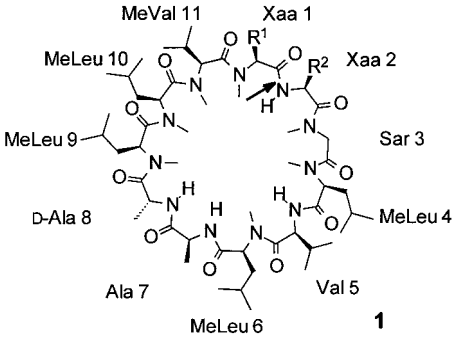
Efficient, Racemization-Free Peptide Coupling of *N*-Alkyl Amino Acids by Using Amino Acid Chlorides Generated In Situ—Total Syntheses of the Cyclopeptides Cyclosporin O and Omphalotin A**

Norbert Sewald*

Linear or cyclic peptides that contain *N*-alkyl amino acids (e.g. *N*-methyl amino acids, MeXaa), frequently occur in nature and display interesting biological activities. The cyclosporins, a family of about 25 cyclic peptides, are produced by the fungus *Beauveria nivea* (previously designated *Tolypocladium inflatum*). Cyclosporin O (**1b**, Scheme 1)[1] and the well-known relative cyclosporin A (**1a**) are sequence-homologous



1

Scheme 1. Structures of cyclosporin A (**1a**, Xaa 1: MeBmt, $R^1 = CH(OH)$-$CH(CH_3)$-CH_2-(*E*)-CH=CH-CH_3; Xaa 2: Abu, $R^2 = CH_2CH_3$) and cyclosporin O (**1b**, Xaa 1: MeLeu, $R^1 = CH_2$-$CH(CH_3)$; Xaa 2: Nva, $R^2 = CH_2CH_2CH_3$). The arrow marks the macrocylization site in the synthesis of cyclosporin O.

cyclic undecapeptides with antifungal, anti-inflammatory, and immunosuppressive activity. Omphalotin A (**2**), which is structurally related to the cyclosporins, belongs to a family of cyclic dodecapeptides formed by *Omphalotus olearius*. It outweighs known nematicides in vitro with respect to activity and selectivity. This has been interpreted as a hint towards a novel biological mode of action.[2] Cyclosporins and omphalotin A contain several *N*-methyl amino acids (cyclosporin A, O: seven MeXaa; omphalotin A: nine MeXaa). *N*-Alkylation results in a decreased number of possible hydrogen bonds and, consequently, in increased conformational flexibility.

Biosynthesis of the cyclosporins occurs non-ribosomally at the multienzyme complex cyclosporin synthetase (~1600 kDa) according to the thiotemplate mechanism.[3] More than 40 single chemical steps are catalyzed by this enzyme complex en route to the cyclosporins.

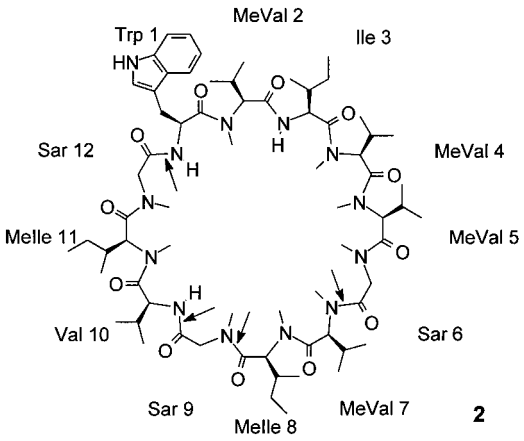
The chemical synthesis of peptides that are rich in *N*-alkyl amino acids faces severe difficulties because of the steric hindrance exerted by the secondary amino groups. Despite the fact that several hundred analogues of cyclosporin A have been obtained chemically after it was synthesized for the first time, there is no generally applicable method for peptide couplings involving *N*-alkyl amino acids. Practically quantitative coupling yields are an indispensable precondition especially in solid-phase peptide synthesis (SPPS), as the separation of the target peptide from core sequences (truncation of the peptide because of unattainable further acylation) or mismatch sequences (omission of amino acids in the sequence because of incomplete peptide coupling) after cleavage from the polymeric support often is impossible even by preparative HPLC.[3] The reactivity of resin-bound secondary amines sometimes is even lower compared to reactions in solution. Slow coupling reactions may be accompanied by undesired side reactions like racemization or diketopiperazine formation. Moreover, peptides containing *N*-alkyl amino acids are acid-labile.

Jung et al. recently reported in two publications on the total syntheses of cyclosporin O[1] (**1b**, Scheme 1) and omphalotin A[2] (**2**, Scheme 2) on solid support in which triphosgene (bis(trichloromethyl)carbonate, BTC) was used as the coupling reagent in accord with a method originally introduced by Gilon et al.[4] Carboxy group activation of N^{α}-protected amino acids by using BTC is likely to succeed through in situ generation of the corresponding acid chloride.[4] Alternatively, mixed anhydrides may be formed. Triphosgene (m.p. 81–83 °C) is a safe replacement of phosgene and diphosgene.[5] Gilon et al. used BTC in combination with collidine as the base in THF at 50 °C to obtain in situ N^{α}-Fmoc-protected amino acid chlorides both from proteinogenic and *N*-alkyl amino acids.[4] The resin-bound peptide is allowed to react with these pre-activated derivatives. These authors also proved that this method is devoid of racemization even for

[*] Prof. Dr. N. Sewald
Organische und Bioorganische Chemie
Universität Bielefeld
Postfach 100131, 33501 Bielefeld (Germany)
Fax: (+49) 521-106-8094
E-mail: norbert.sewald@uni-bielefeld.de

[**] A list of the abbreviations used in this article is located directly before the references.

  1433-7851/02/4124-4661 $ 20.00+.50/0



Scheme 2. Structure of omphalotin A (**2**). The arrows mark the different variants of the macrocyclization.

a sequence containing three consecutive *N*-methyl amino acids.

Amino acid chlorides were used already by Emil Fischer for peptide couplings. However, amino acid halides were reputed for a long time in peptide chemistry to be inappropriate "over-activated" derivatives. During the past decade *N*-urethane-protected amino acid fluorides and chlorides experienced a renaissance as stable and highly efficient building blocks for peptide synthesis due to pioneering contributions by Carpino et al.[6] While the former display a somewhat broader tolerance towards acid-labile functionalities, amino acid chlorides are applied nearly exclusively as N^{α}-Fmoc-protected derivatives.[6] N^{α}-Fmoc-protected amino acid chlorides that require side chain protection by acid-labile groups (Boc, *t*Bu, Trt) for tactical reasons, however, are not sufficiently stable and not accessible in all cases.

The in situ generation of N^{α}-Fmoc amino acid chlorides avoids this problem by the addition of collidine.[4] Jung et al. confirmed that BTC activation is significantly superior to a series of other methods (DCC, DIC/HOAt, in situ generation of amino acid fluorides with TFFH[6]).[2] The formation of diastereomeric peptides by racemization of activated amino acids was not detected. According to these investigations, however, the original protocol by Gilon et al.[4] turned out to be inappropriate for the synthesis of longer peptides. The highly acid-labile TCP resin (trityl anchor on polystyrene, cleavage with hexafluoroisopropyl alcohol) is required because of the pronounced lability of the products towards treatment with stronger acids. It can only be applied if coupling of the in situ generated N^{α}-Fmoc-protected amino acid chloride to the resin-bound peptide is performed in the presence of several equivalents of diisopropylethylamine.[2] Interestingly, elevated reaction temperatures as required according to Gilon et al.[4] are not necessary under these conditions. A similar increase in coupling efficiency by combined application of a weak base for carboxy group activation and a stronger base in the coupling reaction has been described for the system DIC/HOAt.[7]

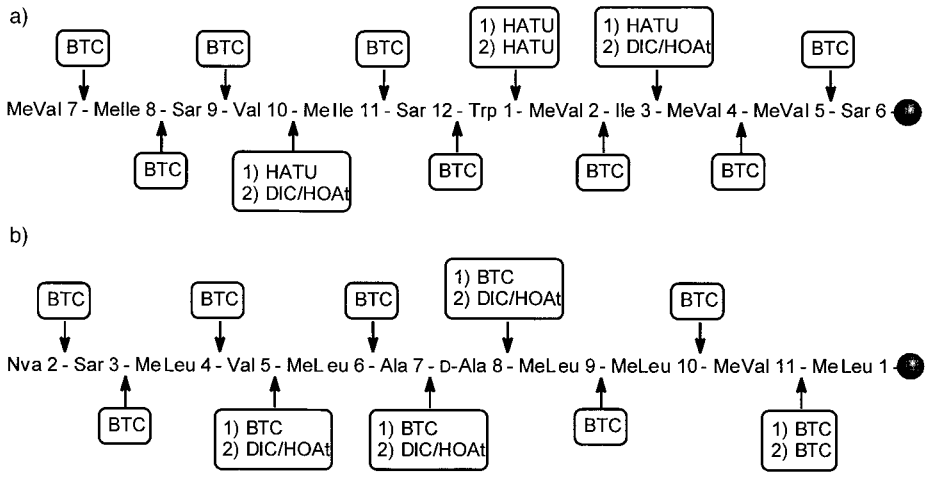
Jung et al. also found that the BTC protocol may be inferior to activation by DIC/HOAt or HATU in coupling of amino acids that are not *N*-alkylated. In such cases double couplings, first with BTC and then with DIC/HOAt or HATU were found to be the method of choice (Scheme 3). The final cyclization of the linear precursors succeeded in all cases with EDC/HOAt.

The coupling steps in the syntheses of **1b** and **2** are displayed in Scheme 3. In the synthesis of omphalotin A (OmA), assembly and cyclization of different linear peptides was examined, starting from resin-bound Sar 12 (OmA(1–12)), Sar 9 (OmA(10–9)), and Sar 6 (OmA(7–6), Scheme 3a), as well as from resin-bound MeIle 8 (OmA(9–8)). Peptide macrocycle formation by activation of a C-terminal sarcosine residue (Sar) has to be preferred, as otherwise considerable epimerization of the C-terminal residue occurs. This was, for instance, observed upon activation of MeIle 8 in the linear peptide OmA(9–8).

In all cases investigated by the authors the BTC method reliably led to quantitative formation of peptide bonds upon activation of N^{α}-Fmoc-protected *N*-methyl amino acids. Thus, Jung et al. have succeeded in establishing a reproducible and generally applicable method for incorporation of these sterically hindered derivatives by modification and optimization of the BTC protocol.[4] This paves the way for a broader application in the synthesis of naturally occurring peptides that contain *N*-methyl amino acids and are characterized by interesting biological properties. Automated parallel syntheses or large-scale syntheses of such peptides and analogues thereof can now be performed reliably.

Abbreviations

Abu: 2-aminobutyric acid; Boc: *tert*-butoxycarbonyl; BTC: bis(trichloromethyl)carbonate; DCC: *N*,*N*′-dicyclohexylcarbodiimide; DIC: *N*,*N*′-diisopropylcarbodiimide; EDC: *N*-



Scheme 3. a) Solid-phase synthesis of the linear precursor of omphalotin A (**2**) suited best for cyclization. b) Solid-phase synthesis of the linear precursor of cyclosporin O (**1b**).

  1433-7851/02/4124-4662 $ 20.00+.50/0

ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide hydrochloride; Fmoc: 9-fluorenylmethoxycarbonyl; HATU: O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (IUPAC: 1-[bis-(dimethylamino)methyliumyl]-1*H*-1,2,3-triazolo[4,5-*b*]pyridin-3-oxide hexafluorophosphate); HOAt: 1-hydroxy-7-azabenzotriazole; MeBmt: (4*R*)-4[(*E*)-2-butenyl-4,*N*-dimethyl-L-threonine; Nva: norvaline; tBu: *tert*-butyl; TFFH: tetramethylfluoroformamidinium hexafluorophosphate; Trt: triphenylmethyl (trityl).

[1] B. Thern, J. Rudolph, G. Jung, *Tetrahedron Lett.* **2002**, *43*, 5013 – 5016.
[2] B. Thern, J. Rudolph, G. Jung, *Angew. Chem.* **2002**, *114*, 2401 – 2403; *Angew. Chem. Int. Ed.* **2002**, *41*, 2307 – 2309.
[3] N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, Wiley-VCH, Weinheim, **2002**.
[4] E. Falb, Y. Yechezkel, Y. Salitra, C. Gilon, *J. Pept. Res.* **1999**, *53*, 507 – 517.
[5] H. Eckert, B. Forster, *Angew. Chem.* **1987**, *99*, 922 – 923; *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 894 – 895.
[6] L. A. Carpino, M. Beyermann, H. Wenschuh, M. Bienert, *Acc. Chem. Res.* **1996**, *29*, 268 – 274.
[7] L. A. Carpino, A. El-Faham, *Tetrahedron* **1999**, *55*, 6813 – 6830.

New Principles of Protein Structure: Nests, Eggs—and What Next?**

Debnath Pal, Jürgen Sühnel,* and Manfred S. Weiss*

Based on results from refolding experiments on urea-denatured ribonuclease conducted more than 40 years ago, the chemistry nobel laureate Anfinsen formulated the still to a large extent valid paradigm of protein folding: "… *it may be concluded that the information … for the assumption of the native secondary and tertiary structures* (of proteins) *is contained in the amino acid sequence itself.*"[1] As a direct consequence of this, one has to postulate that it should be possible to predict the native structure of a protein from the protein's amino acid sequence alone. However, despite much work of many excellent scientists and a database of experimentally determined protein structures that is increasing frighteningly fast on a daily basis,[2] successes in protein structure prediction are scarce, and the current situation is rather disappointing. The reasons for this are not entirely clear. A large body of experimental information has become available with the boost structural biology has experienced in the last decade and much effort has been put into the thorough analysis of these data over the years,[3–5] but, with the predictive power largely lacking, current knowledge of the basic principles of protein structure is still mainly descriptive. One explanation may be that currently known structural principles do not disclose the complete picture and that new concepts and new ideas are necessary to propel the field from the mainly descriptive into a more predictive mode.

[*] Dr. J. Sühnel, Dr. D. Pal
Institut für Molekulare Biotechnologie
Beutenbergstrasse 11, 07745 Jena (Germany)
Fax: (+49) 3641-656210
E-mail: jsuehnel@imb-jena.de
Dr. M. S. Weiss
EMBL Hamburg Outstation
c/o DESY, Notkestrasse 85, 22603 Hamburg (Germany)
Fax: (+49) 40-89902-149
E-mail: msweiss@embl-hamburg.de

[**] The authors are grateful to E. James Milner-White for introducing them to his nest concept and for stimulating discussions.

In this respect, two interesting papers have been published recently in the *Journal of Molecular Biology*.[6, 7] Based on the analysis of main-chain torsion angles of adjacent amino acid residues, Watson and Milner-White discovered that many anion and cation binding sites (where anions and cations can be any atoms exhibiting a full or a partial negative and positive charge, respectively) in proteins are made up by a sequence of three amino acids of which two exhibit "enantiomeric" main-chain conformations. The term "enantiomeric" refers to the fact that the main-chain torsion angles (ϕ,ψ) of the two adjacent amino acids are approximately inverted about the center of the Ramachandran plot.[3] Whereas successive residues with identical or nearly identical main-chain conformations form α helices, β strands, or polyproline type II helices, adjacent residues with enantiomeric main-chain conformations form so-called "nests" when their (ϕ,ψ)-values are close to ($-90°$,$0°$) and ($+90°$,$0°$) or the other way round.

The term "nest" is derived from the fact that the NH groups of three successive residues obeying this torsion angle criterion form a concave depression which can serve as a binding site for an atom or a group of atoms with a full or partial negative charge. Depending on which combination of the two torsion angle pairs is observed, the nests can be divided into RL nests ($\phi_1,\psi_1 = -90°,0°$; $\phi_2,\psi_2 = +90°,0°$) and LR nests ($\phi_1,\psi_1 = +90°,0°$; $\phi_2,\psi_2 = -90°,0°$). Two or more of these nests can also constitute a compound nest, a tandem nest, or a combination of both with up to eight successive residues involved. In the majority of cases the nests bind to an atom or a group of atoms, which we suggest may, as a binding partner of a "nest", be descriptively and conveniently called an "egg". It is intriguing that many structural motifs described previously, such as Schellman loops, the oxyanion holes of serine proteases, and P loops in ATP- or GTP-binding proteins can be subsumed under this nest/egg concept. If dipeptides with different enantiomeric main-chain torsion angle combinations are considered, the nest/egg concept can

  1433-7851/02/4124-4663 $ 20.00+.50/0