ARTICLE

Understanding protease catalysed solid phase peptide synthesis

Rein V. Ulijn,^a Nicola Bisek,^a Peter J. Halling^b and Sabine L. Flitsch *^a

^a Department of Chemistry, The University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, Scotland UK EH9 3JJ; Fax: +44(0)131 650 4737; Tel: +44(0)131 650 4737

^b Department of Chemistry, The University of Strathclyde, Thomas Graham Building, Cathedral Street, Glasgow, Scotland UK G1 1XL; Fax: +44(0)141 552 4822; Tel: +44(0)141 548 2683

Received 5th December 2002, Accepted 4th March 2003 First published as an Advance Article on the web 19th March 2003

A protease (thermolysin) was used to directly synthesise a number of dipeptides from soluble Fmoc-amino acids onto a solid support (PEGA₁₉₀₀) in bulk aqueous media, often in very good yields. This shift in equilibrium toward synthesis is remarkable because for soluble dipeptides in aqueous solution hydrolysis rather than synthesis is observed. Three possible reasons for the equilibrium shift were considered: (i) using a solid support makes it easy to use an excess of reagents, so mass action contributes towards synthesis; (ii) reduction in the unfavourable hydrophobic hydration of the Fmoc group within the solid support compared with the free amino acid in solution and (iii) suppression of the ionization of amino groups linked to the solid phase due to mutual electrostatic repulsion. It was found that under the conditions studied the second effect was most important.

Introduction

Applications of enzyme catalysis on substrates that are immobilised onto insoluble supports are on the increase. Examples are chemo-enzymatic synthesis on solid supports,¹ 'on-bead' screening for enzyme substrates or inhibitors in combinatorial libraries,² enzyme cleavable linkers for release of polymer bound libraries³ and enzyme cleavable protecting groups in solid phase chemistry.⁴ In the future, increased applications are expected in the area of small molecule microarrays,⁵ and solid-phase combinatorial biocatalysis.⁶

It may seem rather surprising that the fundamentals of enzyme reactions on immobilised substrates (kinetics, thermodynamics) have not been studied in detail to date, even though one might predict that these could be significantly different from those of reactions in dilute aqueous solutions. Thus, we have recently reported an example where reaction thermodynamics were altered when the enzymatic peptide synthesis was carried out on solid PEGA₁₉₀₀ support (poly(acrylamide)ethyleneglycol). We found that the amide synthesis/hydrolysis equilibrium was significantly shifted toward synthesis when compared to the reaction in dilute aqueous solution where the hydrolytic reaction predominates.7 This shift from preferred hydrolysis with dissolved amine substrate toward preferred synthesis with immobilised amine substrate allowed for efficient protease catalysed synthesis of a variety of peptides on solid support (Scheme 1).

These findings prompted us to investigate the reasons for the shift in equilibrium towards peptide synthesis when the amine substrate is immobilised. We considered three possible factors that could play a role.⁷ These were i) the effect of using an excess of soluble amino acid in solid phase synthesis. Secondly (ii), the synthesis reaction involves transfer of the protected amino acid 1 from aqueous solution into the polyethylene glycol micro-environment of the solid phase resin, where unfavourable hydrophobic hydration is expected to be reduced. This should increase the synthesis yield of solid phase dipeptide 3, especially when hydrophobic amino acid substrates are used. And finally (iii), ionisation of PEGA₁₉₀₀ immobilised amino groups (in this case the phenylalanine-amine) could be suppressed in comparison to a dissolved amino acid, due to the proximity of neighbouring amines. Suppression of ionisation is well known to shift peptide synthesis/hydrolysis equilibria



Scheme 1 Phenylalanine was coupled *via* the Wang linker to PEGA₁₉₀₀ and treated with excess N-protected amino acid in the presence of thermolysin and a phosphate buffer (i). After 16 hours, reaction products were released from the resin by TFA cleavage of the Wang linker (ii) and quantified by reverse phase HPLC. X is either Fmoc or Z protecting group. R_1 is the amino acid side chain.

as observed when organic (co-) solvents are used in peptide synthesis.

In this article we aim to quantify the relative importance of each of these three factors in order to understand what governs the equilibrium shift that leads to the high yielding peptide synthesis that was observed.

Results and discussion

What causes the equilibrium shift toward peptide synthesis when the substrate is immobilized?

All reactions were performed as shown in Scheme 1. Phenylalanine was chosen as the immobilised amino component of the reaction because of the known substrate preference of thermolysin at an amine site for hydrophobic and aromatic amino acid residues. $PEGA_{1900}$ immobilised phenylalanine was reacted overnight in the presence of thermolysin with a range of N-protected amino acids with unprotected side chains in an aqueous phosphate buffer. i. The effect of the excess soluble amino acids on the equilibrium position. A well-known method of shifting reaction equilibria is to apply larger concentrations of starting materials to push the equilibrium toward synthesis of the end-product. In particular in solid phase chemistry this method is commonly used because unreacted excess starting material can simply be washed from the solid supported product. In order to determine if excess of amino acid 1 was responsible for the high yields of synthesis observed, the solid phase reaction was compared to an equivalent reaction in aqueous solution. In this case the acylation equilibrium is compared for PEGA₁₉₀₀-Phe (2A) and free phenylalanine methyl ester (Phe-OMe) (2B, Scheme 2).



Scheme 2 A: Amide synthesis/hydrolysis on solid supported substrate; B: Amide synthesis in aqueous solution.

The equilibrium position of reaction for the solution phase reaction can be calculated reliably from the initial reactant concentrations, reactant ionisation constants, and the pH of the reaction which are known. This calculation is based on the single reference equilibrium constant that was recently identified for amide hydrolysis/synthesis reactions in aqueous medium. When all reactant concentrations in the amide equilibrium were expressed in their uncharged form, a reference value of $K_{ref}^{0} = 10^{3.6} \text{ M}^{-1}$ was consistently found, regardless of the actual molecular form of the reactants.⁸ Hence, at equilibrium we have for reaction 2B:

$$K_{\rm ref}^{0} = \frac{[X-aa_{2}-{\rm Phe-OMe}]_{\rm eq}}{[X-aa_{2}]_{\rm eq}^{0}[{\rm Phe-OMe}]_{\rm eq}^{0}} = 10^{3.6} \,\,{\rm M}^{-1}$$
(1)

The concentrations of uncharged [Phe-OMe]⁰ and [X-aa₂]⁰ follow from the simple ionisation relationships:

$$[Phe-OMe]^{0} = \frac{[Phe-OMe]^{total}}{(1+10^{pKa,amine-pH})}$$
(2)

$$[X-aa_{2}]^{0} = \frac{[X-aa_{2}]^{total}}{(1+10^{pH-pK_{uacid}})}$$
(3)

For the zwtterionic molecule Fmoc-His and the di-acid Fmoc-Asp with two pK_a values each the equations are slightly more complicated.⁸ The peptide concentration at equilibrium follows from combining eqn. (1)–(3).

$$[X-aa_{2}-Phe-OMe]_{eq}^{0} = \frac{K_{ref}^{0}[X-aa_{2}]_{eq}^{total}[Phe-OMe]_{eq}^{total}}{(1+10^{pH-pKa,acid})(1+10^{pKa,amine-pH})}$$
(4)

After adding the mass balance equation

$$[Phe-OMe]_{initial}^{total} = [X-aa_2-Phe-OMe]_{eq}^0 + [Phe-OMe]_{eq}^{total} (5)$$

we obtain for the fraction equilibrium conversion eqn. 6 below.

This equation shows that the fractional conversion of Phe-OMe to peptide is independent of its initial concentration. However, for the best direct comparison to solid phase experiments, the total initial concentration [Phe-OMe]_{initial} was chosen as equal to the amount of immobilized phenylalanine amine per unit volume present in the solid phase experiments (2 µmol per 2 ml = 1mM). The pK_a of the amine group of Phe-OMe equals 7.0.⁸ The pK_a values of the Fmoc amino acids **1a–1g** were estimated as 3.7. For **1h** they were taken as 3.0 and 8.0 for the macroscopic ionization constants and 3.6 and 7.4 for the microscopic ionization constants. For **1i** they were 3.2 for the α acid and 4.8 for the side chain. (For methods of pK_a estimation, see ref. 8). The pH of the reaction was 7.4 and the total saturated concentrations of **1a–1i** were measured (Table 1). Using eqn. (6) then resulted in the predicted conversions in aqueous solution given in Table 1. These results could be directly compared to those experimentally observed on immobilised Phe (Scheme 2A).

From Table 1 it is clear that the conversions on immobilised Phe-PEGA are in all cases much higher than those expected for the soluble Phe-OMe. These observations clearly show that the high yields of peptide synthesis obtained on immobilised Phe cannot exclusively be explained in terms of the excess of 1 present.

As expected, the amino acids that are present in smaller concentrations (poorly soluble) are predicted to give rise to the lowest yields in solution. By contrast, poor aqueous solubility of 1 leads to generally higher conversion on immobilized Phe. As seen in particular for 3c and 3e as compared to *e.g.* 3g-3i. The apparent inverse correlation between substrate solubility and yield will be discussed in the next section.

Given the finding that a large excess of amino acid was not required for successful synthesis of peptide on solid support, a systematic study relating yield of synthesis to excess of substrate was carried out.

Thus, the synthesis of peptide Fmoc-Nle-Phe (3c) was carried out with increasing ratio of equivalents of protected amino acid over supported amino acid. Fig. 1 shows that a 10-fold excess was sufficient to allow for complete conversion to peptide 3c. This compares favourably with conventional chemical peptide synthesis, where 5–10 fold excess of reagants are usually employed.



Fig. 1 Stoichiometry of thermolysin catalysed synthesis of peptide 3c on PEGA immobilized Phe at pH 8 in 0.1 M K-phosphate buffer. Note that 1 equivalent of Fmoc-Nle corresponds to a 2 mM solution of Fmoc-Nle.

ii. Removal of hydrophobic groups from aqueous solution. The micro-environment within the polymeric $PEGA_{1900}$ support is more hydrophobic than that of the bulk aqueous solution. Thus one might expect a transfer of hydrophobic enzyme substrates (such as Fmoc protected amino acids used here) from the aqueous bulk solution to the polymer support to be favourable. Hence, the synthesis reaction is expected to be more favoured with more hydrophobic enzyme substrates. We have used log *P* values to quantify the hydrophobicity of the

$$f_{\rm eq} \stackrel{\rm peptide}{=} \frac{[X-aa_2-Phe-OMe]_{\rm eq}}{[Phe-OMe]_{\rm initial}} = \frac{[X-aa_2]_{\rm eq}^{\rm total}K_{\rm ref}^0}{(1+10^{pK_{\rm asymine}-pH})(1+10^{pH-pK_{\rm aacid}}) + [X-aa_2]_{\rm eq}^{\rm total}K_{\rm ref}^0}$$
(6)

Table 1 Thermolysin catalysed peptide synthesis yields from protected amino acids in the solid phase reaction on $PEGA_{1900}$ -Phe compared to the calculated yields of the equivalent reaction in aqueous solution on free phenylalanine methyl ester (Phe-OMe). The aqueous solubility and log *P* of the protected amino acids are also given

Peptide (3)	R ₁	Acyl donor (1)	Saturated conc. acyl donor/mM	Observed conversion on solid phase $(\%)^{a,b}$	Calculated conversion in solution (%) ^c	Log P ^a
a	Н	Fmoc-Gly	40	99	2.3	1.7
b	CH ₂ -CH ₂ -(CH ₃) ₂	Fmoc-Leu	16	99	0.9	3.5
c	(CH ₂) ₃ -CH ₃	Fmoc-Nle	5	99	0.3	3.6
d	CH ₂ -C ₆ H ₅	Z-Phe	63	72	2.8	1.6
e	CH ₂ -C ₆ H ₅	Fmoc-Phe	3	99	0.1	3.5
f	CH ₂ -CH ₂ -CONH ₂	Fmoc-Gln	7	84	0.4	0.3
g	CH ₂ -OH	Fmoc-Ser	36	10	2.0	1.0
ĥ	CH_{2}^{-} (imidazole) ⁺	Fmoc-His	24	77	0.3	1.2
i	CH ₂ -COO ⁻	Fmoc-Asp	25	70	0.4	0.7

^{*a*} Conversions (%) determined by HPLC after acid cleavage of the Wang linker. ^{*b*} The conversions were always measured after at least 16 h reaction time. In several cases results were compared to 80 hours and no significant differences were found. Hence, it is reasonable to assume that reactions had approached thermodynamic equilibrium. ^{*c*} If this value exceeds the peptide solubility limit of the product, precipitation of the peptide would be observed leading to overall higher conversions. ^{*d*} Calculated using the *c* log *P* calculator on www. Daylight.com

protected amino acids used. From Table 1 it is clear that substrates with the highest $\log P$ gave rise to the best conversions (1a, 1b, 1c, 1e and to a lesser extent 1d), while less hydrophobic amino acids lead to lower conversion yields (1f-1i). These observations indicate that the hydrophobic effect is an important factor in the equilibrium shift observed when comparing solid phase reactions to those in aqueous solution. However, there is no absolute correlation between $\log P$ and conversion. For example, amino acid 1f has the lowest $\log P$ value but reaches a good conversion of 84%. Thus there are other specific interactions between the polymer and amino acids that contribute to the conversion yields obtained.

In order to obtain more accurate data on the contribution of substrate hydrophobicity to the equilibrium of the reaction, the conversions obtained for the same amino acid with different protecting groups under otherwise identical conditions were investigated. Fmoc-Phe and Z-Phe have the same pK_a of 3.7 and only differ in the hydrophobicity due to their different protecting groups. Fig. 2 shows the synthesis of peptides **3d** and **3e** as a function of pH. For both these two acyl donors the shape of the curve is the same and its shape can be explained in term of ionization effects (see next section). The difference in conversions observed when comparing Z-Phe and Fmoc-Phe therefore directly reflects the contribution of substrate hydrophobicity.



Fig. 2 Effect of substrate hydrophobicity on the synthesis yields of the thermolysin catalysed peptide synthesis on solid phase PEGA₁₉₀₀ as a function of pH: Z-Phe-Phe (squares) and Fmoc-Phe-Phe (diamonds).

Thus, it appears that the use of more hydrophobic protecting groups will allow for a further decrease of the number of equivalents of soluble amino acid needed. It is expected that conversion yields could be improved further by using tailor-made highly hydrophobic protecting groups.

iii. The effect of suppressed amine ionization on the equilibrium position. As a third possible contribution to the equilibrium shift it was suggested that protonation of the amino group of the immobilized phenylalanine may well be significantly suppressed as compared to the amino group of a soluble phenylalanine derivative. Peptide hydrolysis is favored in dilute aqueous solution largely because of favourable ionization of both the amine and acid components. Any suppression of ionization is therefore expected to result in a shift toward peptide synthesis.

When in solution, amino acid amines will protonate independently of each other. By contrast, solid supported amino groups are restricted in their movement, and when close together they are likely to influence each other. When the pH of the bulk solution is lowered, polymer bound amines will start to become protonated leading to local positive charges within the polymer. As the pH drops further, protonation of polymer bound amino groups will become increasingly more difficult due to electrostatic repulsion between the charged amines. Hence, ionization will be more and more suppressed at low pH values when compared to free solution. And suppressed ionisation will result in an equilibrium shift towards peptide synthesis as is also observed when organic (co-) solvents are used in solution phase biocatalysis.

To analyze the behavior quantitatively we have considered the case where excess solid $X-aa_2$ is present, so that $[X-aa_2]^0$ remains constant at its solubility in the reaction mixture. Eqn. (6) may then be re-written as:

$$f_{\rm eq}^{\rm peptide} = \frac{[X-aa_2]^0 K^0}{1+10^{pK_{\rm a,amine}-pH} + [X-aa_2]^0 K^0}$$
(7)

Note, that K_{ref}^0 is now replaced by K^0 because hydrophobic effects may cause a shift from the aqueous solution reference value, as explained above. Eqn. (7) can be rewritten as:

$$\frac{1}{f_{eq}^{peptide}} = 1 + \frac{1}{K^0 [X-aa_2]^0} + \frac{[H^+]}{K^0 K_{amine} [X-aa_2]^0}$$
(8)

If one assumes that there is no electrostatic repulsion between amines, all amines ionise according to a single pK_{amine} value. In that case, eqn. (8) suggests that a plot of $1/f_{eq}^{peptide}$ *versus* [H⁺] will be linear (K^0 and [X-aa_2]⁰ are constants). Unfortunately such a plot gives rise to bunched data points even over the relatively small pH range (2 units) in our experiments. When plotting instead $-\log f_{eq}^{peptide}$ limiting behaviour can be recognised at high and low pH: above the pK_{amine} the [H⁺] term in eqn. (8) becomes negligible and peptide yield will be independent of pH resulting in a straight line of slope 0. At low pH the [H⁺] term dominates, and the equation predicts a linear relationship between $-\log f_{eq}^{peptide}$ and pH, with a gradient of -1. By contrast, if the amine charges did affect each other (*i.e.* no single pK_{amine}), the protonation of the amino groups becomes increasingly more difficult when the pH is lowered. In this case the amino groups would behave as if they had a range of pK_{amine} values. For the plot this would result in a curved line instead of a straight line at pH values below the of pK_{amine} .

Fig. 3 shows the pH dependence of the reactions producing solid supported peptides **3d** and **3e**. Both show the two straight line regions as expected for a system where amines protonate independently. The break between the two occurs at bulk pH around 7.0, which is the pK_a of the amino group of Phe-OMe in free solution. Hence it can be concluded that under the conditions used suppressed amine ionisation is a minor factor in the observed equilibrium shift. The shift of the line at all pH values observed when comparing the two peptides reflects the difference in K^0 through the hydrophobic contribution as discussed in the previous section.



Fig. 3 Effect of solid supported amine ionisation on the synthesis yields of the thermolysin catalysed peptide synthesis on solid phase $PEGA_{1900}$ as a function of pH: Z-Phe-Phe (squares) and Fmoc-Phe-Phe (diamonds).

In these experiments a relatively high ionic strength buffer was used (0.1 M K-phosphate buffer). Electrostatic interaction between the charged groups will be greater at lower ionic strength, and hence ionization of solid supported amines will be suppressed more. It is therefore expected that synthesis will be more favoured at lower ionic strength. We are currently studying the effect of ionic strength on the peptide synthesis/ hydrolysis equilibrium in more detail.

Conclusion

There can be a substantial equilibrium shift in favour of amide synthesis when amines are immobilized on a solid support, compared with preferred hydrolysis in aqueous solution. This allows efficient protease catalysed solid-phase synthesis of a variety of peptides. By comparison with the reaction in aqueous solution it was shown that the shift in equilibrium is not caused just by the use of excess amino acids. The most important factor appears to be substrate hydrophobicity and consequently the highest conversions are observed with amino acid substrates carrying hydrophobic protecting groups.

Experimental

Enzymes, reagents and solvents

Thermolysin was obtained from Sigma (UK) as protease type X. Fmoc-Phe,

Z-Phe, Fmoc-Nle, and Fmoc-Gly were all from Sigma (UK). Fmoc-Asp, Fmoc-His, Fmoc-Ser, Fmoc-Gln were from Bachem (UK). Fmoc-Leu, hydroxymethylphenoxyacetic acid (HMPA), and dimethylaminopyridine (DMAP) were from Novabiochem. PEGA₁₉₀₀ was a kind gift provided by Polymer Laboratories (UK). Diisopropylcarbodimide (DIC), hydroxybenzotriazole (HOBt) and acetic anhydride were from Aldrich. All solvents were of the highest purity available and obtained from Aldrich.

Solid phase peptide synthesis

Solid phase substrates were linked to amino functionalised PEGA₁₉₀₀–NH₂ (Polymer Laboratories) beads *via* a Wang-type linker (hydroxymethylphenoxyacetic acid, HMPA). Attachment of the linker was achieved in DMF with 3 equivalents of HMPA, 4 equivalents diisopropylcarbodimide (DIC), and 6 equivalents hydroxybenzotriazole (HOBt). After overnight rotation on a blood rotator at 20 °C, the resin was washed (5 × 50 : 50 acetonitrile : water, 5 × MeOH, 5 × DMF, 5 × acetonitrile, 5 × DCM, 5 × DMF) and the first amino acid was coupled to the resin. This step used 3 equivalents Fmoc-amino acid, 4 equivalents DIC and 0.1 equivalents of dimethylaminopyridine (DMAP) in DMF. The mixture was incubated on the blood rotator overnight at 20 °C. This was again followed by a washing sequence as above.

Unreacted OH groups were capped by addition of 5 equivalents acetic anhydride in DMF and left on the blood rotator at 20 °C overnight. This was followed by a washing sequence (as above) and removal of the Fmoc group by addition of a solution of 20% piperidine in DMF, and left 30 minutes on blood rotator.

Determining loading of Phe on PEGA₁₉₀₀

A weighed amount of PEGA₁₉₀₀-Wang-Phe-Fmoc resin was cleaved with 95 : 5 (v/v) TFA-water during 2 hours. The cleavage mixture was washed with 10 ml of a mixture of 50 :50 (v/v) water : acetonitrile. The solvent was evaporated in a vacuum centrifuge (Christ, Germany). The residue was redissolved in 1 ml 50 : 50 (v/v) water : acetonitrile containing 0.1% TFA. The loading was then determined by reverse phase HPLC (for conditions, see below).

In addition, a known amount of $PEGA_{1900}$ -Phe-Fmoc was weighed, then dried in a vacuum oven at RT to constant weight. This gave a value for the swelling of the resin and allowed for the dry weight to be calculated. The loading of the resin obtained varied from batch to batch between 70 and 110 µmol g⁻¹.

Enzymatic peptide synthesis on solid support

For the enzymatic reactions 5 mg of thermolysin (protease type X from Sigma) was added to a suspension of 10 mg PEGA₁₉₀₀ resin, 0.2 mmol protected amino acid, and 2 mL 0.1 M potassium phosphate buffer of the appropriate pH. Reactions were briefly mixed, and then incubated overnight at 20 °C on a blood rotator. The next day the resin was washed extensively using 5 mL volumes in the following sequence: $5 \times 50 : 50$ acetonitrile : water, $5 \times$ MeOH, $5 \times$ DMF, $5 \times$ acetonitrile, $5 \times$ DCM, $5 \times$ DMF. The products were cleaved from the resin with 2ml TFA : water 95 : 5 for 2 hours. Resin was then washed with 10 mL of a mixture of 50 : 50 acetonitrile in water, solvent was removed by evaporation and the residue redissolved in 1 mL 50 : 50 mixture of acetonitrile and water.

Analysis

The samples were analysed by HPLC on a Waters 2690 LC system equipped with a Waters 468 UV detector and a reverse phase column (0.46×25 cm Hichrom HIRPB-250A). Mobile phases were acetonitrile and water with 0.1% TFA added to each, a gradient was used that started with 20% acetonitrile increasing to 50% acetonitrile in 30 minutes at a flow rate of 1 ml min⁻¹. LCMS was done on a Waters 2790 LC system coupled with a Micromass Platform II mass spectrometer using Electrospray ionisation mode.

Acknowledgements

The authors gratefully acknowledge financial support from the BBSRC, EC and the Wellcome Trust. We would also like to thank Polymer Laboratories (UK) for the generous supply of $PEGA_{1900}$

References

- 1 N. Bezay, G. Dudziak, A. Liese and H. Kunz, *Angew. Chem., Int. Ed.*, 2001, **40**, 2292–2295.
- 2001, 40, 2252–2253.
 2 M. Meldal, *Biopolymers*, 2002, 66, 93.
 3 R. Reents, D. A. Jeyaraj and H. Waldmann, *Drug Disc. Today*, 2002, 7, 71–76.
- 4 D. W. Kadereit and H. Waldmann, Chem. Rev., 2001, 101, 3367-3396.
- 5 F. G. Kuruvilla, A. F. Shamji, S. M. Sternson, P. J. Hergenrother and
- S. L. Schreiber, *Nature*, 2002, 416, 653–657.
 6 P. C. Michels, Y. L. Khmelnitsky, J. S. Dordick and D. S. Clark, *Trends Biotechnol.*, 1998, 16, 210–215.
- 7 R. V. Ulijn, B. Baragana, P. J. Halling and S. L. Flitsch, *J. Am. Chem. Soc.*, 2002, **124**, 10988–10989.
- 8 R. V. Ulijn, A. E. M. Janssen, B. D. Moore and P. J. Halling, *J. Chem. Soc., Perkin Trans.* 2, 2002, 5, 1024–1028.