COMMUNICATION

Automated synthesis of backbone protected peptides†

Abu-Baker M. Abdel-Aal, George Papageorgiou, Martin Quibell and John Offer*

The synthesis of peptides rich in aggregation prone sequences can be improved with backbone protection. We report the automated introduction of backbone protection to a peptide. This new method was applied in a fully-automated synthesis, giving improved handling, quality and yield of several challenging target sequences.

Solid phase peptide synthesis (SPPS) is a powerful technology for the chemical synthesis of peptides and small proteins. However, access to many targets is often complicated and sometimes precluded by the occurrence of so-called difficult sequences.[1,2] When encountered during SPPS difficult sequences are associated with a collapse of the swollen resin volume, incomplete acylation and in the case of Fmoc/tBu synthesis, incomplete deprotection steps, extending over several residues.[3,4] As the cause of difficult sequences is intermolecular chain association, double coupling provides no improvement. Pioneering work by Sheppard and co-workers demonstrated that introducing proline into an aggregating peptide sequence, before the onset of aggregation prevented inter-chain association by removal of hydrogen bonding.[1] By extension, they also demonstrated that reversible alkylation of the peptide backbone suppressed interchain association and was a general solution to the difficult sequence problem.[5] Reversible substitution of the amide bond was called 'backbone protection' when introduced by Weygand and co-workers because of similarities to the protecting group strategies then in development.[6] However, it is only rarely necessary to protect the amide bond itself.[7] Amongst the many backbone protection groups available the most widely used are the commercially available pseudoprolines (ψ-Pro **1** Scheme 1).[8] Pseudoprolines have revolutionized Fmoc/tBu synthesis by enabling the synthesis of previously intractable peptides, unobtainable even by *in situ* neutralisation Boc protocols.[9] Their key advantage is that they can be introduced with great convenience as dipeptide building blocks[10]

and have been successfully applied to enable synthesis of difficult sequences[11,12] and long peptides.[13] However, their use is unfortunately limited to those sequences containing conveniently positioned X-Ser or X-Thr. Backbone amide protection was first investigated (Dmb **2** Scheme 1) for its dramatic effect on peptide solubility.[6] The effect of Dmb **2** on improving peptide solubility was explored in detail by Narita and co-workers.[14] Current opinion considers that the poor solubility observed for many peptide sequences in solution reappears on the solid phase as difficult peptides and backbone protection acts in both cases by disrupting structure formation.[1] A fully solvated peptide-resin should give the best coupling kinetics and evidence from NMR and I.R. studies on aggregating peptides on solid phase supports this model.[15,16] Many novel backbone protection strategies have been developed but the increased steric hindrance that accompanies the introduction of a secondary amine into a sequence prohibits quantitative coupling using standard conditions (except with glycine) and has limited the wider adoption of these new backbone protection strategies.[17] A solution to overcome this obstacle harnessed an intramolecular acyl transfer. This was achieved by the use of 2-hydroxy-4-methoxybenzyl (Hmb) **3** which can be considered as a simple modification of Dmb **2** (Scheme 1). Acylation of the exposed 2-hydroxy position gave a phenyl ester, positioned for acyl migration through a con-strained, six-membered ring. This procedure gave quantitative coupling under favourable conditions. However, the kinetics of acyl transfer were slow and variable between residues. Additionally, the optimised non-standard conditions used, consisting of a symmetric anhydride in dichloromethane, were difficult to automate.[5] Ideally, for practical convenience the coupling onto the secondary amine

MRC National Institute for Medical Research, The Ridgeway, London, UK.
E-mail: joffer@nimr.mrc.ac.uk; Tel: +44 (0)20 8816 2082

† Electronic supplementary information (ESI) available: Synthesis and characterization of peptides, experimental conditions. See DOI: 10.1039/c4cc03065f
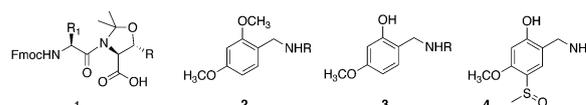
Scheme 1 Peptide backbone protection groups. (1) Pseudoproline(ψ-Pro), R = H, Ser; R = CH₃, Thr. (2) Dimethoxylbenzyl (Dmb); (3) 2-hydroxy-4-methoxybenzyl (Hmb), (4) 2-hydroxy-4-methoxy-5-methylsulfinyl benzyl (Hmsb).
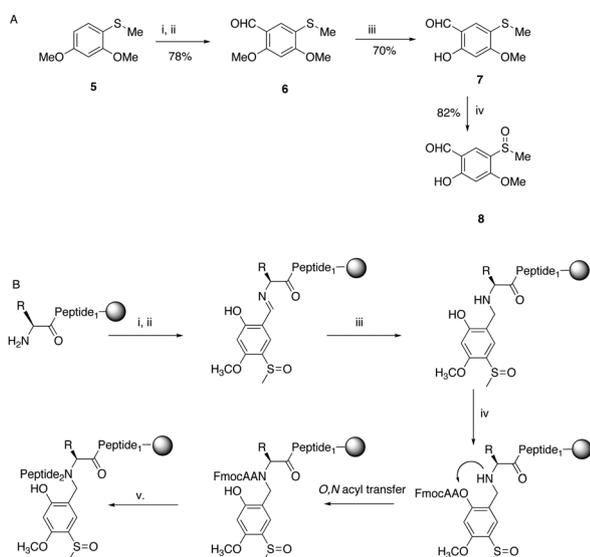
should be performed under standard conditions. Clearly, with a more reactive internal ester the acylation/migration step could be accelerated. This had previously been achieved by modifying Hmb with an electron withdrawing group *para* to the 2-hydroxy position.[18–20] Based on these considerations, we chose a variant of the sulfoxide modified Hmb previously reported by us Hmsb **4** (Scheme 2B).[18,20] The use of sulfoxide, resolves the problem of how to introduce an electron withdrawing group to the backbone protection without making the modification irreversibly stable to acid, as it can be mildly reduced to the acid labile thioether. For a chemical reaction to be generally applicable to the solid phase it has to be quantitative. Ede and co-workers had demonstrated introduction of Hmb to a resin-bound peptide by exploiting the unusually stable Schiff base,[21] formed between a salicylaldehyde and resin-bound primary amine, stable to washing and subsequently quantitatively reduced on resin.[22] Therefore a successful demonstration of quantitative on-resin reductive amination of salicylaldehydes, with a group capable of assisting its own acylation under standard conditions would allow automation of backbone protection installation.

Synthesis of the salicylaldehyde **8** (Scheme 2) was simplified from an earlier route and obtained in a good overall yield from commercial starting materials (ESI†). Chemoselective removal of the methylether by boron tribromide gave salicylaldehyde **7**. For automation to work the sulfoxide group had to be present for imine formation on the solid phase and survive the conditions of reductive amination (Scheme 2B). Previously, we had introduced backbone protection by synthesizing an amino acid building block and oxidizing the thioether to the sulfoxide *after* preparation of the building block.[18] However, oxidation is difficult to perform quantitatively on-resin and is also not compatible with all amino acids. Therefore thioether **7** was cautiously oxidized to the sulfoxide **8** in the presence of the unprotected aldehyde function. With the

sulfoxide substituted salicylaldehyde **8** in hand we first needed to prove that the salicylaldehyde could be installed quantitatively with the sulfoxide intact. In contrast to many aldehydes where imine formation is slow, salicylaldehydes form stable imines rapidly.[21]

Formation of the imine with a single equivalent of **8** gave an intense yellow colouration, the resin was washed to remove excess salicylaldehyde and treated with an additional single equivalent of salicylaldehyde before thoroughly washing with dimethylformamide (DMF). NaBH$_4$ in DMF was added and the strong yellow colouration of the Schiff base quickly faded. Analytical HPLC of a test cleavage showed the presence of a single product with the correct mass of the target peptide with backbone protection attached, bearing intact sulfoxide **4** (ESI,† Fig. S1).

This technology was demonstrated by the improved preparation of a 'difficult sequence' derived from influenza virus Hemagglutinin, reported by Sampson and co-workers (Fig. 1).[11] For comparison, the test sequence was synthesised without backbone protection on rink amide resin (Fig. 1A). The peptide aggregated around the tenth residue (Ser[10]) and provided a poor quality crude product with many deletion impurities (Fig. 1A). The use of microwave gave a marginally improved product, however it still contained major deletion sequences (Fig. 1C). In contrast, fully automated addition of Hmsb **4** backbone protection to Ala[9], followed by a standard coupling cycle for the addition of Lys[8] provided a greatly improved crude product (Fig. 1(B & D)). The results using conventional automated synthesis and peptide coupling agents O-(6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate. (HCTU)/DIEA, 30 min or microwave (DIC/1-hydroxybenzotriazole (HOBt), 10 min) were comparable.

H-Met-Glu-Asp-Ser-Thr-Tyr-Tyr-Lys-***Ala***-Ser-Lys-Gly-Cys-*NH₂*

Fig. 1 Synthesis of a difficult peptide sequence from influenza virus Hemagglutinin.[11] Analytical HPLC traces of crude product prepared using: (A) conventional automated SPPS, (a = target peptide, b = deletion of Met[1]-Glu[2]-Asp[3], c = deletion of Met[1]-Glu[2], d = deletion of Glu[1], e = *t*-butylated product). (B) Conventional automated SPPS with backbone protection at Ala[9]. (C) Automated microwave assisted SPPS. (D) Automated microwave-assisted SPPS backbone protection at Ala[9]. (E) MALDI-MS of target peptide, a, (calculated mass [M + H]$^+$ = 1481.6 m/z) peptide cleavage conditions: TFA/TMSBr/thioanisole/ethanedithiol (1.0 : 0.05 : 0.05 : 0.025 v/v/v), 1.0 h. HPLC conditions: RP-C18, 0–50% (0.1% TFA, 90% CH₃CN) in 30 min, 1 mL min⁻¹.

Scheme 2 (A) Synthesis of Hmsb backbone protection. (i) DMF, POCl₃, 0 °C; (ii) ClCH₂CH₂Cl, 80 °C; (iii) BBr₃·SMe₂, CH₂Cl₂, 0 °C; (iv) *m*-CPBA, CHCl₃, −10 °C. (B) Automated introduction of backbone protection to peptide on solid phase. (i) Imine formation; salicylaldehyde **8** 1.1 equivalent to resin loading; (ii) DMF wash; (iii) NaBH₄/DMF; (iv) FmocAA, HCTU/DIEA, 30 min; (v) SPPS.
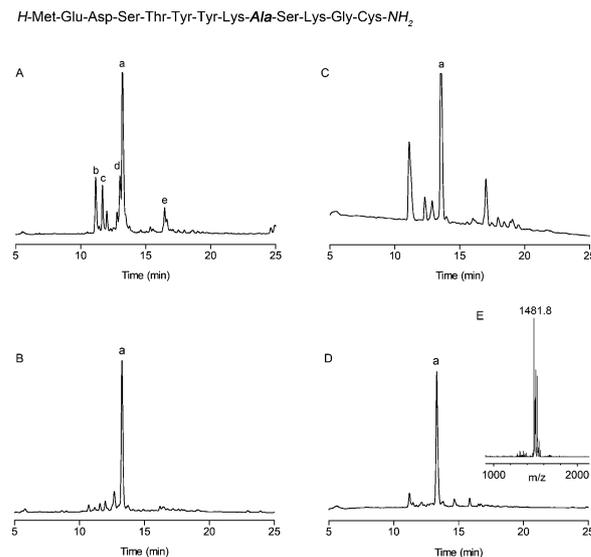
However, with conventional automated synthesis an additional dichloromethane mix/wash step for 1 h gave a noticeable improvement to the initial product purity, presumably by favouring intramolecular acyl migration.[5,23] The recovered yield after preparative HPLC was much higher, approximately 30%, for both cases, in contrast to 3% for conventional SPPS and 8% for microwave assisted SPPS, reflecting the less challenging purification task and the practical advantages of using backbone protection. The Hmsb backbone protecting group requires the sulfoxide to be reduced to a thioether for clean deprotection. Previously we used ammonium iodide and dimethylsulfide,[18] but we found trimethylsilyl bromide (TmsBr) and thioanisole in the cleavage cocktail more convenient.[24] All the peptides were cleaved using the same cleavage cocktail for comparison of the crude product (Scheme 3). The backbone protection can also be retained on the side-chain deprotected peptide for its beneficial solubility effects.

In the absence of an *a priori* method to predict an aggregation-prone sequence, a pragmatic approach would be to include a synthetic cycle to protect every sixth residue and prevent any potential aggregation. Polyalanine (with a C-terminal Val) is the prototypical difficult sequence, identified by Merrifield and co-workers and frequently used since as a benchmark.[25–27] Furthermore, homooligomers of alanine have become the subject of interest because of their biological relevance; as they are one of the most common homopeptide repeats and expanded polyalanine repeats are central to several neurodegenerative diseases.[28] Aggregation begins at the 5th alanine added from C-terminus and the addition of further alanine residues becomes increasingly troublesome.[26] The crude product at any of the later stages would be highly insoluble and difficult to analyse or purify. In contrast, automated introduction of two appropriately cited Hmsb backbone protecting groups (at positions Ala[8] and Ala[14]) using microwave protocols on a Tentagel resin not only prevented aggregation during SPPS, but also by retention on the cleaved crude peptide provided a fully soluble, chemically defined, analogue that is readily analysed and shown to be of high quality with confirmed molecular weight. Circular dichroism of the product gave a spectra characteristic for random coil (Fig. 2C). Cleavage of the Hmsb groups from the polyalanine product removed the solubilising properties afforded by backbone protection, yielding an insoluble, but chemically homogenous product.

Both previous examples had possessed alanine at the position for backbone protection. We therefore synthesised an additional example, the highly conserved epitope of gp41 that binds tightly to the broadly neutralising 4E10 antibody and previously synthesised with *in situ* neutralisation Boc cycles[29] We added backbone
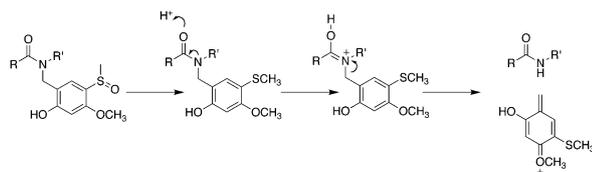
Fig. 2 Automated synthesis of the polyalanine octadecapeptide. Fmoc-Ala$_7$HmsbAla$_6$HmsbAla$_4$Val-OH (A). Analytical HPLC trace of crude product HPLC conditions: RP-C18, 30–50% B (0.1% TFA, 90% CH$_3$CN) in 30 min. 1 mL min^{-1}. (B) MALDI-MS of product (calculated mass [M + Na]$^+$ = 1965.8 m/z [M + K]$^+$ = 1981.8 m/z). (C) Circular dichroism of polyalanine containing backbone protection forming random coil (dotted). Aggregated beta-sheets of polyalanine (solid) without backbone protection.

protection at Leu^{679} using the automated procedure with tryptophan as the following residue. The results demonstrate successful incorporation of backbone protection and improved synthesis (ESI,† Fig. S2).

We have demonstrated that the introduction of backbone protection into a difficult sequence can be fully automated and delivers a significant improvement in yield and quality of previously aggregation prone peptides. The ability to freely add reversible amide protection along the peptide backbone has been a long-term goal for peptide chemists, both to overcome difficult sequences and also to solubilise the peptide in solution. The site of backbone protection is no longer restricted to only two residues. This study has demonstrated the suitability of salicylaldehydes for automated introduction to solid phase by the inclusion of an imine formation/reduction cycle. The Hmbs group actively participates in acylation using an activated acyl transfer so that its reactivity resembles a primary amine more than a secondary amine. With the demonstration of efficient automated introduction, backbone protection can be used preventively at every sixth residue for routine

Scheme 3 Mechanism of removal of Hmsb backbone protection. TFA/TMSBr/thioanisole/ethanedithiol (1.0 : 0.05 : 0.05 : 0.025 v/v).
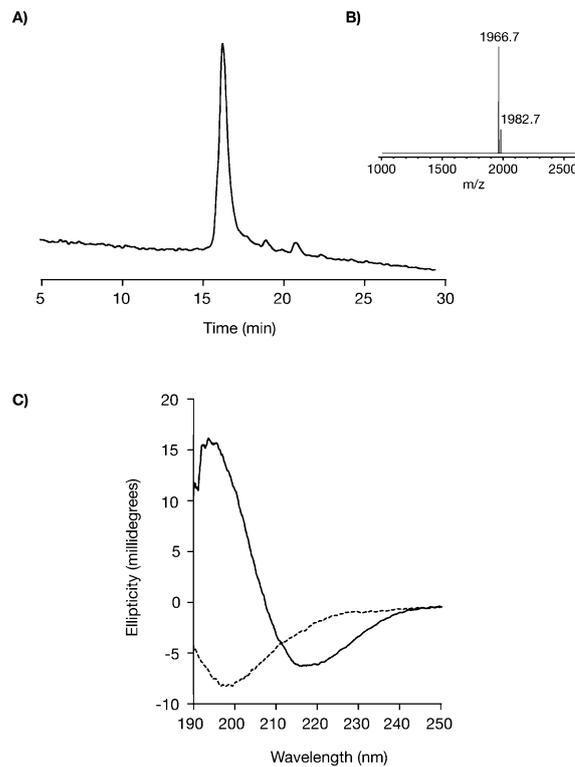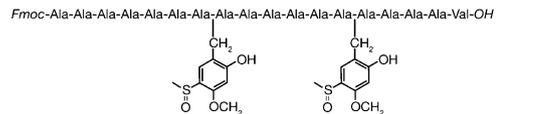
peptide synthesis. The single equivalent of salicylaldehyde required for this approach has potential to be relevant to optimised production of peptides. With the increasing recognition of the importance of peptides as next generation therapeutics and homooligopeptides as interesting clinical materials such a development will meet an urgent need. Of particular interest would be their application to emerging automated flow based techniques for rapid peptide synthesis.[30] We are currently investigating the scope of this modification, especially its tolerance to β-branched amino acids at the substitution site and its application to longer targets.

Notes and references

1 (a) C. Hyde, T. Johnson, D. Owen, M. Quibell and R. C. Sheppard, *Int. J. Pept. Protein Res.*, 1994, **43**, 431–440; (b) M. Quibell and T. Johnson, *Fmoc Solid Phase Pept. Synth.*, 2000, 115–135.
2 A. K. Tickler, A. B. Clippingdale and J. D. Wade, *Protein Pept. Lett.*, 2004, **11**, 377–384.
3 (a) S. B. Kent, *Annu. Rev. Biochem.*, 1988, **57**, 957–989; (b) S. B. H. Kent, D. Alewood, P. Alewood, M. Baca, A. Jones and M. Schnolzer, *Innovation Perspect. Solid Phase Synth. Collect. Pap., Int. Symp.*, 1992, pp. 1–22.
4 I. Coin, M. Beyermann and M. Bienert, *Nat. Protoc.*, 2007, **2**, 3247–3256.
5 T. Johnson, M. Quibell and R. C. Sheppard, *J. Pept. Sci.*, 1995, **1**, 11–25.
6 F. Weygand, W. Steglich, J. Bjarnason, R. Akhtar and N. M. Khan, *Tetrahedron Lett.*, 1966, **7**, 3483–3487.
7 (a) J. Offer, M. Quibell and T. Johnson, *J. Chem. Soc., Perkin Trans. 1*, 1996, 175–182; (b) M. Liley and T. Johnson, *Tetrahedron Lett.*, 2000, **41**, 3983–3985.
8 (a) T. Wohr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. C. Sun and M. Mutter, *J. Am. Chem. Soc.*, 1996, **118**, 9218–9227; (b) P. Dumy, M. Keller, D. E. Ryan, B. Rohwedder, T. Wohr and M. Mutter, *J. Am. Chem. Soc.*, 1997, **119**, 918–925.
9 M. Schnolzer, P. Alewood, A. Jones, D. Alewood and S. B. Kent, *Int. J. Pept. Protein Res.*, 1992, **40**, 180–193.
10 T. Wohr and M. Mutter, *Tetrahedron Lett.*, 1995, **36**, 3847–3848.
11 W. R. Sampson, H. Patsiouras and N. J. Ede, *J. Pept. Sci.*, 1999, **5**, 403–409.
12 G. A. Cremer, H. Tariq and A. F. Delmas, *J. Pept. Sci.*, 2006, **12**, 437–442.
13 (a) P. White, J. W. Keyte, K. Bailey and G. Bloomberg, *J. Pept. Sci.*, 2004, **10**, 18–26; (b) F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. de Jong, H. Hilkmann, T. K. Sixma and H. Ovaa, *Angew. Chem., Int. Ed.*, 2010, **49**, 10149–10153.
14 (a) M. Narita, K. Ishikawa, H. Nakano and S. Isokawa, *Int. J. Pept. Protein Res.*, 1984, **24**, 14–24; (b) M. Narita, K. Ishikawa, J. Y. Chen and Y. Kim, *Int. J. Pept. Protein Res.*, 1984, **24**, 580–587; (c) M. Narita, N. Ohkawa, S. Nagasawa and S. Isokawa, *Int. J. Pept. Protein Res.*, 1984, **24**, 129–134.
15 R. Warrass, J. M. Wieruszeski, C. Boutillon and G. Lippens, *J. Am. Chem. Soc.*, 2000, **122**, 1789–1795.
16 B. Henkel and E. Bayer, *J. Pept. Sci.*, 1998, **4**, 461–470.
17 (a) E. C. B. Johnson and S. B. H. Kent, *Chem. Commun.*, 2006, 1557–1559; (b) A. I. Fernandez-Llamazares, J. Spengler and F. Albericio, *Tetrahedron Lett.*, 2014, **55**, 184–188; (c) 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.
18 (a) J. Offer, T. Johnson and M. Quibell, *Tetrahedron Lett.*, 1997, **38**, 9047–9050; (b) J. Offer, M. Quibell and T. Johnson, *Int. Symp. - Innovation Perspect. Solid Phase Synth. Comb. Libr.*, 1999, 357–360.
19 L. P. Miranda, W. D. F. Meutermans, M. L. Smythe and P. F. Alewood, *J. Org. Chem.*, 2000, **65**, 5460–5468.
20 J. Howe, M. Quibell and T. Johnson, *Tetrahedron Lett.*, 2000, **41**, 3997–4001.
21 J. Crugeiras, A. Rios, E. Riveiros and J. P. Richard, *J. Am. Chem. Soc.*, 2009, **131**, 15815–15824.
22 N. J. Ede, K. H. Ang, I. W. James and A. M. Bray, *Tetrahedron Lett.*, 1996, **37**, 9097–9100.
23 D. S. Kemp, D. J. Kerkman, S. L. Leung and G. Hanson, *J. Org. Chem.*, 1981, **46**, 490–498.
24 N. Fujii, A. Otaka, N. Sugiyama, M. Hatano and H. Yajima, *Chem. Pharm. Bull.*, 1987, **35**, 3880–3883.
25 R. B. Merrifield, J. Singer and B. T. Chait, *Anal. Biochem.*, 1988, **174**, 399–414.
26 T. Johnson, M. Quibell, D. Owen and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1993, 369–372.
27 S. A. Kates, N. A. Sole, M. Beyermann, G. Barany and F. Albericio, *Pept. Res.*, 1996, **9**, 106–113.
28 J. P. Bernacki and R. M. Murphy, *Biochemistry*, 2011, **50**, 9200–9211.
29 (a) R. M. Cardoso, M. B. Zwick, R. L. Stanfield, R. Kunert, J. M. Binley, H. Katinger, D. R. Burton and I. A. Wilson, *Immunity*, 2005, **22**, 163–173; (b) F. M. Brunel, M. B. Zwick, R. M. F. Cardoso, J. D. Nelson, I. A. Wilson, D. R. Burton and P. E. Dawson, *J. Virol.*, 2006, **80**, 1680–1687.
30 M. D. Simon, P. L. Heider, A. Adamo, A. A. Vinogradov, S. K. Mong, X. Li, T. Berger, R. L. Policarpo, C. Zhang, Y. Zou, X. Liao, A. M. Spokoyny, K. F. Jensen and B. L. Pentelute, *ChemBioChem*, 2014, **15**, 713–720.