Total chemical synthesis of deglycosylated human erythropoietin

Nicola Robertson and Robert Ramage *

ALBACHEM Limited and The Edinburgh Centre for Protein Technology, The University of Edinburgh, Joseph Black Building, King's Buildings, West Mains Road, Edinburgh, UK EH9 3JJ

Received (in Glasgow) 15th December 1998, Accepted 18th February 1999

For the first time the total chemical synthesis of deglycosylated human erythropoietin is described. Synthesis of this 166 amino acid protein, molecular weight 18 389 daltons, was achieved by automated synthesis on the solid phase by stepwise elongation with fluoren-9-ylmethoxycarbonyl amino acids. A novel coupling reagent ethyl 1-hydroxy-(1*H*)-1,2,3-triazole-4-carboxylate in conjunction with diisopropylcarbodiimide was employed for carboxy activation and tetrabenzo[*a,c,g,i*]fluoren-17-ylmethoxycarbonyl, a hydrophobic, chromatographic probe which has an affinity for charcoal, aided the purification immensely.

Introduction

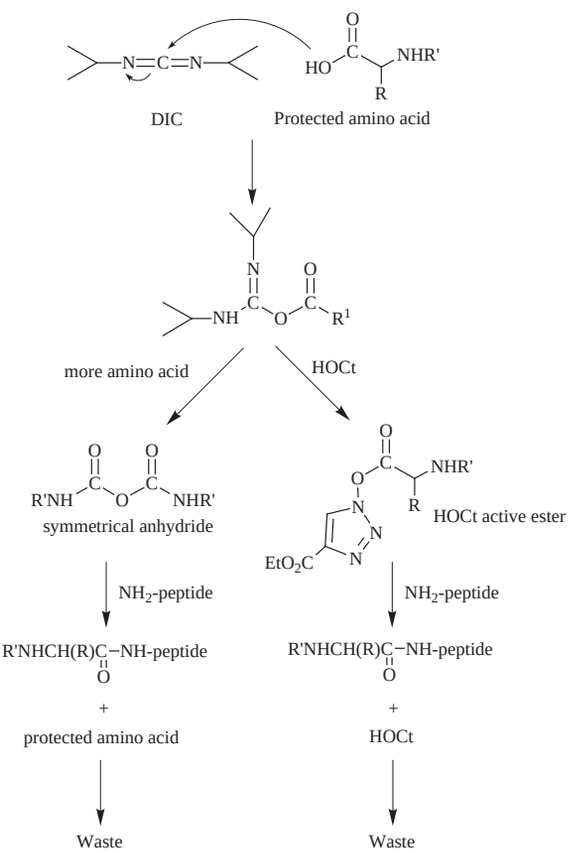
Human erythropoietin (hEPO), a glycoprotein, is the principal hormone involved in promoting red blood cell production in the bone marrow,[1,2] however, when there is progressive destruction of kidney mass, such as in chronic renal failure, an anaemia results due to a decrease in the production of hEPO[3] thus patients suffering from such a condition require hEPO to survive. The glycosylated recombinant protein has therapeutic value and is licensed for clinical use. We therefore chose the deglycosylated form of hEPO (dhEPO) as our protein model to test the methodology we had developed to synthesise large polypeptides *i.e.* 166 amino acids (a.a), molecular weight (M_r) = 18 389.

For the successful stepwise synthesis of long peptide chains (such as dhEPO) it is important that each amino acid be added with close to 100% efficiency. Key to the synthesis of dhEPO was the recently reported 1-hydroxy-(1*H*)-1,2,3-triazole-4-carboxylate (HOCt) which, when used in conjunction with diisopropylcarbodiimide (DIC), has proven itself to be superior to both the symmetrical anhydride and *N*-hydroxybenzotriazole (HOBT)/DIC couplings.[4,5] In fact the only total chemical synthesis of a fully characterised polypeptide in this range was reported by Kent *et al.* in 1986[6] who synthesised Interleukin-3 (140 a.a) utilising *tert*-butoxycarbonyl (Boc) solid phase peptide synthesis (SPPS) and double coupling of each amino acid *via* the symmetrical anhydride. Standard couplings with HOCt/DIC normally requires single coupling cycles and no expensive protected amino acid is effectively 'wasted' as the activating group (Scheme 1).[4,5]

In addition to ensuring effective coupling of each amino acid another major obstacle in the stepwise synthesis of large polypeptides is the separation of the target protein from the truncated sequences. This can often require protracted purification in order to isolate the target protein in pure form,[7] and this in turn can result in a low yield of the isolated product. Ramage *et al.* have developed a base-labile N^{α}-protecting group, tetrabenzo[*a,c,g,i*]fluoren-17-ylmethoxycarbonyl (Tbfmoc), for both affinity purification of polypeptides on porous graphitised carbon (PGC)[8] and also as a hydrophobic chromatographic probe to simplify the purification of peptides by RP-HPLC or gel filtration (Scheme 2).[9] Thus Tbfmoc methodology was applied to the synthetic dhEPO in order to aid the purification.

Results and discussion

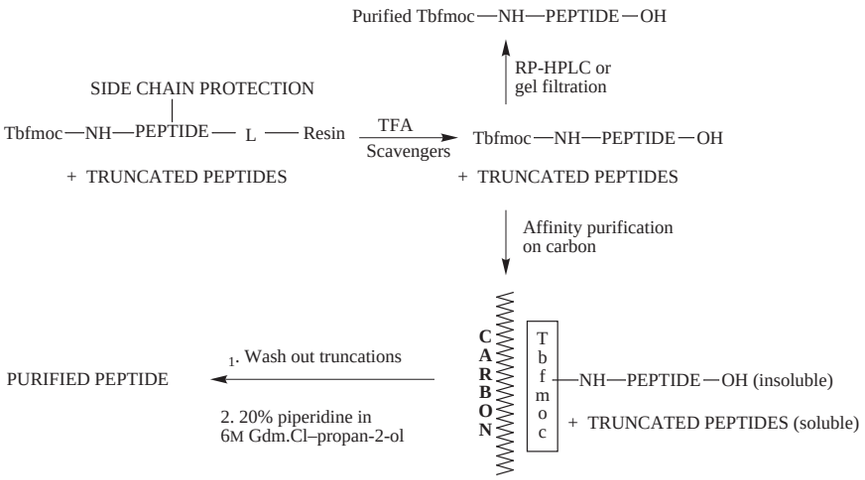
In this paper we show that automated chemical synthesis utilising the Fmoc solid phase synthetic strategy, combined with the appropriate methodology, can be a practical approach towards the production of large polypeptides.



Scheme 1

Chemical synthesis and cleavage

The primary structure of dhEPO deduced from the cDNA nucleotide sequence[10] is shown in Fig. 1. The protected peptide was assembled by solid phase synthesis[11] on a cross-linked polystyrene support, coupled to the 4-alkoxybenzyl alcohol (Wang) linker,[12] starting from the carboxy terminal arginine residue and adding amino acids in a stepwise fashion until the 166-residue chain had been assembled. The synthesis was performed on a fully automated peptide synthesiser (Applied Biosystems 430A) which had been modified for fluoren-9-ylmethoxycarbonyl[13,14] (Fmoc) SPPS and fitted with a monitor-



Scheme 2

```
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys
1               5                   10                  15                  20

Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr
21              25                  30                  35                  40

Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala
41              45                  50                  55                  60

Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
61              65                  70                  75                  80

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
81              85                  90                  95                  100

Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
101             105                 110                 115                 120

Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys
121             125                 130                 135                 140

Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
141             145                 150                 155                 160

Cys Arg Thr Gly Asp Arg
161             165
```

Fig. 1 The amino acid sequence predicted for the dhEPO by translation of nucleotide sequence from the cDNA clone λHEPOFL13.[6]

ing system to estimate the efficiency of each coupling step (Fig. 2).[15] The amino acids were coupled as their highly reactive HOCt esters[4,5] automatically formed *in situ* prior to use on the ABI 430A peptide synthesiser. An initial substitution of 0.1 mmol g^{-1} resin and 1 mmol of reactive amino acid per coupling was used in order to achieve high concentrations of reactants. Each coupling reaction lasted 30 minutes.

In preliminary experiments we found a decrease in coupling efficiency over the last 20 residues, as indicated by the Fmoc deprotection protocol,[15] (Fig. 2) thus double coupling cycles from Lys 20 to Ala 1 were adopted. Fmoc amino acids were used with appropriate side-chain protecting groups stable to the conditions of chain assembly but labile to acid. All the amino acids, reagents and solvents used were of high purity to minimise side reactions.

After assembly, Tbfmoc *via* its chloroformate,[8,9] was introduced to the free N-terminus of the resin bound peptide. The protecting groups and the peptide-resin anchoring bond were cleaved by acidolysis using trifluoroacetic acid (TFA) in the presence of a scavenger cocktail containing ethane-1,2-dithiol (EDT), thioanisole, water and phenol. The crude Tbfmoc-labelled protein was isolated, after removal of the resin by filtration, followed by rapid concentration of the filtrate *in vacuo*, ether precipitation and lyophilisation which afforded 914 mg of crude Tbfmoc-dhEPO from 2 g of resin bound product.

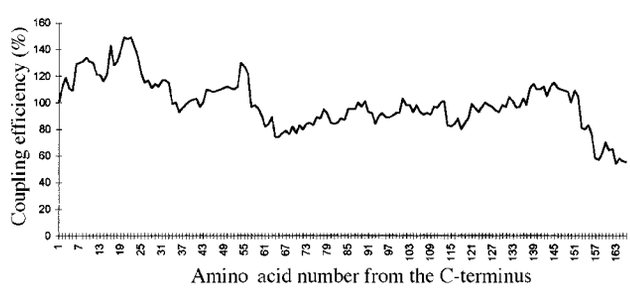


Fig. 2 The deprotection profile for dhEPO calculated from the area under the peak of the UV absorbance of a solution of the Fmoc–piperidine adduct in DMF at 302 nm.[15]

Purification

The peptide was kept in a fully denatured and reduced state throughout each purification procedure in order to avoid aggregation, incorrect formation of the disulfides or polymerisation.

The crude cleaved Tbfmoc-dhEPO obtained was dissolved in 6 M Gdm·Cl–propan-2-ol (1:1) and analysed by RP-HPLC. The trace obtained (Fig. 3) outlined a number of peaks but confirmed that Tbfmoc had been successfully incorporated onto the protein as an absorbance was visible at $\lambda = 214$ nm and $\lambda = 364$ nm which is characteristic of Tbfmoc labelled peptides. The protein solution was vortexed thoroughly with freshly

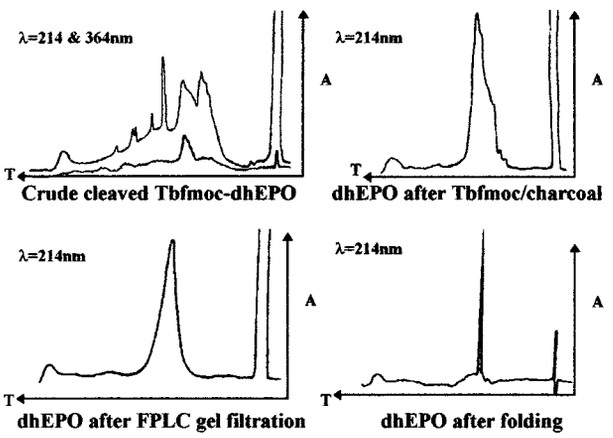


Fig. 3 HPLC (Aquapore RP300 C_4, 4.6 × 100 mm, 7 μm with a flow rate of 1 ml min^{-1} using a 2 ml loop with a linear gradient of 30–90% CH_3CN–H_2O over 20 min. The traces illustrate the purification and folding of synthetic dhEPO.

washed charcoal to which the Tbfmoc-polypeptide should adsorb. Adsorption was monitored by RP-HPLC ($\lambda = 214$, 364 nm) and was assumed to be complete after 30 minutes as no peak was observed at 364 nm after injection of the supernatant from centrifuging the charcoal mix *i.e.* no Tbfmoc-dhEPO remained in solution. The charcoal was then washed with 6 M Gdm·Cl–propan-2-ol (1 : 1) until no absorbance at $\lambda = 214$ nm was observed for the supernatant on RP-HPLC. This indicated that all peptide impurities, resulting from deletion sequences formed during synthesis, had been removed successfully. The protein was then released from Tbfmoc by vortexing the charcoal mixture in 10% piperidine/6 M Gdm·Cl–propan-2-ol (1 : 1), followed by centrifugation then repeating the process. The charcoal was then washed once with 6 M Gdm·Cl–propan-2-ol (1 : 1) and centrifuged to a pellet. The supernatants from these were combined and reduced *in vacuo* to remove the propan-2-ol. The pH of the protein solution was adjusted to pH 8 with acetic acid and stood at 37 °C after addition of dithiothreitol (DTT, 1,4-dimercaptobutane-2,3-diol) to reduce the cysteine side chains. The solution was analysed by RP-HPLC to validate the purity of the protein at this stage. The trace (Fig. 3) shows that the procedure worked extremely well and only a few minor impurities remained (observed as shoulders on the main peak). Scheme 2 outlines the Tbfmoc purification process.

The protein was further purified by applying the above solution to a Superdex™ 75 FPLC gel filtration column and eluting with 6 M Gdm·Cl. The fractions collected were analysed by RP-HPLC and those which contained the resolved protein were pooled and concentrated by ultra filtration (Fig. 3). Fractions containing any unresolved material were recombined, concentrated by ultra filtration and reapplied to the column. The protein content of the final solution (40 ml) was determined by UV analysis[16] to give an approximate yield of 12.6 mg (14% with respect to crude cleaved peptide) of purified dhEPO. Protein folding and oxidation was accomplished by first fully denaturing the above protein solution by buffering to pH 8 with Tris and then adding DTT (0.1 M). The mixture was incubated at 37 °C for 2 h. The fully denatured dhEPO solution (0.32 mg ml^{-1}) was then diluted to 500 ml with 3 M Gdm·Cl containing folding buffer[17] (Tris, 50 mM, pH 8, 5 °C, 40 μM $CuSO_4$, 0.1% *N*-lauroylsarcosine). The resulting solution (0.025 mg ml^{-1}) was stirred overnight at 4 °C. A precipitate formed and was removed by centrifuging the mixture and the pellet obtained was denatured (as above) and stored at 0 °C. The supernatant was dialysed overnight against 1 M Gdm·Cl containing folding buffer and a precipitate formed again. This was treated as before and the supernatant was dialysed against folding buffer only; no protein precipitated this time. The progress of protein folding was monitored by RP-HPLC and when complete a sharp peak was observed in place of the broad hump which is indicative of folded protein (Fig. 3).

Table 1 Amino acid analysis results for synthetic dhEPO taken at each stage of purification[a]

Amino acid	Results found				Results expected
	Resin	Crude	Tbfmoc	FPLC	
Asp/Asn	12.7	12.7	12.9	12.3	12
Thr	10.4	11.2	9.1	10.2	11
Ser	8.1	10.7	8.5	8.5	10
Gln/Glu	20.1	18.8	25.2	20.5	19
Pro	9.4	7.3	6.2	8.3	8
Gly	11.5	13.5	12.8	10.4	9
Ala	20.3	20.1	19.4	19.2	19
Cys	3.6	3.9	14.8Cys/Val	0.1	4
Val	10.4	10.6	not resolved	12.2	11
Met	0.97	1.0	1.5	1.2	1
Ile	3.7	5.0	4.9	4.4	5
Leu	21.7	22.5	21.6	22.0	23
Tyr	4.4	4.8	3.1	4.6	4
Phe	3.8	4.1	4.2	4.1	4
His	1.8	1.5	1.5	2.7	2
Lys	8.71	8.8	8.7	8.7	8
Arg	12.92	12.9	9.7	12.9	13
Trp	—	—	—	—	3

[a] Obtained by hydrolysing 1 mg of protein in 1 ml of 6 M HCl at 110 °C. After removal of the acid the residue was taken into trisodium citrate buffer pH 2.2 and applied to an 8 μm polystyrene sulfonated ion exchange resin. Elution is with trisodium citrate buffer (pH 3.2) for 6.45 min at 52 °C then trisodium citrate buffer (pH 4.25) for 40 min at 6 °C.

Chemical characterisation of synthetic dhEPO

The synthetic product was characterised by a number of highly discriminating, complementary chemical and analytical techniques. First the fidelity of the chain assembly was assessed by passing an aliquot of the piperidine solution, after Fmoc deprotection, through a UV detector set at 302 nm. This corresponds to the absorbance of the released fulvene–piperidine adduct and, due to the strong correlation between deprotection and coupling, the coupling efficiency of each amino acid was calculated.[15] Thus this method quantifies the cumulative efficiency of both the removal of the N^{α}-protecting group and the extent of the amino acid coupling at each cycle of the synthesis. The deprotection profile obtained (Fig. 2) shows coupling efficiencies of greater than 100% for the C-terminal amino acids, this is due to variable resin swelling which indicates that the concentration of the deprotection solution varies with the amount of solvent absorbed by the resin. Overall the profile shows that the synthesis proceeded with approximately 100% coupling efficiency until the last 20 residues where there was a drop, despite the double coupling cycles. The first occurred at Arg 14 to 80% then at Arg 10 to 56%, however when the number of chemical steps involved (approximately 333) is taken into consideration then this is a very good result overall.

After cleavage and deprotection, the synthetic Tbfmoc-dhEPO was first characterised by amino acid analysis. In fact amino acid analysis was carried out at each stage of purification (Table 1) all of which were in good agreement with the expected composition for synthetic dhEPO.

Following Tbfmoc–charcoal purification the molecular weight (M_r) of the peptide was determined by calibrating the FPLC gel filtration column with known M_r standards. A graph was constructed by plotting the elution volume parameter (k_{AV}) for each of the standards *versus* the logarithm of their M_r.[18] The M_r of the synthetic protein was then determined from the graph by cross reference of its calculated k_{AV} value with log M_r. The log M_r found from the graph (Fig. 4), 4.265

Table 2 MALDI TOF MS data for tryptic fragments of synthetic dhEPO[a]

Tryptic fragment	Mass found	Mass expected	Deviation
3–52	5750.64	5652.5	−1.86
96–152	5961.06	5962.88	−1.82
98–140	4439.09	4437.11	+1.98
104–150	5211.28(Na^+K^+)	5212.66	−1.38
122–162	4682.48(Na^+K^+)	4681.35	+1.13

[a] Obtained by incubation of dhEPO in 3 M Gdm·Cl, pH 8 Tris along with trypsin (5% w/v) for 16 h at 37 °C.

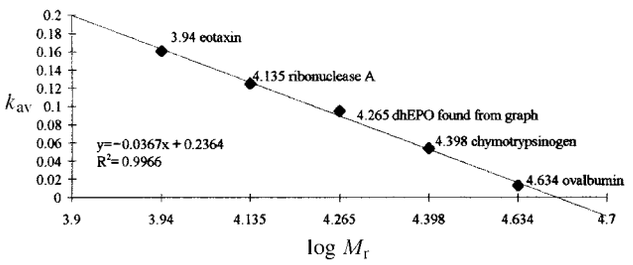


Fig. 4 Plot of k_{AV} *versus* log M_r to determine the M_r of synthetic dhEPO from FPLC performed on Superdex™ 75 (V_t = 24 ml, V_0 = 8.86 ml). k_{AV} calculated for dhEPO = 0.095 (V_e = 10.3 ml), thus the log M_r = 4.265 determined from the graph corresponds to a M_r = 18 407 which is in good agreement with the expected M_r = 18 389.

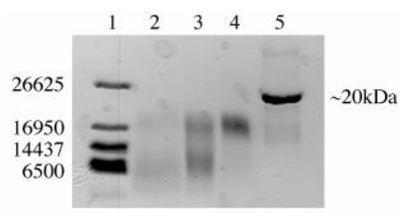


Fig. 5 SDS PAGE (15% gel, Commassie stained). Lane: 1 = M_r standards, 2 = crude dhEPO, 3 = dhEPO after Tbfmoc–charcoal, 4 = dhEPO after FPLC, 5 = His_{10}-rdhEPO.

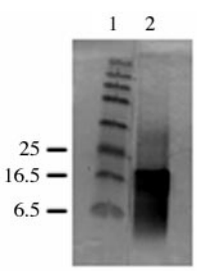


Fig. 6 Western blot Lane 1 = M_r standards (KDa). Lane 2 = dhEPO after FPLC. The anti-EPO antibody bound to the protein.

(M_r = 18407), was in good agreement with the expected M_r of 18 389 for dhEPO.

After FPLC gel filtration the peptide was further analysed by mass spectrometry. Both electrospray and matrix-assisted laser desorption–ionisation (MALDI) time of flight mass spectrometry (TOF MS) were unable to detect any mass for the intact dhEPO sequence. Thus the protein was hydrolysed with the enzyme trypsin and the non-isolated fragments were identified by MALDI TOF MS. Every mass could be accounted for as a tryptic cleavage product of the target sequence (Table 2).

A monoclonal mouse anti-hEPO antibody exists which is specific for the first 26 amino terminal residues of hEPO. Western blotting,[19] using this antibody, was carried out on a nitrocellulose membrane of the pure protein following SDS PAGE[20] (Fig. 5). Fig. 6 shows that the monoclonal anti-EPO antibody successfully bound to the protein. The amino terminus was further confirmed by automated Edman degradation[21] which also gave the expected NH_2-terminal sequence of amino acids from a total of 19 rounds of Edman degradation.

Summary

There have been a number of chemical syntheses of peptides reported in the literature[6,22–25] but no group has reported a synthesis of a polypeptide as great as 166 a.a in length with a molecular weight of 18 389 Da by stepwise solid phase synthesis. This linear synthesis of dhEPO which combined the Merrifield solid phase[11] technique and our contemporary methodology[4,5,7–9] has demonstrated that previously unattainable large polypeptide systems are now amenable to synthesis, purification and characterisation.

Experimental

General

Sonication was carried out in a Decon FS300b sonic bath. UV spectra were recorded on a Varian Carey 210 double beam spectrophotometer in the solvents described in the text. Peptide synthesis grade dimethylformamide (DMF), 1,4-dioxane, piperidine and far UV grade acetonitrile (CH_3CN) were obtained from Rathburn Chemicals, Walkerburn, Scotland. Peptide synthesis grade TFA was purchased from Applied Biosystems (ABI). Acetic anhydride and EDT were from Fluka. Dimethylaminopyridine (DMAP), DIC, DTT and thioanisole were purchased from Aldrich. Phenol was from Sigma. Guanidium hydrochloride (Gdm·Cl) and diisopropylethylamine (DIEA) were purchased from Fluka and Applied Biosystems, respectively. Dialysis was carried out using Spectra/Por 6 membrane tubing (diameter 45 mm, MWCO 10 000) and was supplied by Pierce and Warriner. Solutions of protein were concentrated using an ultra filtration cell and a YM10 membrane (MWCO 10 000) both purchased from Amicon. SDS PAGE was performed in a BIO-RAD Mini-PROTEIN® II Electrophoresis Cell and molecular weight markers were also from BIO-RAD. Western blotting was performed in a BIO-RAD Trans-Blot Cell on Hybond™ ECL™ nitrocellulose membrane supplied by Amersham Life Science. The pre-stained protein markers were from New England BioLabs. Monoclonal mouse anti-human erythropoietin antibody (primary antibody), AE7A5, was obtained from Genzyme Diagnostics, West Malling, UK. Peroxidase-conjugated rabbit IgG to mouse IgG (secondary antibody) was supplied by Dako, High Wycombe, UK. MALDI TOF MS were recorded on a PerSeptive Biosystems Voyager™ Biospectrometry™ Workstation from Vestec mass spectrometry products. Analytical and preparative gel filtrations were carried out on a Pharmacia LKB p500 FPLC system using Superdex™ 75 HR 10/30 or 26/60 respectively. Protein sequencing was performed on an ABI 477A protein sequencer at the Wellmet sequencing facility, University of Edinburgh. HPLC grade TFA was purchased from Fisons and all water used was Milli-Q grade.

High performance liquid chromatography (HPLC)

HPLC was carried out using an ABI system comprising 2 × 1406A solvent delivery systems, a 1480 injector/mixer and a 1783A detector/controller. Protein was eluted from an ABI, Aquapore RP300 C_4, 4.6 × 100 mm, 7 μm with a flow rate of 1 ml min^{-1} using a 2 ml loop. A linear gradient of CH_3CN–H_2O (30–90%) containing 0.1% v/v HPLC grade TFA, over 20 minutes, was used.

Amino acid analysis

Amino acid analyses were performed on an LKB 4150 alpha amino acid analyser on the hydrolysate obtained from heating samples of protein in 6 M HCl at 110 °C for 48 h in sealed Carius tubes. Following evaporation to dryness the remaining

residue was taken into trisodium citrate buffer (pH 2.2) and applied to an 8 micron polystyrene sulfonated ion exchange resin. Elution was achieved with trisodium citrate buffer (pH 3.2) for 6.45 min at 52 °C then trisodium citrate buffer (pH 4.25) for 40 min at 6 °C.

Solid phase synthesis

Peptide synthesis was performed on an ABI 430A automated peptide synthesiser with on line UV monitoring using an ABI 758A detector. The Fmoc strategy of N^{α} protection with the complementary use of acid labile side-chain protection and the acid labile Wang linker on polystyrene resin were used. All were purchased from Bachem or Novabiochem. The side-chain protecting groups were as follows: *tert*-butyl ('Bu) ethers for Ser, Thr and Tyr; 'Bu esters for Asp and Glu; τ-triphenylmethyl (Trt) for His; 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) for Arg. The carboxamide side chains of Asn and Gln were protected with the Trt group. Cys was incorporated as the Trt derivative and the free amine on the side chain of Lys was protected with the Boc group.

Resin loading

Fmoc-Arg(Pmc) (3 equivalents) was dissolved in DMF (10 ml) and to this DIC (1.5 equivalents) was added. The solution was sonicated for 15 minutes to allow formation of the mixed anhydride before adding it to Wang resin (1 g, 0.65 mmol g^{-1}), pre-swollen in a minimum amount of DMF, along with a catalytic amount of DMAP. The mixture was allowed to stand for 1.5 h at room temperature. The resin was then filtered and washed with copious volumes of DMF, DCM, then ether and dried *in vacuo*.

Fmoc loading test

Dry Fmoc-Arg(Pmc)-resin (7–10 mg), was placed in a 10 ml volumetric flask and 20% piperidine–DMF was added to the volumetric level. The mixture was sonicated at room temperature for 15 minutes and the UV absorbance for the released fulvene–piperidine adduct at 302 nm was recorded. The resin functionality was calculated using the Beer–Lambert law (ε_{302} = 15400 for the fulvene–piperidine adduct) and was found to be 0.1 mmol g^{-1}.

Automated solid phase peptide synthesis

Synthetic procedures were pre-programmed into the ABI 430A synthesiser prior to synthesis. Each synthetic cycle, resulting in the coupling of a single amino acid, involved: *Capping*, to block any unreacted amino groups; *Deprotection*, to remove the Fmoc group from the N^{α} position; *Coupling* of the next N^{α}-Fmoc-amino acid. Each step was followed by thorough washing of the resin and each cycle was repeated with the required amino acid in order to build the desired polypeptide sequence. The pre-programmed synthetic cycles are summarised below.

Capping

The peptide–resin was vortexed with a solution of acetic anhydride (0.5 M), DIEA (0.125 M) and HOBT (0.2% w/v) in DMF–1,4-dioxane (1 : 1, 10 ml) for 10 minutes. The capping solution was drained from the reaction vessel and the resin was washed with 6 portions of DMF–1,4-dioxane (1 : 1).

Deprotection

The resin was vortexed with a solution of 20% piperidine in DMF–1,4-dioxane (1 : 1, 10 ml) for 6 minutes then drained and washed with DMF (4 times). The deprotection cycle was repeated twice but vortexing lasted 1.5 minutes on the second deprotection. Finally the resin was washed with 6 portions of DMF–1,4-dioxane (1 : 1) (10 ml).

The percentage of amino acid successfully coupled in the previous cycle was determined by the absorbance of the fulvene–piperidine adduct at 302 nm from the first deprotection. An aliquot of the deprotection filtrate was passed through an UV detector with on-line integration. Comparison of the area of each successive deprotection peak with that of the first amino acid allowed an estimate of the percentage coupling of each residue.[13]

Coupling

Residues were incorporated as their HOCt activated esters[4,5] which were pre-formed in an 'activation vessel' before being transferred to the 'reaction vessel' containing the starting resin (1 g, 0.1 mmol g^{-1}). In the case of histidine, hydroxybenzotriazole (HOBt, 1 mmol) was placed in the cartridge with the protected histidine prior to synthesis to avoid racemisation. The HOCt active esters were pre-formed by adding a solution of HOCt in DMF (0.25 M, 4 ml) to dry Fmoc-amino acid (1 mmol) followed by DIC (0.25 M, 4 ml) and allowed to activate for 5 min. All activated amino acids were single coupled with the exception of the last 20 residues, Lys 20 to Ala 1, which were double coupled.

Approximately half the resin was removed after coupling Ala 114 and the remaining resin was used to continue the synthesis until Val 46 when the resin was split again. The sequence was continued to the end, simultaneously, from these two portions. On completion of the synthesis the resins from both syntheses were combined, filtered, washed with DMF, DCM then ether and dried *in vacuo* to an off white solid.

Yield 2.2 g. Amino acid analysis: (found)expected; Asp/Asn(12.7)12, Thr(10.4)11, Ser(8.07)10, Gln/Glu(20.1)19, Pro(9.4)8, Gly(11.5)9, Ala(20.3)19, Cys(3.6)4, Val(10.4)11, Met(1.0)1, Ile(3.7)5, Leu(21.7)23, Tyr(4.4)4, Phe(3.8)4, His(1.8)2, Lys(8.7)8, Arg(12.9)13.

Tbfmoc loading

The protected resin bound protein (2 g) was given an additional capping by sonicating in DCM containing acetic anhydride–HOBt–DIEA. The resin was then filtered, washed with DCM, DMF, 1,4-dioxane then ether and dried. The dried resin was sonicated in 20% piperidine–DMF for 15 minutes to remove the N-terminal Fmoc group and was filtered, washed with DMF, DCM, then ether and dried.

Tbfmoc chloroformate (~3 equivalents) was added to the dry resin and enough DCM was added to swell the resin and dissolve the Tbfmoc. DIEA (1 equivalent) was then added and the resulting mixture was sonicated at room temperature, in the absence of light, for 3 h after which time the resin was filtered, washed with copious DCM then ether and dried *in vacuo*.

Cleavage of Tbfmoc-dhEPO from Wang resin

Resin bound Tbfmoc-dhEPO (2 g) was stirred in EDT (4 ml), thioanisole (1 ml), H_2O (1 ml) and phenol (1.5 g) for 10 minutes. TFA (20 ml) was added and the mixture was stirred for 4.5 h under dry nitrogen in the absence of light. The resin was filtered, washed with TFA (4 ml) and reduced *in vacuo* to a yellow oil. Ether was added to the residue and the precipitated peptide was isolated by filtration through a fine sinter and washed with ether. The precipitated protein was dissolved in 20% AcOH–H_2O and lyophilised, in the absence of light, to a white fluffy solid.

Yield 914 mg, 0.05 mmol. HPLC: λ_1 = 214 nm, λ_2 = 364 nm, R_t = 12 and 13 min, 60% and 62% CH_3CN–H_2O respectively. Amino acid analysis: (found)expected; Asp/Asn(12.7)12, Thr(11.2)11, Ser(10.7)10, Gln/Glu(18.8)19, Pro(7.3)8, Gly(13.5)9, Ala(20.1)19, Cys(3.9)4, Val(10.6)11, Met(1.0)1, Ile(5.0)5, Leu(22.5)23, Tyr(4.8)4, Phe(4.1)4, His(1.5)2, Lys(8.8)8, Arg(12.9)13.

Affinity purification on charcoal

Charcoal (5 g) was washed with 10% piperidine/6 M Gdm·Cl–propan-2-ol, 1 : 1, (100 ml) followed by 6 M Gdm·Cl–propan-2-ol, 1 : 1, (9 × 100 ml). Crude Tbfmoc-dhEPO (900 mg, 50 μmol) was dissolved in 6 M Gdm·Cl–propan-2-ol, 1 : 1, (60 ml) and added to the freshly washed charcoal (5 g). The mixture was vortexed for 10 minutes then centrifuged to a pellet. The supernatant was decanted and analysed by HPLC $\lambda_1 = 214$ nm, $\lambda_2 = 364$ nm. No absorbance was visible at 364 nm; therefore it was assumed that all the Tbfmoc material had been adsorbed onto the charcoal. The charcoal was washed with 6 M Gdm·Cl–propan-2-ol 1 : 1 until no absorbance at 214 nm was visible by HPLC. The protein was cleaved from Tbfmoc by vortexing the charcoal with 10% piperidine–6 M Gdm·Cl–propan-2-ol 1 : 1 (15 ml) for 10 minutes. The mixture was centrifuged to a pellet and the supernatant removed. This process was repeated once and the combined supernatants were reduced *in vacuo* to remove propan-2-ol. The mixture was then acidified to pH 4 with AcOH, and DTT (0.1 M) was added. The solution was incubated at 37 °C overnight. The solution (200 ml) was dialysed against 6 M Gdm·Cl for 2 days, changing the dialysis solution twice daily. The solution was removed from the dialysis tubing and concentrated by ultra filtration to a volume of 40 ml.

HPLC $\lambda_1 = 214$ nm, $\lambda_2 = 280$ nm, $R_t = 12$ min, 60% CH_3CN–H_2O. Amino acid analysis: (found)expected; Asp/Asn(12.9)12, Thr(9.1)11, Ser(8.5)10, Gln/Glu(25.2)19, Pro(6.2)8, Gly(12.8)9, Ala(19.4)19, unresolved Cys/Val(14.8)15, Met(1.5)1, Ile(4.9)5, Leu(21.6)23, Tyr(3.1)4, Phe(4.2)4, His(1.5)2, Lys(8.7)8, Arg(9.6)13.

FPLC molecular weight determination

Analysis was carried out using a Superdex™ 75 HR 10/30 column which was pre-equilibrated with 6 M Gdm·Cl using a flow rate of 0.5 ml min^{-1}. The column bed volume (V_t) was 24 ml and the column void volume (V_o) was 8.86 ml which was determined from elution of Blue dextrin ($M_r = 2\,000\,000$) from the column with 6 M Gdm·Cl. An aliquot of protein solution from above was loaded onto the column and the elution volume (V_e) was recorded (10.3 ml) by monitoring at $\lambda = 280$ nm. Molecular weight standards were also eluted from the column; these were ovalbumin ($M_r = 43\,000$, $V_e = 9.07$ ml), chymotrypsinogen A ($M_r = 25\,000$, $V_e = 9.94$ ml), ribonuclease A ($M_r = 13\,700$, $V_e = 10.88$ ml) and eotaxin ($M_r = 8\,704$, $V_e = 11.30$ ml).

Preparative FPLC gel filtration

Gel filtration was carried out using a Superdex™ 75 HR 26/60 column. Prior to protein application the column was pre-equilibrated with 6 M Gdm·Cl over a column volume of 335 ml using a flow rate of 2 ml min^{-1}. The solution obtained after affinity purification on charcoal was applied in 5 ml aliquots to the column and a flow rate of 2 ml min^{-1} applied. The protein was eluted from the column in 6 M Gdm·Cl and fractions (4 ml) were collected after 50 ml was allowed to pass through the column. The progress was monitored at 280 nm and fractions absorbing at 280 nm were analysed by HPLC $\lambda_1 = 214$ nm, $\lambda_2 = 280$ nm. Fractions 20–25 eluted with $R_t = 11$ min and were pure. Fractions 26–35 eluted with the same R_t but were impure and therefore were combined, concentrated by ultra filtration, and reapplied to the column. All fractions containing pure material with $R_t = 11$ min were combined and concentrated by ultra filtration to a volume of 40 ml.

Protein content by UV analysis[16] = 12.6 mg, 0.7 μmol. HPLC $\lambda_1 = 214$ nm, $\lambda_2 = 280$ nm, $R_t = 11$ min, 60% CH_3CN–H_2O. Amino acid analysis: (found)expected; Asp/Asn(12.3)12, Thr(10.2)11, Ser(8.5)10, Gln/Glu(20.2)19, Pro(8.3)8, Gly(10.4)9, Ala(19.2)19, Cys(0.1)4, Val(12.2)11, Met(1.2)1, Ile(4.4)5, Leu(22.0)23, Tyr(4.6)4, Phe(4.1)4, His(2.7)2, Lys(8.7)8, Arg(12.9)13.

Oxidation and folding

The pH of the above solution (40 ml) was adjusted to pH 8 by addition of Tris, and DTT (0.1 M) was added. The mixture was incubated at 37 °C for 2 h. The fully denatured dhEPO solution (0.32 mg ml^{-1}) was diluted to 500 ml with 3 M Gdm·Cl containing folding buffer[17] (Tris, 50 mM, pH 8, 5 °C, 40 μM $CuSO_4$, 0.1% *N*-lauroylsarcosine) and the resulting protein solution (0.025 mg ml^{-1}) was stirred overnight. A precipitate formed and was removed by centrifugation, denatured (as above) and stored in the freezer. The supernatant was dialysed overnight against 1 M Gdm·Cl containing folding buffer. Again a precipitate formed and this was removed and treated as above. The supernatant was then dialysed against folding buffer only to give the folded protein.

Protein content by UV analysis[16] = 1 mg per 100 ml. HPLC $\lambda_1 = 214$ nm, $\lambda_2 = 280$ nm, $R_t = 11$ minutes, 60% CH_3CN–H_2O.

Tryptic digest

Denatured protein solution (3 ml containing 0.2 mg by UV analysis[23]) was dialysed against Tris (100 ml, 50 mM, pH 8 at 37 °C) in 3 M Gdm·Cl. Trypsin (5% w/v) was added and the mixture was incubated at 37 °C for 8 h before a second aliquot of trypsin (5% w/v) was added. The resulting mixture was incubated at 37 °C for a further 16 h.

HPLC (ABI Aquapore RP300 C_{18}, 220 × 4.6 mm, 7 μ, 10–90% CH_3CN–H_2O over 30 minutes) $\lambda = 214$ nm, $R_t = 10$–22 minutes, 24–60% CH_3CN–H_2O. MS (MALDI TOF) m/z = 4437.11 (MH^+, fragment 98–140), 4681.35 (MH^+, fragment 122–162, K^+ Na^+ salt), 5211.28 (MH^+, fragment 104–150, K^+ Na^+ salt), 5750.64 (MH^+, fragment 3–52) and 5961.06 (MH^+, fragment 96–152).

Sample preparation for protein sequencing

Denatured protein solution (3 ml containing approximately 0.2 mg by UV analysis[16]) was dialysed against H_2O (100 ml) for 2 weeks changing the dialysis solution twice daily. Some protein precipitated from solution and was dissolved in 100% CH_3CN then diluted with H_2O, combined with soluble protein from the dialysis and lyophilised to an off-white solid.

Yield (~0.2 mg). Edman degradation 19 rounds from the N-terminus (found)expected: cycle1 (Ala)Ala, cycle2 (Pro)Pro, cycle3 (Pro)Pro, cycle4 (Arg)Arg, cycle5 (Leu)Leu, cycle6 (Ile)Ile, cycle7 (not found)Cys, cycle8 (Asp)Asp, cycle9 (Ser)Ser, cycle10 (Arg)Arg, cycle11 (Val)Val, cycle12 (Leu)Leu, cycle13 (Glu)Glu, cycle14 (Arg)Arg, cycle15 (Tyr)Tyr, cycle16 (Leu)-Leu, cycle17 (Leu)Leu, cycle18 (Glu)Glu, cycle19 (Ala)Ala.

Western blotting

Following SDS PAGE,[20] the pure protein was transferred onto a nitrocellulose membrane by a potential difference of 40 V for 2.5 h in transfer buffer (25 mM Tris, pH 8 (HCl), 192 mM glycine, 15% (v/v) methanol). The membrane was then incubated overnight with mixing in blocking buffer (20 mM Tris, pH 8 (HCl), 37 mM NaCl, 0.1% (v/v) Tween-20, 5% (w/v) non fat milk powder) at room temperature. Hence any sites on the membrane which were not bound by protein were blocked, thus reducing potential non-specific binding of the antibody in the next step. Blocking was followed by a 1 h incubation with the primary antibody at a dilution of 1 : 2000 in blocking solution. The membrane was washed with blocking buffer (3 × 15 min) and then incubated for 1 h with the labelled secondary antibody at a dilution of 1 : 1000 in blocking buffer. The labelled secondary antibody bound to any immobilised primary antibody present and allowed its subsequent detection. After washing the membrane with blocking buffer (3 × 15 min) it was rinsed with water.

For detection the membrane was incubated in the following solution: *o*-dianisidine (10 mg), H_2O (19 ml), imidazole (0.1 M,

1 ml) and 30% H_2O_2 (0.2 ml). The membrane was incubated in the above solution until a clear strong band was observed. The solution was removed and the membrane was washed with copious water.

Acknowledgements

This paper is dedicated to R. A. Raphael, Regius Professor of Chemistry at The University of Glasgow who was an inspiring teacher, enthusiastic research supervisor and friend.

Mr Kevin Shaw for peptide synthesis, Mr Brian Whigham for MALDI TOF MS and amino acid analysis, Emmanuelle Suarez for the provision of a dhEPO sample, Professor Sabine Flitsch and Derek MacMillan for help with Western blotting analysis, Dr Andrew Cronshaw for sequencing analysis, Parke Davis, BBSRC and DTI for financial support.

References

1 E. Goldwasser, *Fed. Proc.*, 1975, **34**, 2285.
2 S. E. Graber and S. B. Krantz, *Annu. Rev. Med.*, 1978, **29**, 51.
3 A. J. Erslev, J. Wilson and J. Caro, *J. Lab. Clin. Med.*, 1987, **109**, 429.
4 L. Jiang, A. Davison, G. Tennant and R. Ramage, *Tetrahedron*, 1998, **54**, 14233.
5 N. Robertson, L. Jiang and R. Ramage, *Tetrahedron*, 1998, in press.
6 I. Clark-Lewis, R. Aebersold, H. Ziltener, J. W. Schrader, L. E. Hood and S. B. H. Kent, *Science*, 1986, **231**, 134.
7 R. Ramage, J. Green, T. W. Muir, O. M. Ogunjobi, S. Love and K. Shaw, *Biochem. J.*, 1994, **299**, 151.
8 R. Ramage and G. Raphy, *Tetrahedron Lett.*, 1992, **33**, 385.
9 A. R. Brown, S. L. Irving and R. Ramage, *Tetrahedron Lett.*, 1993, **34**, 7129.
10 K. Jacobs, C. Shoemaker, R. Rudersdorf, S. D. Neill, R. J. Kaufman, A. Mufson, J. Seehra, S. S. Jones, R. Hewick, E. F. Fritsch, M. Kawakita, T. Shmizu and T. Miyake, *Nature*, 1985, **313**, 806.
11 R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149.
12 S. S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328.
13 L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404.
14 E. Atherton, C. J. Logan and R. C. Sheppard, *J. Chem. Soc.*, *Perkin Trans. 1*, 1981, 538.
15 K. M. Otteson, R. L. Nobel, P. D. Hoeprich, Jr., K. T. Shaw and R. Ramage, *Applied Biosystems Res. News*, June 1993.
16 H. Edeloch, *Biochemistry*, 1967, **6**, 1948.
17 J.-P. Boissel, W. R. Lee, S. R. Presnell, F. E. Cohen and H. F. Bunn, *J. Biol. Chem.*, 1993, **268**, 15983.
18 M. Le Maire, A. Viel and J. Moller, *Anal. Biochem.*, 1989, **177**, 50.
19 H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 4350.
20 U. K. Laemmli, *Nature*, 1970, **227**, 680.
21 J. D. Hayes, L. A. Kerr and A. D. Cronshaw, *Biochem. J.*, 1989, **264**, 437.
22 C. H. Li, D. Yamashiro, D. Gospodarowicz, S. L. Kaplan and G. van Vliet, *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 2216.
23 W. F. Heath and R. B. Merrifield, *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 6367.
24 I. Clark-Lewis, B. Moser, A. Walz, M. Baggiolini, G. J. Scott and R. Abersold, *Biochemistry*, 1991, **30**, 3128.
25 R. F. Nutt, S. F. Brady, P. L. Darke, T. M. Ciccarone, D. Colton, E. M. Nutt, J. A. Rodkey, C. D. Bennett, L. H. Waxman, I. S. Sigal, P. S. Anderson and A. F. Veber, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 7129.

Paper 9/01371G