Multifunctional Folded Polypeptides from Peptide Synthesis and Site-Selective Self-Functionalization—Practical Scaffolds in Aqueous Solution

Linda K. Andersson,[b] Gunnar T. Dolphin,[a] and Lars Baltzer*[a]

The site selectivity of His-mediated lysine and ornithine side-chain acylation in a designed four-helix bundle protein scaffold was mapped by reaction of several polypeptides with one equivalent of mono-p-nitrophenyl fumarate in aqueous solution at pH 5.9 and room temperature followed by an analysis of the degrees and sites of acylation. Integration of the HPLC chromatograms of the acylated polypeptides and trypsin cleavage followed by mass spectrometry analysis of the tryptic fragments provided the experimental evidence. Based on these and previously published results a strategy was developed for the site-selective and stepwise incorporation of three residues into a folded polypeptide in aqueous solution at room temperature. The first substituent was incorporated by reaction of a 1.7-fold excess of the corresponding active ester with the polypeptide at pH 5.9, the second substituent was introduced in a 3-fold excess after the pH value was raised to 8, and the third substituent was incorporated by reaction of a 10-fold excess with the polypeptide at pH 5.9. No intermediate steps of purification were taken and the overall yield was 30% or more. Examples of the substituents included are carbohydrates, an enzyme inhibitor, a fumarate, and an acetate group. The introduction of different substituents into three individually addressable positions in a stepwise, efficient, and controllable reaction demonstrates that designed folded polypeptides are practically useful scaffolds that can be functionalized by using very simple chemistry in aqueous solution. Predicted applications include designed receptors, biosensors, and molecular devices.

KEYWORDS:

bioorganic chemistry · protein design · protein folding · protein modifications · scaffolds

Introduction

Biomolecular supramolecular chemistry in aqueous solution is the cornerstone of the life processes. The self-assembly of linear peptides into folded proteins is the pathway by which complex structures for catalysis and binding are formed that are capable of discrimination between the components of the vast biological pool of biomacromolecules and metabolites with almost perfect precision. In addition to the complexity that arises from the naturally occurring amino acids and the large number of available folding motifs,[1] covalent posttranslational modifications add considerable structural and functional variability.[2–5] In spite of the opportunities provided by the diversity of protein scaffolds they have, so far, not been explored by chemists for manmade purposes to any significant degree, probably because of the difficulties encountered in understanding protein folding. Recent advances in de novo protein design[6–10] suggest that new proteins can be designed from scratch and exciting opportunities for the design of novel proteins for tailor-made purposes are now becoming apparent in chemistry, medicine, and biotechnology.

The protein scaffold is a versatile building block with well-defined distances and geometries between amino acid residues. In a helical segment, the distance between α carbon atoms is 5.2 or 6.3 Å if the residues are three or four residues apart in the sequence, respectively, and several residues along the face of a helix can be used together to form sites of great complexity. Larger motifs that combine several secondary structure elements increase the number of addressable functional sites as well as the range of interresidue distances. The size and complexity of proteins, even small ones, compare favorably with what is currently achievable in organic compounds designed to self-assemble, especially in aqueous solution. Designed proteins therefore have the potential to become practically useful vehicles for a large variety of purposes. The difficulties encountered in understanding the so-called protein folding problem should not be underestimated but an increased understanding of the relationship between sequence and structure has emerged from de novo design of proteins. Several de novo designed proteins that fold into structures like those of

[a] *Prof. L. Baltzer, Dr. G. T. Dolphin*
Department of Chemistry
Linköping University, 58183 Linköping (Sweden)
Fax: (+46) 13-122587

[b] *Dr. L. K. Andersson*
Department of Chemistry
Göteborg University, 41296 Göteborg (Sweden)
Fax: (+46) 31-772-3840

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

the native proteins have been reported, together with their high-resolution NMR structures,[8, 11–16] which supports the conclusion that we now know how to design proteins that approach and even surpass a hundred residues in size.

Native proteins are posttranslationally modified in enzyme-catalyzed reactions with high efficiency[17] but few chemical reactions exist that provide the precision needed for the site-specific functionalization of manmade proteins. In order for chemists to be able to make full use of designed protein scaffolds, chemical reactions are needed that make it possible to address site-selectively several positions for the introduction of multiple functions in controlled geometries. Classical protein chemistry provides many reactions capable of chemoselectively addressing specified amino acid residue side chains, but without site selectivity. Chemoselective reactions[18–25] based on the reactivities of functional groups in artificial amino acid residue side chains show great promise, especially in combination with protein synthesis through chemical ligation.[21, 26] There are, however, advantages in site-selective functionalization strategies that are based on the exclusive use of naturally occurring amino acids. The availability of molecular biological methods for selection and screening makes it possible to refine structures and functions. Self-catalyzed functionalization reactions based on the reactivities of the naturally occurring amino acids are very economical in terms of the cost of the introduction of new functions. While the solid-phase synthesis of peptides requires addition of a large excess of any amino acid derivative to be introduced, typically several tenths of millimoles of material, self-catalyzed reactions are readily carried out in high yields at micromolar reagent concentrations.

We have previously reported site-selective functionalization reactions of lysine side chains based on the cooperativity of His–Lys pairs in helical sequences,[27] and also strategies for addressing lysine residues in four-helix bundle proteins directly by using activated ester substrates.[28] In the former functionalization process, His residues react with an ester in a two-step reaction.[29] The first and rate-limiting step is the formation of an acyl intermediate at the His side chain upon the release of the leaving group. In the second step, the acyl group is transferred in a fast intramolecular reaction to form an amide at the side chain of the Lys residue. This amide formation occurs even at a pH value below 6, where Lys side chains are predominantly protonated. The pKa value of a solvent-exposed Lys residue in aqueous solution is 10.4[30] and the efficiency of the acylation reaction is ensured by the fact that it is intramolecular and its cooperativity between the His and Lys residues. If the His residue is flanked by more than one Lys residue, intramolecular competition determines which Lys residue is acylated.[31] In a helix, the Lys residue four residues towards the C terminus ($i,i+4$) from the position of the His (*i*) is the preferred site of acylation, in comparison with that in the position three residues towards the N terminus ($i,i-3$).

In the case of direct acylation of lysine residues, we have recently reported that the introduction of Lys residues at positions designed to form the hydrophobic core of a folded four-helix bundle protein forms an efficient strategy for site-selective Lys acylation at pH 8.[28] We have thus established two complementary strategies for the acylation of Lys side chains. Combined, these strategies may be used to site-selectively introduce several substituents at the side chains of ornithine (Orn) or Lys residues in folded proteins in aqueous solution without the need for side-chain protection groups. Here, we report the efficient and stepwise incorporation of three different acyl groups into a designed, folded four-helix bundle protein scaffold without intermediate steps of purification. The acyl groups used here include sugar derivatives, a high-affinity enzyme ligand, an acetyl, and a fumaryl group, but the reactions apply to all active ester substrates. The demonstration for the first time of a strategy for the stepwise and site-selective introduction of several different substituents into a folded protein shows that protein scaffolds are versatile and practically useful building blocks in the design and synthesis of supramolecular compounds for the purposes of biomolecular recognition and interaction. Predicted areas of application are biosensors, receptors, catalysts, and molecular devices.

Results

Design of the helix–loop–helix dimers

Sixteen 42-residue peptides were designed to fold into hairpin helix–loop–helix motifs and dimerize to form four-helix bundles (Figure 1, Table 1). These peptides were synthesized on an Applied Biosystems Pioneer automated peptide synthesizer by using the 9-fluorenylmethoxycarbonyl (Fmoc) protection group strategy, purified by reversed-phase (RP) HPLC and identified by using a Voyager DE-STR MALDI-TOF mass spectrometer from Applied Biosystems. The peptide designs were based on the structures of the 42-residue polypeptides SA-42,[32] KO-42,[33]

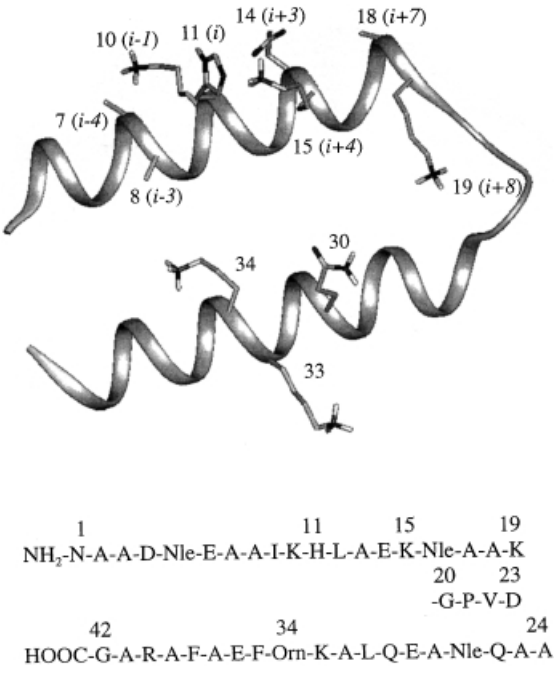


Figure 1. *The modeled structure and amino acid sequence of LA-42b showing the side chains of the residues involved in the acylation reaction. The dimer is the active peptide but for reasons of clarity only the monomer is shown. The one letter code for the amino acids is used.*

Table 1. *Amino acid residue substitutions used to determine the hierarchy of reactivities in histidine-mediated acylation of the 42-residue polypeptides.*[a]

Peptide	7 $i-4$	8 $i-3$	10 $i-1$	11 i	14 $i+3$	15 $i+4$	18 $i+7$	19 $i+8$	30 helix II	33 helix II	34 helix II
LA-42b	A	A	**K**	H	E	***K*** *(60)*	A	**K**	Q	**K**	**Orn**
P1	A	A	***K***[c]	H	E	S	A	***K***[c]	Q	**K**	***Orn*** *(49)*
P2	A	A	***K*** *(11)*	H	E	A	A	**K**	Q	**K**	A
P3[b]	A	A	**K**	H	***K*** *(4)*	S	A	***K***[c]	Q	**K**	***Orn*** *(43)*
P4[b]	A	A	***K*** *(12)*	H	***K*** *(12)*	S	A	**K**	Q	**K**	A
P5	***K*** *(18)*	S	***K*** *(10)*	H	E	A	A	**K**	Q	**K**	A
P6	A	***K*** *(20)*	***K*** *(13)*	H	E	A	A	**K**	Q	**K**	A
P7	A	A	***K*** *(13)*	H	E	A	A	**K**	***K*** *(22)*	**K**	A
P8	A	A	***K*** *(3)*	H	E	S	**K**	A	Q	**K**	A
P9	A	A	***K*** *(5)*	H	E	S	**K**	S	Q	**K**	A
P10[b]	A	S	***K***[c]	H	E	A	A	***K***[c]	Q	**K**	***Orn*** *(38)*
P11[b]	A	S	**K**	H	***K*** *(4)*	A	A	**K**	Q	**K**	***Orn*** *(34)*
P12	A	S	***K*** *(12)*	H	***K*** *(6)*	A	A	**K**	Q	**K**	A
P13	A	A	***K*** *(25)*	H	**K**	S	**K**	S	Q	**K**	A
P14	A	A	***K*** *(6)*	H	E	A	A	**K**	Q	**K**	S
Pref[b]	A	**K** (7)	**K**	A	**K**	A	A	**K** (41)	**K** (7)	**K**	A

[a] The degree (%) of monomodification is given in brackets. All lysine and ornithine residues are shown in bold and the positions of monomodification in italics. Residues not shown remain the same as those in LA-Y2b. [b] Besides monomodification, a small amount of dimodified product was detected (< 6%). The positions of dimodification have not been identified. [c] These residues were also monomodified to an extent (< 5% of the total monomodification) estimated from the MS spectra of the tryptic digests. The monomodified peptides were not separable by analytical HPLC.

LA-42b,[34] and KA-I,[28] which have been described in great detail previously. The sequences were shown previously by NMR and circuluar dichroism (CD) spectroscopy and by equilibrium sedimentation ultracentrifugation (SA-42 and KO-42) to fold into hairpin helix – loop – helix motifs that dimerize in an antiparallel way to form four-helix bundles. The polypeptides were designed to form amphiphilic helical segments linked by a short loop and residues capable of capping, of stabilization of the helix dipole moment, and of formation of salt bridges were introduced to stabilize the folded helices.[6] In comparison with the sequence of LA-42b, six residues or less were changed in the sixteen sequences reported here and there is little reason to suppose that major structural changes occur as a result of these modifications. Nevertheless, the solution structure of one of the polypeptides (LA-42h, Figure 2) was investigated in detail by CD spectroscopy as well as by NMR spectroscopy. The sequences of the remaining fifteen peptides were studied by CD spectroscopy to probe in a qualitative way whether they all folded into the designed motif.

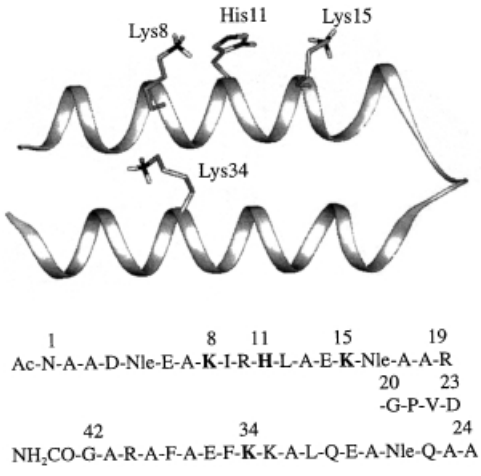


Figure 2. *The modeled structure and amino acid sequence of LA-42h. The side chains involved in the functionalization are shown. The dimer is the active peptide but for reasons of clarity only the monomer is shown.*

In order to develop a system in which the helix-I – loop – helix-II dimer can be used as a scaffold for site-selective functionalization, the hierarchy of the reactivities of the lysine residues that flank the histidine was investigated in a series of peptides based on the sequence of LA-42b (Figure 1, Table 1). The most reactive residue, which was preferentially acylated in a sequence, was replaced by an alanine or serine residue to form a peptide from which the second most reactive lysine could be identified, and so on. LA-42b contains a His residue in position 11 (i) in helix I and in total four Lys residues in positions 10 ($i-1$), 15 ($i+4$), 19 ($i+8$), and 33 (helix II), and an Orn residue in position 34 (helix II). In addition to the positions in which the lysine and ornithine residues are found in LA-42b, the reactivity hierarchy of six other positions were investigated. His-mediated acylation of Lys residues was expected to require proximity between His and Lys residues and the Lys residues were therefore incorporated at positions that flanked His11 either in the same turn of the helix or in helical turns before or after that of His11. To test the hypothesis that His-mediated acylation requires spatial proximity, lysine residues were also introduced into positions 18 and 19, two turns away from His11, at a distance of more than 10 Å between the α carbon atoms. Thus, modifications were carried out in helix I in positions 7 ($i-4$), 8 ($i-3$), 10 ($i-1$), 14 ($i+3$), 15 ($i+4$), 18 ($i+7$), and 19 ($i+8$). Positions in helix II were also probed although the geometrical relationship between the His11 residue and the residues of helix II is not very clearly defined. We have previously reported catalytic sites for ester hydrolysis where cooperative effects between residues in positions 11 and 34 have been observed.[35] His11 was therefore concluded to be in close proximity to the residue in position 34

and lysines were introduced in positions 30 and 34. The preference for acylation of each lysine residue was determined by measuring the degree and site of acylation of each polypeptide, and by replacing each one of the more reactive lysines in turn by Ala residues (Table 1). The preferentially acylated lysine was replaced by an Ala or Ser residue so that the acylation of the second most reactive lysine could be observed, and the hierarchy of reactivities determined.

Sequences to probe the role of serine residues in Lys acylation were also designed, with Ser introduced in positions 8, 15, 19, and 34. The side chain of Ser is an alcohol and the intention of the incorporation of Ser residues was to determine whether Ser acylation could also be achieved by the His-mediated pathway. No evidence of intermediates or reaction products that result from Ser acylation was obtained. However, serine residues were found to affect the site selectivity, as discussed below.

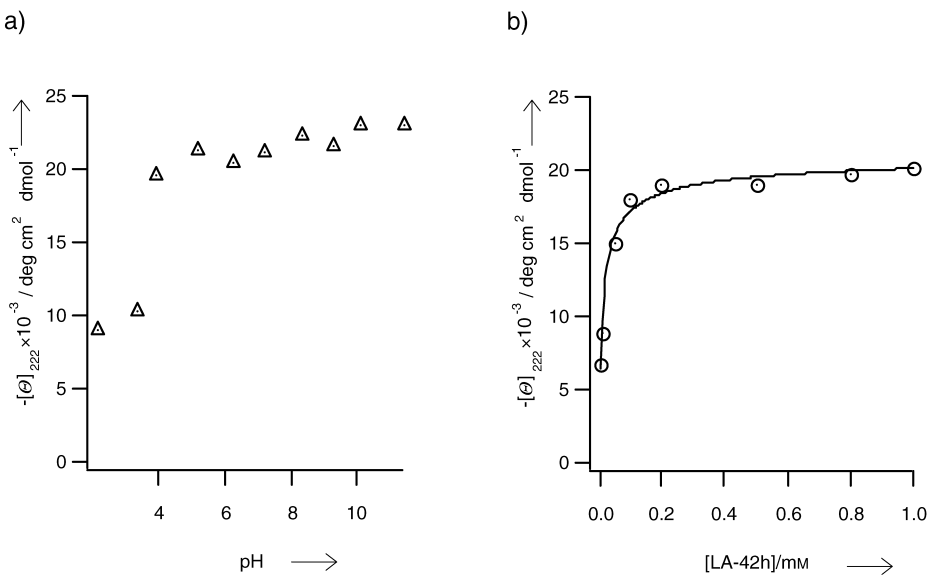


Figure 3. *The pH (a) and concentration (b) dependence of the mean residue ellipticity of LA-42h at 222 nm. At a pH value below 4 and concentrations below 0.2 mM the dimer dissociates and forms a monomer with low helical content.*

The sequence of LA-42h shown in Figure 2 was designed to provide a scaffold for the introduction of three different substituents in a stepwise reaction in aqueous solution at room temperature without intermediate purification to demonstrate the usefulness of the folded sequence in the synthesis of a complex structure with ease and in high yield. The design was based on the relative reactivities of individual lysine and ornithine residues determined both previously and in the present series of polypeptides. His11 was known to mediate the preferential acylation of Lys15 and Orn34 was shown to be the most reactive residue in the direct acylation reaction at pH 8. Since Lys residues in positions 8, 10, and 14 were found in the present investigation to be of approximately equal reactivity, Lys8 was introduced to carry the third substituent because of the spatial proximity between residues 8, 15, and 34.

The structure of the folded helix – loop – helix dimers

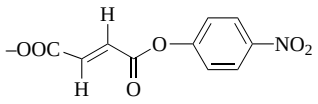
The secondary structure of LA-42h was determined by CD spectroscopy and the spectrum showed the characteristics of an α-helical protein, with minima at 208 and 222 nm. The mean residue ellipticity of LA-42h at 222 nm was $-20\,100$ deg cm^2 dmol^{-1} in 100 mM bis[2-hydroxyethyl]imino – tris(hydroxymethyl)aminomethane (Bis – Tris)/Tris buffer at pH 5.9 and room temperature. The mean residue ellipticity of LA-42h showed no pH dependence in the interval pH 4 – 9 and was independent of concentration in the interval 0.2 – 1 mM (Figure 3).

The secondary and tertiary structures of LA-42h were investigated by ^{1}H NMR spectroscopy. The ^{1}H NMR spectrum of LA-42h was assigned from the TOCSY and NOESY spectra recorded in an H_2O/D_2O (90:10 v/v) solution that contained 4 vol% trifluoroethanol (TFE)-d_3 at 308 K by using methods previously described in detail.[36] Most of the spin systems of the amino acids were identified from the αH – NH region in the TOCSY spectrum and the sequential assignments of residues 5 – 19 and 23 – 41 were obtained from the NH – NH region in the NOESY spectrum. An extra set of resonances with low intensity from the amino acids near the proline in the loop region was identified as a result of the *cis* – *trans* equilibrium of Pro21. The αH chemical shifts are diagnostic of secondary structure formation, and the measured values were compared to tabulated values for random coil conformations.[37] Two helical segments from norleucin 5 (Nle5) to Nle16 and from Ala25 to Arg40 were identified in the sequence of LA-42h from their upfield shifts relative to the random coil values. A loop region was also assigned from the downfield shifts or absence of deviations relative to random coil conformation of the α protons. Medium-range NOEs typical of α helix formation, αH – NH $i,i+3$ and $i,i+4$, were found in the sequence Nle5 – Arg19 in helix I and Ala24 – Ala41 in helix II (for details, see the Supporting Information). The NOE connectivities between the aromatic protons of the phenylalanine side chains of residues 35 and 38 and the methyl protons of Ile9 and Leu12 showed that the peptide folds into a helix – loop – helix motif and the connectivity between the aromatic ring protons and the methyl group of Nle16 suggests that the peptide dimerizes in an antiparallel manner to form a four-helix bundle.

The secondary structures of the folded peptides other than LA-42h were investigated by CD spectroscopy and the spectra showed the characteristics of α-helical proteins, with minima at 208 and 222 nm. The mean residue ellipticities of the peptides at 222 nm were in the range $-17\,500$ to $-28\,000$ deg cm^2 dmol^{-1} at pH 5.9 in 50 mM Bis – Tris buffer, which is well within the range seen for previously designed helix – loop – helix dimers (see the Supporting Information). We conclude that all polypeptides reported here fold into a hairpin helix – loop – helix motif.

Site-selectivity in histidine-mediated acylation of lysine residues

In order to explore the selectivity of lysine acylation, the peptides were reacted at approximately 1 mM concentration with one equivalent of the active ester, *p*-nitrophenyl fumarate (**I**; Scheme 1), at room temperature and pH 5.9. The reaction



I

Scheme 1. *The structure of the active ester* p*-nitrophenyl fumarate.*

mixture was analyzed by analytical HPLC and the peaks in the chromatogram that resulted were integrated. The peptides were identified as unmodified, monomodified, or dimodified by MALDI-TOF mass spectrometry. Tryptic digestion of each peptide was used to determine the sites of modification. The tryptic fragments were identified by MALDI-TOF MS and the analysis was based on the known capacity of trypsin for cleavage of peptides on the C-terminal side of the positively charged residues lysine and arginine. No cleavage occurs if the side chain of Lys has been modified and in addition to preventing cleavage, acylation increases the weight of the fragment by the mass of the acyl group. The degrees and sites of modification are shown in Table 1.

The reaction between the peptide and the substrate competes with the spontaneous hydrolysis of the substrate to form the corresponding carboxylic acid and thus reaction with one equivalent of **I** leads to incomplete modification since some of the substrate is hydrolyzed. The fraction of substrate expected to be incorporated can be estimated from the pseudo-first-order rate constant of the peptide self-catalyzed reaction and the first-order rate constant of the spontaneous hydrolysis. For preparative purposes compensating amounts of substrate can thus be added to ensure optimal incorporation. In the determination of the degree of modification no compensation was undertaken since the purpose was to identify the site that was the most reactive in competition with other sites.

The hierarchy of reactivities of Lys15, Orn34, and Lys10

LA-42b has a histidine residue in position 11 (*i*), lysine residues in positions 10 ($i-1$), 15 ($i+4$), 19 ($i+8$), and 33 (helix II), and an ornithine residue in position 34 (helix II). This peptide was 60% monoacylated at position 15 after reaction with one equivalent of **I** and no other reaction products were observed (Table 1). When the lysine residue in position 15 was replaced by a serine residue in the peptide P1, the new main site of modification was Orn34 and the degree of monomodification was 49%, although minor amounts of peptides acylated at positions 10 and 19 were also identified. Lys15 therefore competes favorably with Orn34 to capture the acyl group of **I**. The selectivity is very high since monoacylation at Orn34 was never observed in peptides with His11, Lys15, and Orn34 in the sequence. Orn34 was only acylated in peptides that contained a Lys15 when Lys15 itself had already been acylated. The removal of Lys15 reduced the reactivity of the peptide towards the ester, which shows that Lys15 enhances the reactivity of His11, probably by depression of the pK_a value of the His residue through an electrostatic effect and by stabilization of the tetrahedral transition state by the same mechanism. The second-order rate constant (Table 2)

Table 2. *Second-order rate constants for reaction with* ***I*** *at pH 5.9 and 298K.*

Peptide	k_2 [$M^{-1}s^{-1}$]
LA-42b	0.18
P1	0.04
P3	0.07
P10	0.04
P11	0.07

of the reaction of LA-42b with **I** decreased from 0.18 $M^{-1}s^{-1}$ to 0.04 $M^{-1}s^{-1}$ upon replacement of Lys15 by a Ser residue. The reduced overall reactivity does not affect the site selectivity of the polypeptides but the competition between spontaneous hydrolysis and functionalization becomes less favorable. This situation is, however, easily improved by an increase in the concentration of peptide and thereby the rate of incorporation.

Replacement of both Lys15 and Orn34 by Ala residues in P2 led to a peptide with low reactivity that was only acylated at the side chain of Lys10 and only to a low degree. In LA-42b the reactivity hierarchy was therefore Lys15, Orn34, Lys10, with very little acylation of other lysine residues in the sequence. The low reactivity of Lys10 is advantageous because it makes the selectivity high and because the degree of incorporation can be improved by the addition of excess substrate if functionalization of Lys10 is attempted after functionalization of Lys15 and Orn34.

The relative reactivities of Lys10 and Lys14

The reactivity was increased in peptide P3 by the introduction of Lys14, which is known to depress the pK_a value of His11. The second-order rate constant was increased from 0.04 to 0.07 $M^{-1}s^{-1}$. Again, Orn34 was the preferred site of modification; competition between Lys14 and Orn34 results in approximately ten times more acylation of Orn34 than of Lys14. The degree of incorporation was 4% at Lys14 and 43% at Orn34. No acylation of Lys10 was observed in P3, which suggests that Lys14 competes favorably with Lys10 for the acyl group at the side chain of His11. However, the low degree of modification made quantitative comparisons difficult. Orn34 was therefore replaced by an Ala residue in the peptide P4, and the degree of acylation was found to be equal between Lys10 and Lys14, with a yield of 12% at each side chain. Lys10 and Lys14 are therefore equally reactive.

The reactivity of Lys7 and Lys8

Positions 7 and 8 are also in the proximity of His11. The introduction of Lys7 into P5 and Lys8 into P6 led to substrate incorporation in low yields at both positions. The degree of incorporation at Lys7 and Lys8 was approximately 50% higher than at Lys10, which was also present in the sequences. This result means that Lys7 and Lys8 are sites of comparable reactivity to those of Lys10 and Lys14. In the synthesis of a multiply-modified protein scaffold, Lys7, Lys8, Lys10, and Lys14 are therefore alternative positions for incorporation of functional groups, depending on the required geometry. The small difference in the degrees of acylation of the lysine residues in positions 7, 8, 10, and 14 suggests that these different positions can be used equally well in combination with Lys15 and Orn34 for the formation of trisubstituted polypeptide scaffolds.

The reactivity of lysines in helix II

The acylation of Lys or Orn residues in helix II was investigated by the incorporation of Lys30 to probe whether His-mediated acylation could be extended further into the folded motif. Orn34 in helix II was selectively acylated in the absence of Lys15. The reaction of a peptide that contained a lysine residue in position 30 as well as in positions 10, 19, and 33, but alanine residues in positions 15 and 34, provided a probe of the reactivity of position 30. Upon reaction of this polypeptide P7 with one equivalent of **I**, Lys30 was 22% monoacylated but 13% acylation was also observed in position 10 and the intramolecular reactivity of Lys30 was therefore less than a factor two higher than that of Lys10. Nevertheless, the possibility of acylation of Lys30 expands the available incorporation geometries.

The reactivity of Lys18 and Lys19 compared to that of Lys10 and Lys14

To probe whether His-mediated acylation could be extended beyond one turn of a helix, the peptides P8 and P9 (Table 1) were synthesized for which Lys18 was introduced and the competition from Lys15 and Orn34 was eliminated by replacement with a Ser and an Ala residue, respectively. Only trace amounts of acylation were detected and only at Lys10. Acylation of Lys19 was only observed in trace amounts in the peptides P1, P3, P10, and P11 and possibly occured by a direct pathway as a result of a pK_a reduction of the Lys19 residue that enhances its reactivity. We conclude that only residues in close proximity to His11 were acylated by the His-mediated pathway. The substrate is consumed by reaction with the His residue to form amide bonds or to be hydrolyzed, and by spontaneous hydrolysis. Direct acylation of lysine residues can not in general compete with His-mediated reactions at low pH values because the nucleophilicity of histidine is higher than that of lysine.

Orn acylation has been shown previously to proceed efficiently even when there are no His residues in the vicinity because position 34 is formally in a hydrophobic core position according to the pattern of the heptad repeat. However, at pH 5.9 the degree of incorporation into P1 is higher than that observed in polypeptides without flanking His residues and it is likely that Orn34 is acylated both through a His-mediated pathway and by a direct pathway.

The effect of Ser residues on the degree of acylation

In order to begin to understand the role of residues not directly involved in bond-forming or bond-breaking steps, sequences were designed in which serine residues were introduced (Figure 1, Table 1). The sequence of P10 was the same as that of P1 except that Ala8 and Ser15 in P1 were changed to Ala15 and Ser8 in P10. As a result, the efficiency of the acylation of Orn34 was decreased; the degree of modification of P1 was 49% whereas the degree of modification of P10 was only 38%. Similarly, peptide P3 was compared with peptide P11, in which the same residues had been subjected to an identical exchange of positions. Here the effect was again to reduce the degree of acylation of Orn34 and to reduce the efficiency and selectivity of the site-selective acylation. In peptide P12 the incorporation of Ser8 and the removal of Ser15, both exchanged with Ala residues, led to a substantial increase in the site selectivity. Lys10 was favored by a factor of two over Lys14 in P12, although the residues were acylated to an equal extent in P4. In peptide P13, Ser15 and Ser19 were introduced in an attempt to affect the possible acylation of Lys18 but surprisingly the introduction of Ser19 appeared to affect only the acylation of Lys10. In spite of the fact that Lys10 and Lys14 were both present, and had previously been shown to be of equal reactivity, P13 was selectively acylated at Lys10 and in relatively good yield (25%). In peptide P14, the Ala34 residue of P2 was replaced by a Ser residue and the degree of acylation of Lys10 was reduced by approximately a factor of two.

The role of the serine residues is not understood. The modification pattern is affected by serine substitutions but the mechanism remains unclear. A possible explanation may be that the serine is involved in hydrogen bonding to the acyl intermediate and that therefore reaction pathways are favored that lead to acylation of specific lysine residues. The serine might also disrupt the structure of the helix and affect the relative positions of residues involved in acyl transfer, although no effects on helical content caused by the incorporation of serine residues could be detected. No evidence of serine-acylated intermediates was obtained, which suggests that covalent catalysis is not involved and Ser incorporation did not lead to alternative sites of acylation, only the site selectivity was affected. The incorporation of residues other than those directly involved in acyl transfer can therefore be expected to have a profound influence on site selectivity and the efficiency of substrate incorporation and to expand the usefulness of protein scaffolds.

His-mediated acyl transfer versus direct acylation of Lys residues

To determine if all of the positions discussed were acylated through the His–Lys pathway, a reference peptide, Pref, was

designed (Table 1). The histidine residue was removed and lysine residues were introduced into positions 8, 10, 14, 19, 30, and 33. Lys15 and Orn34 were both excluded. After reaction with one equivalent of **I** at room temperature and pH 5.9, the reaction mixture was analyzed by RP HPLC. The five peaks that resulted were identified by MALDI-TOF MS to be from unmodified peptide (*m/z*: found: 4264.4; calcd: 4265.0), monomodified peptide (three peaks; *m/z*: found: 4363.3; calcd: 4363.0), and dimodified peptide (*m/z*: found: 4461.3; calcd: 4461.1). Lys19 was the main acylation site with a degree of modification of 41%. Lys8 and Lys30 were modified to a degree of 7% each. In the presence of a His residue all esters react with the His side chain and only Orn34 can compete with the His-mediated pathway. Lysine residues that are not close to a His residue are not acylated and the site selectivity is high. In the absence of a His residue, Lys and Orn residues compete with the spontaneous hydrolysis reaction and even though reaction rates are low, lysine residues are acylated, although with little site selectivity. The exception is Orn34 or lysine residues in hydrophobic environments with depressed pK_a values.

The hierarchy of the reactivities of lysine residues

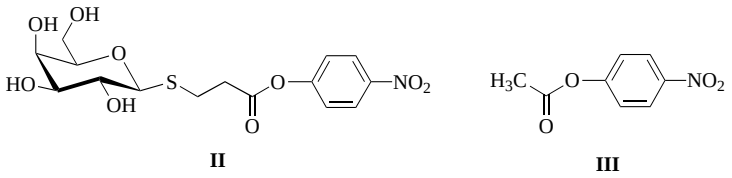
His-mediated acylation transfers were extraordinarily efficient between His11 and Lys15 and of a significantly lower efficiency between His11 and other flanking lysine residues. Orn34 was mainly acylated through a direct mechanism because of the low pK_a value induced by the hydrophobic character of position 34. The efficiency of the His-mediated pathway is therefore a key factor in four-helix bundle functionalization. The lysine residue in position $i+4$ is strongly favored over all other flanking lysine residues in helix I as well as over the lysine and ornithine residues in helix II. The more remote lysine residues in helix I, Lys18, Lys19 and Lys33, were not acylated by His11. Orn34 was the second most efficient site of modification and was acylated by a combination of His-mediated and direct acylation pathways. Site-selective incorporation of three different residues into positions 15, 34, and 8 was therefore considered to be optimal for the synthesis of a triple-substituted four-helix bundle.

The site-selective triple functionalization of a folded four-helix bundle protein in aqueous solution

Based on the results presented above, LA-42h was designed for the selective incorporation of three different functional groups in a one-pot reaction in aqueous solution. The strategy included the His – Lys-mediated pathway, which is favored at pH 5.9, and the direct acylation pathway, which is most efficient at pH 8. The sequence of LA-42h was based on that of LA-42b with a histidine residue in position 11, a lysine residue in position 15, and the incorporation of a lysine residue in position 34 to replace the ornithine (Figure 2). The third residue to be functionalized, Lys8, was chosen because of its position on the surface of the peptide. It was shown above that the side chains of Lys8 and Lys10 are of comparable reactivity and Lys10 was therefore replaced by an arginine residue to avoid competition for acylation. There is a lysine residue in position 19 of LA-42b to stabilize the macroscopic dipole moment of the helix. This amino acid was replaced by an arginine residue since it was shown to compete with Orn34 in KA-I.[28] The strategy was to direct the first functional group to position 15 by the His – Lys-mediated reaction pathway at pH 5.9, then to raise the pH value to 8 and direct the next group to position 34 with the direct acylation reaction, and then to lower the pH value to 5.9 again to address the final position and introduce the third group. The His – Lys-mediated reaction is efficient at pH 5.9 and the intramolecular competition between the lysine residues in positions 15 and 8 ensures that the first substituent will be incorporated with high selectivity at the side chain of Lys15.

The stepwise multifunctionalization reaction

Three substrates *p*-nitrophenyl fumarate (**I**), *p*-nitrophenyl 3-(β-D-galactopyranosyl-1-thio)propionate (**II**; Scheme 2),[38, 39] and *p*-nitrophenyl acetate (**III**) were chosen to demonstrate the stepwise introduction of three different substrates into a folded helix – loop – helix polypeptide motif. In order to simplify the change in pH value from 5.9 to 8, a mixture of the two buffers Bis – Tris and Tris was used, each at a concentration of 100 mM. In the first step, 1.7 equivalents of **I** were reacted with LA-42h at pH 5.9 and room temperature. After three days a sample from the reaction mixture was analyzed by analytical HPLC and 67% monomodification at the side chain of Lys15 was observed. The reaction mixture was also analyzed by MALDI-TOF MS and it was confirmed that the peptide was mainly monomodified, (*m/z*: found: 4600.1, calcd: 4600.2), but unmodified peptide remained and a small amount of dimodified peptide was also observed. The pH value of the reaction solution was adjusted to 8 and three equivalents of **II** were added. After 24 hours, HPLC analysis showed that 35% of the peptide was dimodified. MS analysis of the reaction mixture showed that the main product was peptide modified by **I** and **II** (*m/z*: found: 4849.8, calcd: 4850.5). The pH value was then lowered to 5.9 and 10 equivalents of **III** were added. HPLC analysis followed by MALDI-TOF analysis showed that 30% of the peptide was trimodified with the three different substituents (*m/z*: found: 4891.4, calcd: 4892.5) and the sites of modification were identified by tryptic digestion followed by HPLC – ESMS and MALDI-TOF MS to show that the peptide had been modified by **I** at Lys15, **II** at Lys34, and **III** at Lys8 (see the Supporting Information).

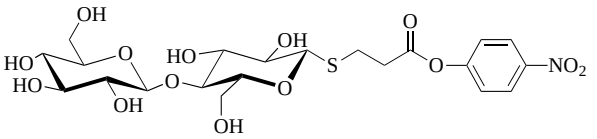


Scheme 2. p-*nitrophenyl 3-(*β-D-*galactopyranosyl-1-thio)propionate (**II**) and* p-*nitrophenyl acetate (**III**), two of the substrates used in the stepwise functionalization of the peptide scaffold.*

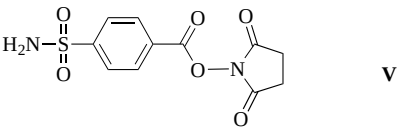
The incorporation was also performed in the reverse order by starting with modification of Lys34 at pH 8. In the first step 1.5 equivalents of **I** were added to the peptide at pH 8 and HPLC analysis showed that the peptide had become monomodified to

a level of 50% but 33% of dimodification was also observed. In the second step the pH value was lowered to 5.9 and 2.1 equivalents of **II** were added. HPLC analysis showed that 30% of the peptide was dimodified with **I** and **II**. At the same pH value, 10 equivalents of **III** were then added to the reaction mixture and analysis by HPLC showed 28% trimodified peptide. The modification sites were identified by tryptic digestion of the purified peptide and the resulting fragments were identified by LC–ESMS. The masses found corresponded to the peptide segments (see the Supporting Information) and the identity of the target peptide was unequivocally established.

To demonstrate the general applicability of the protein scaffold, a second incorporation with different substituents was carried out with the galactose derivative **II**, a cellobiose derivative **IV** (Scheme 3),[39, 40] and a high-affinity ligand for the



IV



V

Scheme 3. *Further substrates used in the peptide functionalization reactions to demonstrate the general applicability of the functionalization method described in the text.* ***IV*** *= A cellobiose derivative,* ***V*** *= the* N*-hydroxysuccinimide ester of 4-carboxy benzenesulfonamide.*

protein human carbonic anhydrase II, the *N*-hydroxysuccinimide ester of 4-carboxy benzenesulfonamide, **V**.[41] The example demonstrates the design of a scaffold that can be used to study the effect of glycosylation on the interplay between two proteins driven by a high affinity interaction. However, the number of possible applications in the study of protein–protein interactions and the inhibition of such interactions is virtually endless and a variety of intramolecular competition experiments can be envisioned where ligands of different affinity can compete for the active site of the enzyme.

LA-42h at 1 mM concentration was treated with 1.7 equivalents of **II** at pH 5.9, 1.3 equivalents of **V** at pH 8, and 10 equivalents of **IV** at pH 5.9. HPLC analysis and MS analysis of the reaction product showed that the desired functionalized folded four-helix bundle was obtained in 30% yield (Figure 4). The reaction is therefore of general applicability.

Discussion

A four-helix bundle protein scaffold ensures water solubility for covalently linked groups, tremendous versatility in design, and the opportunity to present a wide variety of functionalities in a large number of well-defined geometries. In spite of these

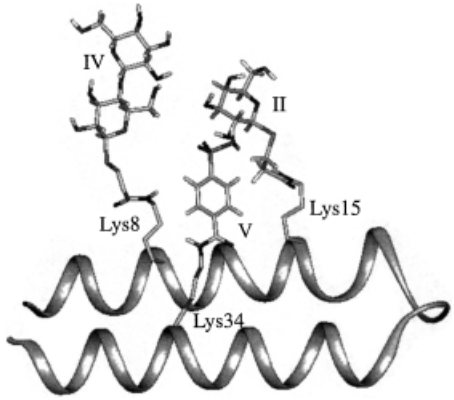


Figure 4. *The modeled structure of LA-42h functionalized with galactose, cellobiose, and benzenesulfonamide groups.*

advantages, designed proteins have not been used to any great extent as scaffolds in supramolecular chemistry or for the purpose of studying biomolecular interactions, most likely because of our poor understanding of how to construct new proteins, the so-called protein folding problem. Major advances in protein design have recently led to a number of nonnatural sequences that fold according to prediction but in order to exploit these novel scaffolds methods must be available that make it possible to introduce new functions easily and efficiently by design. While solid-phase peptide synthesis methods have been developed to a high level of efficiency and selectivity, the introduction of expensive groups on the solid phase remains a cost-inefficient alternative. Derivatives of amino acids that carry nonnatural functional groups require nontrivial synthesis efforts and the protected, chiral amino acid must be available in typically 2.5- to 4-fold excess over the growing polymer chain. Orthogonal protection groups have been developed that allow the site selective modification of side chains of the naturally occurring amino acids in an important development of solid phase synthesis schemes and provide the best alternative for generation of molecular diversity during solid-phase synthesis. However, self-catalyzed reactions represent a very easy alternative that allows the incorporation of very small amounts of substrate in aqueous solution under mild pH conditions and at room temperature. Little control over reaction conditions is required by the chemist because control of reactivity and selectivity resides in the amino acid sequence and the reaction can be left on the lab bench until complete. The incorporation depends on competition between polypeptide residues and spontaneous hydrolysis and the competitive situation is relatively insensitive to external disturbances.

With the advent of site-selective acylation reactions, designed proteins may be functionalized in a large number of ways by using very simple chemistry. We previously reported the mechanisms for site-selective acylation of lysine residues, directly or by a His-mediated pathway. We have now mapped the site-selectivities and determined the relative reactivities of individual lysine residues that flank the nucleophilic histidine beyond those in positions $i,i-3$, and $i,i+4$ in a helical segment reported previously. We have also probed the possibility that

residues that appear not to participate in the making and breaking of bonds, in this case serine residues, can affect the site selectivities of the functionalization reactions. The mapping of all sites that surround the His residue has provided an understanding of how to multifunctionalize the four-helix bundle in several steps by addition of active esters in a stepwise fashion.

The opportunity to address individual lysine residues to incorporate virtually any functional group that can be transformed into an active ester provides a vehicle for the construction of receptors for biomacromolecules and small organic compounds, but also for the introduction of functions beyond binding. For example, the incorporation of fluorescent dyes or attachment to solid supports paves the way for reporting and biosensing. The protein scaffold can be reacted with any functional group that is water soluble and that can be presented in the form of an active ester. We have demonstrated the use of sugar derivatives,[39] enzyme inhibitors,[42] fluorescent probes,[42] and acetyl[39] and fumaryl[27] groups but clearly peptide esters, DNA, PNA (nucleic acids with peptide backbones), and RNA derivatives, steroids, and small molecule libraries can also be incorporated to provide a large variety of receptors, catalysts, and molecular devices. Insoluble groups can be introduced on the solid phase by using orthogonal protection group strategies and when these methodologies are used in combination, the number of functionalized proteins that can be produced is limited only by the number of applications.

A critical issue is the structure of the functionalized protein as the designed function of the peptide may depend upon an unmodified fold upon functionalization. The helix – loop – helix dimer motif folds mainly as a result of hydrophobic interactions between the nonpolar residues of amphiphilic helices but it is known to dissociate at low concentrations and form unstructured monomers. However, the polypeptide KE-I, which was designed for biosensor applications, was shown by CD spectroscopy to remain highly folded even at 1 μM concentration and furthermore, it remained folded even when bound to the target protein, human carbonic anhydrase.[42] Should the need arise to have access to fully folded motifs at nanomolar concentrations, the monomeric subunits can be ligated to form a single-chain four-helix bundle by using the strategies of chemical ligation, in which case dissociation does not occur.

The sequences reported here do not represent separate scaffolds but a single one for which, to ensure optimum selectivity and yields of synthesis, some lysines are introduced while others are replaced by nonreactive residues, depending on the application. Alanine residues are introduced where a nonreactive site is required and arginine residues where pK_a depressions are of interest and to ensure solubility and a practical overall charge. Therefore the helix – loop – helix scaffold is a very versatile one, readily synthesized and functionalized. Applications can be envisioned in the biomedicinal area, in the field of proteomics, in drug development, and in functional devices such as biosensors.

The four-helix bundle that can be site-selectively functionalized is a practically useful scaffold for several reasons. There are a large number of individually addressable amino acid side chains for functionalization, which provides versatility with regards to the number and relative geometries of groups that can be incorporated. Self-catalyzed reactions are efficient and require little material to be incorporated. No protection groups are needed and the stepwise incorporation of several groups is readily achieved without intermediate purification by an automated process, for example in microtitre plate format. The scaffold controls the directionality of the substituents and the final construct is very robust. The product can be lyophilized and stored, and when dissolved in buffer it regains its inherent structure and function instantaneously.

Conclusions

We have demonstrated the simple and practical introduction of three different residues into a folded four-helix bundle polypeptide motif site-selectively and without intermediate purification in aqueous solution at pH 5.9 and pH 8. This is the first demonstration of the site-selective incorporation of three different residues into a folded protein by using only the reactivities of the naturally occurring amino acids to control reactivity and selectivity. The simple procedure required to perform the functionalization suggests that it may open the way for the use of designed proteins as scaffolds for the study of biomolecular interactions and for a number of applications. The overall yield of incorporation was shown to be 30% or better for a range of widely different substituents and the scaffold is therefore expected to be useful for the incorporation of virtually any water-soluble ester derivative.

Experimental Section

Mass spectrometry: MALDI-TOF MS analyses were performed on an Applied Biosystems Voyager DE-STR mass spectrometer with α-cyano-4-hydroxycinnamic acid as the matrix. The mass spectrometer was calibrated with calibration mixture 2, which contained angiotensin I, adrenocorticotropic hormone corticotropin-like intermediate lobe peptide (ACTH clip) 1 – 17, ACTH clip 18 – 39, ACTH clip 7 – 38, and bovine insulin. Measured molecular weights corresponded to calculated ones within 1.1 mass units in all cases and typically within 0.5 mass units. ESMS analyses were performed on a VG ZabSpec magnetic sector instrument. For LC – ESMS analyses, the peptides were eluted with a MeOH/H_2O mixture that contained 1% HOAc by using a gradient of 10 – 90% MeOH over 20 minutes on a 5μm C-18 Kromasil column with a flow rate of 75 μL min^{-}. CsI was used for calibration.

NMR spectroscopy: NMR spectra of the peptides were recorded on a Varian Inova 600 spectrometer at 308 K in H_2O/D_2O (90:10 vol%) with TFE-d_3 (4 vol%) at pH 5.1. Water suppression was accomplished by preirradiation of the water resonance. The 90° pulses were 7.5 μs for both 1D, NOESY, and TOCSY spectra, and the spinlock pulse in the TOCSY experiment was 18.5 μs with a window function of 30 μs. The mixing times were 200 ms for the NOESY experiments and 80 ms for the TOCSY experiments. 2×256 increments were recorded with 32 transients in each increment. The data were processed with linear prediction algorithms.

CD spectroscopy: CD spectra were recorded on a Jasco J-715 CD spectropolarimeter, in 0.1, 0.5, or 1 mm cuvettes in the interval from 280 – 190 nm at room temperature. Each spectrum was an average of

six scans and the background was subtracted from the spectrum before the mean residue ellipticity at 222 nm was measured. A stock peptide solution (1 mM) was prepared by dissolving the peptide in buffer under the assumption that the lyophilized peptide contains 25% water. The pH value was adjusted by addition of small amounts of 0.1 M – 1 M NaOH or HCl. The stock peptide solution was diluted to the desired concentrations for the concentration dependence studies by use of a pipette. A stock peptide solution (0.3 mM) in water was prepared for the pH dependence studies and the pH value was adjusted by small additions of HCl or NaOH. The concentration of the stock solution for the concentration dependence measurements was determined by quantitative amino acid analysis. Bis – Tris buffer (50 mM; pH 5.9) was used for all peptides except for LA-42h where Bis – Tris/Tris (100 mM; 1:1) at pH 5.9 was used.

General procedure for peptide synthesis and purification: The polypeptides were synthesized on an Applied Biosystems Pioneer automated peptide synthesizer on a 0.1 mmol scale by using standard Fmoc chemistry. A peptide – amide – linker – polyethylene glycol – polystyrene (PAL – PEG – PS; PAL is formally 5(aminomethyl-3,5-dimethoxyphenoxy) valeric acid) polymer was used with a substitution level of 0.16 – 0.23 mmol g^{-1}. A mixture of *O*-(7-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborat (TBTU; 0.5 M in dimethylformamide (DMF)) and diisopropylethylamine (DIPEA; 1 M in DMF) was used together with an excess of four equiv amino acid in each coupling. A standard coupling time of 60 minutes was used, except in the cases of Nle and Leu, where a 30-minute coupling time was used and for Gln, Arg, and Asn, where 90-minute couplings were used. The side chains of the amino acids were protected with base-stable groups: *t*butyl ester for Asp and Glu, *tert*-butoxymethyl (Boc) for Lys and Orn, trityl (Trt) for His, Asn, and Gln, and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) for Arg. The Fmoc group was removed from the amino acid by treatment with 20% piperidine in DMF. The N terminus of the peptide was capped by using a solution of acetic anhydride (0.3 M) in DMF. After the completed synthesis, the resin was rinsed with dichloromethane (DCM) and dried under vacuum. The peptide was cleaved from the resin to create an amide at the C terminus and deprotected at room temperature by treatment with a TFA/H_20/ethanedithiol/triisopropyl silane (94:2.5:2.5:1 v/v) mixture (10 mL per gram of polymer) for three hours. After filtration and concentration, the peptide was precipitated by addition of cold diethyl ether, centrifuged, then resuspended three times in diethyl ether and lyophilized.

The synthesis of LA-42b was performed as described above except that a Gly – 4-hydroxymethylphenoxyacetic acid – PEG – PS (Gly – PAC – PEG – PS) polymer was used with a substitution level of 0.17 mmol g^{-1} and no capping was performed. This polymer does not create an amide at the C terminus. The crude products were purified by RP HPLC on a semi-preparative C-8 HICHROM column, eluted isocratically with isopropanol (36 – 43%) and TFA (0.1%) in water at a flow rate of 10 mL min^{-1}. The purity was checked by analytical HPLC and one symmetric peak was found and identified by mass spectrometry.

General procedure for determination of the site selectivity of the acylation reaction: A stock peptide solution (1 mM) was prepared by dissolving the weighed peptide in Bis – Tris buffer (50 mM; pH 5.9) under the assumption that the peptide contains 25% water, and by adjusting the pH value with small additions of 1 – 2 M NaOH and HCl. A stock solution (15 mM) was also prepared of the substrate *p*-nitrophenyl fumarate in Bis – Tris buffer/AcCN (1:1 v/v). One equivalent of substrate was added to the peptide solution and in a typical acylation experiment 2 μL of the substrate solution had to be added to 30 μL of peptide solution. After three days at room temperature the reaction mixture was analyzed by RP HPLC on an analytical C-8 HICHROM column eluted isocratically with isopropanol (38 – 43%) and TFA (0.1%) in water at a flow rate of 0.6 mL min^{-1}. The resulting peaks were identified by mass spectrometry. The degree of modification was estimated from the analytical RP HPLC chromatogram as the area of each peak divided by the total area of peptide-fraction peaks. It was assumed that all peptides, modified and unmodified, have the same extinction coefficient.

Tryptic digestion: A peptide solution (1 mM) in $NH_4^+HCO_3^-$ (0.1 M; pH 8.0) was prepared and trypsin (0.5 mg) was dissolved in HCl (50 μL, 0.1 mM) and added to the peptide solution to give a final trypsin concentration of 200 g trypsin per mole peptide. After 3 hours at 37 °C the reaction was quenched by addition of HCl (20 μL) and the reaction solution was lyophilized. The resulting fragments were identified by either MALDI-TOF MS or LC – ESMS.

Kinetic measurements: The kinetic studies were performed on a Varian CARY 100 Bio UV/Visible or a CARY 5E UV-Vis-NIR spectrophotometer equipped with a CARY temperature controller. All measurements were performed at 298 K. A peptide stock solution (1 mM) was prepared in Bis – Tris buffer (50 mM; pH 5.9) and the pH value was adjusted to the correct value by addition of small amounts of 1 M NaOH and HCl. The stock solution was diluted with Bis – Tris buffer to the desired concentration (0.4 mM, 0.3 mM, or 0.2 mM) and 270 μL was transferred to a cuvette. After 15 minutes of temperature equilibration, a substrate solution (5 μL, 5 mM) was added to give a final substrate concentration of 0.1 mM. The *p*-nitrophenyl fumarate substrate was dissolved in Bis – Tris buffer/AcCN (1:1 v/v). The reaction was followed for more than 3 half-lives at 320 nm and the data was then processed with IGOR Pro software. A plot of the absorbance versus time gives the pseudo-first-order rate constant k_{obs} and the second-order rate constant was obtained by fitting a straight line to a plot of the pseudo-first-order rate constants versus the peptide concentration.

General procedure for the incorporation of three substituents: A peptide solution (1 mM) was prepared in Bis – Tris/Tris buffer (100 mM, 1:1) and the pH value was adjusted to 5.9 by small additions of 1 – 5 M NaOH and HCl. The substrate was dissolved in the same buffer/AcCN (1:1) to a final concentration of 40 mM and 1.7 – 3 equivalents were added to the peptide solution. After three days at room temperature a small sample of the reaction mixture was analyzed by analytical RP HPLC and MALDI-TOF MS. The pH value was adjusted to 8 by small additions of 1 – 5 M NaOH and HCl and the second substrate (1.3 – 3 equiv) was added. After one day at room temperature a small sample of the reaction mixture was analyzed by analytical RP HPLC and MALDI-TOF MS. The pH value was adjusted to 5.9 by small additions of 1 – 5 M NaOH and HCl and the final substrate (10 equiv) was added. When the reaction was complete, the reaction mixture was analyzed by analytical RP HPLC and MALDI-TOF MS. The sites of modification were determined by tryptic digestion followed by MS analysis.

We are indebted to the Swedish Research Council for financial support, to the Swedish NMR Centre for NMR spectrometer time, and to Charlotta Damberg for expert assistance.

[1] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Publishing, Inc., New York, **1991**.
[2] H. Lis, N. Sharon, *Eur. J. Biochem.* **1993**, *218*, 1 – 27.
[3] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683 – 720.
[4] T. Hunter, *Cell* **2000**, *100*, 113 – 127.
[5] D. J. Sweatt, *Curr. Biol.* **2001**, *11*, R391 – R394.

[6] J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neil, W. F. DeGrado, *Science* **1995**, *270*, 935 – 941.
[7] C. K. Smith, L. Regan, *Acc. Chem. Res.* **1997**, *30*, 153 – 161.
[8] B. I. Dahiyat, S. L. Mayo, *Science* **1997**, *278*, 82 – 87.
[9] C. Micklatcher, J. Chmielewski, *Curr. Opin. Chem. Biol.* **1999**, *3*, 724 – 729.
[10] L. Baltzer, H. Nilsson, J. Nilsson, *Chem. Rev.* **2001**, *101*, 3153 – 3163.
[11] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342 – 345.
[12] C. Schafmeister, S. LaPorte, L. J. W. Miercke, R. M. Stroud, *Nat. Struct. Biol.* **1997**, *4*, 1039.
[13] A. J. Maynard, M. S. Searle, *Chem. Commun.* **1997**, 1297 – 1298.
[14] R. B. Hill, W. F. DeGrado, *J. Am. Chem. Soc.* **1998**, *120*, 1138 – 1145.
[15] J. W. Bryson, J. R. Desjarlais, T. M. Handel, W. F. DeGrado, *Protein Sci.* **1998**, *7*, 1404 – 1414.
[16] T. Kortemme, M. Ramirez-Alvarado, L. Serrano, *Science* **1998**, *281*, 253 – 256.
[17] G. R. Krishna, F. Wold, in *Proteins: Analysis and Design* (Ed.: R. H. Angeletti), Academic Press, **1998**, pp. 121 – 206.
[18] D. S. Kemp, R. I. Carey, *Tetrahedron Lett.* **1991**, *32*, 2845 – 2848.
[19] P. E. Dawson, S. B. Kent, *J. Am. Chem. Soc.* **1993**, *115*, 7263 – 7266.
[20] G. Tuchscherer, *Tetrahedron Lett.* **1993**, *34*, 8419 – 8422.
[21] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, 776 – 779.
[22] O. Melnyk, J.-A. Fehrentz, J. Martinez, H. Gras-Masse, *Biopolymers* **2000**, *55*, 165 – 186.
[23] S. J. Bark, S. Schmid, K. M. Hahn, *J. Am. Chem. Soc.* **2000**, *122*, 3567 – 3573.
[24] P. Durieux, J. Fernandez-Carneado, G. Tuchscherer, *Tetrahedron Lett.* **2001**, *42*, 2297 – 2299.
[25] C. Grandjean, H. Gras-Masse, O. Melnyk, *Chem.-Eur. J.* **2001**, *7*, 230 – 239.
[26] J. P. Tam, Y. Lu, C. Liu, J. Shao, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12485 – 12489.
[27] L. Baltzer, A.-C. Lundh, K. Broo, S. Olofsson, P. Ahlberg, *J. Chem. Soc., Perkin Trans. 2* **1996**, 1671 – 1676.
[28] L. K. Andersson, M. Caspersson, L. Baltzer, *Chem.-Eur. J.* in press.
[29] K. Broo, A.-C. Lundh, P. Ahlberg, L. Baltzer, *J. Am. Chem. Soc.* **1996**, *118*, 8172 – 8173.
[30] C. Tanford, *Adv. Protein Chem.* **1962**, *17*, 69 – 165.
[31] K. Broo, M. Allert, L. Andersson, P. Erlandsson, G. Stenhagen, J. Wigström, P. Ahlberg, L. Baltzer, *J. Chem. Soc., Perkin Trans. 2* **1997**, 397 – 398.
[32] S. Olofsson, G. Johanson, L. Baltzer, *J. Chem. Soc., Perkin Trans. 2* **1995**, 2047 – 2056.
[33] K. S. Broo, L. Brive, P. Ahlberg, L. Baltzer, *J. Am. Chem. Soc.* **1997**, *119*, 11362 – 11372.
[34] L. Andersson, G. Stenhagen, L. Baltzer, *J. Org. Chem.* **1998**, *63*, 1366 – 1367.
[35] K. S. Broo, H. Nilsson, J. Nilsson, A. Flodberg, L. Baltzer, *J. Am. Chem. Soc.* **1998**, *120*, 4063 – 4068.
[36] S. Olofsson, L. Baltzer, *Folding Des.* **1996**, 347 – 356.
[37] D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **1991**, *222*, 311 – 333.
[38] M. Elofsson, B. Walse, J. Kihlberg, *Tetrahedron Lett.* **1991**, *32*, 7613 – 7616.
[39] L. K. Andersson, G. T. Dolphin, J. Kihlberg, L. Baltzer, *J. Chem. Soc., Perkin Trans. 2* **2000**, 459 – 464.
[40] M. Elofsson, S. Roy, B. Walse, J. Kihlberg, *Carbohydr. Res.* **1993**, *246*, 89 – 103.
[41] A. E. Eriksson, P. M. Kylsten, T. A. Jones, A. Liljas, *Proteins: Struct., Funct., Genet.* **1988**, *4*, 283 – 293.
[42] K. Enander, G. T. Dolphin, L. K. Andersson, B. Liedberg, I. Lundström, L. Baltzer, *J. Org. Chem.* **2002**, *67*, 3120 – 3123.

Received: October 2, 2002
Revised: May 16, 2002 [Z 302]