Non-covalent control of site-selective incorporation of the pyridoxal phosphate cofactor into a folded polypeptide motif—mimicking a key step in enzymatic transamination

Malin Allert, Martin Kjellstrand, Klas Broo, Åke Nilsson and Lars Baltzer*†

Department of Chemistry, Göteborg University, S-412 96, Göteborg, Sweden

The site-selective incorporation of the pyridoxal phosphate cofactor into a designed polypeptide motif has been achieved and shown to be controlled by the non-covalent interactions between the phosphate group of the cofactor and a single arginine residue on the surface of the folded polypeptide.

The *de novo* design of functionalized proteins is an attractive alternative in the construction of new catalysts because of the inherent capacity of polypeptides for folding into complex three-dimensional structures determined by non-covalent bonds. Folded polypeptide catalysts, with reactive sites formed by naturally occurring amino acids, are now emerging that catalyse multistep reactions with rate enhancements that rival those of typical catalytic antibodies.[1–3] Coenzyme-based catalysts have, perhaps, an even greater potential because of their intrinsic reactivities, provided that efficient strategies can be found to embed the cofactors into reactive sites in template peptides. The site-selective incorporation of residues for which no codons exist has so far been accomplished by the cumbersome synthesis of artificial amino acids[4,5] or by the site-selective functionalization of lysine, ornithine, diaminobutyric acid[6–8] and cysteine residues.[9,10] Here we report that for the first time non-covalent interactions have been used to control the incorporation of a complex cofactor into a folded polypeptide. Understanding the principles that govern the recognition of functional groups by folded polypeptides opens up efficient routes for the engineering of new proteins with tailor-made specificities.

PP-42, a polypeptide with 42 residues, was designed to fold into a helix-loop-helix motif and dimerise in aqueous solution to form a four-helix bundle protein (Fig. 1). Its design was based on the sequence and solution structure of SA-42[11,12] with the objective of engineering a template polypeptide into which a pyridoxal cofactor could be introduced. PP-42 was synthesised by solid-phase peptide synthesis using Fmoc chemistry, purified by reversed-phase HPLC and identified from the electrospray mass spectrum (calc. 4532.1; found 4531.5).‡ The mean residue ellipticity of PP-42 at a concentration of 0.60 mM is -24500 ± 1000 deg cm^2 dmol^{-1} in aqueous solution at pH 7.0 and ambient temperature, which is comparable to those of other designed helix-loop-helix dimers.[13] The helical content of amphiphilic helix-loop-helix dimers is known to decrease substantially upon dissociation, therefore the measured mean residue ellipticity of PP-42 provides strong evidence that it folds into the designed four-helix bundle motif.

The mechanism of transamination in native enzymes includes the formation of an aldimine from the pyridoxal phosphate cofactor **I** and the side chain of a lysine residue in the active site. Upon introduction of the amino acid substrate into the active site, the lysine side chain is replaced by the amino group of the amino acid, which is transformed to an α-keto acid in a multistep reaction.[14] In order to incorporate the cofactor into the folded helix-loop-helix motif, to mimic the initial step of the enzymatic reaction, a lysine residue was targeted for aldimine formation by placing an arginine side chain in a position to interact with the phosphate group of the cofactor. In the computer modelled structure the guanidino group of Arg-19 is ideally positioned to bind the phosphate group upon aldimine formation at the side chain of Lys-30. Two more lysine residues were introduced into PP-42 to probe whether the phosphate-arginine interaction would be strong enough to make the incorporation site-selective. The sequence of PP-42 contains three lysine and four arginine residues, Lys-11, Lys-15, Lys-30, Arg-10, Arg-19, Arg-33 and Arg-40 (Fig. 1).

The reaction between **I** and PP-42 to form the aldimine was followed by the increase in absorbance at 390 nm, using an extinction coefficient for the aldimine of 4000 M^{-1} cm^{-1} (Fig. 2).§ The degree of aldimine formation was determined from the 400 MHz ^{1}H NMR spectrum recorded in D_2O at pH 4.4 by measuring the reduction of the intensity of the aldehyde proton of **I** at δ 10.43 upon addition of PP-42 under these conditions. More than 95% of the monofunctionalised peptide was obtained using approximately 0.6 mM PP-42 and a five-fold excess of **I**. At pH 4.4 the reaction is rapid and equilibrium is established within a few minutes. The UV spectrum of a solution of 0.6 mM PP-42 and 3.1 mM of **I** at 25 °C and pH 4.4 is shown in Fig. 2 and shows that more than 95% of the PP-42 is bound. In contrast, under identical reaction conditions, the equilibrium

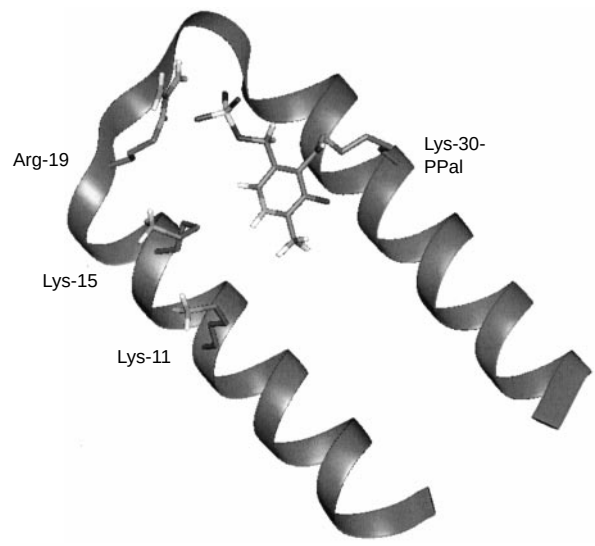

Fig. 1 The amino acid sequence and the modelled structure of PP-42 showing the positions of arginine and lysine side chains on the surface of the folded motif. The amino acid residues are given in the one-letter code where Nle is the artificial amino acid norleucine. Only the monomer is shown for simplicity.

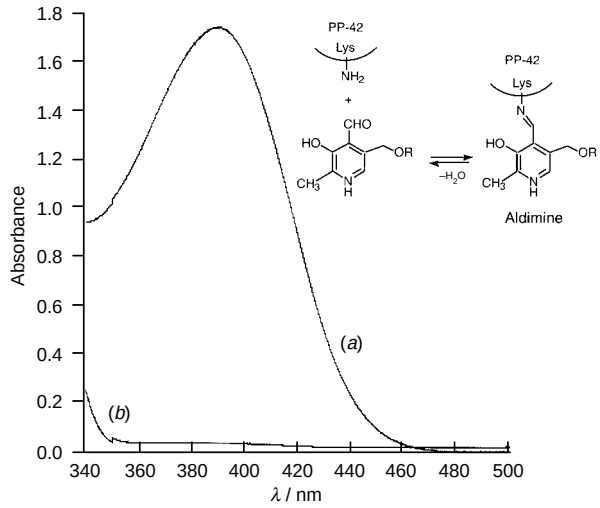

Fig. 2 The UV spectrum recorded in aqueous solution, at 25 °C and pH 4.4, of 0.6 mM PP-42 in the presence of (*a*) 3.1 mM pyridoxal phosphate **I** (R = PO_3^{2-}) and (*b*) 3.1 mM pyridoxal II (R = H). The aldimine formation was followed at 390 nm.

concentration of aldimine upon reaction between PP-42 and pyridoxal **II**, that has no phosphate residue, is insignificant. It can be estimated from the ratio of the absorbances of the two aldimines that the equilibrium constants differ by more than two orders of magnitude, although the absorbance of the pyridoxal-containing mixture is too weak to permit an accurate determination. A strong interaction between the negatively charged phosphate residue and the peptide is thus demonstrated.

The site of incorporation of pyridoxal phosphate into PP-42 was determined by $NaBH_4$ reduction of the aldimine in aqueous solution to form the corresponding secondary amine, followed by trypsin cleavage of the functionalised peptide. The site of functionalization was determined by LC-ESMS of the polypeptide digest. Trypsin cleaves polypeptides on the C-terminal side of basic residues. No cleavage was expected at the lysine side chain that had been functionalized.

A fragment from the reduced form of the functionalized peptide (dication 835.91 and trication 557.61), which corresponded to the molecular mass of the sequence from Gly-20 to Arg-33 (1438.64) plus the mass of pyridoxal phosphate (247.2) and 2H (2.02), less the weight of water (18.02) amounting to approximately 85% of the total amount of reduced functionalized peptide, were found in the LC-ESMS spectrum. The dominant site of aldimine formation was therefore the side chain of Lys-30. A minor amount of aldimine formation at the side chain of Lys-11 was also observed, whereas no aldimine formation involving the side chain of Lys-15 could be detected.

Arg-19 is apparently in a good position relative to Lys-30 to control the functionalization of its side chain, but the proximity of Arg-10 to Lys-11 in a helical conformation appears to give rise to small amounts of aldimine, too, under conditions of excess **I** over peptide. The absence of detectable amounts of aldimine at the side chain of Lys-15 shows that a lysine side chain that is not flanked by an arginine in PP-42 does not form an aldimine with **I** to a measurable extent.

Arg-19 and Lys-30 thus form a two-residue site that competes favourably with other lysine residues in PP-42, and the aldimine formation with **I** is controlled by the binding of phosphate by arginine. The engineering of a Lys-Arg two-residue site is therefore sufficient to ensure the incorporation of the cofactor provided that the arginine side chain has a similar orientation towards the lysine side chain as that of Arg-19 towards Lys-30. The incorporation of **I** can thus be made site-selective in the presence of other lysine residues. To enhance the selectivity further Arg-(*i*)-Lys-(*i*+1) configurations in helical segments should be avoided. In the second step of the catalytic cycle the amino group of an amino acid will replace the lysine side chain in an exchange reaction. Further design of the polypeptide motif is now under way so that the second intermediate, too, will be bound to the peptide catalyst. It can for example be envisioned that the trianion formed in the reaction between **I** and Asp will be bound by arginine residues.

The demonstration of non-covalent control of the incorporation of **I** shows that the interaction between the phosphate group and an arginine side chain in aqueous solution at pH 4.4 is strong enough to ensure the site-selectivite funtionalization of the folded polypeptide. A well-defined two-residue site on the surface of the folded polypeptide can therefore be used to create a molecule of high complexity in a one-step reaction in aqueous solution. The use of the arginine–phosphate bond should be of general interest in a wide range of applications involving phosphates and phospho esters. The results suggest that understanding the interplay between functional amino acid side chains, organised in three-dimensional space by polypeptide templates, is a key element in the rational design of functionalised polypeptides with tailor-made specificities.

Financial support from the Swedish Natural Science Research Council and from Carl Tryggers Stiftelse is gratefully acknowledged. We are indebted to Dr Gunnar Stenhagen for mass spectrometric determinations.

Notes and References

† E-mail: baltzer@oc.chalmers.se

‡ The solid-phase peptide synthesis, cleavage and purification strategies have been described in detail for the four-helix bundle proteins SA-42[11,12] and KO-42[1] that have similar sequences.

§ The extinction coefficient at pH 4.4 and 25 °C was estimated from the absorbance of an aqueous solution containing 0.3 mM PP-42 and 0.3 mM **I**.

1 K. S. Broo, L. Brive, P. Ahlberg and L. Baltzer, *J. Am. Chem. Soc.*, 1997, **119,** 11362.
2 K. Johnsson, R. K. Allemann, H. Widmer and S. A. Benner, *Nature*, 1993, **365**, 530.
3 K. Severin, D. H. Lee, A. J. Kennan and R. M. Ghadiri, *Nature*, 1997, **389**, 706.
4 B. Imperiali and R. Sinha Roy, *J. Org. Chem.*, 1995, **60**, 1891.
5 R. Sinha Roy and B. Imperiali, *Tetrahedron Lett.*, 1996, **37**, 2129.
6 L. Baltzer, A.-C. Lundh, K. Broo, S. Olofsson and P. Ahlberg. *J. Chem. Soc., Perkin Trans. 2*, 1996, 1671.
7 K. Broo, L. Brive, A.-C. Lundh, P. Ahlberg and L. Baltzer, *J. Am. Chem. Soc.*, 1996, **118**, 8172.
8 K. Broo, M. Allert, L. Andersson, P. Erlandsson, G. Stenhagen, J. Wigström, P. Ahlberg and L. Baltzer, *J. Chem. Soc., Perkin Trans. 2*, 1997, 397.
9 H. Mihara, Y. Tanaka, Y. Fujimoto and N. Nishino, *J. Chem. Soc., Perkin Trans. 2*, 1995, 1915.
10 H. Mihara, K. Tomizaki, N. Nishino and T. Fujimoto, *Chem. Lett.*, 1993, 1533.
11 S. Olofsson, G. Johansson and L. Baltzer, *J. Chem. Soc., Perkin Trans. 2*, 1995, 2047.
12 S. Olofsson and L. Baltzer, *Folding Des.*, 1996, **1**, 347.
13 S. Ho and W. DeGrado, *J. Am. Chem. Soc.*, 1987, **109**, 6751.
14 H. Dugas, *Bioorganic Chemistry*, Springer-Verlag, New York, 3rd edn., 1996, p. 523.

Received in Glasgow, UK, 2nd March 1998; 8/01741G