaction promoted by a long acylation. Final cleavage of the protecting group is completed after a reductive-acidolysis treatment. Along these lines, some years later, Quibell and Johnson presented the *N*-(3-methylsulfinyl-4-methoxy-6-hydroxybenzyl (SiMB, also named **Hmsb**)[180] (Table 4), a new generation of "safety-catch" protecting groups that

was easier to synthesize and showed equivalent acylation kinetics to the previous one. Quibell and collaborators have recently described new applications for the Hmsb group that support the efficiency of synthesizing "difficult peptides" when this protecting group is used.[181] More recently, one research group has also focused their attention on the Hmsb amide backbone protection and has studied the optimization of coupling conditions and also the application of this protection scheme to some difficult large sequences.[20] Furthermore, the Hmsb group provides an additional advantage attributed to the fact that being acid-resistant confers the possibility of obtaining the Hmsb-containing sequences after TFA cleavages. The acidic cleavage resistance of the Hmsb and the 1,3-benzoxathiole-3-oxide derivative groups places them in the "safety-catch" protecting group type. This concept was introduced in 1971 by professor Kenner and collaborators, who defined the safety-catch term as a principle applicable to the synthesis of peptides on solid-phase.[182] The same concept was also mentioned by Merrifield and Barany in a book section.[183] The safety-catch idea described by Kenner is assigned to those molecules that fit the situation where a "...stable bond is eventually labilised at the appropriate moment by a specific chemical modification". Therefore, various kinds of structures are suitable for this description, the only requirement being that the removal of this linked molecule occurs after a chemoselective transformation. Although the first safety-catch structures for SPPS were proposed as sulfonamide linkers (*Kenner's safety-catch linkers*), they were later described as amino acid protecting groups and even as backbone amide protectors.

2.2.1.1.3. ortho-Mercaptobenzyl structure-based

Regarding the same *O*→*N*-acyl transfer mechanism associated with Hmb-based protectors to favor the amide bond formation, several researchers described similar protecting groups, but to address a different proposal, specifically to assist the synthesis of large peptides/proteins through the chemical ligation of unprotected peptide fragments in solution[184] (Fig. 10). Among all the described auxiliaries for the *N*-terminus, here we have highlighted two based on a mercaptobenzyl scaffold. Although these mercaptobenzyl derivatives were not designed as a solubilizing tool, we have considered them backbone protecting groups because, after the fragment ligation, they become amide bond protectors.

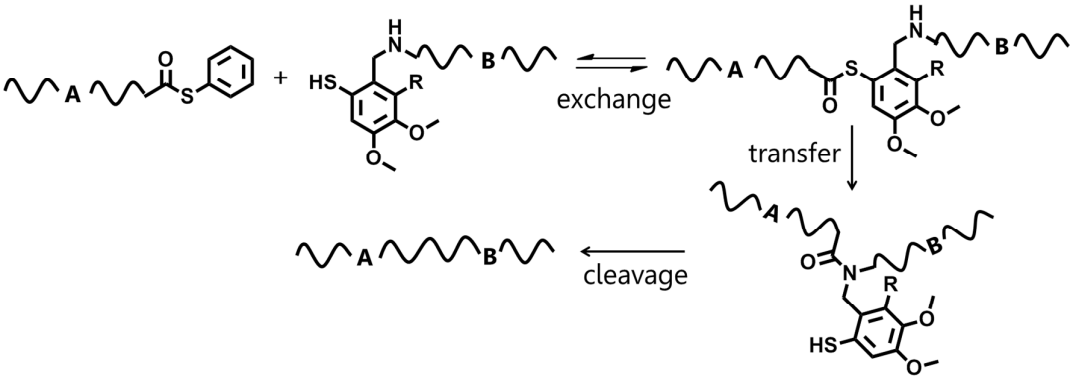


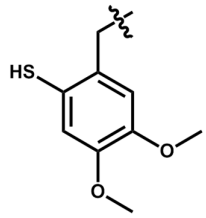
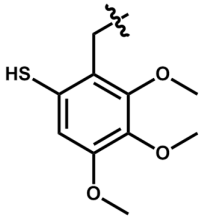
Fig. 10 Auxiliary-mediated peptide fragment ligation. Peptides A and B are unprotected sequences and R can be a hydrogen or a methoxy group.

In the earlier 2000s, Dawson[185,186] and Aimoto,[187] in parallel, developed the mercaptobenzyl moiety as an *N*-terminal protecting group. In particular, they defined two analogs, the di-methoxy derivative 4,5-dimethoxy-2-mercaptobenzyl (**Dmmb**,[186,187] also named Dmb, abbreviation that unfortunately was also selected for another protector, see 2.2.1.1.4. section) (Table 5), and the tri-substituted 4,5,6-trimethoxy-2-

mercaptobenzyl (**Tmb**,[186] abbreviation that unfortunately was also selected for another protector, see 2.2.1.1.4. section) (Table 5).

The introduction of the Dmmb group on the peptidyl-resin can be performed by two methods based on a reductive amination mechanism. In one case, the reaction occurs between the amino group contained in the mercaptobenzyl moiety[187] and the aldehyde present on the *N*-terminal peptidyl-resin. In another case, the introduction involves an aldehyde group in the mercaptobenzyl moiety[185] and an amino group present on the *N*-terminal peptidyl-resin. On the other hand, the Tmb group is introduced into the sequence by reacting the amino group in the mercaptobenzyl moiety through a nucleophilic attack on an acid bromide *N*-terminal peptidyl-resin.[186]

Table 5 Backbone amide protecting groups found in the literature based on *o*-mercaptobenzyl structure

		
	Dmmb	**Tmb**
REMOVAL CONDITIONS	1 M TFMSA/1 M thioanisole in TFA 1 h, 0 °C	TFA/TIS (95:5 v/v), 2 h, rt
REF.	186, 187	186

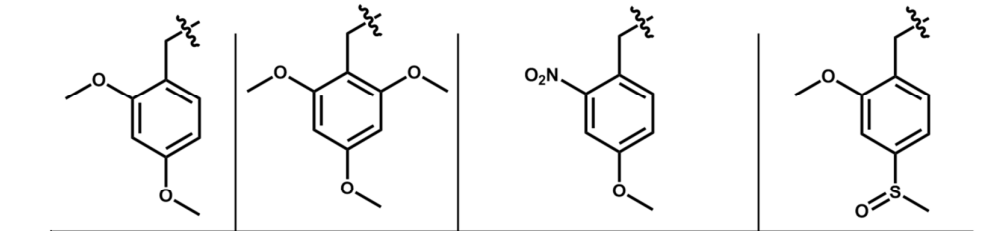
The incoming AA is coupled onto the *N*(Hmb)-peptidyl-resin through an *S*→*N*-acyl transfer, which occurs in an equivalent manner as the *O*→*N*-acyl shift in *o*-hydroxybenzyl-based protectors. Dawson and collaborators demonstrated that, for chemical ligation, the Dmmb group allows the acylation of the incoming AA faster than

when Hmb derivative is used.[186] The final Dmmb and Tmb removal (Table 5), which affords the desired peptide target, together with other acid-labile protecting groups, is carried out during the peptide acidolytic cleavage from the resin. It is worth highlighting that Dmmb removal is slower than that of Tmb, the former requiring strong acids, such as TFMSA or HF.

2.2.1.1.4. Methoxybenzyl structure-based

Several backbone amide protectors that are not based on the *ortho*-hydroxyl moiety are composed by methoxy-substituted benzyl structures addressed to enhance the TFA lability to remove them during peptide cleavage. The first of these groups proposed was the 2,4-dimethoxybenzyl (Dmob, also named **Dmb**),[188] defined initially as a backbone amide protecting group for solution-phase synthesis in 1966 (Table 6). The *ortho* and *para* electron-donating substitution of these protectors allows their rapid removal at high concentrations of TFA. The proton of the amino group forms an internal hydrogen bond with the oxygen from the *o*-methoxy function, equivalently to the *o*-hydroxy moiety in the Hmb derivative (see Fig. 9c), which favors acylation. Nevertheless, the bulkiness of the methoxy group induces severe steric hindrance that, in some cases, may preclude the introduction of the following AA. In spite of this limitation, Zahariev and White, independently, demonstrated the efficiency of Dmb in preventing aspartimide side-reactions.[189,190]

Table 6 Backbone amide protecting groups found in the literature based on methoxybenzyl structure-based



	Dmb	Tmob	4-methoxy-2-Nb	Mmsb
REMOVAL CONDITIONS	TFA/scavengers (95:5 v/v), 1 h, rt	TFA/scavengers (95:5 v/v), 1 h, rt	366 nm UV light (in MeCN/H_2O + 0.1% TFA (pH 2) + 200 eq L-Cys·HCl·H_2O) 1 h, rt	NH_4I (20 eq) in TFA 1 h, rt
REF.	188-191	158, 161, 192	193	198, 199

To minimize the steric hindrance drawback associated with this group, the same authors described the introduction of the pre-synthesized Fmoc-AA_2-*N*(Dmob)-AA_1-OH building block in the sequence to allow the synthesis of large peptides.[190,191] However, they specified that this dipeptide method is restricted to AA-Gly dipeptides because steric hindrance may preclude their preparation when using other amino acids at the *N*-terminus. These dipeptides are commercially available and, in fact, Dmb is one of the most used backbone protecting groups, together with pseudoprolines and Hmb derivatives.

The less studied protecting group with three methoxy units, the 2,4,6-trimethoxybenzyl (Tmb, also named **Tmob**) (Table 6),[192] has also been used as an amide backbone protector for peptides. The introduction of an extra methoxy group introduces more acid lability to the group compared with its di-substituted analog. Although some comparative studies revealed that Tmob can be removed with only 5% TFA in dichloromethane,[161] standard Tmob removal is performed under the same acidic

proportion as common cleavages in Fmoc strategies (Table 6). In spite of the preference for the di-substitution over the substituted trimethoxy, Tmob resulted in faster acylation than that achieved with Dmb.[158,161,192] The facilitated acylation observed for this protecting group is due to a similar hydrogen bonding effect as that described for Hmb (Fig. 9c). The greater the number of methoxy groups in *ortho* position, the more favored is the acylation, thus influencing hydrogen bonding more than the steric hindrance of the methoxy group. However, bulkiness is the main reason why sometimes neither Dmb nor Tmob is selected as backbone amide protecting groups, although some Dmb- or Tmob-protected AAs are commercially available.

A family of photolabile backbone protecting groups also based on methoxybenzyl moieties has been explored by Kent and co-workers, namely the 2-nitrobenzyl group (2-Nb) and the two methoxy versions (**4-methoxy-2-Nb** (Table 6) and 4,5-dimethoxy-2-Nb)[193], for use in Boc/Bzl SPPS chemistry. They based their studies on the photolytic nitro properties previously reported[178,194] and also on the photolytic cleavage analysis.[195,196] The introduction of methoxy groups to increase acid lability (previously demonstrated for nitrobenzyl derivative linkers)[197] led Kent to confirm that the introduction of a second methoxy in *meta* position with respect to the benzyl carbon atom does not favor photolytic cleavage. This finding is explained by the extra methoxy position, which represents an electron-withdrawing group with respect to the benzyl carbon atom, and these less electron-rich benzyl carbons lead to a slower decomposition than that of their more electron-rich analogs. Thus using the mono-methoxy-substituted nitrobenzyl 4-methoxy-2-Nb analog, they found the most promising backbone amide protecting group. On the other hand, in addition to being cleaved under mild removal conditions, nitrobenzyl derivatives have the advantage of being acid-resistant, thus

allowing the obtaining of 4,5-dimethoxy-2-Nb-containing sequences after HF or TFA cleavages. This property allows them to be considered a "safety-catch" protecting group.

Recently, our laboratory developed the 2-methoxy-4-methylsulfinylbenzyl (**Mmsb**)[198] (Table 6) as a backbone amide protecting group. Although this group was first described by Liu and collaborators[199] as an amine protector or linker, we describe its application to facilitate the synthesis of "difficult peptides". The peculiarity of this group is its design as a "safety-catch" protector, which allows the obtaining of unprotected side-chain peptides with the Mmsb remaining on the sequence after elongation. Therefore Mmsb facilitates the synthesis, characterization, and purification of aggregation-prone peptides. The sulfoxide in *para* position makes the Mmsb resistant under acidic conditions and labile after the reduction of this function to the electron-donating thioether by a reductive acidolytic treatment.[200]

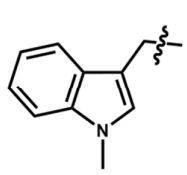
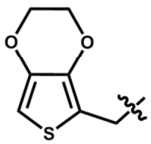
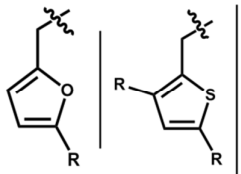
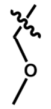
2.2.1.1.5. Other structures

Improved alternative protecting groups include the 1-methyl-3-indolylmethyl (**MIM**)[201] or the 3,4-ethylenedioxy-2-thenyl (**EDOTn**)[201] (Table 7), two electron-rich systems, both described by Albericio and collaborators and designed for Fmoc/*t*Bu strategies and maintaining the same range of acid lability as Dmb. The less sterically hindered properties of EDOTn is an advantage, allowing faster acylation than Dmb and thus overcoming the limitations associated with this protector. In 2009, Carpino proposed the dicyclopropylmethyl (**Dcpm**)[202,203] (Table 7) as a backbone protecting group labile to TFA. This group is based on the analog dimethylcyclopropyl (Dmcp)[204], previously described by the same author as an amide side-chain protecting group for Asn and Gln, as well as *C*-terminal amide protector. First, the author described the

synthesis of the building block as Fmoc-*N*(Dcpm)-AA-OH for unhindered AAs and later proposed the pre-synthesized dipeptide building block Fmoc-AA_2-*N*(Dcpm)-AA_1-OH to overcome sterical hindrance,[205] a strategy also proposed and mentioned previously for other backbone amide protectors.

Two other backbone amide protecting groups not based on benzyl structures are the substituted **2-furfuryl** and **2-thienylmethyl**,[161] both proposed by Quibell in 1999 (Table 7) and designed for Fmoc/*t*Bu SPPS strategy. These two benzyl structures show greater acid lability than Hmb, and only in the case of the 5-methoxythienylmethyl derivative is the lability equivalent to the Tmob group. The introduction of the substituted furfuryl and thienylmethyl moieties is performed by preparing the Fmoc-*N*(furfuryl derivative)AA-OH or the Fmoc-*N*(thienyl derivative)-AA-OH. Their coupling onto the peptidyl-resin is performed under common solid-phase conditions; however, the most significant limitation of these groups lies in the extremely low yields obtained when the AA is coupled onto the *N*-(furfuryl/thienyl)peptidyl-resin.

Table 7 Backbone amide protecting groups found in the literature that are not based on any common structure

	MIM	EDOTn	Dcpm	derivatives of: 2-furfuryl | 2-thienylmethyl	Etom
				 R = H, Me or OMe	
REMOVAL CONDITIONS	$TFA/CH_2Cl_2/H_2O$ (95:2.5:2.5 v/v), 1 h, rt	$TFA/CH_2Cl_2/H_2O$ (95:2.5:2.5 v/v), 1 h, rt	$TFA/H_2O/TIS$ (95:3:2 v/v), 1 h, rt	$TFA/H_2O/TIS$ (95:5 v/v) R = H (2-3 h), rt R = Me or OMe (2 min), rt	TFA/CH_2Cl_2 or H_2O (50:50 v/v), 1 h, rt
REF.	201	201	202, 203, 205	161	206

More recently, the *N*-alkoxymethyl-based amide backbone protecting group, the ethyloxymethyl (**Etom**)[206] (Table 7), was proposed by Spengler and Albericio and its efficiency on enhancing solubility was proved for the Fmoc/*t*Bu strategy by SPPS. This group is incorporated into the sequence under standard conditions as the dipeptide Alloc-*N*-(Etom)AA_2-*N*(Etom)-AA_1-OH, previously synthesized in solution-phase. The removal of Etom is carried out under 50% TFA/DCM; however, it is also conceivable under mild acidic TFA treatment (5%).

2.2.1.2. Isopeptides

Depsipeptides, also known as *O*-acyl isopeptides or merely isopeptides, are suitable for use as solubilizing peptide strategy because the temporary ester bond in the sequence disrupts the hydrogen bond continuity of backbone amides to the same extent as backbone protectors, thus preventing β-sheet interactions and consequently aggregations. This strategy is not mediated by the introduction of an extra protecting group. Indeed, the sequence is synthesized by SPPS, and the ester bond is performed through the hydroxyl group of side-chain of Ser or Thr, instead of performing the

common amide through the α-amino. The peptide is elongated and then cleaved from the resin, affording the *O*-acyl isopeptide (Fig. 11).

After the acidic cleavage of the depsipeptide, the unprotected primary amino group of the depsi unit is protonated, thus the intramolecular attack of that amino on the carbonyl of the ester is not produced and it is possible to obtain a stable depsipeptide in solution. In order to achieve the native peptide sequence, once the peptide is released from the resin and unprotected, the last step consists of the *O*→*N*-acyl shift, which is mediated by mild basic conditions in aqueous media. Thus, the *O*→*N* arrangement occurs through a five-member ring intermediate by controlling the pH (Fig. 11).

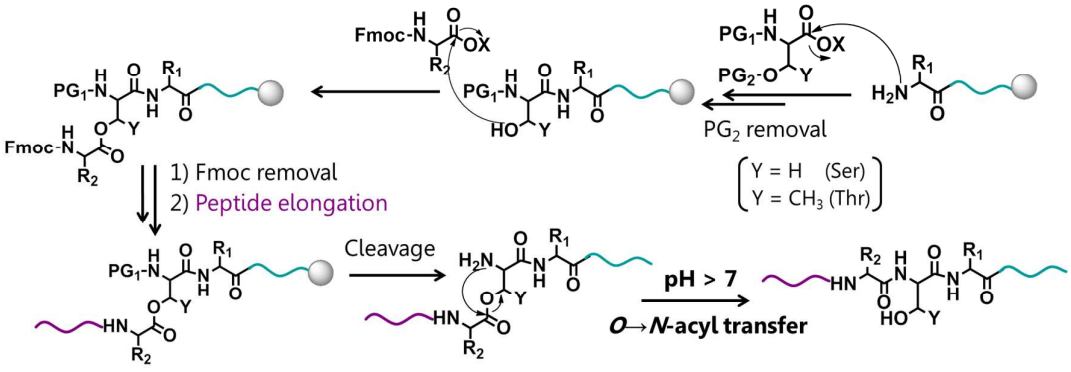


Fig. 11 Isopeptide concept: synthesis by SPPS and the *O*→*N* shift in solution to achieve the native sequence.

The first peptides synthesized by SPPS following the *O*→*N*-acyl transfer strategy were published in 1998 by Horikawa and Ohfune[207,208], based on earlier studies with peptides synthesized in solution.[209–212] However, it was in 2004 when the application of isopeptides as a β-sheet disrupting strategy was exploited, in parallel, by Carpino,[213] Kiso,[214,215] Mutter[216] and Aubagnac[217] to synthesize "difficult peptides", where the ester bond was produced through the Ser/Thr-hydroxyl (Fig. 11). Initially, the solid-phase

synthesis of these isopeptides was performed in a stepwise manner, including the esterification step; however, some side-reactions were detected after introduction of the depsi unit. These non-desired reactions are associated with racemization during ester formation or the well-known DKP formation after deprotection of the amino group from the second AA after the depsi unit. Various solutions were proposed in order to minimize or prevent these drawbacks in "difficult peptide" synthesis. Some of them were based mainly on the use of pre-synthesized isodipeptides in solution to prevent racemization;[218,219] and others on the replacement of the Fmoc of the second AA after the depsi unit by other amino protecting groups labile to milder basic conditions.[218,220] The same idea of acyl migration has been reported for *O*→*N*- and *S*→*N*-acyl isopeptides, depending on the AA involved in the migration step. Therefore, the preparation of new isodipeptides units[221,222] has opened up a diversity of building blocks, in the beginning subjected only to Ser/Thr, but now extended to any AA.[223] During recent decades, several published syntheses have demonstrated the efficiency of depsipeptide strategies.[224,225] In this regard, these strategies favor the cyclization of peptides[226] and also peptide elongations by MW-assisted SPPS.[227]

Although this strategy was merely developed to address the solubility of peptides during the solid-phase elongation, the stability of the isopeptide to cleavage conditions allows the preservation of its optimized solubility properties, thus facilitating its manipulation and purification. Other recent variations of the isopeptide method have focused on the *O*→*N*-acyl shift step, proposed by Mutter and named "switch-peptides",[216] or the "click-peptides"[228] proposed by Kiso, where the modulation of the acyl-transfer is mediated by a last selective and controlled reaction, not only by the pH. Thus, in order to favor the stability of isopeptides in solution under neutral conditions, a number of groups have

proposed distinct protecting groups for the isopeptide site (PG_1 in Fig. 11), assigned as *switch elements*, which induce the acyl-transfer via a specific trigger. Some examples of proposed switch elements are based on photocleavable protecting groups such as 6-nitroveratryloxycarbonyl (Nvoc), which is removed under photolytic conditions, thus leading to the *O*→*N*-acyl transfer.[228,229] The azide temporary protecting group, one of the most widely used switch elements, is preserved after cleavage of the peptide, and once the azide is reduced to amino, the *O*→*N* arrangement takes place, affording the native sequence.[230,231]

2.2.2. Solubilization in Solution

2.2.2.1. "Pegylation" or Glycosylation

The two most studied chemical modifications to increase the aqueous solubility of non-polar peptides are "pegylation" and glycosylation. Both strategies are addressed to enhance solubility, thus facilitating not only the manipulation of peptides, but also their reactions in solution. “Pegylation” involves the conjugation of a peptide to a polyethylene glycol (PEG) tag and glycosylation to a sugar moiety. The "pegylation" strategy was initially introduced in proteins by Davis and collaborators[232] in the '70s and subsequently adopted by other authors for smaller molecules or peptides.[233,234] The most significant property associated with the PEG conjugation to a peptide[235] is that it provides a substantial increase in solubility of hydrophobic sequences in water. This mechanism relies on the formation of more hydrogen bonds with water through ethylene oxide units[236,237] Another advantage, apart from the enhancement of solubility, lies in the synthetic facilities to link them to a peptide sequence.[238] The most frequent active site in peptides for the attachment of PEG moieties are the amino groups of lysines (α or ε) or the *N*-terminus of the peptide,[239] although new strategies have

emerged expanding the versatility of "pegylation" (Fig. 12).[240–243] The introduction of PEG tags into peptide sequences has been extensively studied and some authors have summarized the strategies most used to for this purpose in solution and solid-phase peptide synthesis.[238,244]

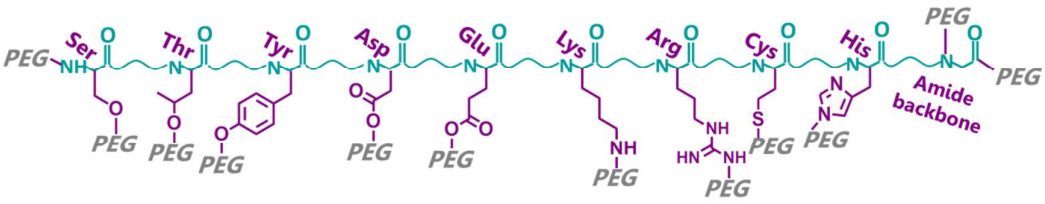


Fig. 12 Peptide positions at which it is possible to introduce a polyethylene glycol chain.

Several examples in the literature have confirmed the efficiency of conjugated PEGs to modify both the properties and architectures of peptides.[245,246] The PEG moieties most widely used have a molecular weight average of M_w<1000 Da. In addition, many kinds of PEG moieties are commercially available. In this regard, monodisperse PEG units are the polymer of choice to overcome complex characterizations or even to tackle as yet unachievable purifications, generally associated with the introduction of polydisperse PEG units.[239] Various authors have linked the PEG unit through an orthogonal cleavable linker, whereby the PEG moiety temporarily preserves the solubility of certain sequences, thereby facilitating the characterization and purification of the peptide.[138] The PEG moiety is then detached from the peptide by a final selective reaction that allows recovery of the native sequence. In addition to these appreciated properties, the "pegylation" strategy has additional advantages over other strategies, namely its lack of toxicity, which has brought about approval by the Food and Drug Administration (FDA). Furthermore, conjugation of drugs to a PEG unit enhances the half-life of these

therapeutic agents, thus expanding the number of "pegylated" peptide applications.[234,247,248] Recent studies further support the use of "pegylation" strategies as attractive options to modify non-polar peptides in order to increase their solubility.[249,250]

In a parallel manner, glycosylation, in which the peptide sequence is attached to sugar units, has also been developed as a strategy to enhance solubility of peptides and reduce aggregations.[251] In contrast to "pegylation", the introduction of sugar moieties into a sequence is a reaction that occurs in biological systems when glycoproteins are produced by glycosyltranferase enzymes.[252] Sialic acid from carbohydrate moieties is responsible for increasing the solubility of the connected hydrophobic molecule, as well as for introducing other therapeutic properties.[253]

Appropriate orthogonal hydroxyl protecting groups are required for the solution and solid-phase synthesis of glycopeptides. In spite of the difficulties encountered when attaching a sugar moiety to a sequence, several authors have further addressed this strategy in recent years (Fig. 13a).[254,255] The main linkage of sugar to a peptide sequence is formed through the side-chain of certain AAs, with the modified AA building block generally being previously synthesized and then sequentially introduced into the sequence (Fig. 13b). Depending on the functional group involved in the linkage between the peptide and the saccharide, the sequence afforded is known as *N*-, *O*-, or *C*-glycopeptide. Improved syntheses of glycopeptides both in solution[256–258] and on solid-phase[97,259–261] have been described even for long peptide sequences. In both strategies, peptides containing conjugated PEG moieties or saccharides require the condensation of large molecules, which may lead to difficult reactions. In this case, the combination of various solubilizing strategies is mandatory.

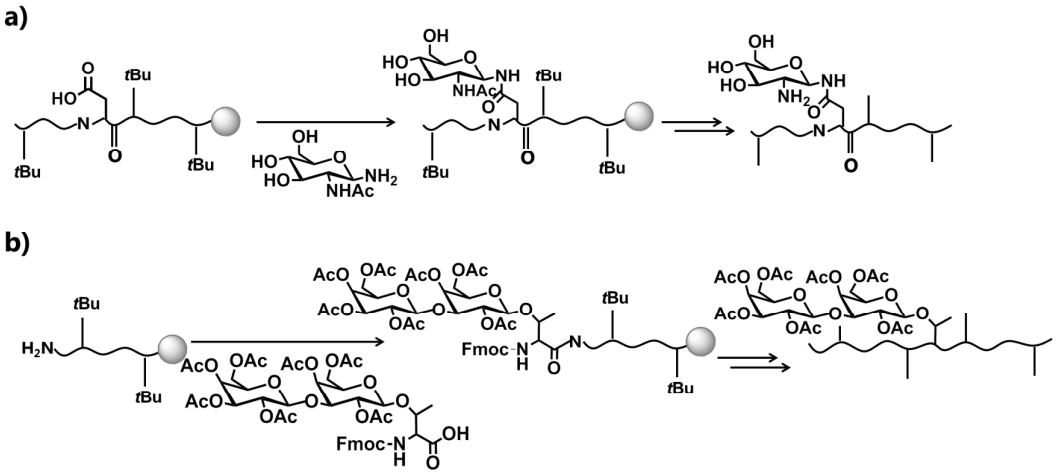


Fig. 13 Two main exemplified strategies to introduce sugar moieties on solid-phase to afford glycopeptides by: (a) stepwise introduction; and (b) pre-synthesized sugar building block introduction.

2.2.2.2. Solubilizing Short Peptide-Tags

Given the need to understand the interactions and behavior of biological systems, recent years have witnessed an increase in the demand for the synthesis of proteins with high purity. The field devoted to the design, development, and optimization of methods to achieve the synthesis of proteins is experiencing constant progress. However, the use of recombinant methods to achieve proteins has significant bottlenecks with regard to obtaining proteins of considerable purity. In order to overcome purification issues, several authors have supported the strategy based on peptide fusion, developed initially in the '70s and '80s. These strategies are probably among those most widely used nowadays to purify proteins obtained by recombinant methods. Fusion between a protein and a peptide sequence is achieved by expression systems that introduce these peptides into the *N*- or *C*-terminus of target proteins.[262] These peptides, which act as property modulators, are named *tags*. The principal

concept underlying such fusion is based on the induction of certain physico-chemical changes on the proteins, specifically addressed to the purification of proteins. After isolation of the protein, when necessary removing the previously attached peptide tag, the molecules are detached by various mechanisms, for example by the action of specific enzymes.

The most commonly used tag is the oligo-histidine-tag [usually $(His)_6$],[263,264] an AA that shows high affinity to metal ions, thus allowing protein purification by columns composed of a combination of metals that immobilize proteins at a certain pH.[265] Other peptide tags bind specifically to certain antibodies, such as FLAG-tag.[266] Furthermore, proteins have also been combined with peptides with high affinity to bind streptavidin protein (Strep-tag or SBP-tag),[267] thus purification relies on this affinity to temporarily immobilize the tagged proteins to be isolated from other non-tagged impurities. One of the shortest peptide-tag sequences used for recombinant protein purification is the oligo-arginine-tag [usually $(Arg)_5$],[268,269] thus cationic exchange columns are used to isolate proteins carrying this tag.

In parallel, although less studied than peptide tags for recombinant proteins, some short peptide tags have been conjugated to peptide sequences. In this case, the attachment of the tag to a peptide is not addressed to facilitate the peptide separation by column affinity methods but rather focused on enhancing the solubility of the non-polar peptide targets. Hydrophobic peptide sequences are examples of molecules synthesized following this strategy, especially taking into account the relevance of solubilizing peptides to perform accurate HPLC characterization or peptide purification. In this group of peptide tags, short sequences are preferred, generally comprising 5-10 residues, rather than small proteins or large peptides, which are extensively selected for

protein purification purposes. The short peptide tags used to increase solubility are formed by poly-AAs containing one type of AA (homooligo-peptides), some of which are polar and with charged side-chains. Therefore, the basic AAs disposed as poly-arginine and poly-lysine[270] and the acidic sequences poly-glutamic and poly-aspartic are the peptide tags used for this purpose. Synthetic strategies to afford the non-polar peptide connected to the poly-AA tag can be performed easily on solid-phase because the tag is also formed by AAs. Major examples described in the literature are those solubilizing tags introduced on the *C*-terminus of a target peptide, although some references also report a straightforward synthesis by linking the tag through an AA side-chain.[271]

The introduction of ionic short peptide-based solubilizing tags into a non-polar peptide through the *C*-terminus can be achieved by in two ways. The two molecules can be linked through a non-hydrolysable bond, such as an amide bond, or through a bifunctional linker that is stable under cleavage conditions but labile to other treatments. Concerning the first method, after the SPPS and the cleavage, the tag persists on the peptide target and it is not possible to detach the two molecules, thus resulting in a permanent conjugation (Fig. 14a). In this case the peptide-tag conjugate is isolated and used directly for the desired purpose; examples of nanoparticles containing peptide-based solubilizing tags have been described as drug carriers or metal chelators, among others, thanks to their biocompatibility. In contrast, in the second approach, the tag is temporarily conjugated to achieve a soluble peptide for further selective release (Fig. 14b). However, after cleavage of the peptide from the resin in both strategies, the peptide and the solubilizing tag remain linked, thus allowing the manipulation of the non-soluble peptide and facilitating characterization and also purification. The isolated

peptide-tag conjugate has temporarily modified physico-chemical properties, thus facilitating not only its manipulation in solution, but also its characterization and purification. Finally, the short peptide tag is detached from the non-polar peptide by a specific chemical treatment that removes the linker, rendering the native peptide sequence.

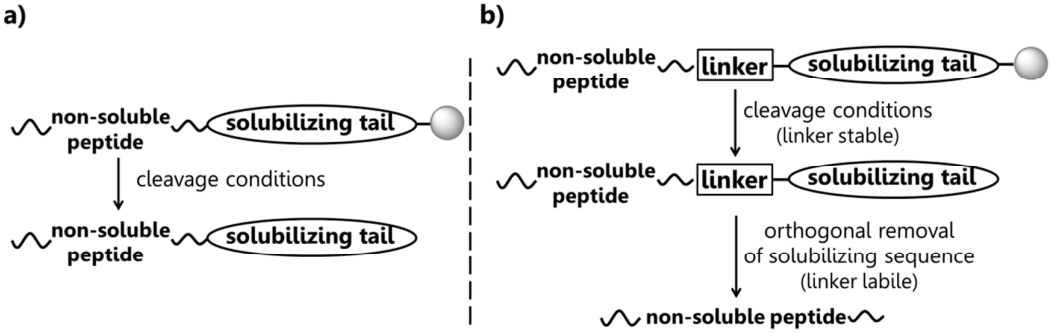
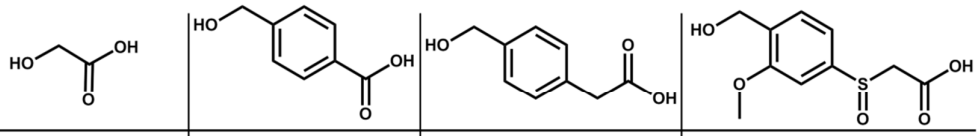


Fig. 14 Schematic concept of the "solubilizing tail" strategies by conjugation: (a) permanently; and (b) temporarily.

Although temporary conjugations of proteins to solubilizing tags have been extensively reported in the literature, only a few articles address the solubility enhancement of non-polar peptides. In 1996, Englebretsen[272] and co-workers proposed the SPPS of $(Gly\text{-}Arg)_4$ sequence as solubilizing tag connected to the hydrophobic dodecaalanine peptide through the basic-labile hydroxyacetic acid linker to be used in the Boc/Bzl strategy (Table 8). After the cleavage from the resin by HF, the peptide remains connected to the sequence, thus making the purification feasible. Finally, a last treatment with aqueous triethylamine renders the dodecaalanine sequence. Later, in 1998, the same author proposed,[273] for the Fmoc/*t*Bu methodology, the basic-labile 4-hydroxymethylbenzoic acid (HMBA) linker (Table 8) for the same purpose, this time with various series of $Gly(Lys/Arg\text{-}Gly)_n$ tags. Again, after the cleavage from the resin,

the peptide remains connected to the solubilizing tag, thus allowing its purification, and subsequently, under 0.1 M NaOH, the tag is hydrolyzed, yielding the desired peptide. Later, in 2009, two parallel studies developed by Wade[274] and Brimble,[275] based on the strategy described in 1998, used the same HMBA linker to conjugate a non-soluble peptide to the $(Arg)_6$ or $(Lys)_5$ solubilizing tags, respectively. One of the limitations associated with these linkers is that cleavages are performed under basic conditions, which commonly favors the aspartimide side-reaction. More recently, in 2014,[276] a $(Lys)_4$ tag connected to a non-polar peptide through the acid-labile 4-(hydroxymethyl)phenylacetic linker (Table 8) was described to be compatible only with the Fmoc/*t*Bu strategy, but with the advantage of being totally free of aspartimide formation. Recently, in our laboratory, a new linker for temporary conjugations through the Mmsb-OH[277] linker (Table 8) was described, which had been previously reported by Liu[199] and co-workers for a different purpose. This linker was proposed to be applied to connect peptides to solubilizing tags by taking advantage of the "safety-catch" moiety that allows the synthesis of conjugates stable to acidic cleavage but labile after a reductive-acidolytic treatment.

Table 8 Linkers described for temporary conjugation of non-polar peptides to a solubilizing tag



	hydroxyacetic acid	HMBA	4-(hydroxymethyl)-phenylacetic	Mmsb-OH
REMOVAL CONDITIONS	Et_3N aq.	0.1 M NaOH	HF	NH_4I/TFA
REF.	272	273-275	276	199, 277

3. CONCLUSIONS

The potential of "difficult peptides" as drugs and nanomaterials, or even in research on structural peptide behaviours has encouraged lines of investigation into methodologies to facilitate their synthesis. The present review summarizes the strategies described and developed in the literature to obtain "difficult peptides". The main factors responsible for the complexity of this kind of sequences rely on their insolubility during synthesis. In this regard, the most relevant synthetic advances proposed to date, and classified depending on the strategy addressed to solubilize sequences during the synthesis, are exposed herein. The first attempt to enhance the solubility of sequences is based on alterations of certain external factors that favor solubility. However, more elaborate methods are required for the synthesis of some sequences. These methods focus on introducing chemical modifications on peptide chains, thus causing structural disruptions that promote peptide solubilization during the synthesis. Although significant advances have been made regarding the synthesis of "difficult peptides", emerging methods and new protocols are required in order to open the range of applicability and/or to reduce the synthetic cost. Such breakthroughs would be translated into general methods to attain any "difficult peptide". This review provides an extensive guide to published strategies that could be contemplated for the synthesis of "difficult peptides".

This manuscript also covers other groups of peptides that present similar solubility issues but not during the synthesis but once in solution. Some strategies address the manipulation of these aggregation-prone sequences by enhancing their solubility after synthesis, when the sequences are fully unprotected. The combination of solubility

enhancement, both during and after the elongation, emerges as an interesting strategy to overcome aggregation at both stages.

ACKNOWLEDGMENTS

The work undertaken in the laboratories of the authors was funded in part by the following: CICYT (CTQ2012-30930); the *Generalitat de Catalunya* (2014 SGR 137); and the Institute for Research in Biomedicine Barcelona (IRB Barcelona). J. T.-P acknowledges a Ramon y Cajal contract (MINECO).

REFERENCES

1. R. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
2. A. Isidro-Llobet, M. Alvarez, and F. Albericio, *Chem. Rev.*, 2009, **109**, 2455–2504.
3. J. Tulla-Puche and F. Albericio (eds.), *The power of functional resins in organic synthesis*, Wiley-VCH, Weinheim, Germany, 2008.
4. M. Góngora-Benítez, J. Tulla-Puche, and F. Albericio, *ACS Comb. Sci.*, 2013, **15**, 217–228.
5. A. El-Faham and F. Albericio, *Chem. Rev.*, 2011, **111**, 6557–6602.
6. S. B. H. Kent, in *Proceedings of the 9th American Peptide Symposium*, K. D. Deber, C. M., Hruby, V. J., Kopple (eds.), Pierce Chemical Company Co., Rockford, Illinois, 1985, pp. 407–414.
7. C. Hyde, T. Johnson, and R. C. Sheppard, *J. Chem. Soc. Chem. Commun.*, 1992, 1573–1575.
8. M. Mutter, V. N. R. Pillai, H. Anzinger, E. Bayer, and C. Toniolo, in *Proceedings of the 16th European Peptide Symposium.*, K. Brunfeldt, Brunfeldt K (eds.), Scriptor, Copenhagen, 1981, pp. 660–665.
9. S. B. H. Kent, *Annu. Rev. Biochem.*, 1988, **57**, 957–989.
10. R. C. de L. Milton, S. C. F. Milton, and P. A. Adams, *J. Am. Chem. Soc.*, 1990, **112**, 6039–6046.

11. K. C. Pugh, E. J. York, and J. M. Stewart, *Int. J. Pept. Protein Res.*, 1992, 208–213.

12. E. Atherton, V. Woolley, and R. C. Sheppard, *J. Chem. Soc. Chem. Commun.*, 1980, 970–971.

13. J. Bedford, C. Hyde, T. Johnson, W. Jun, D. Owen, M. Quibell, and R. C. Sheppard, *Int. J. Pept. Protein Res.*, 1992, **40**, 300–307.

14. C. Hyde, T. Johnson, D. Owen, M. Quibell, and R. C. Sheppard, *Int. J. Pept. Protein Res.*, 1994, **43**, 431–440.

15. J. P. Schneider and J. W. Kelly, *Chem. Rev.*, 1995, **95**, 2169–2187.

16. A.-M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, and L. Serrano, *Nat. Biotechnol.*, 2004, **22**, 1302–1306.

17. D. J. Adams, D. Atkins, A. I. Cooper, S. Furzeland, A. Trewin, and I. Young, *Biomacromolecules*, 2008, **9**, 2997–3003.

18. S. Chen, V. Berthelier, W. Yang, and R. Wetzel, *J. Mol. Biol.*, 2001, **311**, 173–182.

19. S. Chen, F. A. Ferrone, and R. Wetzel, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11884–11889.

20. Y.-C. Huang, C.-J. Guan, X.-L. Tan, C.-C. Chen, Q.-X. Guo, and Y.-M. Li, *Org. Biomol. Chem.*, 2015, **13**, 1500–1506.

21. W. B. Gratzer and P. Doty, *J. Am. Chem. Soc.*, 1963, **85**, 1193–1197.

22. R. T. Ingwall, H. A. Scheraga, N. Lotan, A. Berger, and E. Katchalski, *Biopolymers*, 1968, **6**, 331–368.

23. R. B. Merrifield, J. Singer, and B. T. Chait, *Anal. Biochem.*, 1988, **174**, 399–414.

24. J. W. Kelly, *Nat. Struct. Biol.*, 2000, **7**, 824–826.

25. Z. Luo and S. Zhang, *Chem. Soc. Rev.*, 2012, **41**, 4736–4754.

26. H. Hosseinkhani, P.-D. Hong, and D.-S. Yu, *Chem. Rev.*, 2013, **113**, 4837–4861.

27. S. Zhang, *Nat. Biotechnol.*, 2003, **21**, 1171–1178.

28. K. E. Marshall, R. Marchante, W.-F. Xue, and L. C. Serpell, *Prion*, 2014, **8**, 192–196.

29. A. B. Clippingdale, J. D. Wade, and C. J. Barrow, *J. Pept. Sci.*, 2001, **7**, 227–249.

30. H. Jang, T. Arce, S. Ramachandran, B. L. Kagan, R. Lal, and R. Nussinov, *Chem. Soc. Rev.*, 2014, **43**, 6750–6764.

31. G. G. Glenner and C. W. Wong, *Biochem. Biophys. Res. Commun.*, 1984, **120**, 885–890.

32. J. Kang, H.-G. Lemaire, A. Unterbeck, J. M. Salbaum, C. L. Masters, K.-H. Grzeschik, G. Multhaup, K. Beyreuther, and B. Müller-Hill, *Nature*, 1987, **325**, 733–736.

33. D. A. Butterfield, A. M. Swomley, and R. Sultana, *Antioxid. Redox Signal.*, 2013, **19**, 823–835.

34. G. J. S. Cooper, A. C. Willis, A. Clark, R. C. Turner, R. B. Sim, and K. B. M. Reid, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 8628–8632.

35. G. J. S. Cooper, B. Leighton, G. D. Dimitriadis, M. Parry-Billings, J. M. Kowalchuk, K. Howland, J. B. Rothbard, A. C. Willis, and K. B. M. Reid, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 7763–7766.

36. K. E. Marshall, M. R. Hicks, T. L. Williams, S. V. Hoffmann, A. Rodger, T. R. Dafforn, and L. C. Serpell, *Biophys. J.*, 2010, **98**, 330–338.

37. A. Portillo, M. Hashemi, Y. Zhang, L. Breydo, V. N. Uversky, and Y. L. Lyubchenko, *Biochim. Biophys. Acta*, 2015, **1854**, 218–228.

38. C. J. Bowerman and B. L. Nilsson, *Biopolymers*, 2012, **98**, 169–184.

39. J. D. Hartgerink, E. Beniash, and S. I. Stupp, *Science*, 2001, **294**, 1684–1688.

40. S. Vauthey, S. Santoso, H. Gong, N. Watson, and S. Zhang, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5355–5360.

41. P. Kiley, X. Zhao, M. Vaughn, M. A. Baldo, B. D. Bruce, and S. Zhang, *PLoS Biol.*, 2005, **3**, 1180–1186.

42. S. Santoso, W. Hwang, H. Hartman, and S. Zhang, *Nano Lett.*, 2002, **2**, 687–691.

43. G. Von Maltzahn, S. Vauthey, S. Santoso, and S. Zhang, *Langmuir*, 2003, **19**, 4332–4337.

44. A. Aggeli, I. A. Nyrkova, M. Bell, R. Harding, L. Carrick, T. C. B. Mcleish, A. N. Semenov, and N. Boden, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 11857–11862.

45. A. Aggeli, M. Bell, N. Boden, L. M. Carrick, and A. E. Strong, *Angew. Chem. Int. Ed.*, 2003, **42**, 5603–5606.

46. A. L. Sieminski, C. E. Semino, H. Gong, and R. D. Kamm, *J. Biomed. Mater. Res. Part A*, 2008, **87**, 494–504.

47. J. Kisiday, M. Jin, B. Kurz, H. Hung, C. Semino, S. Zhang, and A. . Grodzinsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9996–10001.

48. R. Fairman and K. S. Akerfeldt, *Curr. Opin. Struct. Biol.*, 2005, **15**, 453–463.

49. A. Brack and L. E. Orgel, *Nature*, 1975, **256**, 383–387.

50. D. G. Osterman and E. T. Kaiser, *J. Cell. Biochem.*, 1985, **29**, 57–72.

51. J. H. Collier and P. B. Messersmith, *Adv. Mater.*, 2004, **16**, 907–910.

52. J. P. Jung, J. L. Jones, S. A. Cronier, and J. H. Collier, *Biomaterials*, 2008, **29**, 2143–2151.

53. J. P. Jung, A. K. Nagaraj, E. K. Fox, J. S. Rudra, J. M. Devgun, and J. H. Collier, *Biomaterials*, 2009, **30**, 2400–2410.

54. D. M. Marini, W. Hwang, D. A. Lauffenburger, S. Zhang, and R. D. Kamm, *Nano Lett.*, 2002, **2**, 295–299.

55. S. Zhang, L. Yan, M. Altman, M. La, H. Nugent, F. Frankel, D. A. Lau, G. M. Whitesides, and A. Rich, *Biomaterials*, 1999, **20**, 1213–1220.

56. S. Zhang and M. Altman, *React. Funct. Polym.*, 1999, **41**, 91–102.

57. S. Zhang, C. Lockshin, R. Cook, and A. Rich, *Biopolymers*, 1994, **34**, 663–672.

58. S. Zhang, T. Holmest, C. Lockshin, and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 3334–3338.

59. T. C. Holmes, S. de Lacalle, X. Su, G. Liu, A. Rich, and S. Zhang, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 6728–6733.

60. S. Zhang, *Biotechnol. Adv.*, 2002, **20**, 321–339.

61. H. Yokoi, T. Kinoshita, and S. Zhang, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **12**, 8414–8419.

62. P. Arosio, M. Owczarz, H. Wu, A. Butté, and M. Morbidelli, *Biophys. J.*, 2012, **102**, 1617–26.

63. A. R. Cormier, X. Pang, M. I. Zimmerman, H.-X. Zhou, and A. K. Paravastu, *ACS Nano*, 2014, **7**, 7562–7572.

64. S. Zhang, R. Ellis-Behnke, X. Zhao, and L. Spirio, *Scaffolding in tissue engineering. PuraMatrix: self-assembling peptide nanofiber scaffolds*, CRC Press, Boca Raton, FL, 2005.

65. M. Paradís-Bas, J. Tulla-Puche, A. A. Zompra, and F. Albericio, *European J. Org. Chem.*, 2013, 5871–5878.

66. E. Bayer, *Angew. Chem. Int. Ed.*, 1991, **30**, 113–216.

67. V. Krchňák, Z. Flegelová, and J. Vágner, *Int. J. Pept. Protein Res.*, 1993, **42**, 450–454.

68. L. P. Miranda and P. F. Alewood, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 1181–1186.

69. C. Arunan and V. N. R. Pillai, *Tetrahedron*, 2000, **56**, 3005–3011.

70. A. L. Goldstein, T. L. K. Low, M. McAdoo, J. McClure, G. B. Thurman, J. Rossio, C.-Y. Lai, D. Chang, S.-S. Wang, C. Harvey, A. H. Ramel, and J. Meienhofer, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 725–729.

71. A. Ciancio and M. Rizzetto, *Ann. N. Y. Acad. Sci.*, 2010, **1194**, 141–146.

72. J. Li, C. H. Liu, and F. S. Wang, *Peptides*, 2010, **31**, 2151–2158.

73. Y. García-Ramos, M. Giraud, J. Tulla-Puche, and F. Albericio, *Biopolymers (Pept. Sci.)*, 2009, **92**, 565–572.

74. I. Coin, M. Beyermann, and M. Bienert, *Nat. Protoc.*, 2007, **2**, 3247–56.

75. M. T. M. Miranda, C. W. Liria, and C. Remuzgo, *Amino acids, peptides and proteins in organic chemistry. Difficult peptides*, Wiley-VCH Verlag GmbH & Co., KGaA, Weinheim, Germany, 2011, vol. 3.

76. J. P. Tam and Y.-A. Lu, *J. Am. Chem. Soc.*, 1995, **117**, 12058–12063.

77. C. Dhalluin, C. Boutillon, A. Tartar, and G. Lippens, *J. Am. Chem. Soc.*, 1997, **119**, 10494–10500.

78. A. Berkessel, J. A. Adrio, D. Hüttenhain, and J. M. Neudörfl, *J. Am. Chem. Soc.*, 2006, **128**, 8421–8426.

79. N. H. Andersen, R. B. Dyer, R. M. Fesinmeyer, F. Gai, Z. Liu, J. W. Neidigh, and H. Tong, *J. Am. Chem. Soc.*, 1999, **121**, 9879–9880.

80. G. Ebert and Y. Kuroyanagi, *Int. J. Biol. Macromol.*, 1983, **5**, 106–110.

81. K. P. Fears, S. J. Photiadis, J. L. Kulp, and T. D. Clark, *J. Pept. Sci.*, 2014, **20**, 366–374.

82. M. Mutter and R. Hersperger, *Angew. Chem. Int. Ed.*, 1990, **29**, 185–187.

83. H. Reiersen and A. R. Rees, *Protein Eng.*, 2000, **13**, 739–743.

84. S. C. F. Milton and R. C. D. L. Milton, *Int. J. Pept. Protein Res.*, 1990, **36**, 193–196.

85. H. Hinou, K. Hyugaji, F. Garcia-Martin, S.-I. Nishimura, and F. Albericio, *RSC Adv.*, 2012, **2**, 2729–2731.

86. G. B. Fields and C. G. Fields, *J. Am. Chem. Soc.*, 1991, **113**, 4202–4207.

87. H. Kuroda, Y.-N. Chen, T. Kimura, and S. Sakakibara, *Int. J. Pept. Protein Res.*, 1992, **40**, 294–299.

88. G. A. Acosta, M. del Fresno, M. Paradís-Bas, M. Rigau-DeLlobet, S. Côté, M. Royo, and F. Albericio, *J. Pept. Sci.*, 2009, **15**, 629–633.

89. L. M. Varanda and T. M. Miranda, *J. Pept. Res.*, 1997, **50**, 102–108.

90. M. P. Souza, M. F. M. Tavares, and M. T. M. Miranda, *Tetrahedron*, 2004, **60**, 4671–4681.

91. K. Barlos, D. Papaioannou, S. Patrianakou, and T. Tsegenidis, *Liebigs Ann. Der Chemie*, 1986, 1950–1955.

92. C. Remuzgo, G. F. S. Andrade, M. L. A. Temperini, and M. T. M. Miranda, *Biopolymers (Pept. Sci.)*, 2009, **92**, 65–75.

93. Z. Luo, B. Åkerman, S. Zhang, and B. Nordén, *Soft Matter*, 2010, **6**, 2260–2270.

94. A. Lew, P. O. Krutzik, M. E. Hart, and A. R. Chamberlin, *J. Comb. Chem.*, 2002, **4**, 95–105.

95. B. Bacsa, K. Horva, S. Bo, F. Andreae, and C. O. Kappe, *J. Org. Chem.*, 2008, **73**, 7532–7542.

96. A. S. Galanis, F. Albericio, and M. Grøtli, *Org. Lett.*, 2009, **11**, 4488–4491.

97. R. Izumi, T. Matsushita, N. Fujitani, K. Naruchi, H. Shimizu, S. Tsuda, H. Hinou, and S.-I. Nishimura, *Chem. Eur. J.*, 2013, **19**, 3913–3920.

98. M. Mutter, R. Gassmann, U. Buttkus, and K.-H. Altmann, *Angew. Chem. Int. Ed.*, 1991, **30**, 1990–1992.

99. J. Tan, W. Leung, and A. G. Fincham, *Connect. Tissue Res.*, 1998, **38**, 215–221.

100. R. Li, C. Horgan, B. Long, A. Rodriguez, L. Mather, C. J. Barrow, D. R. Nisbet, and R. J. Williams, *RSC Adv.*, 2014, **5**, 301–307.

101. X. Qin, W. Xie, S. Tian, J. Cai, H. Yuan, Z. Yu, G. L. Butterfoss, A. C. Khuong, and R. A. Gross, *Chem. Commun.*, 2013, **49**, 4839–4841.

102. P. Ponnumallayan and C. J. Fee, *Langmuir*, 2014, **30**, 14250–14256.

103. A. Thaler, D. Seebach, and F. Cardinaux, *Helv. Chim. Acta*, 1991, **74**, 617–627.

104. A. Thaler, D. Seebach, and F. Cardinaux, *Helv. Chim. Acta*, 1991, **74**, 628–643.

105. V. Madison, C. M. Deber, and E. R. Blout, *J. Am. Chem. Soc.*, 1977, **99**, 4788–4798.

106. H. Kessler, W. Hehlein, and R. Schuck, *J. Am. Chem. Soc.*, 1982, **104**, 4534–4540.

107. D. Seebach, A. Thaler, and A. K. Beck, *Helv. Chim. Acta*, 1989, **72**, 857–867.

108. D. R. Robinson and W. P. Jencks, *J. Am. Chem. Soc.*, 1965, **87**, 2462–2470.

109. R. Balambika, T. Inui, H. Sargsyan, B. Arshava, L. S. Cohen, F. X. Ding, J. M. Becker, and F. Naider, *Int. J. Pept. Res. Ther.*, 2007, **13**, 251–263.

110. L. R. Malins, K. M. Cergol, and R. J. Payne, *ChemBioChem*, 2013, **14**, 559–563.

111. J. S. Smith and J. M. Scholtz, *Biochemistry*, 1996, **35**, 7292–7297.

112. K. W. Plaxco, C. J. Morton, S. B. Grimshaw, J. A. Jones, M. Pitkeathly, I. D. Campbell, and C. M. Dobson, *J. Biomol. NMR*, 1997, **10**, 221–230.

113. P. K. Nandi and D. R. Robinson, *Biochemistry*, 1984, **23**, 6661–6668.

114. K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, and T. Arakawa, *Biochem. Biophys. Res. Commun.*, 2003, **312**, 1383–1386.

115. K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J. S. Philo, and T. Arakawa, *Biotechnol. Prog.*, 2004, **20**, 1301–1308.

116. R. DeLa Cruz, F. G. Whitby, O. Buczek, and G. Bulaj, *J. Pept. Res.*, 2003, **61**, 202–212.

117. R. D. A. Alvares, D. V Tulumello, P. M. Macdonald, C. M. Deber, and R. S. Prosser, *Biochim. Biophys. Acta*, 2013, **1828**, 373–381.

118. V. Rangachari, D. K. Reed, B. D. Moore, and T. L. Rosenberry, *Biochemistry*, 2006, **45**, 8639–8648.

119. C. Katsuda, K. Niiyama, E. Obana, T. Yamamoto, Y. Katou, M. Kataoka, K. Ohkura, and Y. Shinohara, *Biochim. Biophys. Acta*, 2010, **1798**, 2090–2093.

120. A. Thapa, W. Han, R. H. Simons, A. Chilkoti, E. Y. Chi, and G. P. López, *Biopolymers*, 2013, **99**, 55–62.

121. M. le Maire, P. Champeil, and J. V. Møller, *Biochim. Biophys. Acta*, 2000, **1508**, 86–111.

122. A. C. Chadwick, D. R. Jensen, F. C. Peterson, B. F. Volkman, and D. Sahoo, *Protein Expr. Purif.*, 2015, **107**, 35–42.

123. R. R. O. Loo, N. Dales, and P. C. Andrews, *Protein Sci.*, 1994, **3**, 1975–1983.

124. H. Zischka, C. J. Gloeckner, C. Klein, S. Willmann, M. Swiatek-de Lange, and M. Ueffing, *Proteomics*, 2004, **4**, 3776–3782.

125. Y. Lin, Y. Liu, J. Li, Y. Zhao, Q. He, W. Han, P. Chen, X. Wang, and S. Liang, *Electrophoresis*, 2010, **31**, 2705–2713.

126. B. Schlager, A. Straessle, and E. Hafen, *BMC Biotechnol.*, 2012, **12**, 95–101.

127. L. Zhang, C. Goldammer, B. Henkel, F. Zuhl, G. Panhaus, G. Jung, and E. Bayer, in *In Innovations and Perspectives in Solid Phase Synthesis. 3rd International Symposium*, ed. R. Epton, Mayflower Scientific Ltd., Birmingham, 1994, pp. 711–716.

128. J. W. Stewart and W. A. Klis, in *In Innovations and Perspectives in Solid Phase Synthesis. 3rd International Symposium*, ed. R. Epton, SPCC (UK) Ltd, Birmingham, 1990, pp. 1–9.

129. Y. Sohma, T. Yoshiya, A. Taniguchi, T. Kimura, Y. Hayashi, and Y. Kiso, *Biopolymers (Pept. Sci.)*, 2007, **88**, 253–262.

130. A. Thakkar, T. B. Trinh, and D. Pei, *ACS Comb. Sci.*, 2013, **15**, 120–129.

131. J. F. McNamara, H. Lombardo, S. K. Pillai, I. Jensen, F. Albericio, and S. A. Kates, *J. Pept. Sci.*, 2000, **6**, 512–518.

132. P. M. Fischer, *J. Pept. Sci.*, 2003, **9**, 9–35.

133. N. Bayó-Puxan, J. Tulla-Puche, and F. Albericio, *European J. Org. Chem.*, 2009, 2957–2974.

134. M. Bodanszky and S. Natarajan, *J. Org. Chem.*, 1974, **40**, 2495–2499.

135. M. Mergler, F. Dick, B. Sax, P. Weiler, and T. Vorherr, *J. Pept. Sci.*, 2003, **9**, 36–46.

136. R. Subirós-Funosas, A. El-Faham, and F. Albericio, *Tetrahedron*, 2011, **67**, 8595–8606.

137. T. Haack and M. Mutter, *Tetrahedron Lett.*, 1992, **33**, 1589–1592.

138. M. Mutter, H. Oppliger, and A. Zier, *Die Makromol. Chemie, Rapid Commun.*, 1992, **13**, 151–157.

139. M. Mutter, A. Nefzi, T. Sato, X. Sun, F. Wahl, and T. Wöhr, *Pept. Res.*, 1995, **8**, 145–153.

140. T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun, and M. Mutter, *J. Am. Chem. Soc.*, 1996, **118**, 9218–9227.

141. P. Dumy, M. Keller, D. E. Ryan, B. Rohwedder, T. Wöhr, and M. Mutter, *J. Am. Chem. Soc.*, 1997, **119**, 918–925.

142. C. Toniolo, G. M. Bonora, M. Mutter, and V. N. R. Pillai, *Die Makromol. Chemie*, 1981, **182**, 1997–2005.

143. C. Toniolo, G. M. Bonora, M. Mutter, and V. N. R. Pillai, *Die Makromol. Chemie*, 1981, **182**, 2007–2014.

144. L. Moroder, C. Renner, J. J. Lopez, G. Tuchscherer, and M. Mutter, *Tailoring the cis-trans isomerization of amides. Cis-trans isomerization in biochemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2006.

145. Y. K. Kang and H. S. Park, *J. Phys. Chem.*, 2007, **111**, 12551–12562.

146. G. Chaume, O. Barbeau, P. Lesot, and T. Brigaud, *J. Org. Chem.*, 2010, **75**, 4135–4145.

147. G. Chaume, J. Simon, C. Caupène, N. Lensen, E. Miclet, and T. Brigaud, *J. Org. Chem.*, 2013, **78**, 10144–10153.

148. D. Skropeta, K. A. Jolliffe, and P. Turner, *J. Org. Chem.*, 2004, **69**, 8804–8809.

149. K. A. Fairweather, N. Sayyadi, I. J. Luck, J. K. Clegg, and K. A. Jolliffe, *Org. Lett.*, 2010, **12**, 3136–3139.

150. T. M. Postma and F. Albericio, *Org. Lett.*, 2014, **16**, 1772–1775.

151. M. S. Y. Wong, A. D. Taleski, and K. A. Jolliffe, *Aust. J. Chem.*, 2015, **68**, 627–634.

152. P. W. R. Harris, R. Kowalczyk, D. L. Hay, and M. A. Brimble, *Int. J. Pept. Res. Ther.*, 2013, **19**, 147–155.

153. N. Sayyadi, D. Taleski, S. Leesch, and K. A. Jolliffe, *Tetrahedron*, 2014, **70**, 7700–7706.

154. E. Vernieri, J. Valle, D. Andreu, and B. G de la Torre, *Amino Acids*, 2014, **46**, 395–400.

155. A. Fernández-Tejada, P. A. Vadola, and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2014, **136**, 8450–8458.

156. V. Ullmann, M. Rädisch, I. Boos, J. Freund, C. Pöhner, S. Schwarzinger, and C. Unverzagt, *Angew. Chem. Int. Ed.*, 2012, **51**, 11566–11570.

157. R. Nabika, S. Oishi, R. Misu, H. Ohno, and N. Fujii, *Bioorg. Med. Chem.*, 2014, **22**, 6156–6162.

158. T. Johnson, M. Quibell, D. Owen, and R. C. Sheppard, *J. Chem. Soc. Chem. Commun.*, 1993, 369–372.

159. W. Zeng, P. O. Regamey, K. Rose, Y. Wang, and E. Bayer, *J. Pept. Res.*, 1997, **49**, 273–279.

160. E. Nicolás, M. Pujades, J. Bacardit, E. Giralt, and F. Albericio, *Tetrahedron Lett.*, 1997, **38**, 2317–2320.

161. T. Johnson, M. Quibell, and R. C. Sheppard, *J. Pept. Sci.*, 1995, **1**, 11–25.

162. R. G. Simmonds, *Int. J. Pept. Protein Res.*, 1996, **47**, 36–41.

163. T. Johnson, L. C. Packman, C. B. Hyde, D. Owen, and M. Quibell, *J. Chem. Soc. Perkin Trans. 1*, 1996, 719–728.

164. M. Quibell, W. Turnell, and T. Johnson, *J. Org. Chem.*, 1994, **59**, 1745–1750.

165. M. Quibell, L. C. Packman, and T. Johnson, *J. Am. Chem. Soc.*, 1995, **117**, 11656–11668.

166. A. B. Clippingdale, M. Macris, J. D. Wade, and C. J. Barrow, *J. Pept. Res.*, 1999, **53**, 665–672.

167. K. Wahlström, O. Planstedt, and A. Undén, *Tetrahedron Lett.*, 2008, **49**, 3921–3924.

168. S. A. Kates, B. G. de la Torre, R. Eritja, and F. Albericio, *Tetrahedron Lett.*, 1994, **35**, 1033–1034.

169. K. Wahlström, O. Planstedt, and A. Undén, *Tetrahedron Lett.*, 2008, **49**, 3779–3781.

170. M. Quibell, D. Owen, L. C. Packman, and T. Johnson, *J. Chem. Soc. Chem. Commun.*, 1994, 2343–2344.

171. M. Mergler, F. Dick, B. Sax, C. Stähelin, and T. Vorherr, *J. Pept. Sci.*, 2003, **9**, 518–526.

172. M. Liley and T. Johnson, *Tetrahedron Lett.*, 2000, **41**, 3983–3985.

173. S. Abdel Rahman, A. El-Kafrawy, A. Hattaba, and M. F. Anwer, *Amino Acids*, 2007, **33**, 531–536.

174. T. Johnson and M. Quibell, *Tetrahedron Lett.*, 1994, **35**, 463–466.

175. W. D. F. Meutermans, S. W. Golding, G. T. Bourne, L. P. Miranda, M. J. Dooley, P. F. Alewood, and M. L. Smythe, *J. Am. Chem. Soc.*, 1999, **121**, 9790–9796.

176. L. P. Miranda, W. D. F. Meutermans, M. L. Smythe, and P. F. Alewood, *J. Org. Chem.*, 2000, **65**, 5460–5468.

177. W. D. F. Meutermans, G. T. Bourne, S. W. Golding, D. A. Horton, M. R. Campitelli, D. Craik, M. Scanlon, and M. L. Smythe, *Org. Lett.*, 2003, **5**, 2711–2714.

178. J. A. Barltrop, P. J. Plant, and P. Schofield, *Chem. Commun.*, 1966, **22**, 822–823.

179. J. Offer, T. Johnson, and M. Quibell, *Tetrahedron Lett.*, 1997, **38**, 9047–9050.

180. J. Howe, M. Quibell, and T. Johnson, *Tetrahedron Lett.*, 2000, **41**, 3997–4001.

181. A.-B. M. Abdel-Aal, G. Papageorgiou, M. Quibell, and J. Offer, *Chem. Commun.*, 2014, **50**, 8316–8319.

182. G. Kenner, J. McDermott, and R. Sheppard, *Chem. Commun.*, 1971, **426**, 636–637.

183. G. Barany and R. B. Merrifield, *The peptides*, Academic Press, New York, Vol. 2., 1979.

184. J. Offer, *Biopolymers (Pept. Sci.)*, 2010, **94**, 530–541.

185. J. Offer and P. E. Dawson, *Org. Lett.*, 2000, **2**, 23–26.

186. J. Offer, C. N. C. Boddy, and P. E. Dawson, *J. Am. Chem. Soc.*, 2002, **124**, 4642–4646.

187. T. Kawakami, K. Akaji, and S. Aimoto, *Org. Lett.*, 2001, **3**, 1403–1405.

188. F. Weygang, W. Steglich, J. Bjarnason, R. Akhtar, and N. M. Khan, *Tetrahedron Lett.*, 1966, **29**, 3483–3487.

189. S. Zahariev, C. Guarnaccia, C. I. Pongor, L. Quaroni, M. Čemažar, and S. Pongor, *Tetrahedron Lett.*, 2006, **47**, 4121–4124.

190. V. Cardona, I. Eberle, S. Barthélémy, J. Beythien, B. Doerner, P. Schneeberger, J. Keyte, and P. D. White, *Int. J. Pept. Res. Ther.*, 2008, **14**, 285–292.

191. S. Zahariev, C. Guarnaccia, F. Zanuttin, A. Pintar, G. Esposito, G. Maravić, B. Krust, A. G. Hovanessian, and S. Pongor, *J. Pept. Sci.*, 2005, **11**, 17–28.

192. N. Clausen, C. Goldammer, K. Jaunch, and E. Bayer, in *Proceedings of the 14th American Peptide Symposium*, eds. P. T. P. Kaumaya and R. S. Hodeges, Mayflower Scientific Ltd., Kingswinford, United Kingdom, 1996, pp. 71–72.

193. E. C. B. Johnson and S. B. H. Kent, *Chem. Commun.*, 2006, **14**, 1557–1559.

194. A. R. Katritzky, Y.-J. Xu, and A. V Vakulenko, *J. Org. Chem.*, 2003, **68**, 9100–9104.

195. C. G. Bochet, *Tetrahedron Lett.*, 2000, **41**, 6341–6346.

196. Y. V Ll'ichev, M. A. Schwörer, and J. Wirz, *J. Am. Chem. Soc.*, 2004, **126**, 4581–4595.

197. C. P. Holmes, *J. Org. Chem.*, 1997, **62**, 2370–2380.

198. M. Paradís-Bas, J. Tulla-Puche, and F. Albericio, *Chem. Eur. J.*, 2014, **20**, 15031–15039.

199. S. Thennarasu and C.-F. Liu, *Tetrahedron Lett.*, 2010, **51**, 3218–3220.

200. E. Nicolás, M. Vilaseca, and E. Giralt, *Tetrahedron*, 1995, **51**, 5701–5710.

201. A. Isidro-Llobet, X. Just-Baringo, M. Alvarez, and F. Albericio, *Biopolymers (Pept. Sci.)*, 2008, **90**, 444–449.

202. L. A. Carpino, K. Nasr, A. A. Abdel-Maksoud, A. El-Faham, D. Ionescu, P. Henklein, H. Wenschuh, M. Beyermann, E. Krause, and M. Bienert, *Org. Lett.*, 2009, **11**, 3718–3721.

203. P. Henklein, R. Röder, H. Weißhoff, P. Henklein, and L. A. Carpino, in *The Proceedings of the 20th American Peptide Symposium*, eds. S. Del Valle, E. Escher, and W. D. Lubell, Springer, New York, 2009, pp. 247–248.

204. L. A. Carpino, H.-G. Chao, C. Riemer, R. Warrass, D. Sadat-Aalaee, G. A. Truran, H. Imazumi, A. El-Faham, D. Ionescu, H. Wenschuh, M. Beyermann, M. Bienert, H. Schroff, F. Albericio, S. A. Triolo, N. A. Sole, and S. A. Kates, *J. Org. Chem.*, 1995, **60**, 7718–7719.

205. R. Röder, P. Henklein, H. Weisshoff, C. Mügge, M. Pätzel, U. Schubert, L. A. Carpino, and P. Henklein, *J. Pept. Sci.*, 2010, **16**, 65–70.

206. A. I. Fernández-Llamazares, J. Spengler, and F. Albericio, *Tetrahedron Lett.*, 2014, **55**, 184–188.

207. M. Horikawa, T. Nakajima, and Y. Ohfune, *Synlett*, 1998, **6**, 609–610.

208. M. Horikawa and T. Nakajima, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 2027–2032.

209. D. F. Elliott, *Biochem. J.*, 1952, **50**, 542–550.

210. S. Sakakibara, K. H. Shin, and G. P. Hess, *J. Am. Chem. Soc.*, 1962, **84**, 4921–4928.

211. D. S. Kemp, R. I. Carey, J. C. Dewan, N. G. Galakatos, and D. Kerkman, *J. Org. Chem.*, 1989, **54**, 1589–1603.

212. C.-F. Liu, C. Rao, and J. P. Tam, *Tetrahedron Lett.*, 1996, **37**, 933–936.

213. L. A. Carpino, E. Krause, C. D. Sferdean, M. Schümann, H. Fabian, M. Bienert, and M. Beyermann, *Tetrahedron Lett.*, 2004, **45**, 7519–7523.

214. Y. Sohma, M. Sasaki, Y. Hayashi, T. Kimura, and Y. Kiso, *Chem. Commun.*, 2004, 124–125.

215. Y. Sohma, M. Sasaki, Y. Hayashi, T. Kimura, and Y. Kiso, *Tetrahedron Lett.*, 2004, **45**, 5965–5968.

216. M. Mutter, A. Chandravarkar, C. Boyat, J. Lopez, S. Dos Santos, B. Mandal, R. Mimna, K. Murat, L. Patiny, L. Saucède, and G. Tuchscherer, *Angew. Chem. Int. Ed.*, 2004, **43**, 4172–4178.

217. L. Mouls, G. Subra, C. Enjalbal, J. Martinez, and J.-L. Aubagnac, *Tetrahedron Lett.*, 2004, **45**, 1173–1178.

218. I. Coin, R. Dölling, E. Krause, M. Bienert, M. Beyermann, C. D. Sferdean, and L. A. Carpino, *J. Org. Chem.*, 2006, **71**, 6171–6177.

219. T. Yoshiya, Y. Sohma, T. Kimura, Y. Hayashi, and Y. Kiso, *Tetrahedron Lett.*, 2006, **47**, 7905–7909.

220. L. A. Carpino, M. Ismail, G. A. Truran, E. M. E. Mansour, S. Iguchi, D. Ionescu, A. El-Faham, C. Riemer, and R. Warrass, *J. Org. Chem.*, 1999, **64**, 4324–4338.

221. S. Liaqat, S. Panda, A. Rauf, A. Al-Youbi, and A. Katritzky, *Synthesis*, 2013, **46**, 67–72.

222. S. Biswas, R. Kayaleh, G. G. Pillai, C. Seon, I. Roberts, V. Popov, K. A. Alamry, and A. R. Katritzky, *Chem. Eur. J.*, 2014, **20**, 8189–8198.

223. T. Yoshiya, A. Taniguchi, Y. Sohma, F. Fukao, S. Nakamura, N. Abe, N. Ito, M. Skwarczynski, T. Kimura, Y. Hayashi, and Y. Kiso, *Org. Biomol. Chem.*, 2007, **5**, 1720–1730.

224. I. Coin, *J. Pept. Sci.*, 2010, **16**, 223–230.

225. Y. Sohma and Y. Kiso, *Chem. Rec.*, 2013, **13**, 218–223.

226. J. Lécaillon, P. Gilles, G. Subra, J. Martinez, and M. Amblard, *Tetrahedron Lett.*, 2008, **49**, 4674–4676.

227. W. M. Hussein, T.-Y. Liu, I. Toth, and M. Skwarczynski, *Org. Biomol. Chem.*, 2013, **11**, 2370–2376.

228. Y. Sohma and Y. Kiso, *ChemBioChem*, 2006, **7**, 1549–1557.

229. S. Dos Santos, A. Chandravarkar, B. Mandal, R. Mimna, K. Murat, L. Saucède, P. Tella, G. Tuchscherer, and M. Mutter, *J. Am. Chem. Soc.*, 2005, **127**, 11888–11889.

230. N. Nepomniaschiy, V. Grimminger, A. Cohen, S. Digiovanni, H. A. Lashuel, and A. Brik, *Org. Lett.*, 2008, **10**, 5243–5246.

231. K. Hojo, M. Maeda, and K. Kawasaki, *J. Pept. Sci.*, 2001, **7**, 615–618.

232. A. Abuchowski, T. Van Es, N. C. Palczuk, and F. F. Davis, *J. Biol. Chem.*, 1977, **252**, 3578–3581.

233. R. V. N. Pillai and M. Mutter, *Acc. Chem. Res.*, 1981, **14**, 122–130.

234. R. B. Greenwald, *Expert Opin. Ther. Pat.*, 1997, **7**, 601–609.

235. S. Zalipsky and C. Lee, *Use of functionalized polyethylene glycols for modification of polypeptides. Poly-(ethylene glycol) chemistry: biotechnical and biomedical applications*, Plenum Press, New York, 1992.

236. K.-J. Liu and J. L. Parsons, *Macromolecules*, 1969, **2**, 529–533.

237. J. M. Harris and R. B. Chess, *Nat. Rev. Drug Discov.*, 2003, **2**, 214–221.

238. G. Pasut and F. M. Veronese, *J. Control. Release*, 2012, **161**, 461–472.

239. M. S. Kim, E. J. Park, and D. H. Na, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 38–42.

240. J.-F. Lutz and H. G. Börner, *Prog. Polym. Sci.*, 2008, **33**, 1–39.

241. M. J. Roberts, M. D. Bentley, and J. M. Harris, *Adv. Drug Deliv. Rev.*, 2012, **64**, 116–127.

242. A. I. Fernández-Llamazares, J. García, V. Soto-Cerrato, R. Pérez-Tomás, J. Spengler, and F. Albericio, *Chem. Commun.*, 2013, **49**, 6430–6432.

243. I. W. Hamley, *Biomacromolecules*, 2014, **15**, 1543–1559.

244. L. a Canalle, D. W. P. M. Löwik, and J. C. M. van Hest, *Chem. Soc. Rev.*, 2010, **39**, 329–353.

245. H. G. Börner, R. I. Sütterlin, P. Theato, and K. T. Wiss, *Macromol. Rapid Commun.*, 2014, **35**, 180–185.

246. J. Zou, X. He, J. Fan, J. E. Raymond, and K. L. Wooley, *Chem. Eur. J.*, 2014, **20**, 8842–8847.

247. S. M. Lim, H. N. Eom, H. H. Jiang, M. Sohn, and K. C. Lee, *J. Pharm. Sci.*, 2015, 1–9.

248. V. Mäde, S. Els-Heindl, and A. G. Beck-Sickinger, *Beilstein J. Org. Chem.*, 2014, **10**, 1197–1212.

249. K. Susumu, E. Oh, J. B. Delehanty, F. Pinaud, K. B. Gemmill, S. Walper, J. Breger, M. J. Schroeder, M. H. Stewart, V. Jain, C. M. Whitaker, A. L. Huston, and I. L. Medintz, *Chem. Mater.*, 2014, **26**, 5327–5344.

250. J. Loureiro, R. Crespo, H. Börner, P. M. Martins, F. A. Rocha, M. Coelho, M. C. Pereira, and S. Rocha, *J. Mater. Chem. B*, 2014, **2**, 2259–2264.

251. D. Yuan, X. Du, J. Shi, N. Zhou, A. A. Baoum, and B. Xu, *Beilstein J. Org. Chem.*, 2014, **10**, 2406–2413.

252. P. Lafite and R. Daniellou, *Nat. Prod. Rep.*, 2012, **29**, 729–738.

253. A. M. Sinclair and S. Elliott, *J. Pharm. Sci.*, 2005, **94**, 1626–1635.

254. M. A. Brimble, N. Miller, and G. M. Williams, *Amino acids, peptides and proteins in organic chemistry. Protection reactions, medicinal chemistry, combinatorial synthesis. Recent developments in neoglycopeptide synthesis*, Weinheim, Germany, Wiley-VCH, 2011, vol. 4.

255. R. Peltier, M. A. Brimble, J. M. Wojnar, D. E. Williams, C. W. Evans, and A. L. DeVries, *Chem. Sci.*, 2010, **1**, 538–551.

256. P. Wang, B. Aussedat, Y. Vohra, and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2012, **51**, 11571–11575.

257. J. Liu, C. Luo, P. A. Smith, J. K. Chin, M. G. P. Page, M. Paetzel, and F. E. Romesberg, *J. Am. Chem. Soc.*, 2011, **133**, 17869–17877.

258. D. P. Galonić, W. A. van der Donk, and D. Y. Gin, *Chem. Eur. J.*, 2003, **9**, 5997–6006.

259. Y. Asahina, M. Kanda, A. Suzuki, H. Katayama, Y. Nakahara, and H. Hojo, *Org. Biomol. Chem.*, 2013, **11**, 7199–7207.

260. S. Koch, D. Schollmeyer, H. Löwe, and H. Kunz, *Chem. Eur. J.*, 2013, **19**, 7020–7041.

261. R. Chen, M. A. Pawlicki, and T. J. Tolbert, *Carbohydr. Res.*, 2014, **383**, 69–75.

262. K. Terpe, *Appl. Microbiol. Biotechnol.*, 2003, **60**, 523–533.

263. J. E. Grundy, L. Y. Wirtanen, and M. Beauregard, *Protein Expr. Purif.*, 1998, **13**, 61–66.

264. D. Bang and S. B. H. Kent, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 5014–5019.

265. J. Porath, J. Carlsson, I. Olsson, and G. Belfrage, *Nature*, 1975, **258**, 598–599.

266. T. P. Hopp, K. S. Prickett, V. L. Price, R. T. Libby, C. J. March, D. P. Cerretti, D. L. Urdal, and P. J. Colon, *Nat. Biotechnol.*, 1988, **6**, 1204–1210.

267. T. G. M. Schmidt, J. Koepke, R. Frank, and A. Skerra, *J. Mol. Biol.*, 1996, **255**, 753–766.

268. H. M. Sassenfeld and S. J. Brewer, *Nat. Biotechnol.*, 1984, **2**, 76–81.

269. A. Kato, K. Maki, T. Ebina, K. Kuwajima, K. Soda, and Y. Kuroda, *Biopolymers*, 2006, **85**, 12–18.

270. M. Paradís-Bas, M. Albert-Soriano, J. Tulla-Puche, and F. Albericio, *Org. Biomol. Chem.*, 2014, **12**, 7194–7196.

271. Y.-C. Huang, Y.-M. Li, Y. Chen, M. Pan, Y.-T. Li, L. Yu, Q.-X. Guo, and L. Liu, *Angew. Chem. Int. Ed.*, 2013, **52**, 4858–4862.

272. D. Englebretsen and P. Alewood, *Tetrahedron Lett.*, 1996, **37**, 8431–8434.

273. C. T. Choma, G. T. Robillard, and D. R. Englebretsen, *Tetrahedron Lett.*, 1998, **39**, 2417–2420.

274. M. A. Hossain, A. Belgi, F. Lin, S. Zhang, F. Shabanpoor, L. Chan, C. Belyea, H.-T. Truong, A. R. Blair, S. Andrikopoulos, G. W. Tregear, and J. D. Wade, *Bioconjugate Chem.*, 2009, **20**, 1390–1396.

275. P. W. R. Harris and M. A. Brimble, *Synthesis*, 2009, **20**, 3460–3466.

276. S. Peigneur, M. Paolini-Bertrand, H. Gaertner, D. Biass, A. Violette, R. Stöcklin, P. Favreau, J. Tytgat, and O. Hartley, *J. Biol. Chem.*, 2014, **289**, 35341–35350.

277. M. Paradís-Bas, J. Tulla-Puche, and F. Albericio, *Org. Lett.*, 2015, **17**, 294–297.

Table of Contents Graphic