JOURNAL OF
PEPTIDE SCIENCE

Synthesis of full length PB1-F2 influenza A virus proteins from 'Spanish flu' and 'bird flu'

RENÉ RÖDER,[a,b] KARSTEN BRUNS,[a,c] ALOK SHARMA,[a,c,d] ANDRÉ EISSMANN,[a] FRIEDRICH HAHN,[a] NICOLE STUDTRUCKER,[a,d] TORGILS FOSSEN,[e] VICTOR WRAY,[c] PETER HENKLEIN[b]* and ULRICH SCHUBERT[a]

[a] Institute of Clinical and Molecular Virology, University of Erlangen-Nürnberg, D-91054 Erlangen, Germany
[b] Institute of Biochemistry, Charité Universitätsmedizin-Berlin, D-10117 Berlin, Germany
[c] Department of Structural Biology, Helmholtz Center for Infection Research, D-38124 Braunschweig, Germany
[d] ViroLogik GmbH, Innovationszentrum Medizintechnik und Pharma, 91052 Erlangen, Germany
[e] Department of Chemistry, University of Bergen, N-5007 Bergen, Norway

Abstract: The proapoptotic influenza A virus PB1-F2 protein contributes to viral pathogenicity and is present in most human and avian isolates. Previous synthetic protocols have been improved to provide a synthetic full length H1N1 type PB1-F2 protein that is encoded by the 'Spanish flu' isolate and an equivalent protein from an avian host that is representative of a highly pathogenic H5N1 'bird flu' isolate, termed SF2 and BF2, respectively. Full length SF2, different mutants of BF2 and a number of fragments of these peptides have been synthesized by either the standard solid-phase peptide synthesis method or by native chemical ligation of unprotected *N*- and *C*-terminal peptide fragments. For SF2 chemical ligation made use of the histidine and the cysteine residues located in positions 41 and 42 of the native sequence, respectively, to afford a highly efficient synthesis of SF2 compared to the standard SPPS elongation method. By-product formation at the aspartic acid residue in position 23 was prevented by specific modifications of the SPPS protocol. As the native sequence of BF2 does not contain a cysteine residue two different mutants of BF2 (Y42C) and BF2 (S47C) with appropriate cysteine exchanges were produced. In addition to the full length molecules, fragments of the native sequences were synthesized for comparison of their physical characteristics with those from the H1N1 human isolate A/Puerto Rico/8/34 (H1N1). All peptides were analyzed by mass spectrometry, ^{1}H NMR spectroscopy, and SDS-PAGE. The protocols allow the synthesis of significant amounts of PB1-F2 and its related peptides. Copyright © 2008 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: proapoptotic influenza A virus proteins; PB1-F2; Spanish flu; avian influenza; protein synthesis; chemical ligation

INTRODUCTION

Influenza A virus (IAV) is a pathogen in humans and animals, which caused at least three pandemics during the last century, the most serious being the outbreak of 'Spanish flu' in 1918 that claimed up to 40 million victims [1,2]. Specifically aquatic birds have been shown to be the primary reservoir of IAV, in which occasionally new avian strains or reassortants with the capability of infecting other mammals or even humans can be created. This is caused upon genetic reassortment by a process termed antigenic shift [3,4,5].

The understanding of the biological mechanisms how IAV mediates distinct pathogenicity to different hosts is a major focus of research. An 11th IAV protein, termed PB1-F2, encoded by an alternative (+1) open reading frame of PB1, has been characterized as a 87aa protein [6].

PB1-F2 contains a C-terminal mitochondrial targeting sequence [7], is localized in the mitochondria of infected or transfected cells and it can stimulate the mitochondrial apoptosis pathway [8,9]. Recently it has been reported that PB1-F2 increases the viral pathogenicity in mice, indicated by a delay of viral clearance from lungs [10,11]. The latest results indicate an important role of PB1-F2 in the regulation of IAV virulence in as much as a single amino acid exchange in the sequence of PB1-F2 from an isolate of the Hong Kong 1997 H5N1 outbreak contributes to a significant increase in viral pathogenicity [12]. In addition the expression of the 1918 PB1-F2 enhances secondary bacterial pneumonia [13,14], which is of high significance as most deaths arising from IAV are caused by a bacterial super-infection [15].

A prerequisite for detailed investigations of the structural and functional studies of PB1-F2 is the availability of suitable quantities of either recombinant or synthetic material of full length proteins corresponding to different isolates. Here we describe the first synthesis of the 90aa protein from the 'Spanish flu' variant, SF2, whose sequence was obtained by RNA isolation from the lung of a victim buried in the permafrost since 1918 [16]. In addition we have selected an analogous 90aa protein as a representative of a H5N1 'bird flu' isolate, BF2. We have established synthesis protocols for the production of both proteins and their related fragments.

* Correspondence to: Peter Henklein, Institute of Biochemistry, Charité Universitätsmedizin-Berlin, Monbijoustrasse 2, D-10117 Berlin, Germany; e-mail: peter.henklein@charite.de

MATERIAL AND METHODS

Solid Phase Peptide Synthesis

The SPPS was performed on an ABI 433A automated peptide synthesizer equipped with UV-detector from Applied Biosystems on a 0.1 mM scale using the Fmoc/*tert*-butyl strategy. The following side chain protecting groups were used during the automated synthesis: 2,2,4,6,7-pentamethyl-dihydro-benzofurane-5-sulfonyl (Arg), *tert*-butyloxycarbonyl (Trp, Lys), *tert*-butyl ether (Thr, Ser, Tyr), *tert*-butyl ester (Asp, Glu), and trityl (Asn, Cys, Gln, His).

Protein Sequences

The sequences used for the full length H1N1 PB1-F2 protein are those encoded by the 'Spanish flu' H1N1 human isolate A/Brevig Mission/1/1918 [16] (Figure 1) and 'bird flu' H5N1 avian isolate A/duck/Guangdong/12/2000 [17] (Figure 3).

1H NMR Spectroscopy

^{1}H TOCSY and NOESY NMR spectra of BF2 and SF2 were recorded on a Bruker Avance 600 MHz instrument equipped

Figure 1 RP-HPLC monitoring of the synthesis of full-length SF2 (* thiolester of SF2(1–41), § SF2(41–90), Ω SF2): (A) start of the ligation reaction; (B) after 45 min; (C) purified protein SF2. (D) Primary sequence of SF2 derived from the human H1N1 'Spanish flu' IAV isolate A/Brevig Mission/1/1918. (E) Schematic depiction of the three synthesized overlapping fragments. (F) Schematic depiction of the two fragments used for chemical ligation to the full-length SF2.

with an UltraShield Plus magnet and a triple resonance cryoprobe head with gradient unit. Samples were dissolved without pH adjustment (pH ~3.0) at concentrations between 0.3 and 0.8 mM (proteins), and 0.7 and 2 mM (peptide fragments) in H_2O/CF_3CD_2OH (1:1). Measurements were carried out at 300 K without spinning with mixing times of 110 ms for the TOCSY and 250 ms for the NOESY experiment, respectively. Data acquisition, processing and spectral analysis were performed with standard Bruker software. All spectra were internally referenced to the residual TFE-d_2 methylene signal at 3.95 ppm.

Mass Spectrometry

Matrix assisted laser desorption ionization mass spectra (MALDI-MS) were recorded on a Voyager-DE PRO BioSpectrometry Workstation from Applied Biosystems. Samples were dissolved in 50% aqueous acetonitrile and α-cyano-4-hydroxycinnamic acid was used as matrix. Positive ion electrospray ionization mass spectra (ESI-MS) were recorded on a micromass Q-Tof-2 mass spectrometer. Samples were dissolved in 70% aqueous methanol and infused into the electrospray chamber with a needle voltage of 0.9 kV at a flow rate of 40 nl/min.

SDS-PAGE

Detection of synthesized peptides and proteins by Coomassie staining: SF2, the overlapping fragments SF2(1–40), SF2(30–70), SF2(50–90), and BF2 variants BF2 Y42C and BF2 S47C and the overlapping fragments BF2(1–40), BF2(30–70), BF2(50–90) were diluted in phosphate buffered saline to a final protein concentration of 100 µg/ml. The samples were heated in SDS buffer (0.2% SDS, 10% glycerol, 65 mM Tris/HCl, pH 6.8, with 5% β-mercaptoethanol) at 55 °C for 5 min, and resolved on a 15% PAA gel by SDS-PAGE. After Coomassie staining (0.5% Coomassie Brilliant Blue G-250, 10% acetic acid, 25% isopropyl alcohol), the gel was destained (10% acetic acid, 25% methanol), and processed for imaging.

Synthesis and Purification of Overlapping Fragments of SF2 and BF2

Three overlapping fragments corresponding to the *N*-terminus SF2(1–40), the central section SF2(30–70), and the *C*-terminus SF2(50–90) of SF2 were synthesized for structural and biological analyses using a step-by-step protocol similar to that used previously by us for sPB1-F2 [18] derived from the H1N1 human isolate A/Puerto Rico/8/34, termed PR8 [6]. SF2(1–40) and SF2(30–70) were synthesized using TentaGel R-RAM amide resin (330 mg, capacity 0.19 mmol/g, RAPP Polymere) while TentaGel R-PHB-Ser(tBu)-Fmoc resin (280 mg, capacity 0.19 mmol/g, RAPP Polymere) was used for SF2(50–90). The couplings were performed with 2(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU). Deprotection of the Fmoc group was performed with 20% piperidine in DMF. The dipeptide derivative Fmoc-Asp(OtBu)-(Hmb)Gly-OH was used during the synthesis of SF2(1–40) to prevent aspartimide formation at residues Asp-23/Gly-24, as described [18]. The three peptides were cleaved from the resin with trifluoroacetic acid/water/triisopropylsilane (95:3:2 v/v/v) for 3.5 h at rt.
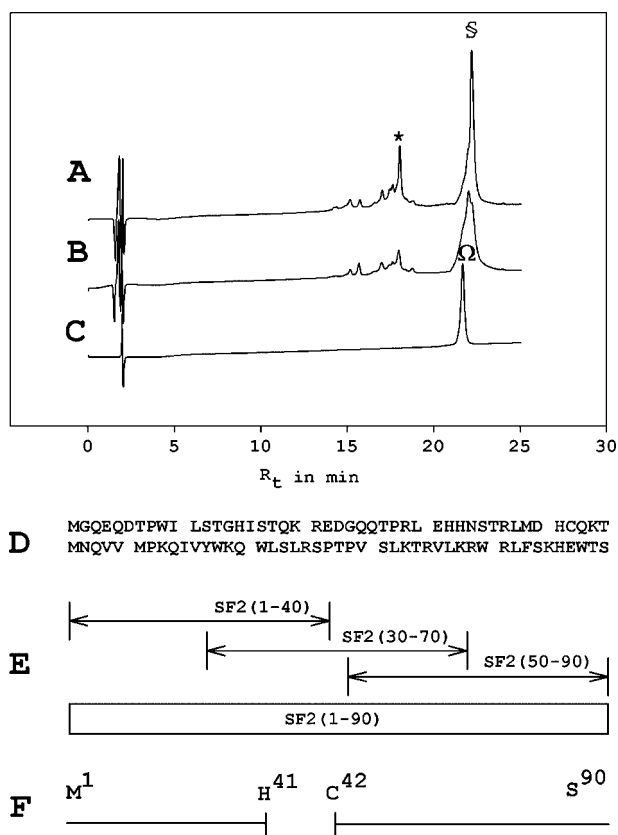
After concentration of the cleavage solution and addition of ice-cold diethyl ether, the crude product was obtained as a colorless precipitate. In spite of using double coupling steps to increase the efficiency for all three fragments the synthesis of SF2(1–40) afforded a crude product with only 17% yield compared to 44% for SF2(30–70) and 51% for SF2(50–90) (HPLC data not shown). The crude peptides were purified by reverse phase high performance liquid chromatography (RP-HPLC) on a Shimadzu LC8 preparative HPLC using a Zorbax SB C18 column (21.2 × 250 mm, 7 µm, Agilent) with a linear gradient of 28% B–48% B in 50 min for SF2(1–40) and 39% B–60% B in 50 min for both SF2(30–70) and SF2(50–90) (A: 2500 ml water, 5 ml TFA; B: 2000 ml acetonitrile, 500 ml water, 5 ml TFA) at a flow rate of 10 ml/min with spectrometric monitoring at $\lambda = 220$ nm to give the final pure products. The purity of the peptides was analyzed by RP-HPLC (Shimadzu LC10) on a Nucleosil 300 C18 (4.6 × 125 mm, 5 µm) with a linear gradient of 10% B–100% B over 45 min.

The three fragments BF2(1–40), BF2(30–70), and BF2(50–90) were prepared using a standard SPPS protocol with double coupling steps and HBTU as coupling agent. For BF2(1–40) the preloaded resin TCP-Asp(OtBu)-Fmoc (250 mg, capacity 0.53 mmol/g, PepChem, Tübingen), for BF2(30–70) TCP-Gly-Fmoc (260 mg, capacity 0.44 mmol/g, PepChem, Tübingen), and for BF2(50–90) TCP-Asn(Trt)-Fmoc (250 mg, capacity 0.48 mmol/g, PepChem, Tübingen) were used. The peptides were cleaved from the resin as described above and then purified by RP-HPLC on a Shimadzu LC8 system. A Kromasil C18 column (250.0 × 50 mm, 10 µm, AKZO Nobel, Sweden) with a linear gradient of 25% B–75% B in 50 min at a flow rate of 60 ml/min was used for BF2(1–40). A Zorbax SB C18 column (21.2 × 250 mm, 7 µm, Agilent) was employed for both BF2(30–70) and BF2(50–90) with linear gradients of 40% B–65% B and 35% B–60% B, respectively, in 50 min at a flow rate of 10 ml/min and spectrometric monitoring at $\lambda = 220$ nm. The peptides were analyzed by RP-HPLC as described above for the three SF2 fragments.

Synthesis and Purification of Full-Length SF2

Unfortunately, the use of a 0.05 M Tris/3 M urea buffer system at pH 8.5 in the native chemical ligation (NCL) method [18,19] was not successful for either SF2 or the mutants of BF2. Hence a 6 M guanidine hydrochloride system at pH 6.5 was used as described by Dawson and Kent [20].

The *N*-terminal 41-residue fully protected fragment SF2(1–41) with Boc-methionine at the *N*-terminus was synthesized with HBTU and utilized double coupling steps by the Fmoc/*tert*-butyl solid phase strategy with an automated peptide synthesizer on 2-chlorotrityl (Cl-Trt) resin (250 mg, capacity 0.65 mmol/g, Iris Biotech), which was preloaded with the first *C*-terminal amino acid histidine. The fully protected peptide SF2(1–41) was then cleaved from the resin with acetic acid/trifluoroethanol/dichloromethane (2:2:6, v/v/v) for 2 h at rt. To control the quality of the protected peptide an aliquot was fully deprotected with trifluoroacetic acid/water/triisopropylsilane (95:3:2, v/v/v) for 3.5 h at rt. The MALDI-TOF mass spectrum showed the correct mass ([M]$_{calc.}$: 4797.3 Da, [M]$_{found}$: 4797 Da), although HPLC analysis revealed only a very limited amount of the desired peptide in the crude product. Therefore the strategy was changed and

H-His(Trt)-2-Cl-Trt resin (200 mg, capacity 0.57 mmol/g, Novabiochem) was applied and *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) as coupling reagent instead of HBTU. After coupling of the dipeptide Fmoc-Asp(OtBu)-(Hmb)Gly-OH, the pseudoproline Fmoc-Ser(tBu)-Thr($\Psi^{Me,Me}$,pro)-OH was employed for introduction of residues Ser-12/Thr-13 and Ser-17/Thr-18 to prevent secondary structure formation of the protected peptide SF2(1–41) on the resin. The correct mass was found in the MALDI-TOF mass spectrum ([M]$_{calc.}$: 4797.3 Da, [M]$_{found}$: 4797 Da) after full deprotection with trifluoroacetic acid/water/triisopropylsilane (95:3:2, v/v/v) for 3.5 h at rt. RP-HPLC analysis indicated a 52% yield of the crude product (HPLC data not shown). The protected peptide SF2(1–41) was then cleaved from the resin as described above and used to synthesize the thiolester by treating the crude protected peptide (30 mg, 3.31 µmol, 9057 g/mol) with diisopropylcarbodiimide (DIPCDI) (1.5 eq, 0.8 µl), HOBt.H$_2$O (1.5 eq, 0.8 mg) and *p*-acetamidothiophenol (15 eq, 8.3 mg) in dichloromethane (1.5 ml) overnight [21]. The solvent was removed and the residue deprotected as described above. After precipitation with ice-cold diethyl ether, the deprotected peptide thiolester was centrifuged and washed with diethyl ether and acetonitrile to remove all traces of *p*-acetamidothiophenol and its oxidized dimer. MALDI-TOF-MS indicated the correct mass ([M]$_{calc.}$: 4946 Da, [M]$_{found}$: 4946 Da). The peptide thiolester was obtained in a yield of approximately 90% based on the amount of peptide in the crude product. The thiolester was used for the NCL without further purification.

SF2(42–90), the 49-mer *C*-terminal peptide of SF2 bearing a cysteine at its *N*-terminus, was synthesized via SPPS on H-Ser(tBu)-HMPB resin (250 mg, capacity 0.21 mmol/g, ChemMatrix) and HBTU as coupling agent, and double coupling steps were used to optimize the yield. MALDI-TOF-MS indicated the correct mass ([M]$_{calc.}$: 6030.3 Da, [M]$_{found}$: 6031 Da). Before further use in the NCL reaction the peptide was purified by RP-HPLC on a Shimadzu LC8 preparative HPLC using a C18 column (30.0 × 250 mm, 10 µm, Agilent) with a linear gradient of 40% B–60% B in 50 min at a flow rate of 20 ml/min with spectrometric monitoring at $\lambda = 220$ nm.

For the NCL reaction both fragments, the thiolester of SF2(1–41) (21 mg, 4.25 µmol) and the *C*-terminal fragment SF2(42–90) (18.9 mg, 3.13 µmol), were dissolved separately in 0.5 ml of 6 M guanidine hydrochloride solution. Both solutions were combined and an equimolar amount of tris-(2-carboxyethyl)-phosphine-hydrochloride (1 mg, 3.49 µmol) was added to prevent the oxidation of the cysteine residue. The pH was adjusted to 6–7 with 0.1 M NaHCO$_3$ solution and the reaction was monitored by RP-HPLC with the signal of the thiolester as indicator. The reaction was complete after 2 h, and the ligation product was then purified by RP-HPLC on a Shimadzu LC8 preparative HPLC using a C18 column (30.0 × 250 mm, 10 µm, Agilent) with a linear gradient of 35% B–55% B in 50 min at a flow rate of 20 ml/min with spectrometric monitoring at $\lambda = 220$ nm (Figure 1). ESI-MS (Figure 2) and MALDI-TOF-MS confirmed the correct mass of the ligation product ([M]$_{calc.}$: 10 809.5 Da, [M]$_{found}$: 10 810 Da). We obtained 5.4 mg (0.5 µmol) of the purified peptide SF2 with a yield of approximately 16% based on the amount of SF2(42–90).
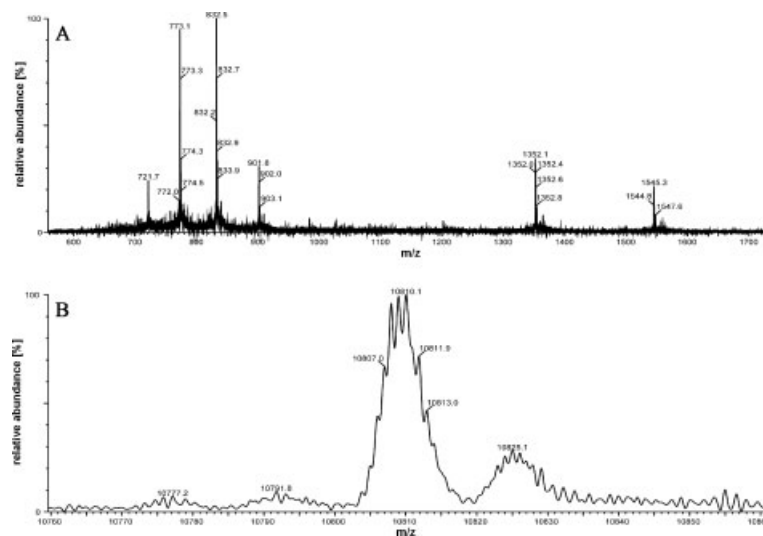
Figure 2 Positive ion ESI mass spectrum of purified SF2 showing the experimental mass spectrum (A) and the deconvoluted mass spectrum with the intense envelope of the molecular ion centered at 10 810.1 Da (B).

Synthetic Approaches to Full-length Native BF2

For the synthesis of the full-length BF2 the preloaded resin H-Asn(Trt)-HMPB (200 mg, capacity 0.26 mmol/g, ChemMatrix) was used and the coupling agent was changed from HBTU to HATU before amino acid Trp-61. After the resin cleavage step, none of the expected full-length peptide could be detected by MALDI-TOF-MS ($[M]_{calc.}$: 11 028.8 Da).

In a second approach the fragment condensation strategy on solid support was attempted. For the synthesis of the *C*-terminus, the preloaded resin H-Asn-(Trt)-HMPB (300 mg, capacity 0.26 mmol/g, ChemMatrix) was employed with HBTU as coupling agent. In order to optimize the synthesis double coupling steps were used after amino acid Glu-87 and the pseudoproline derivative Fmoc-Leu-Ser($\Psi^{Me,Me}$,pro)-OH was employed for Leu-62/Ser-63. The correct mass of the peptide was confirmed by MALDI-TOF-MS ($[M]_{calc.}$: 5029 Da, $[M + H]_{found}$: 5030 Da) after cleavage of an aliquot of the resin. The middle fragment BF2(25–50) was prepared on the preloaded resin H-Gly-2-Cl-Trt (150 mg, capacity 0.77 mmol/g, Novabiochem) with double coupling steps and HBTU. The fully protected peptide was then cleaved from the resin as described above. Control via MALDI-TOF-MS after cleavage/deprotection of an aliquot of the peptide showed the expected mass with the Fmoc protection group at the *N*-terminus ($[M + Fmoc]_{calc.}$: 3430 Da, $[M + Fmoc]_{found}$: 3430 Da). Finally, the *N*-terminal peptide BF2(1–24) was synthesized on the same resin (250 mg) and applying the identical coupling reagent. After the fourth amino acid, double coupling steps were performed to increase the synthesis efficiency. The peptide was cleaved as reported above and MALDI-TOF-MS showed the expected mass ($[M]_{calc.}$: 2828 Da, $[M - H]_{found}$: 2827 Da). As the *C*-terminal fragment BF2(51–90) was still linked to the resin we tried to conjugate it with the middle fragment BF2(25–50). For this purpose, the fully protected middle fragment BF2(25–50) (368.6 mg, 2 eq, 64.74 µmol) was added to 124.5 mg of the resin with the linked *C*-terminus BF2(51–90), which had been swollen before use in 5 ml dimethylformamide. Subsequently HATU (18.5 mg, 1.5 eq, 48.56 µmol) and DIPEA (16.5 µl, 3 eq, 97.11 µmol) were added and the reaction mixture was stirred at rt for 1 week. The

resin was then washed with dichloromethane and an aliquot was cleaved from the resin as described above for analysis by MALDI-TOF-MS. The mass of the expected peptide BF2(25–90) with the Fmoc protection group at its *N*-terminus could not be detected only the original fragments BF2(25–50) with Fmoc group and BF2(51–90) ($[M + Fmoc]_{calc.}$: 8442 Da, $[M1]_{found}$: 3430 Da, $[M2]_{found}$: 5079 Da) were found.

Synthesis and Purification of Mutants of Full-length BF2

The fully protected *N*-terminal fragment BF2(1–41) with Boc-methionine at the *N*-terminus was synthesized on the preloaded H-His(Trt)-2-Cl-Trt resin (200 mg, capacity 0.57 mmol/g, Novabiochem) by double coupling steps and HBTU as coupling reagent. The pseudoproline derivatives Fmoc-Asp(OtBu)-Thr($\Psi^{Me,Me}$,pro)-OH for Asp-6/Thr-7 and Fmoc-Ser-(tBu)-Thr($\Psi^{Me,Me}$,pro)-OH for Ser-12/Thr-13 were used. In order to minimize any secondary structure formation and to increase the synthesis efficiency the sterically less hindered Fmoc-Asn(Dmcp)-OH was used at position 17 instead of the trityl derivative. The protected peptide was then cleaved from the resin as described above. The α-thiolester of BF2(1–41) was produced by treating the crude protected peptide (200 mg, 20.73 µmol, 9649 g/mol) with DIPCDI (2 eq, 6.5 µl), HOBt.H_2O (2 eq, 6.4 mg) and *p*-acetamidothiophenol (15 eq, 52 mg) in dichloromethane (8 ml) overnight. The solvent was removed and the residue deprotected with trifluoroacetic acid/water/triisopropylsilane (95:3:2 v/v/v) for 3.5 h at rt. After precipitation with ice-cold diethyl ether, the deprotected peptide thiolester was centrifuged and washed with diethyl ether and then with acetonitrile to remove all traces of *p*-acetamidothiophenol and its oxidized dimer. The correct mass of the thiolester of the peptide BF2(1–41) was found by MALDI-TOF-MS ($[M]_{calc.}$: 5090 Da, $[M + H]_{found}$: 5091 Da). The crude thiolester was then purified by RP-HPLC on a Shimadzu LC8 system by using a VDS Optilab Kromasil C18 column (250.0 × 50 mm, 10 µm, AKZO Nobel, Sweden) with a linear gradient of 30% B–80% B in 50 min at a flow rate of 50 ml/min with spectrometric monitoring at $\lambda = 220$ nm.

For the peptide BF2(42–90) Y42C, the 49-mer bearing a cysteine residue at its *N*-terminus was synthesized using SPPS with the H-Asn(Trt)-HMPB resin (220 mg, capacity 0.26 mmol/g, ChemMatrix) and HBTU as coupling agent. For optimized synthesis double coupling steps and the pseudoproline Fmoc-Leu-Ser($\Psi^{Me,Me}$,pro)-OH for Leu-62/Ser-63 were employed. After cleavage/deprotection the correct mass of the peptide BF2(42–90) Y42C ($[M]_{calc.}$: 6044.3 Da, $[M]_{found}$: 6044 Da) was observed. Before further use in NCL the peptide was purified by RP-HPLC as described for SF2(42–90) with a linear gradient of 40% B–65% B in 50 min at a flow rate of 15 ml/min.

The first mutant BF2 Y42C was then assembled according to the NCL procedure described above for SF2 using the thiolester of BF2(1–41) (12.0 mg, 2.36 μmol) and the *C*-terminal fragment BF2(42–90) Y42C (12.8 mg, 2.12 μmol). The ligation product was purified by RP-HPLC as described above with a linear gradient of 38% B–58% B in 50 min at a flow rate of 15 ml/min (HPLC data not shown). ESI-MS and MALDI-TOF-MS confirmed the correct mass of the ligation product ($[M]_{calc.}$: 10 967.8 Da, $[M + H]_{found}$: 10 970 Da, data not shown). We obtained 2.8 mg (0.26 μmol) of the purified peptide BF2 Y42C with a yield of approximately 12% based on the amount of BF2(42–90) Y42C.

For the second mutant the fully protected *N*-terminal fragment BF2(1–46) with Boc-methionine at the *N*-terminus was obtained via SPPS with HBTU using 2-chlorotrityl resin (250 mg) from CBL Patras, which was initially preloaded with the *C*-terminal methionine residue (M46), and had a capacity of 0.124 mmol/g as determined by UV-spectroscopy at $\lambda = 301$ nm after cleavage of the Fmoc protecting group. For the synthesis of this peptide we used the same protocol as described for BF2(1–41). The α-thiolester of BF2(1–46) was obtained by treating the crude peptide (80 mg, 7.54 μmol, 10 611.3 g/mol) with DIPCDI (1.5 eq, 1.8 μl), HOBt.H_2O (1.5 eq, 1.7 mg) and *p*-acetamidothiophenol (15 eq, 18.9 mg) in 3 ml dichloromethane overnight. After cleavage/deprotection of an aliquot, the correct mass of the thiolester was confirmed ($[M]_{calc.}$: 5067.6 Da, $[M]_{found}$: 5068 Da), and the crude thiolester was purified by RP-HPLC on a Shimadzu LC8 preparative HPLC using a C18 column (30.0 × 250 mm, 10 μm, Agilent) with a linear gradient of 40% B–65% B in 50 min at a flow rate of 15 ml/min with spectrometric monitoring at $\lambda = 220$ nm.

Peptide BF2(47–90) S47C, the 44-mer bearing a cysteine at its *N*-terminus, was synthesized via SPPS with HBTU using the H-Asn(Trt)-HMPB resin (150 mg, capacity 0.26 mmol/g, ChemMatrix), double coupling steps and the pseudoproline derivative Fmoc-Leu-Ser($\Psi^{Me,Me}$,pro)-OH for positions 62/63. The peptide showed the correct mass ($[M]_{calc.}$: 5443.6 Da, $[M]_{found}$: 5443 Da) and was purified prior to NCL by RP-HPLC on a Shimadzu LC8 preparative HPLC using a Kromasil C18 column (250.0 × 50 mm, 10 μm, AKZO Nobel, Sweden) with a linear gradient of 37% B–87% B in 50 min at a flow rate of 50 ml/min with spectrometric monitoring at $\lambda = 220$ nm.

The second mutant BF2 S47C was obtained as described above for BF2 Y42C after NCL of the thiolester of BF2(1–46) (13.3 mg, 2.31 μmol) and the *C*-terminal fragment BF2(47–90) S47C (12.5 mg, 2.30 μmol), and the reaction was complete within 4 h. The ligation product was purified by RP-HPLC as described for SF2 with a linear gradient of 40% B–60% B in 50 min at a flow rate of 15 ml/min (Figure 3). ESI-MS (Figure 4) and MALDI-TOF-MS confirmed the correct mass of

the ligation product ($[M]_{calc.}$: 11 043.5 Da, $[M]_{found}$: 11 043 Da). We obtained 5.2 mg (0.47 μmol) of the purified peptide BF2 S47C with a yield of approximately 20% based on the amount of BF2(47–90) S47C.

RESULTS AND DISCUSSION

For the synthesis of the full-length SF2, we have applied the NCL method as this strategy has been successful the protein *s*PB1-F2 derived from the H1N1 human isolate PR8 [18]. The NCL method involves the condensation of two unprotected peptide segments, one bearing a *C*-terminal α-thiolester and the other a *N*-terminal cysteine residue, to afford a full-length polypeptide with a native amide linkage at the site of ligation [20,22,23]. Conveniently, a cysteine residue is present at position 42 of SF2, which could be exploited for this ligation approach.

As the native sequence of BF2 does not contain a cysteine residue for use in the NCL method, synthesis of
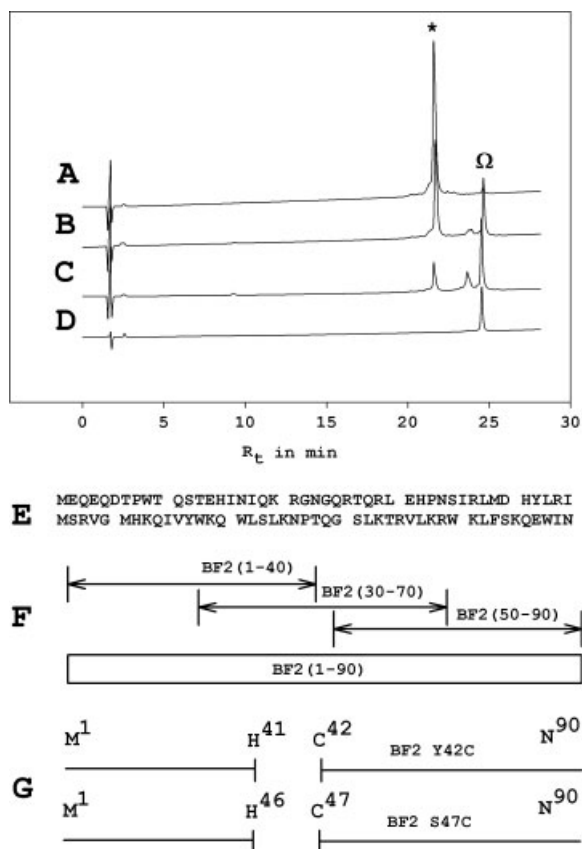

Figure 3 RP-HPLC monitoring of the synthesis of BF2 S47C. (* thiolester of BF2(1–46) and BF2(47–90) S47C, Ω BF2 S47C): (A) start of the ligation reaction; (B) after 45 min, and (C) after 3 h; (D) purified peptide BF2 S47C. (E) Primary sequence of BF2 derived from the avian H5N1 IAV isolate A/duck/Guangdong/12/2000. (F) Schematic depiction of the three synthesized overlapping fragments. (G) Schematic depiction of the two fragments used for chemical ligation to both full-length mutants.
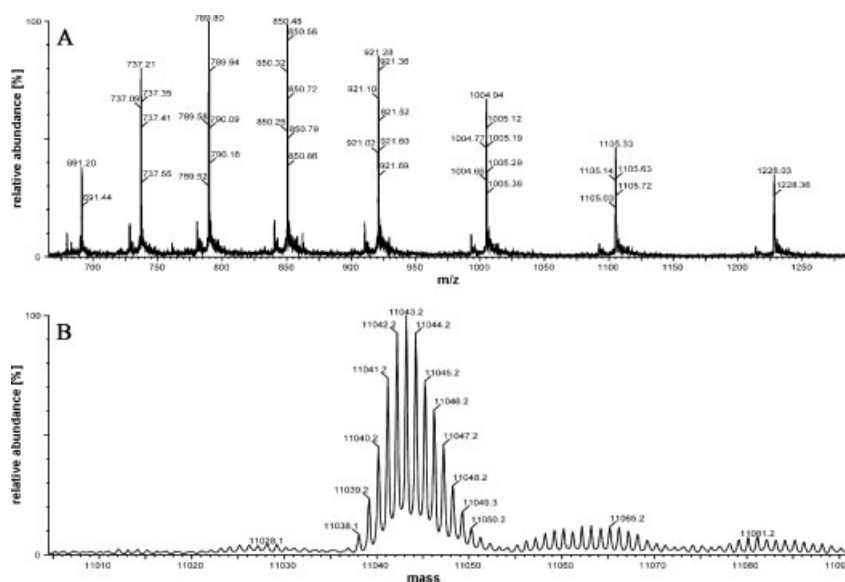
Figure 4 Positive ion ESI mass spectrum of purified BF2 S47C showing the experimental mass spectrum (A) and the deconvoluted mass spectrum with the intense envelope of the molecular ion centered at 11 043.2 Da (B).

the full-length molecule via the standard SPPS protocol was attempted. A new PEG-resin from ChemMatrix, with better swelling features than normal polystyrene resins [24], was selected. After cleavage from the resin, however, the desired BF2 peptide could not be detected in the crude product. In a second approach, we attempted the convergent fragment condensation procedure on solid support by coupling the fully protected peptide BF2(25–50) to the resin linked *C*-terminus BF2(51–90) and subsequent condensation of the fully protected peptide BF2(1–24) to obtain the full-length BF2. In order to prevent the risk of racemization the couplings were performed at glycine residues 24 and 50 of the native sequence. Again the native full-length BF2 could not be produced and only fragments were detectable in the mass spectrum.

As a consequence of our success with the NCL method for SF2 reported above and previously for *s*PB1-F2 [18], we decided to introduce cysteine residues into the central region of the molecule and thus to generate variants of BF2. As first variant the peptide BF2 Y42C was synthesized as this mutation occurs in nearly 30% of native avian influenza A PB1-F2 sequences; it is located in the same position as that in SF2 synthesized above and *s*PB1-F2 synthesized previously. For a second variant an exchange S47C was selected in order to preserve a residue of similar size and properties. The time for this ligation was longer than that of the first mutant BF2 Y42C and of SF2 as a methionine thiolester was used instead of the histidine thiolester. One reason for the different reaction times could be the differences in the steric properties of the two amino acids [25]. However, the strategy was successful and is currently the only method that produces significant amounts of the mutants of BF2.

Molecular Characterization of the Full-Length Proteins and their Fragments

The experimental ESI mass spectra of both BF2 mutants (Y42C, MS data not shown and S47C, Figure 4) showed well-defined peaks for the multiply charged (9–16 fold positively charged) molecular ions (Figure 4(A)) which after deconvolution resulted in molecular ion clusters of isotope-resolved peaks at the expected molecular masses (Figure 4(B)) with only traces of by-products that could not be removed by further purification steps.

The correct mass (average 10 810.4 Da) was also established for SF2 (Figure 2(A)) but in contrast to the BF2 peptides the measured spectrum showed a less regular distribution of molecular ion peaks here (Figure 2(A)). It is widely accepted that such a pattern of peaks is indicative of a heterogeneous population of a peptide with the highly charged peaks (12–15 fold in this case) representing a freely accessible (unfolded, denatured) and the lower charged peaks (7–8 fold) representing a more compact form of the peptide that can be surface protonated. Further, the deconvoluted spectra of SF2(1–90) show small amounts of a +16 product most likely representing a population of the peptide oxidized at one of its Met or Trp residues.

In addition to the reverse phase acetonitrile gradient HPLC chromatography and mass spectrometric analyses, the purity of the peptides and proteins was assessed by SDS-PAGE under reducing conditions (Figure 5). In each case single bands were detectable, thus confirming their purity after HPLC purification.

In order to provide more direct evidence for the correct synthesis of the various peptides, sets of 2D ^{1}H NMR spectra were recorded in 50% aqueous trifluoroethanol
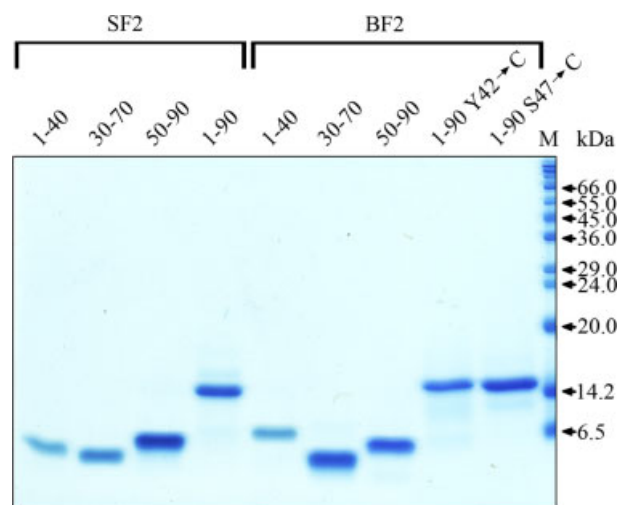
Figure 5 Quality control of peptides and proteins by SDS-PAGE: 1 μg of synthesized peptide/protein per lane were resolved on a 15% PAA gel by SDS-PAGE under reducing conditions, followed by staining with Coomassie blue. molecular weight marker proteins are shown at the right: 6.5 kDa aprotinin; 14.2 kDa α-lactalbumin; 20 kDa trypsin inhibitor; 24 kDa trypsinogen; 29 kDa carbonic anhydrase; 36 kDa glyceraldehydes-3-phosphate dehydrogenase; 45 kDa ovalbumin; 55 kDa glutamic dehydrogenase; 66 kDa albumin. This figure is available in colour online at www.interscience.wiley.com/journal/jpepsci.

under conditions published previously for this kind of protein [26]. Standard procedures allowed the unambiguous amino acid spin system and sequence assignment of all the peptides corresponding to the *N*-, central, and *C*-terminal fragments of BF2 and SF2

(data not shown). Thus, the NMR data confirmed the correct sequence composition and no significant signals corresponding to either by-products or impurities were detected.

It was furthermore important to establish the correctness of the full-length products SF2, BF2 Y42C, and BF2 S47C. The qualitiy of the peptides can be conveniently illustrated with the protein BF2 Y42C. As expected homonuclear ^{1}H spectra of this peptide revealed considerable signal overlap at 600 MHz as well as differential signal broadening in the 1D ^{1}H NMR spectrum (Figure 6). In the NH/aromatic region of the 2D TOCSY spectrum between 6.4 and 8.8 ppm only distinguishable traces from the NH signals are observed for residues at the termini of the molecule with the intensities of the cross peaks decreasing as one proceeds into the interior of the sequence. Similarly, the side-chain signals display sharper signals than those in the backbone. These phenomena are observed for the five low-field ring NH signals of the tryptophan residues between 9.5 and 10 ppm (inset Figure 6) where the signals of residues 9, 80, and 88 are sharper than those of the interior residues 58 and 61. Detailed comparison of the 2D ^{1}H COSY, TOCSY, and NOESY spectra in the NH/aromatic region of the protein and of these with those of the fragments allowed the amino acid spin system and sequence assignments of the first 25 *N*-terminal residues and 19 of the 25 *C*-terminal residues (Figure 7). NOE interactions of the δ ring protons of the aromatic residues to the α and β protons afforded partial assignment of these residues in the central section of the molecule. In total these data confirmed the
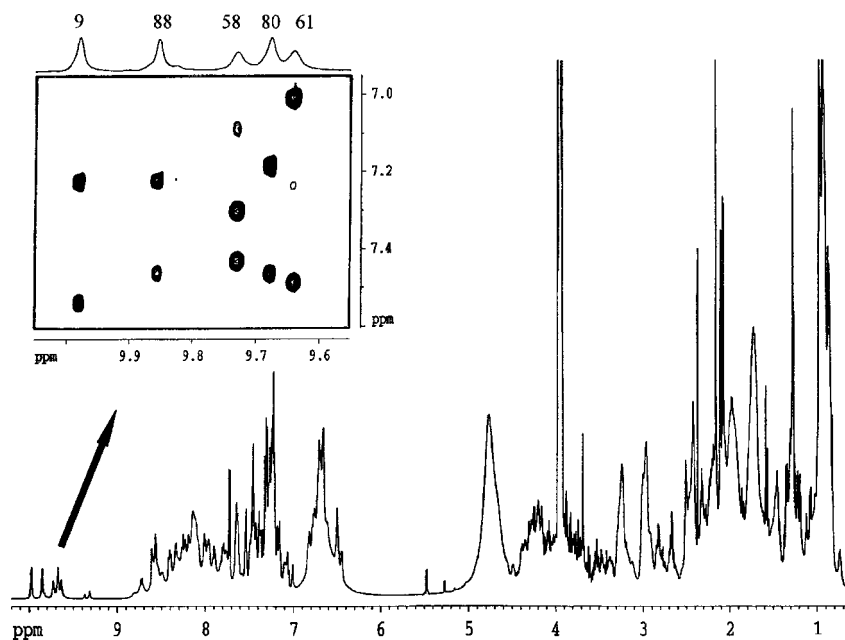

Figure 6 1D ^{1}H NMR spectrum of BF2 Y42C recorded in 50% aqueous TFE at 300 K and pH 3. The inset region of the 2D NOESY spectrum shows the five Trp ring NH intra-residue interactions with H2 (7.0–7.3 ppm) and H7 (7.4–7.6 ppm).
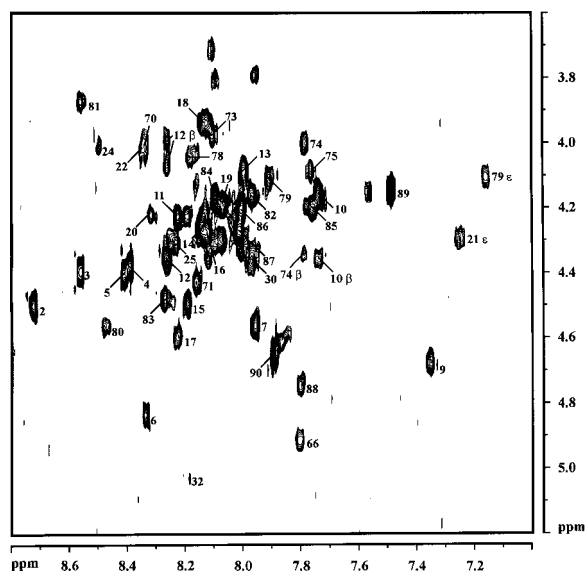
Figure 7 Assigned 'Fingerprint region' of the 2D TOCSY spectrum of BF2 Y42C recorded in 50% aqueous TFE at 300 K and pH 3. The strongest signals are observed for residues at the termini of the molecule with intensities decreasing for residues in the interior of the sequence.

molecular identity of the synthetic material and offered a first partial insight into the folding of the *N*- and *C*-termini.

CONCLUSIONS

The present study of the synthesis of three related full-length PB1-F2 molecules demonstrates that (i) medium-size fragments of up to approximately 40 residues can be conveniently synthesized by SPPS using the step-by-step method provided that side-chain reactions involving Asp–Gly and folding problems are prevented by the use of Hmb-protected glycine residue and the pseudo-proline strategy, and (ii) for larger peptides the step-by-step method is too inefficient and alternative procedures are required. Natural chemical ligation is one of the methods of choice for such full-length polypeptides, but requires the presence of a cysteine residue preferably in the center of the molecule. Because of the absence of a cysteine residue in the native BF2 protein, this problem could be bypassed with the variants BF2 Y42C and S47C, the first in analogy to common mutations observed for the related protein PB1-F2 and the second as a conserved amino acid exchange. This study also shows that in all cases a careful choice of commercially available resins and reaction conditions exerts a significant impact on the efficiency of a particular synthesis.

Acknowledgements

REFERENCES

1. Lamb RA, Takeda M. Death by influenza virus protein. *Nat. Med.* 2001; **7**: 1286–1288.
2. Yewdell J, Garcia-Sastre A. Influenza virus still surprises. *Curr. Opin. Microbiol.* 2002; **5**: 414–418.
3. Capua I, Alexander DJ. Avian influenza infections in birds–a moving target. *Influenza* 2007; **1**: 11–18.
4. Webby RJ, Webster RG. Emergence of influenza A virus. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2001; **356**: 1817–1828.
5. Webster RG. Influenza virus: transmission between species and relevance to emergence of the next human pandemic. *Arch. Virol. Suppl.* 1997; **13**: 105–113.
6. Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, Basta S, O'Neill R, Schickli J, Palese P, Henklein P, Bennik JR, Yewdell JW. A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* 2001; **7**: 1306–1312.
7. Gibbs JS, Malide D, Hornung F, Bennink JR, Yewdell JW. The influenza A virus PB1-F2 protein targets the inner mitochondrial membrane via a predicted basic amphipathic helix that disrupts mitochondrial function. *J. Virol.* 2003; **77**: 7214–7224.
8. Lowy RJ. Influenza virus induction of apoptosis by intrinsic and extrinsic mechanisms. *Int. Rev. Immunol.* 2003; **22**: 425–449.
9. Zamarin D, Garcia-Sastre A, Xiao X, Wang R, Palese P. Influenza virus PB1-F2 protein induces cell death through mitochondrial ANT3 and VDAC1. *PLoS Pathog.* 2005; **1**: e4.
10. Zamarin D, Ortigoza MB, Palese P. Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice. *J. Virol.* 2006; **80**: 7976–7983.
11. Coleman JR. The PB1-F2 protein of Influenza A virus: increasing pathogenicity by disrupting alveolar macrophages. *Virol. J.* 2007; **4**: 9.
12. Conenello GM, Zamarin D, Perrone LA, Tumpey T, Palese P. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. *PLoS Pathog* 2007; **3**: 1414–1421.
13. McAuley JL, Hornung F, Boyd KL, Smith AM, McKeon R, Bennink J, Yewdell JW, McCullers JA. Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Microbe* 2007; **2**: 240–249.
14. Conenello GM, Palese P. Influenza A virus PB1-F2: a small protein with a big punch. *Cell Host Microbe* 2007; **2**: 207–209.
15. McCullers JA. Insights into the interaction between influenza virus and pneumococcus. *Clin. Microbiol. Rev.* 2006; **19**: 571–582.
16. Reid AH, Fanning TG, Hultin JV, Taubenberger JK. Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. *Proc. Natl. Acad. Sci. U.S.A.* 1999; **96**: 1651–1656.
17. Chen H, Deng G, Li Z, Tian G, Li Y, Jiao P, Zhang L, Liu Z, Webster RG, Yu K. The evolution of H5N1 influenza viruses in ducks in southern China. *Proc. Natl. Acad. Sci. U.S.A.* 2004; **101**: 10 452–10 457.
18. Henklein P, Bruns K, Nimitz M, Wray V, Tessmer U, Schubert U. Influenza A virus protein PB1-F2: synthesis and characterization of the biologically active full length protein and related peptides. *J. Pept. Sci.* 2005; **11**: 481–490.
19. Macmillan D. Strategien zur Proteinsynthese vereinigen sich mit der nativen chemischen Ligation. *Angew. Chem.* 2006; **118**: 7830–7834.
20. Dawson PE, Kent SBH. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* 2000; **69**: 923–960.
21. von Eggelkraut-Gottanka R, Klose A, Beck-Sickinger AG, Beyermann M. Peptide α-thioester formation using standard Fmoc chemistry. *Tetrahedron Lett.* 2003; **44**: 3551–3554.
22. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Synthesis of proteins by native chemical ligation. *Science* 1994; **266**: 776–779.
23. Johnson ECB, Kent SBH. Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* 2006; **128**: 6640–6646.

24. García-Martín F, Quintanar-Audelo M, García-Ramos Y, Cruz LJ, Gravel C, Furic R, Côté S, Tulla-Puche J, Albericio F. ChemMatrix, a poly(ethylene glycol)-based support for the solid phase synthesis of complex peptides. *J. Comb. Chem.* 2006; **8**: 213–220.

25. Hackeng TM, Griffin JH, Dawson PE. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U.S.A.* 1999; **96**: 10068–10073.

26. Bruns K, Studtrucker N, Sharma A, Fossen T, Mitzner D, Eissmann A, Tessmer U, Röder R, Henklein P, Wray V, Schubert U. Structural characterization and oligomerization of PB1-F2, a proapoptotic influenza A virus protein. *J. Biol. Chem.* 2007; **282**: 353–363.