Control of Lysine Reactivity in Four-Helix Bundle Proteins by Site-Selective pK_a Depression: Expanding the Versatility of Proteins by Postsynthetic Functionalisation

Linda K. Andersson,[a] Maud Caspersson,[a] and Lars Baltzer*[b]

Abstract: Five 42-residue polypeptides have been designed to fold into hairpin helix – loop – helix motifs that dimerise to form four-helix bundles, and to serve as protein scaffolds for the elucidation at the molecular level of the principles that control and fine-tune lysine and ornithine reactivities in a protein context. Site-selective control of Lys and Orn reactivity provides a mechanism for addressing directly individual residues and is a prerequisite for the site-selective functionalisation of folded proteins. Several lysine and one ornithine residues were introduced on the surface and in the hydrophobic core of the folded motif. The reactivity of each residue was determined by measuring the degree of acylation of the trypsin cleaved fragments by HPLC and mass spectrometry. The most reactive residues were Orn34 and Lys19, both of which were located in d positions in the heptad repeat, and therefore in hydrophobic environments. Upon reaction of the helix – loop – helix dimer KA-I with one equivalent of mono-*p*-nitrophenyl fumarate, Orn34 was acylated approximately three times more efficiently than Lys19, whereas Lys10 (b position), Lys15 (g position), and Lys33 (c position) remained unmodified. In the sequence KA-I-A_{15} Lys15 was replaced by an alanine residue and the selectivity of Orn34 over Lys19 increased to approximately a factor of six, probably because Lys15 had the capacity to reduce the pK_a value of Lys19 and 85% of site-selectively monoacylated product was obtained. The pH dependence of the acylation reaction was determined and showed that the pK_a of the reactive residues were 9.3, more than a pK_a unit below the magnitude of the corresponding residue in a solvent exposed position. Introducing Lys and Orn residues into a or d positions of the heptad repeat therefore serves as a mechanism of depressing their pK_a to increase their reactivity site selectively. Extensive NMR and CD spectroscopic analyses showed that the sequences fold according to prediction.

Keywords: de novo design · helical structures · lysine reactivity · peptides · site-selective functionalisation

Introduction

The functional richness of proteins, based on artificial as well as the naturally occurring amino acids, is now beginning to be explored in designed proteins for purposes of catalysis and binding.[1–4] The magnitude of the available binding energy that arises from charge – charge, hydrogen bonding and hydrophobic interactions in aqueous solution[5a] as well as the capacity of polypeptides to form a wide range of well-defined tertiary structures make proteins unrivalled molecular scaffolds for probing and exploiting molecular recognition and interactions.

Functional diversity beyond that of folded linear sequences is created in native proteins by enzyme-mediated posttranslational modifications, where site-specific phosphorylations[6, 7] and glycosylations[8, 9] are key events in signal transduction, energy storage, immune responses, and protein folding. The understanding of how to functionalise folded proteins by controlled reactions would enhance also the repertoire of designed proteins but chemical methods for site-selective functionalisation of folded proteins have significant limitations. Substituted thiols can be incorporated if there are Cys residue in the sequence, but if there is more than one cysteine residue the sites of incorporation are statistically controlled and site selectivity is not achieved. Lack of site selectivity characterises all methods for protein labelling in classical protein chemistry.[5b] Chemoselective targeting of artificial amino acids[10–13] is an attractive approach particularly in combination with recent advances in protein synthesis through chemical ligation.[13, 14] Template-assisted synthesis

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

[b] Prof. L. Baltzer
Department of Chemistry, Linköping University
58183 Linköping (Sweden)
Fax: (+46) 13-122587
E-mail: lars.baltzer@ifm.liu.se

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

  0947-6539/02/0816-3687 $ 20.00+.50/0

of proteins (TASP) combines chemoselectivity and orthogonal protection group strategies with the concept of a peptide scaffold for the formation of designed topologies and self-assembled protein structures.[15]

Chemical approaches for the site-selective functionalisation of folded proteins based on the reactivities of the naturally occurring amino acids are attractive as they provide opportunities for using the powerful methods of molecular biology for selection and screening. Lessons learnt from the modification of model proteins are likely to be of use in functionalising biologically relevant proteins that cannot be synthesised. Self-catalysed reactions are very efficient with regards to the incorporation of functional groups and the cost and effort of introducing expensive substituents are considerably less than by synthetic routes, because the amounts needed in the direct reaction with folded proteins are much smaller. We have previously reported on a His-Lys mediated site selective functionalisation reaction where the side chains of flanking lysine residues in a designed four-helix bundle protein were acylated at pH 5.9 in aqueous solution upon reaction of the peptide with activated esters.[16–18] In the first, and rate-limiting step of the reaction the unprotonated form of the histidine attacks the ester to form an acyl intermediate. The acyl group is then transferred to the flanking lysine in a fast intramolecular reaction and an amide is formed at the lysine side chain. If several lysine residues are available then at low pH, the ones that flank His residues are acylated, whereas those that are far from His residues remain unmodified. If there is more than one lysine in close proximity to the His residue the site of modification is determined by intramolecular competition between the flanking lysines.[19, 20] This is to our knowledge the only known reaction that is based exclusively on the reactivity of the naturally occurring amino acids that will permit the site-selective incorporation of several substituents in a controlled fashion into folded proteins. An understanding of the reactivity of surface exposed amino acids, both in terms of how to control the pK_a of ionizable residues and in terms of how to obtain cooperativity between groups of amino acids, made it possible to develop this functionalisation strategy.

Unprotonated lysine residues are more efficient nucleophiles than unprotonated His residues[21] and in order to enhance the repertoire of the four-helix bundle protein scaffold we have investigated strategies for controlling lysine pK_a values to explore the site selectivity of the direct acylation of Lys residues. Here we wish to report on the successful exploitation of the properties of four-helix bundle proteins structurally developed because of hydrophobic interactions between amphiphilic helices. The microenvironment provided by residues designed to form the hydrophobic core was probed with regards to its effect on the ionization constants of lysine residues. The results shed considerable light on the factors responsible for lysine reactivity in surface exposed positions and form the basis for a new strategy for the site-selective functionalisation of folded proteins in aqueous solution.

Results

Design and structure: Five 42-residue sequences have been designed to fold into helix – loop – helix motifs and dimerise to form four-helix bundles, see Figure 1, Table 1. They were

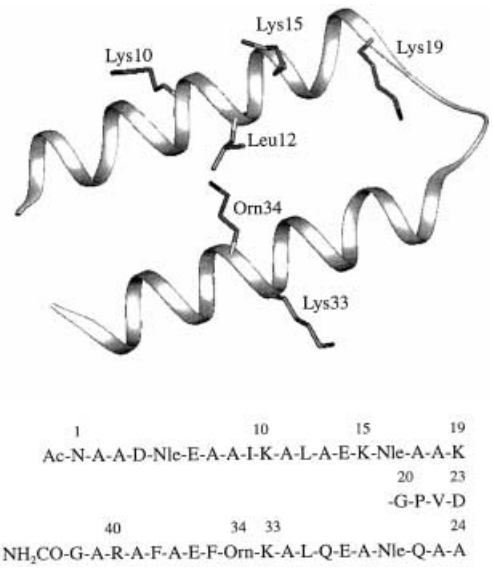


Figure 1. Modelled structure and amino acid sequence of KA-I. The side chains involved in the acylation reaction are shown together with all possible acylation positions. The active peptide is the dimer but for reasons of clarity only the monomer is shown. The one letter code for the amino acids is used, where A is Ala, D is Asp, E is Glu, F is Phe, G is Gly, H is His, I is Ile, K is Lys, L is Leu, N is Asn, P is Pro, Q is Gln, R is Arg, V is Val, Nle is norleucin and Orn is ornithine.

Table 1. The amino acid sequences for the peptides designed for determination of the site selectivity of direct acylation. The derived sequences are based on that of KA-I and the positions not shown here are the same as those of KA-I.

Peptide	10	12	13	15	19	33	34
KA-I	K	L	A	K	K	K	Orn
KA-I-A_{15}	K	L	A	A	K	K	Orn
KA-I-A_{33}	K	L	A	K	K	A	Orn
KA-I-R_{19}	K	L	A	K	R	K	Orn
KA-II	K	K	L	K	R	K	A

synthesised on the solid phase using the Fmoc protection group strategy. The peptides were purified by reversed-phase (RP) HPLC and identified by MALDI-TOF mass spectrometry (MS). The design of the peptides reported here was based on those of the peptide sequences SA-42,[22] KO-42,[23] and LA-42b[19] that have been described previously in detail. In short, these peptides consist of 42 amino acid residues designed to fold in aqueous solution into two amphiphilical helical segments connected by a short loop and dimerise to form four-helix bundles. The amino acid residues of the helical sequences were selected according to their propensity for helix formation; the helices were further stabilised by the introduction of residues capable of interacting with the partial charges of the helix dipole moment, residues capable of salt

0947-6539/02/0816-3688 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 16

bridge formation and residues capable of N- and C-terminal capping.[24] The residues introduced for their structure stabilising properties were conserved in the design of the polypeptides reported here (Figure 1, Table 1), with the minor exception of Lys19 that was replaced by Arg19 in the peptide KA-I-R_{19}. The structure formation is mainly driven by hydrophobic interactions between the hydrophobic faces of the amphiphilic helices; the helical content is lost almost completely as the dimers dissociate at low concentration to form monomers.[25]

The design of the four-helix bundle motif is best described in terms of the heptad repeat pattern $(a-b-c-d-e-f-g)_n$, Figure 2. In an antiparallel four-helix bundle the a and

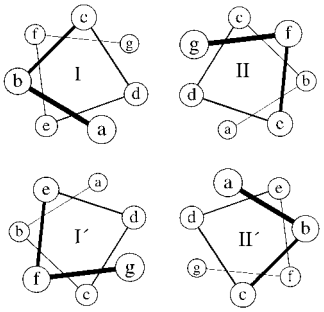


Figure 2. Schematic representation of interactions between amphiphilic helices based on the pattern of the heptad repeat. The dimeric structure is folded in an anitiparallel mode and the hydrophobic core consists of the residues in a and d positions.

d positions are the nonpolar amino acids that form the hydrophobic core, the g and c positions form the "top" and "bottom" surfaces and the b and e positions are at the interface between the hairpin subunits. Several of the solvent exposed residues in the b, c, e, f and g positions are charged and polar to ensure amphiphilicity and solubility. The a and d residues that make up the hydrophobic cores of SA-42, LA-42b, and KO-42 are identical and these residues were also used in the sequences of the peptides reported here, with the exception of KA-II, for reasons explained in detail below. The sequence homology between the sequences KA-I, KA-I-A_{15}, KA-I-A_{33}, and KA-I-R_{19} and that of LA-42b is more than 95% in each case and the sequence homology between KA-II and LA-42b is more than 88%. Several Lys and one Orn residues were introduced in different positions in order to probe the relationship between site and reactivity and to determine the mechanism of acylation.

The sequence of KA-I is the same as that of LA-42b with the exception that His11 was replaced by Ala11 (see Figure 1), to avoid histidine-mediated acylation of lysine residues. In order to determine the hierarchy of reactivities of the Lys and Orn residues of KA-I three peptides were designed in which one Lys was replaced by an Ala or an Arg residue. The sequence KA-I-A_{15} was the same as that of KA-I with the exception that Lys15 was replaced by Ala15. In the sequence KA-I-A_{33} Lys33 was replaced by Ala33 and in the sequence KA-I-R_{19} Lys19 was replaced by Arg19.

To determine whether the enhancement of reactivity observed for Orn34, due to its location in the hydrophobic core d position, was of general applicability, the polypeptide KA-II was designed with an amino acid sequence that was the same as that of KA-I except that Orn34 was replaced by Ala34, Leu12 was replaced by Lys12, Ala13 was replaced by Leu13 and Lys19 was replaced by Arg19. KA-II was designed to probe whether position 12, a d position, would become more reactive than any surface exposed position due to its hydrophobic environment. In order to avoid competition between hydrophobic core positions, Orn34 was replaced by Ala34, to ensure that only one Lys or Orn residue was in an a or a d position.

The structures of SA-42, LA-42b and KO-42 were determined, previously, by NMR and CD spectroscopy[19, 22, 23, 26] and the states of aggregation of SA-42 and KO-42 were determined by analytical ultracentrifugation. The polypeptides were found to be highly helical; the helical contents were determined from the mean residue ellipticities at 222 nm, $[\Theta]_{222}$. The negative deviation of α-proton chemical shifts from those of random coils, and medium-range NOE values were used to identify helical segments and positive chemical shift deviations were used to identify loop regions. Long-range NOE values were observed that were only compatible with helix–loop–helix hairpin formation and antiparallel hairpin dimerisation. The concentration dependences of $[\Theta]_{222}$ showed that the peptides aggregated to form dimers or higher states of aggregation. In the case of SA-42 and KO-42 the sedimentation equilibrium showed that a dimer was formed with only a slight tendency for further aggregation at close to, and above mM concentrations. We assume this to be true also for other, similar sequences. High-resolution NMR structures could, however, not be obtained since the hydrophobic cores were disordered and in fast exchange on the NMR time scale. A solid understanding of the structural features of these folded helix–loop–helix dimers was nevertheless obtained and the dominant fold was shown in each case to be in good agreement with the designed structure. In spite of the fact that the peptides in this investigation are highly homologous to the previously designed sequences, a determination of the structure of KA-I was undertaken by CD and ^{1}H NMR spectroscopy.

The CD spectrum of KA-I showed the signature characteristic of an α-helical protein with minima at 222 and 208 nm. The mean residue ellipticity at 222 nm was $-22\,200$ deg cm^2 dmol^{-1} at pH 5.9 in 50 mM Bis-Tris buffer. The ^{1}H NMR spectrum of KA-I was assigned from the TOCSY and NOESY spectra recorded in H_2O:D_2O (90:10 *v*/*v*) containing 4 vol% of [D_3]TFE at 308 K using methods described previously.[26] Extensive studies of the effect of small amounts of trifluoroethanol (TFE) on the NMR spectrum and the solution structure of SA-42[26] showed that the NH exchange rates, and the resonance line widths were reduced without changing the overall fold of the folded dimer.

The spin systems of most of the amino acids of KA-I (5–22, 24–42) were identified from the TOCSY spectrum. The *cis*–*trans* equilibrium of Pro21 in the loop region gives rise to a conformational equilibrium involving some of the flanking residues and a subset of spin systems with low intensity was therefore observed for residues in, and close to, the loop sequence. The sequential assignment was obtained from the

  0947-6539/02/0816-3689 $ 20.00+.50/0

NH – NH connectivities determined from the NOESY spectrum. The deviations of the αH chemical shifts from tabulated values of random coil conformations[27] are readily used for identification of secondary structure, and the sequences from Nle5 to Ala17 and from Gln26 to Arg40 were identified as helical regions (Figure 3). Medium-range NOEs typical of α-

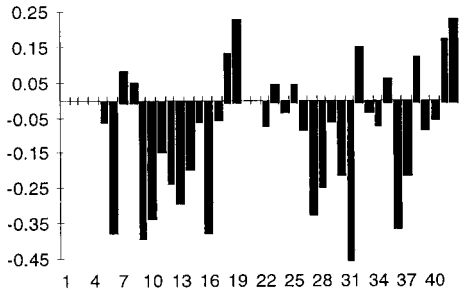


Figure 3. The αH chemical shift deviations from random coil values of the amino acid residues in KA-I. Helical segments are indicated by negative (upfield) shifts and shown between residues 5 and 17, and between residues 26 and 40. The positive or negligible shifts between residues 18 and 25 suggest a disordered structure and correspond to the residues of the loop region.

helix formation (αH-NH i, i+3 and i, i+4) were found in the sequence from Ala8 to Lys19 in helix I and from Ala24 to Gly42 in helix II, see Supporting Information. Positive chemical shift deviations, or the absence of deviations from random coil shifts were observed in the sequence from Ala18 to Ala25 which was concluded to form a disordered loop. NOE connectivities between the aromatic protons of the phenylalanine residues and the methyl groups of Ile9, Leu12 and Leu31 showed that the peptide formed a hairpin helix – loop – helix motif. The observation of NOE connectivities between Phe35 and Nle16, and between Nle5 and Lys19, suggested that the peptide dimerised in an antiparallel mode to form the four-helix bundle. This is because NOEs only arise between nuclei separated by less than 5 Å[28] and the distances between Phe35 and Nle16, and between Nle5 and Lys19, are too long to give rise to NOEs within the monomer and too long to give rise to NOEs in a parallel dimer. The helical content of KA-I was found to be constant in the pH range 4.5 – 9 and in the concentration range from 0.2 – 1 mM, Figure 4a, b, under the conditions used for the functionalisation reactions. We conclude that KA-I forms an antiparallel dimer of helix – loop – helix hairpin motifs, and that residues in a and d positions are positioned in the hydrophobic core, whereas all other residues are solvent exposed.

The secondary structures of the peptides KA-I-A_{15}, KA-I-A_{33}, KA-I-R_{19} and KA-II were investigated by CD spectroscopy. The mean residue ellipticities at 222 nm at 1 mM concentrations were in the range from -21000 to -25500 deg cm^2 dmol^{-1} at pH 5.9 and pH 8.0, Table 2. Complete NMR spectroscopic analyses were not carried out for these peptides due to the high degree of sequence homologies with KA-I but qualitative evaluations of structures were obtained from their one-dimensional ^{1}H NMR spectra (Figure 5). The ^{1}H NMR spectra of native proteins in contrast to those of random coil conformations are characterised by large chemical shift dispersions. For folded polypeptides that are partly disordered the chemical shift dispersion is intermediate between those of native proteins and random coils but resonances are severely broadened in comparison with those of both native proteins and random coils. An inspection of the one-dimensional spectra of closely related sequences therefore provides a qualitative view of whether substantial structural changes have occurred as a result of the sequence modification. The 600 MHz NMR spectra of all four polypeptides at 1 mM concentration in 4 vol% [D_3]TFE in H_2O:D_2O (90:10 v/v) at pH 5.2 and 308 K offered no indications that the sequence modifications have affected the polypeptide fold. The line widths and chemical shift dispersions were largely unaffected.

The effects of acylation on the structures of the polypeptides KA-I-Fum and KA-II-Fum_{12} were probed by inspection

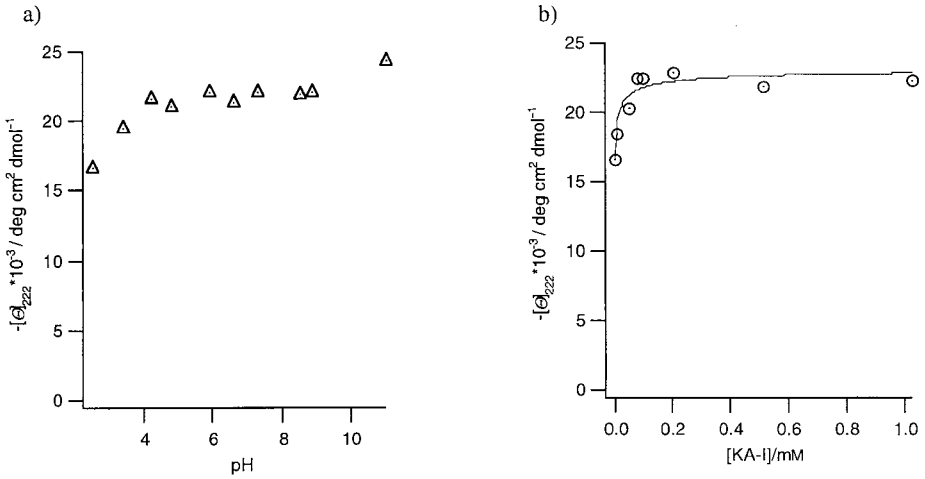


Figure 4. The pH (a) and concentration (b) dependence of the mean residue ellipticity at 222 nm of KA-I. At low pH and low concentration the dimer dissociates to form monomers with low helical contents.

of their ^{1}H NMR and CD spectra (see Figure 6 and Table 2). The incorporation of fumaryl residues was found not to influence the structures of KA-I and KA-II, significantly, in the qualitative sense described above.

The acylation reaction: The peptides were treated with mono-*p*-nitrophenyl fumarate (**1**) at room temperature in aqueous solution at pH 5.9 and at pH 8.0. The reaction mixtures were purified by RP HPLC and the reaction products were shown to be polypeptides amidated by fumaric acid and identified as unmodified, mono- or dimodified peptides by MALDI-TOF

  0947-6539/02/0816-3690 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 16

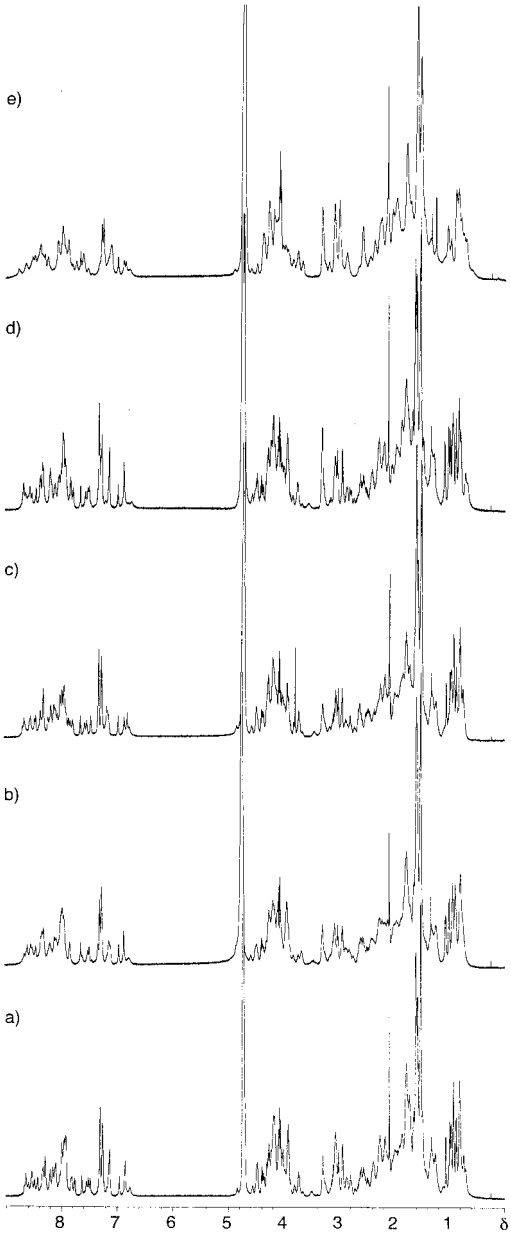


Figure 5. One-dimensional 600 MHz ^{1}H NMR spectra of a) KA-I, b) KA-I-A_{15}, c) KA-I-A_{33}, d) KA-I-R_{19}, e) KA-II recorded in H_2O:D_2O (90:10 *v/v*) with 4 vol% [D_3]TFE added at 308 K and pH 5.2. The chemical shift dispersions and resonance line widths of spectra a–d are comparable for the series of polypeptides; this suggests in a qualitative sense that the folds are not affected to any large extent by the sequence modifications. In KA-II residues in the hydrophobic core have been modified and as a consequence the resonances have become broadened and less dispersed.

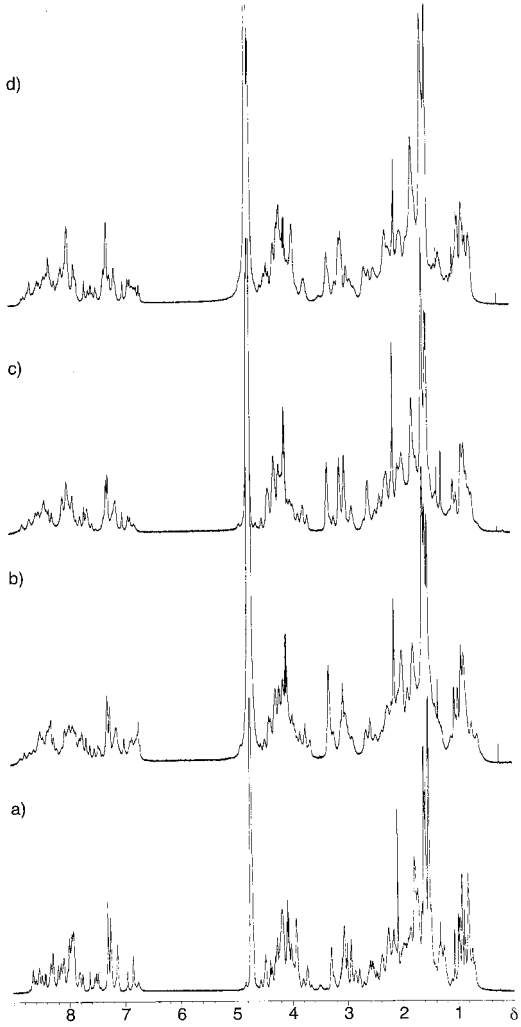
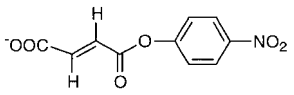


Figure 6. 600 MHz ^{1}H NMR spectra of a) KA-I, b) KA-I-Fum, c) KA-II, d) KA-II-Fum_{12} recorded in H_2O:D_2O (90:10 *v/v*) with 4 vol% [D_3]TFE added at 308 K and pH 5.2. The effect of introducing substituents into the folded peptide on the chemical shift dispersion and line width is small.



1

Table 2. The mean residue ellipticity at 222 nm $[\Theta]_{222}$ at pH 5.9 and 8.0.

Peptide	$[\Theta]_{222}$ at pH 5.9 [deg cm^2 dmol^{-1}]	$[\Theta]_{222}$ at pH 8.0 [deg cm^2 dmol^{-1}]
KA-I	−22200	−24100
KA-I-Fum	−23600	–
KA-I-A_{15}	−23700	−25400
KA-I-A_{33}	−22300	−24800
KA-I-R_{19}	–	−24700
KA-II	−22400	−21200
KA-II-Fum	−23700	–

MS. The observed molecular weights corresponded well to the calculated values. Typically the difference was less than 0.5 mass units, and in no case was it larger than 1.1 mass units. No other modifications were detected but some unreacted starting material remained as the ester **1** was partly consumed by competing background hydrolysis to form fumaric acid and *p*-nitrophenol. The reactivities of individual lysine residues were monitored by measuring the degree and position of acylation in the reaction between one equivalent of **1** and each of the polypeptides at pH 5.9 and at pH 8.0 at room temperature (Table 3). The degrees of acylation were estimated by integration of the peaks in the analytical RP HPLC chromatogram; the sites of modification in each peptide were identified by MALDI-TOF MS after tryptic digestion of the modified

  0947-6539/02/0816-3691 $ 20.00+.50/0

peptides. The degree of acylation of each position of the polypeptides are presented in Table 3.

At pH 8.0, 66% of KA-I was monoacylated according to the accurate integration of the HPLC chromatogram, corresponding to 90% of the total amount of acylated peptide since some polypeptide remained unmodified. KA-I was monoacylated at position 34 as well as at position 19, both of which were d positions, but the two monoacylated peptides could not be separated by RP HPLC; their relative amounts had to be roughly estimated from the intensities of the tryptic fragments in the mass spectrum. The fraction of KA-I that was monoacylated at Orn34 provided approximately 75% of the tryptic fragments from monoacylated peptide, and the fraction of KA-I monoacylated at Lys19 provided approximately 25%. An analogous analysis of the acylation pattern of KA-I-A_{15}, suggested that approximately 85% of the monoacylated products were acylated at Orn34, and only 15% were acylated at Lys19, demonstrating a high degree of site selectivity.

Table 3. Site of modification and estimated degree of mono- and dimodification at pH 5.9 and 8.0 after reaction with one equivalent of **1**. Due to the competing background hydrolysis all peptides are not modified after the reaction. The relative yields, the mono- or dimodified peptide divided by the total amount of modified peptide, are given in brackets.

Peptide	Site of modification	Modification	Degree of modification pH 5.9		pH 8.0	
KA-I	34 and 19	mono	45	(94)	66	(90)
	34 and 19	di	3	(6)	7	(10)
KA-I-A_{15}	34 and 19	mono	36	(100)	73	(94)
	34 and 19	di	–	–	5	(6)
KA-I-A_{33}	34	mono	32	(64)	45	(61)
	15	mono	13	(26)	22	(30)
	34 and 15	di	35	(6)	7	(9)
KA-I-R_{19}	34	mono	42	(93)	58	(85)
	15	mono	3	(7)	8	(12)
	34 and 15	di	–	–	2	(3)
KA-II	12	mono	27	(75)	57	(70)
	15	mono	9	(25)	19	(23)
	12 and 15	di	–	–	5	(6)

The mechanism of the acylation reaction: The degree of mono- and dimodification of KA-I was studied as a function of pH as depicted in Figure 7. One equivalent of **1** was reacted with KA-I at 8 different pH values in the range from 5.9 to 10.2. The reaction mixtures were analysed by analytical RP HPLC and the degrees of modification were determined from the resulting peaks after identification by MALDI-TOF MS. Monomodification of either Orn34 or Lys19 dominates but at high pH dimodification of both residues is significant. No other lysines were modified and it appears that the pK_a values of Lys19 and Orn34 are significantly lower than those of the competing lysine residues. The degree of modification was also determined in aqueous solution at pH 8.0 as a function of the amount of TFE added (see Figure 8), to probe the role of the folded structure on lysine reactivity. The overall degree of modification was reduced substantially already at low fractions of TFE. The degree of monomodification was reduced by a factor of two at 2 vol% TFE, and to less than one third of the value in aqueous solution at 5 vol% TFE. TFE is known to denature proteins by strengthening helical sequences and disrupting hydrophobic interactions.[29] It is also expected to decrease the pK_a of primary amines through a solvent effect, and to decrease the rate of chemical reactions that go through charged transition states. The decreased incorporation of fumaryl substituents was therefore not attributable to a single factor.

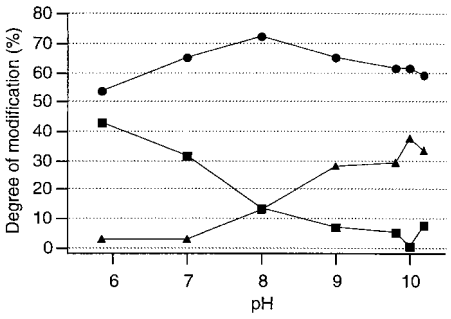


Figure 7. The pH dependence of the degree of modification of KA-I after reaction between KA-I and one equivalent of **1**. Unmodified (■), monomodified (●) and dimodified (▲) peptides are observed, after compensation for background hydrolysis.

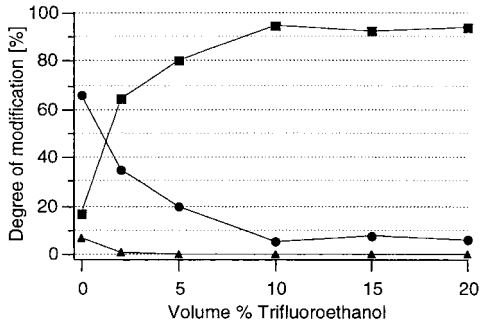


Figure 8. The TFE dependence of the degree of modification of KA-I after reaction between KA-I and one equivalent of **1** at pH 5.9 and room temperature. Unmodified (■), monomodified (●) and dimodified (▲) products were determined.

The degrees of mono- and dimodification of KA-I and KA-II were also studied as a function of the number of equivalents of **1** added. At pH 5.9 and at pH 8.0 one to ten equivalents of **1** were added to a buffered solution of KA-I and the resulting reaction mixtures were analysed by RP HPLC (Figure 9 a, b). The site selectivity was high at low pH, where more than 70% of monomodification was obtained after addition of five equivalents of **1**, and the only other modification was the introduction of two fumaryl groups. At pH 8.0 70% of monomodification was observed after the addition of 1.3 equivalents. At pH 8.0 one to ten equivalents of **1** were also added to a buffered solution of KA-II and the resulting reaction mixtures were analysed by RP HPLC, Figure 9c. The selectivity was very high as the fraction of monomodified KA-II approached 70% for the addition of 2–3 equivalents of substrate, showing that Lys12, a d position, in KA-II was the most reactive Lys residue.

The second-order rate constant for the reaction between KA-I and **1** was determined at 308 K as a function of pH from

  0947-6539/02/0816-3692 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 16

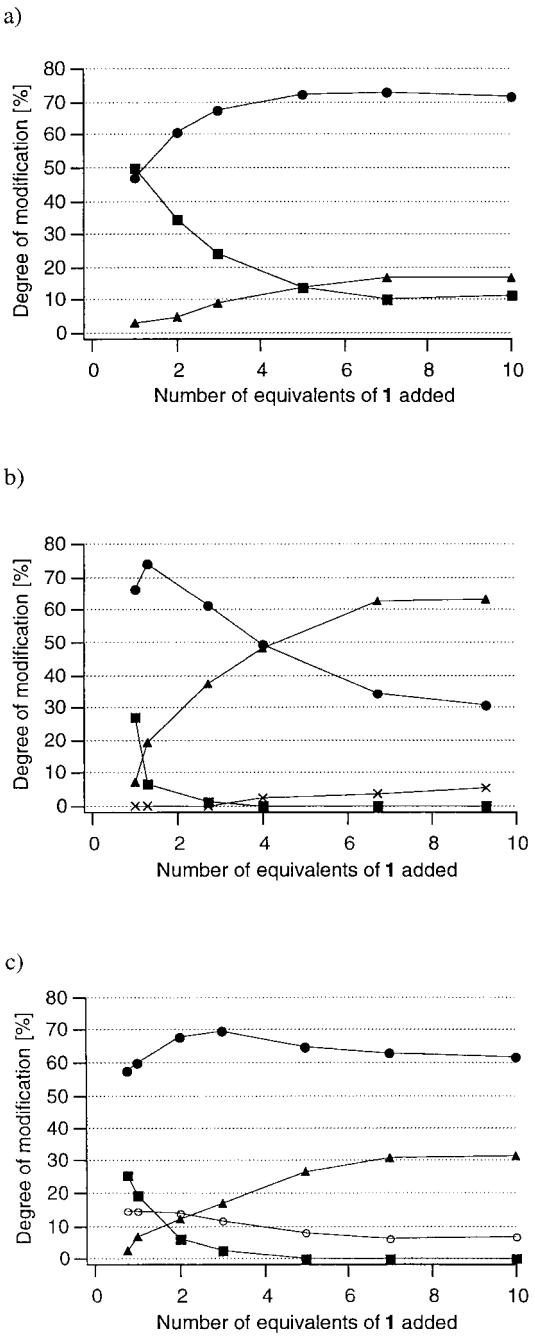


Figure 9. The degree of modification of (a) KA-I at pH 5.9, (b) KA-I at pH 8.0 and (c) KA-II at pH 8.0 as a function of the number of equivalents **1** added. Unmodified (■), monomodified (●), dimodified (▲), and trimodified products (×) were observed. ● denote the monomodification of Lys12, ○ denote the monomodification of Lys15.

pH 5.1 to 10.2. The plot of the logarithm of the second-order rate constant versus pH shows a linear dependence on pH in the range from 5.1 to slightly less than 9, where the plot levels off and approaches pH independence at pH 10.2, Figure 10a. The pH profile shows that the reaction is dependent on an amino acid in its unprotonated form. A function describing the dissociation of a monoprotonic acid with a pK_a of 9.3 gave the best fit to the experimental results as shown in Figure 10b. The expected pK_a of a lysine side chain in a solvent exposed position is 10.4.[30]

Discussion

The understanding of the reactivity of amino acid residues exposed on protein surfaces is a prerequisite for the engineering of protein scaffolds capable of site selective self-functionalisation. We have previously reported on factors that control histidine reactivity on the surface of a four-helix bundle scaffold[31] and His–Lys pairs used for self-functionalisation by forming thermodynamically stable amide bonds at the side chains of Lys residues.[16–18] The possibility of addressing Lys residues directly considerably broadens the scope of protein self-functionalisation as more groups can be introduced over a larger protein surface area and because the site selectivity may be enhanced when used in combination with the His mediated acylation reaction. The control of lysine reactivity also has important implications in catalysis as primary amines are directly involved in imine formation and thus in aldol condensations and Michael additions. While the measurements of His reactivity depended on the direct determination of pK_a values by NMR methods, the corresponding measurement of lysine dissociation constants is difficult mainly because changes in the observed chemical shifts are not directly related to the state of protonation. At the high pH necessary to significantly deprotonate a Lys side chain the structure of the folded polypeptide also changes due to the loss of stabilising charge–charge interactions between protonated Lys residues and negatively charged aspartate and glutamate residues, and possibly also due to the drastic change in overall charge. For this reason chemical shift changes are not only a result of dissociation but also a result of changes in structure. The investigation of lysine reactivity is therefore based on the determination of lysine acylation, in reactions between the folded polypeptides and active esters that lead to irreversible amide bond formations at the lysine side chains, readily detectable by mass spectrometry. In addition to the increased understanding of protein reactivity that is expected from the identification of the principles that control lysine acylation, the reactions reported here provide an efficient strategy for the site-selective functionalisation of the folded four-helix bundle protein in aqueous solution. It provides the opportunity to introduce multiple substituents stepwise into the folded protein without any need for protection groups or artificial amino acids.

The quantification of reactivity at a pH that is lower than the pK_a of the reactive residue requires that the Brønsted coefficient β is known. The state of protonation of a basic residue depends on the relationship between the pH and the pK_a according to the Henderson–Hasselbalch equation, and the relationship between reactivity and pK_a follows the Brønsted equation, where the coefficient β is unique for each reaction. The coefficient β for the reaction between primary amines and *p*-nitrophenyl acetate is 0.8[21] and at a pH lower than the pK_a values of the Lys and Orn residues the one with the lowest pK_a will therefore be the most efficient nucleophile. Because the reaction product in the direct reaction between a primary amine and an ester is an amide, the reaction is irreversible and the degree of acylation in a competition experiment is a direct measure of the difference in reactivity between the competing residues. At pH values above those of the pK_a the relative reactivities will be

  0947-6539/02/0816-3693 $ 20.00+.50/0

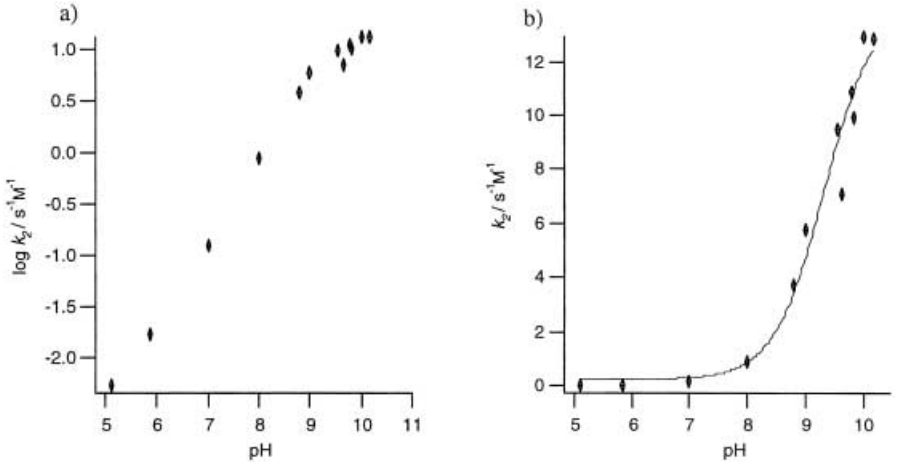


Figure 10. The pH dependence of the logarithm of the second order rate constant (a) and the pH dependence of the second order rate constant (b) for the reaction of *p*-nitrophenyl fumarate, **1**, with KA-I in aqueous solution at 298 K. The solid line in (b) represents an equation describing the dissociation of a monoprotonic acid. The best fit was obtained for a pK_a of 9.3.

reversed since the primary amines with the highest pK_a values will be the most efficient nucleophiles, but protein chemistry at such high pH has little practical use.

The sequence KA-I contains four lysine and one ornithine residues; the structure of the folded polypeptide was determined by NMR and CD spectroscopy. It folds into a helix–loop–helix motif that dimerises to form a four-helix bundle and the residues can be considered to follow the pattern of the heptad repeat. The experimental system was designed to provide a comparison between the reactivities of several different lysine positions and four modified sequences, Table 1, where lysines were replaced by alanine or arginine residues, were synthesised for comparison of their reactivities with that of KA-I. Their structures were studied to a lesser extent than that of KA-I, but the results from a qualitative NMR and CD spectroscopic analysis, and the fact that they are highly homologous to the sequence of KA-I, support the assumption that they, too, form four-helix bundles. In terms of the heptad repeat pattern, Figure 2, Lys10 is in a b position, Lys 15 is in a g position and Lys33 is in a c position, all positions exposed to the solvent. Lys19 is formally in a d position although it was initially incorporated as a cap for the C-terminus of helix I.[22] The chemical shift deviation of the αH of Lys19, Figure 3, suggests that it is not in a helical conformation but in a more disordered loop. The observation of fingerprint NOE connectivities between Lys19 and Lys15, and between Lys19 and Nle16 (see Supporting Information), however, supports the conclusion that is in fact in a helical conformation, at least part of the time. Lys19 therefore has to be considered to be in a d position. In these peptides Orn was used in position 34, a d position, for comparison with LA42b, which also contains an Orn residue in position 34. The reactivity of Orn is the same as that of Lys as they are both primary amines. The only difference between them is the number of methylene groups in the side chain, and the conclusions about Orn reactivity also apply to that of Lys. In the peptide KA-II Orn34 was replaced by an Ala and Leu12 was replaced by a Lys residue, in order to probe whether position 12 would make a lysine more reactive because of the hydrophobic microenvironment.

The peptides were reacted at room temperature with one equivalent of mono-*p*-nitrophenyl fumarate (**1**), in aqueous solution at pH 5.9 and at pH 8.0 so that the Lys and Orn residues were forced to compete for the substrate. Acylation was observed in all cases, but only at two positions in each protein and only at the side chains of lysine or ornithine residues in positions 34, 19 and 15, in the polypeptides KA-I (34 and 19), KA-I-A_{15} (34 and 19), KA-I-A_{33} (34 and 15) and KA-R_{19} (34 and 15). In KA-II, the dominant position of acylation was position 12 but Lys15 was also modified. Lys10 and Lys33 were not modified in any sequence and there is therefore a difference in reactivity between individual Lys and Orn residues depending on structure. Orn34 is preferentially acylated in all peptides, with the obvious exception of KA-II, and position 34 is therefore the most reactive one. Orn34 is in a d position, and therefore in the hydrophobic core; this suggests that it has a depressed pK_a due to the low dielectric constant of the a and d positions. The pH dependence of the acylation of KA-I was determined, Figure 7, and monoacylation was the only detectable reaction product at pH levels below 8, whereas at higher pH several lysines were acylated simultaneously. Two monoacylated products were obtained, one where Orn34 was selectively acylated and one where Lys19 was acylated, in a very approximate ratio of 3 to 1. According to the αH chemical shifts Lys19 is not located in a well-defined helical structure, although measured medium range NOE values suggest that it is, at least part of the time. It is formally in a d position and may have the properties of a partly hydrophobic environment although the influence of Lys15 may contribute to the pK_a depression through an electrostatic effect, in analogy to effects observed in His pK_a values.[31] Selectively depressed pK_a values are compatible with the observed pH dependence and it is suggested that pK_a values of Lys residues may be depressed by introducing them into a and d positions to increase their reactivities towards carbonyl groups and other electrophiles.

The second-order rate constant was determined as a function of pH in order to determine the pK_a of the reactive residues as shown in Figure 10. The best fit to the experimental results corresponded to a pK_a of 9.3, which amounts to a pK_a depression of 1.1 pK_a units in comparison with a solvent exposed lysine residue. The reaction was also carried out in TFE to determine the effect on reactivity of disrupting the tertiary structure while retaining a helical conformation (see Figure 8). The overall degree of acylation was dramatically reduced in spite of the fact that lysine pK_a values would be expected to decrease as a result of the lower polarity of the medium. The observed loss in incorporation efficiency is complex and may not be explained in a simple way, but the

0947-6539/02/0816-3694 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 16

results are in agreement with a model where the reactivity is reduced when the tertiary structure, which is responsible for the pK_a depression, is disrupted. Enhanced reactivities of Lys and Orn residues in a and d positions in comparison with those of residues in other positions are therefore established in aqueous solution. Any reaction that depends on the availability of unprotonated Lys residues is therefore a target that can be addressed using this strategy. The fact that no a position has been probed is mainly due to practical considerations. There is no difference in the hydrophobic character between an a or a d position and the same considerations apply.

In order to provide a critical test for whether position 34, a d position, was more reactive because of pK_a depression due to the hydrophobic environment, a helix–loop–helix motif was designed where an alanine residue was incorporated into position 34, and a second d position was selected for lysine incorporation. In KA-II Leu12 was replaced by Lys12, and Lys12 was preferentially acylated by **1**. Consequently, a and d positions provide environments that are sufficiently hydrophobic to reduce the pK_a value of lysine residues and make the incorporation of acyl groups site selective. The incorporation of lysine residues in positions that according to the heptad repeat pattern are hydrophobic is therefore a viable design strategy for site-selective functionalisation of folded proteins. Comparable positions in native proteins should be able to provide sites that can be functionalised in a similar way.

In KA-I the selectivity between Orn34 and Lys19 was probed but it was not possible to separate the monoacylated peptides KA-I-Fum$_{19}$ from KA-I-Fum$_{34}$ by analytical RP HPLC. An analysis of the tryptic digests by MALDI-TOF MS suggests, very approximately, that the level of acylation was 46–51% at position 34 was and 15–20% at position 19, after reaction at pH 8.0. The preference for Orn34 over that of Lys19 is therefore approximately a factor of 3 at pH 8.0. These numbers are clearly approximate but there is little doubt that Orn34 is the most reactive position. Lys19 is, however, more reactive towards **1** than all other lysines, no other functionalised peptide was detected. Lys19 is also in a d position in the heptad repeat of amino acids and may also be partly included in the hydrophobic core, although its position in close proximity to the more flexible loop region of the peptide suggests that the pK_a might not be affected as much as that of Orn34. Its position as a C-cap would be expected to raise its pK_a due to the interaction with the helix dipole moment, and make it less reactive. The reason may instead be that Lys15 depresses its pK_a because of the charge repulsion between the protonated forms of the lysine residues. The role of Lys15 is not understood in detail but substituting an alanine for Lys15 considerably improved the selectivity in favor of Orn34, approximately 85% of all incorporated fumaryl substituents were introduced at the side chain of Orn34 in the sequence KA-I-A$_{15}$. This is in good agreement with a model where Lys15 reduces the pK_a of Lys19, and when Lys15 is replaced by Ala the pK_a of Lys19 is increased and the reactivity is decreased, leading to enhanced selectivity of Orn34 over Lys19.

Understanding the molecular basis for the reactivity of Lys and Orn residues made it possible to improve selectivity by rational design. Replacing Lys19 by an Arg residue, as in KA-I-R$_{19}$, enhanced the selectivity in favor of Orn34 acylation to a level where 85% of the incorporated fumaryl substituents were introduced at Orn34 to form a monoacylated polypeptide. In the absence of Lys19, Lys15 became the second most efficient nucleophile but it was monoacylated only to a small degree. Possibly Arg19 reduced the pK_a of Lys15 to make it more reactive. A more complex relationship was unravelled in the interplay between Lys33 and Lys19. As Lys33 was replaced by Ala, Lys19 was no longer reactive enough to compete with Lys15 for the substrate. No acylation of Lys19 was observed and it may be that subtle structural effects play a role although the CD spectrum is largely unaffected by the change in sequence.

The pH dependence of the reaction between KA-I and **1** showed that the reaction depends on the unprotonated form of a residue with a pK_a of approximately 9.3 suggesting lysine or ornithine residues. Amide formation is an irreversible reaction under the reaction conditions and the fact that acyl groups were found exclusively at the side chains of Orn34 and Lys19 therefore provided strong evidence in favor of direct acylation of the side chain primary amines. The pH dependence is in principle compatible with a model where flanking Lys and Orn residues carry out general base catalysis in the formation of the amide bond, but this reaction mechanism is unlikely due to the small fraction of unprotonated lysines and ornithines that is available at low pH. We conclude that Lys and Orn residues are acylated in a one-step reaction and that the reactivity is determined mainly by the position in the sequence where a and d positions provide enhanced reactivity enough to make Lys and Orn residues dominant sites for the incorporation of acyl groups. In addition, close proximity of flanking charged residues appear to affect the pK_a of Lys and Orn residues in an analogous manner to what was previously described for His residues.[31]

The addition of only one equivalent of **1** makes competitive measurements of acylation more readily interpretable but it leaves some polypeptide unreacted due to competition from background hydrolysis. In order to optimise selectivity KA-I was reacted with up to 10 equivalents of **1** and the degree of acylation and the site selectivity were analysed, Figure 9a, b. The maximum level, 70%, of monoacylated protein was obtained after approximately five equivalents were reacted at pH 5.9 and 1.3 equivalents at pH 8.0. The ratio of monoacylated to diacylated protein was then approximately 72:13 at pH 5.9 and 73:19 at pH 8.0. Since acylation times are dramatically shorter at the higher pH for practical purposes a high pH is preferable in protein functionalisation. Furthermore, the selectivity does not depend on pH as long as it is less than the value of the pK_a of the reactive residues as discussed above. In KA-II Lys12 was acylated very efficiently by three equivalents of ester.

Conclusions

The reactivity of Lys and Orn residues may be controlled to a level where they become the dominant nucleophiles of the protein scaffold in reactions with active esters to form functionalised side chains. The underlying principles are

  0947-6539/02/0816-3695 $ 20.00+.50/0

therefore of fundamental interest in understanding protein reactivity but also of practical use in developing covalently modified four-helix bundle model proteins with applications as artificial receptors, biosensors and in glycobiology. It was possible after determining the relationship between structure and reactivity to design a polypeptide KA-I-R_{19} where 85% of site-selectively monoacylated product was obtained and readily purified. It may be possible to design even more selective sites as our understanding of the chemistry of protein surfaces develops. As vehicles for site-selective acylation four-helix bundle proteins have thus been shown to be versatile and useful scaffolds.

Understanding the reactivity of lysine side chains also provides a strategy for designing sites capable of imine formation, the key reaction step in aldol reactions and Michael additions. The main factor in pK_a control is the position in the structure of the folded four-helix bundle motif, where their incorporation into a or d positions of the heptad repeat provides a molecular environment capable of depressing the pK_a value by 1.1 units in comparison with those of solvent exposed residues. This corresponds to an increase in reactivity by an order of magnitude in the unprotonated state, which is sufficient to ensure a high degree of selectivity in the incorporation of new functionality into folded proteins.

Experimental Section

General procedure for MALDI-TOF MS analysis: The mass spectra were recorded on an Applied Biosystems Voyager DE-STR MALDI-TOF mass spectrometer. Peptide solutions were prepared at concentrations of approximately 1 mg mL^{-1} in trifluoroacetic acid (TFA) (0.1%). A solution of α-cyano-4-hydroxycinnamic acid (10 mg mL^{-1}) in acetonitrile:TFA (0.1%) (1:1) was used as matrix. Typically 2 μL of the peptide solution was mixed with 18 μL of the matrix, and 1 μL of the mixture was applied to the plate for crystallisation. A mixture of angiotensin I, ACTH clip1–17, ACTH clip 18–39, ACTH clip 7–38 and insulin (bovine) was used for calibration. Measured molecular weights corresponded to the calculated ones within 1.1 mass units in all cases, and typically within 0.5 mass units.

CD spectroscopy: CD spectra were recorded on a Jasco J-715 CD spectropolarimeter, routinely calibrated with D-(+)-10-camphorsulphonic acid, in 0.1, 0.5, 1 or 5 mm cuvettes in the interval from 280–190 nm. Each spectrum is an average of six scans and the background was subtracted from the spectrum before the mean residue ellipticity at 222 nm was measured. For the concentration dependent studies a 50 mM stock peptide solution in BIS-Tris buffer at pH 5.9 was prepared and diluted to the desired concentrations by pipetting. For pH dependence measurements a 0.3 mM aqueous stock solution was prepared and the pH was adjusted by small additions of 0.1 or 1M HCl or NaOH.

NMR spectroscopy: NMR spectra of peptides in 4 vol% [D_3]TFE in H_2O:D_2O (90:10 *v*/*v*) were recorded on a Varian Inova 600 MHz NMR spectrometer at 308 K and pH 5.2 using preirradiation of the water resonance. Typical 90° degree pulses were 9.2 μs for both 1D, NOESY and TOCSY spectra, and the spinlock pulse in the TOCSY experiment was 22 ms with a window function of 30 ms. The mixing times were 200 ms for the NOESY experiments and 80 ms for the TOCSY experiments. 2×256 increments were recorded and the data were processed using linear prediction algorithms.

Peptide synthesis and purification: The peptides were synthesised on an Applied Biosystems automated Pioneer peptide synthesiser using standard Fmoc chemistry. A typical synthesis was performed on a 0.1 mmol scale using a PAL-PEG-PS polymer (Perseptive Biosystems), that forms an amide at the C-terminal of the peptide after cleavage. The substitution level of the polymer was 0.17–0.23 mmol g^{-1}. The side chains of the amino acids were protected using base stable groups; *tert*-butoxy (O*t*Bu) for Asp and Glu, *tert*-butyloxycarbonyl (Boc) for Lys and Orn, trityl (Trt) for 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 20% piperidine in *N*,*N*-dimethylformamide (DMF, peptide synthesis grade). TBTU (*O*-(7-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate) (0.5M in DMF) was used together with DIPEA (diisopropylethylamine) (1M in DMF) as the activating reagent and a four-fold excess of amino acid was used in each coupling. A standard coupling time of 60 minutes was used except for Nle and Leu where 30 minutes was used and for Gln, Arg, Asn and some difficult couplings in the sequence where 90 minutes were used. The N-terminal of the peptide was capped with acetic anhydride in DMF (0.3M). When the synthesis was completed the polymer was rinsed with dichloromethane and dried under vacuo. The peptide was cleaved from the resin and deprotected by treatment with a mixture of TFA:H_2O:ethanedithiol:triisopropyl silane (94:2.5:2.5:1 *v*/*v*), 10 mL per gram of polymer, for three hours at room temperature. After filtration and concentration, the peptide was precipitated by the addition of cold diethyl ether, centrifuged, and resuspended three times in diethyl ether and lyophilised. The crude products were purified by reversed phase HPLC on a semi-preparative C-8 HICHROM column, eluted isocratically with 36–43% isopropanol and 0.1% aqueous TFA at a flow rate of 10 mL min^{-1}. The purity was checked by analytical HPLC and the polypeptides were identified by MALDI-TOF mass spectrometry.

The acylation reaction: Stock peptide solutions (1 mM) were prepared by weighing and dissolving the peptide in buffer solution, assuming a water-content of the lyophilised peptide of 25%, and adjusting the pH by small additions of NaOH and HCl (1–2M). A stock substrate solution was also prepared by dissolving *p*-nitrophenylfumarate in equal volumes of acetonitrile and buffer to the final concentration of 15 mM. The buffers used at pH 5.9 and at pH 8.0 were Bis-Tris and Tris, respectively. The acylation reaction was carried out by the addition of one equivalent of substrate (2 μL) to the peptide solution (30 μL). At the addition of 1–10 equivalents of **1** to the peptide, 2 μL of substrate solution and 18 μL of buffer were added to 30 μL of the peptide solution (1 mM) in the case of one equivalent. In the case of 10 equivalents 20 μL of substrate solution was added to 30 μL of peptide solution (1 mM) to give a final peptide concentration of 0.6 mM. After three days at room temperature and pH 5.9, or after one day at room temperature and pH 8.0, the reaction mixtures were analysed by RP HPLC on an analytical C-8 HICHROM column, eluted isocratically with 36–43% isopropanol and 0.1% aqueous TFA at a flow rate of 0.6 mL min^{-1}. The polypeptides were identified by mass spectrometry. The degree of modification was measured by integration of the peaks of the analytical RP HPLC chromatogram, and calculated from the area of each peak divided by the total area of peptide fractions under the assumption that all peptides, modified and unmodified, have the same extinction coefficient.

Trypsin digestion: A peptide solution (1 mM) in $NH_4^+HCO_3^-$ (0.1M, pH 8.0) was prepared. Trypsin (1 mg) was dissolved in HCl (100 μL, 0.1 mM) and added to the peptide solution (200 g trypsin per mol peptide). After 3 h at 37 °C the reaction was quenched by the addition of HCl (20 μL) and the reaction solution was lyophilised. The resulting fragments were identified by mass spectrometry.

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 at 298 K. A peptide stock solution (1 mM) was prepared in buffer and the pH was adjusted to the correct value by the addition of small amounts of 1M NaOH or HCl. The stock solution was diluted with buffer to the desired concentrations (0.4 mM, 0.3 mM and 0.2 mM) and transferred (270 μL) to the cuvettes. After 15 minutes of temperature equilibration the substrate was added (5 μL, 5 mM) to give a final concentration of 0.1 mM. The *p*-nitrophenylfumarate was dissolved in a 50% (*v*/*v*) mixture of buffer and acetonitrile. The buffers used were sodium acetate (100 mM) at pH 4.1 and 5.1, Bis-Tris (50 mM) at pH 5.85, and pH 7.0, Tris (50 mM) at pH 8.0 and 9.0 and CAPS at pH 9.7–10.2. The reaction was followed for more than three half-lives and the data was processed using IGOR Pro software. The rate constants are the results from fitting a straight line to the pseudo first order rate constants plotted versus the peptide concentration.

0947-6539/02/0816-3696 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 16

Acknowledgement

Financial support from the Swedish Natural Science Research Council is gratefully acknowledged. We thank the Swedish NMR Centre at Göteborg University for use of NMR spectrometers and Charlotta Damberg for expert assistance.

[1] W. F. DeGrado, C. M. Summa, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
[2] L. Baltzer, *Curr. Opin. Struct. Biol.* **1998**, *8*, 466–470.
[3] C. Micklatcher, J. Chmielewski, *Curr. Opin. Chem. Biol.* **1999**, *3*, 724–729.
[4] L. Baltzer, J. Nilsson, *Curr. Opin. Biotechnol.* **2001**, *12*, 355–360.
[5] a) A. Fersht, *Structure and Mechanism in Protein Science*: *A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman, New York, **1999**, Chapter 11; b) A. Fersht, *Structure and Mechanism in Protein Science*: *A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman, New York, **1999**, Chapter 9.
[6] D. J. Sweatt, *Curr. Biol.* **2001**, *11*, R391–R394.
[7] T. Hunter, *Cell* **2000**, *100*, 113–127.
[8] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.
[9] H. Lis, N. Sharon, *Eur. J. Biochem.* **1993**, *218*, 1–27.
[10] D. S. Kemp, R. I. Carey, *Tetrahedron Lett.* **1991**, *32*, 2845–2848.
[11] P. E. Dawson, S. B. Kent, *J. Am. Chem. Soc.* **1993**, *115*, 7263–7266.
[12] G. Tuchscherer, *Tetrahedron Lett.* **1993**, *34*, 8419–8422.
[13] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.
[14] J. P. Tam, Y. Lu, C. Liu, J. Shao, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12485–12489.
[15] M. Mutter, S. Vuilleumier, *Angew. Chem.* **1989**, *101*, 551; *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 535–554.
[16] L. Baltzer, A.-C. Lundh, K. Broo, S. Olofsson, P. Ahlberg, *J. Chem. Soc. Perkin Trans. 2* **1996**, 1671–1676.
[17] K. Broo, A.-C. Lundh, P. Ahlberg, L. Baltzer, *J. Am. Chem. Soc.* **1996**, *118*, 8172–8173.
[18] L. Andersson, G. Stenhagen, L. Baltzer, *J. Org. Chem.* **1998**, *63*, 1366–1367.
[19] L. K. Andersson, G. T. Dolphin, J. Kihlberg, L. Baltzer, *J. Chem. Soc. Perkin Trans. 2* **2000**, 459–464.
[20] L. K. Andersson, G. T. Dolphin, L. Baltzer, *ChemBioChem*, in press.
[21] W. P. Jencks, J. Carriuolo, *J. Am. Chem. Soc.* **1960**, *82*, 1778–1786.
[22] S. Olofsson, G. Johanson, L. Baltzer, *J. Chem. Soc. Perkin Trans. 2* **1995**, 2047–2056.
[23] K. S. Broo, L. Brive, P. Ahlberg, L. Baltzer, *J. Am. Chem. Soc.* **1997**, *119*, 11362–11372.
[24] 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.
[25] P. S. Ho, W. F. DeGrado, *J. Am. Chem. Soc.* **1987**, *109*, 6751–6758.
[26] S. Olofsson, L. Baltzer, *Folding Design* **1996**, 347–356.
[27] D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **1991**, *222*, 311–333.
[28] K. Wüthrich, *NMR of proteins and nucleic acids*, Wiley, **1986**.
[29] M. Buck, *Q. Rev. Biophy.* **1998**, *31*, 297–355.
[30] C. Tanford, *Advan. Protein Chem.* **1962**, *17*, 69–165.
[31] K. S. Broo, L. Brive, R. S. Sott, L. Baltzer, *Folding Design* **1998**, *3*, 303–312.

Received: September 10, 2001
Revised: March 25, 2002 [F3544]

  0947-6539/02/0816-3697 $ 20.00+.50/0