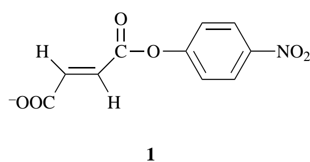
The relationship between structure and reactivity in RA-42, a designed helix–loop–helix motif

Ann-Christin Lundh, Klas Broo and Lars Baltzer *

Department of Organic Chemistry, Göteborg University, 412 96 Göteborg, Sweden

RA-42, a polypeptide with 42 amino acids, has previously been shown to react with *p*-nitrophenyl fumarate (1) and site selectively form an amide at the side chain of ornithine-15 in aqueous solution and in 10 vol% 2,2,2-trifluoroethanol (TFE). In order to investigate the role of the medium on structure and on reactivity the mean residue ellipticity at 222 nm has been measured by circular dichroism spectroscopy as a function of TFE concentration and the second-order rate constant for the reaction between RA-42 and 1 has been measured in 30 vol% TFE. The correlation was poor between the mean residue ellipticity and the second-order rate constants, and the effect of increasing the TFE concentration was therefore most likely to increase the strength of electrostatic interactions in the transition state due to the decreased polarity of the solvent. The ratio of second-order rate constants between that of RA-42 and that of 4-methylimidazole (4-MeIm), k_2(RA-42)/k_2(4-MeIm), increased with TFE concentration suggesting that there was little contribution to transition state binding in helix II since there is little hairpin conformation left in 30 vol% TFE. Finally, the effect on structure of the introduced ornithines was determined by measuring the free energy of unfolding. Although the helix–loop–helix motif was destabilised by the incorporation of positively charged residues, the effect on reactivity was small, suggesting that polypeptides are useful as templates in the rational design of catalysts.

Designed proteins and polypeptides with supersecondary structures[1] have great potential in the engineering of novel catalysts, mainly because of the versatility of the amino acids and the capacity of polypeptides for forming complex structures. With the advances that have been made in NMR spectroscopy with regard to structure determination of biomolecules this potential can now begin to be exploited since it is possible to determine the relationship between structure and reactivity. We recently reported that RA-42, a designed polypeptide with 42 amino acids that folds in solution into a hairpin helix–loop–helix motif, reacts with mono-*p*-nitrophenyl fumarate **1**, to form



1

an amide at the side chain of Orn-15 (ornithine-15) in a self-catalysed site-selective functionalization reaction.[2] The reaction rate was determined in aqueous buffer solution[3] at pH 5.8 and 290 K, as well as in 10 vol% 2,2,2-trifluoroethanol (TFE)[2] at pH 5.8 and 290 K and the rate enhancement over that of the 4-methylimidazole (4-MeIm) catalysed reaction [k_2 (RA-42)/k_2 (4-MeIm)] was found to be solvent dependent. The reaction mechanism was determined by kinetic measurements and product studies[3] and the first step was found to be a rate limiting attack on the substrate by the side chain of His-11 to form an acyl intermediate with the release of *p*-nitrophenol. This step was followed by a fast intramolecular acyl-transfer reaction to give the reaction product, an amidated ornithine side chain, with regeneration of the free imidazole group of His-11.

In view of the complexity of the folded polypeptide the results were gratifying and appeared to be in agreement with the designed reactivity. However, a number of issues remained to be clarified. First, the observed solvent effect raised the question of whether the increase in reactivity [k_2(RA-42)/k_2(4-MeIm)] was due to a more well-ordered helical structure induced by TFE,[4] or whether it arose because of stronger electrostatic interactions in the transition state induced by the less polar

N-Aib-A-D-Nle-E-A-A-I-K-H-L-A-E-Orn-Nle-Aib-A-K
1 19
G-P-V-D
20 23

G-Aib-R-A-F-A-E-F-Orn-K-A-L-Q-E-A-Nle-Q-A-Aib
42 RA-42 24

Fig. 1 The amino acid sequence of RA-42. The residues presented underlined in bold in RA-42 are the only ones that have been changed in comparison with those in SA-42. The one letter code for 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. Aib is α-aminoisobutyric acid and Nle is norleucine.

medium. This distinction is important as it clarifies the significance of charge–charge interactions on the surface of folded polypeptides and the need for well-ordered polypeptide structures.

The second question was related to the structure of the reactive site. The design of RA-42 was based on reaction mechanistic principles of transition state stabilisation and the solution structure of SA-42, a designed polypeptide with 42 amino acids that folds into a helix–loop–helix motif and dimerises in solution in an antiparallel way.[5,6] Only three amino acid residues were replaced in the amino acid sequence of SA-42 in the design of RA-42, Fig. 1. Although it was shown that the introduction of positively charged hydrogen bond donors in the vicinity of the histidine enhanced the reactivity of the histidine it was not clear if one or both of the ornithines participated in the reaction.

Finally, the measured helical content of RA-42 was lower than that of SA-42 and the effect of the substitutions in the amino acid sequence of the template polypeptide on the free energy of unfolding of RA-42 had to be determined. The potential of the template approach depends on the extent to which a folded motif can be modified without losing its structure and is of general interest in protein design.

Results and discussion

The design, synthesis and solution structure of RA-42 have

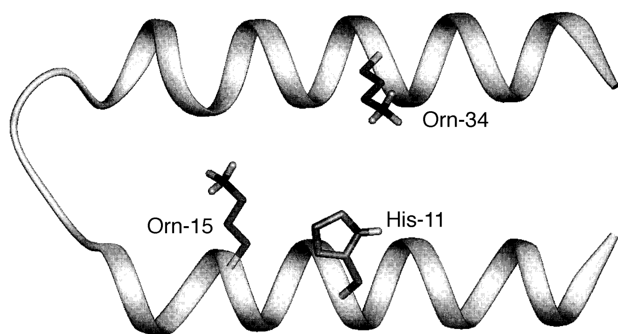


Fig. 2 The designed reactive site of RA-42 showing only the side chains of His-11, Orn-15 and Orn-34. His-11 functions as a nucleophilic catalyst and Orn-15 was designed to bind to the developing oxyanion in the transition state whereas Orn-34 was designed to bind to the carboxylate anion of the substrate.

Table 1 Second-order rate constants (k_2/10^{-3} dm^3 mol^{-1} s^{-1}) and relative rates for the reactions of mono-*p*-nitrophenyl fumarate with RA-42 and 4-MeIm at pH 5.85, 290.2 K in TFE: water mixtures

Solvent	RA-42	4-MeIm	Rel. rate
Aqueous buffer [a]	51	10.5	4.9
10 vol% TFE [b]	28	3.38	8.3
30 vol% TFE	42	1.20	35

[a] See ref. 3. [b] See ref. 2.

been described in detail previously.[2] In short, the design was based on the solution structure of the template polypeptide SA-42, a helix–loop–helix dimer,[5] and RA-42 was shown by NMR and CD spectroscopy to also fold into the designed hairpin helix–loop–helix motif.[2] The reactive site was engineered on the surface of the folded polypeptide, where ornithine residues were introduced in positions that in the folded state would place them in the vicinity of the histidine to provide transition state binding, Fig. 2. Orn-15 was designed to stabilise the developing oxyanion in the tetrahedral transition state by hydrogen bonding or other electrostatic interactions and Orn-34 was designed to bind to the carboxylate residue of **1**, to provide extra transition state binding.

The second-order rate constants for the reactions between RA-42 and **1** and between 4-MeIm and **1**, in aqueous solution[2] and in 10 vol% TFE[3] as well as between SA-42 and **1** in 10 vol% TFE have been reported previously, Table 1. The rate constants were determined by following the rate-limiting release of *p*-nitrophenol by UV spectroscopy at 320 nm. The reaction between SA-42 and **1** was shown to have almost the same second-order rate constant (5.3×10^{-3} dm^3 mol^{-1} s^{-1})[2] as that for the reaction between 4-MeIm and **1** (3.4×10^{-3} dm^3 mol^{-1} s^{-1}) when the difference in pK_a between His-15 of SA-42 (6.5) and that of 4-MeIm (8.0) and the Brønsted coefficient β of 0.8 was taken into account.[3] Apparently, SA-42 catalyses the reaction simply through nucleophilic catalysis by the side chain of His-15. In RA-42 the histidine is flanked by ornithine residues and the increased reactivity of RA-42 relative to that of SA-42 shows that the flanking residues are capable of binding. No saturation kinetics were observed and the observed rate enhancements upon introduction of flanking ornithine residues was probably due to transition state binding by the positively charged Orn residues.

We have now determined the second-order rate constants for the reaction between RA-42 and **1**, and between 4-MeIm and **1**, in 30 vol% TFE at pH 5.8 and 290 K, Table 1. The second-order rate constant for the reaction of RA-42 first decreases and then increases with increased TFE concentration. For 4-MeIm the rate constant decreases with increased TFE concentration,

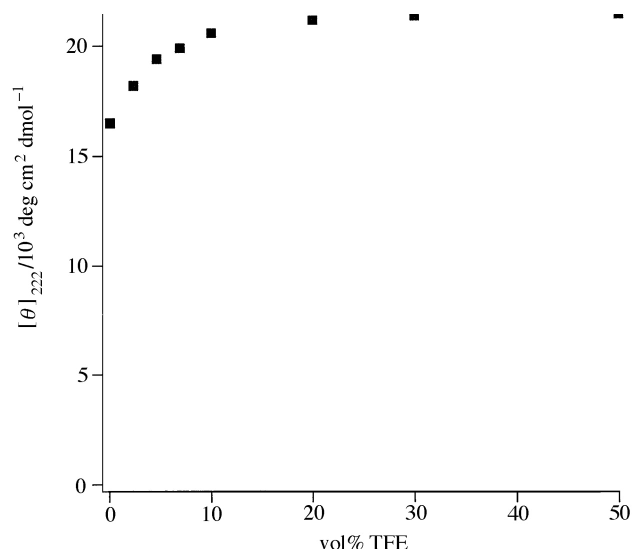


Fig. 3 The mean residue ellipticity at 222 nm, θ_{222}, of RA-42 as a function of TFE concentration at room temperature

probably due to the decreased polarity of the solvent. But the rate constant ratio [k_2 (RA-42)/k_2 (4-MeIm)] which is a measure of transition state binding provided by the flanking residues in RA-42 was increased.

In order to probe whether the measured reactivity difference was due to a more ordered helical structure the CD spectrum was recorded as a function of TFE concentration, Fig. 3. The mean residue ellipticity at 222 nm, θ_{222}, is a measure of the helical content of polypeptides and proteins[7] and its negative value is expected to increase with TFE concentration as TFE is known to induce helical structure in peptides with helical propensity.[4,8] For RA-42 the main part of the increase in helical content took place in the interval 0–10 vol% TFE and the increase in helical content was essentially complete at 30 vol% TFE, with the increase between 10 and 30 vol% being only *ca.* 20% of the total. In contrast, the increase in rate enhancement [k_2 (RA-42)/k_2 (4-MeIm)], Table 1, between that in aqueous solution and that in 10 vol% TFE was less than a factor of 1.7, whereas a further factor of 4.2 was observed upon increasing the TFE content from 10–30 vol%. The correlation between the increased reactivity of RA-42 relative to that of 4-MeIm and the helical structure was therefore poor. The results suggest that the increased reactivity of RA-42 relative to that of 4-MeIm is not due to a more well-defined geometrical relationship between the side chains of His-11 and Orn-15 but mainly due to stronger electrostatic interactions between the developing oxyanion and the flanking residues in the less polar solvent. Apparently, the organisation of residues in the reactive site is good enough in aqueous solution although the helical content is reduced upon incorporation of ornithine residues. If there is an effect on reactivity by increased rigidity and order in the helical secondary structure it would have to be small. Such a correlation has previously been reported[9] but no analysis of the influence of the polarity of the medium was presented. The observed solvent effects show that electrostatic effects are important on the surface of folded polypeptides and the introduction of hydrophobic pockets may become an important design principle in polypeptide catalysts.

The observed solvent dependence of reaction rates could also be used to determine whether Orn-15, or Orn-34 or both Orn-15 and Orn-34 participated in transition state binding. TFE is not only known to increase the helical content of polypeptides with helical propensity but also to disrupt the tertiary structure of proteins.[8] The effect by TFE on the structure of SA-42 was elucidated previously.[6] In 30 vol% TFE SA-42 no longer forms a hairpin helix–loop–helix dimer but exists predominantly as a monomer with two noninteracting helices con-

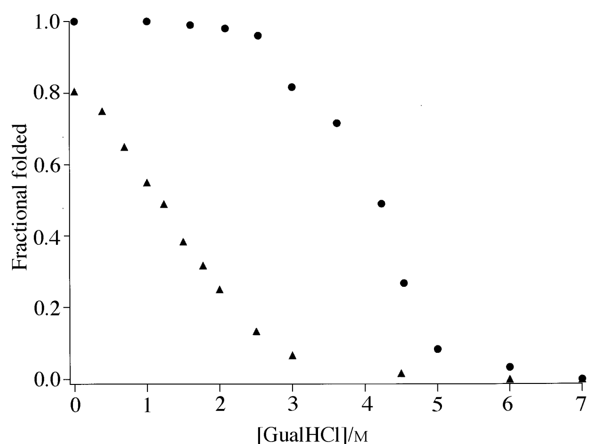


Fig. 4 Guanidine hydrochloride (GuaHCl) denaturation of SA-42 at a concentration of 75×10^{-6} mol dm^{-3} (●) and RA-42 at a concentration of 75×10^{-6} mol dm^{-3} (△) presented as fraction folded, *i.e.* the mean residue ellipticity at the given GuaHCl concentration divided by the mean residue ellipticity in aqueous solution. Assuming a two-state model for unfolding an equilibrium constant can be calculated for each point and the free energy of unfolding extrapolated to aqueous solution.[10] The free energy of unfolding of SA-42 in aqueous solution was 12.8 kcal mol^{-1}. No limiting plateau at low GuaHCl concentration was observed for RA-42 and the mean residue ellipticities were therefore divided by the mean residue ellipticity of SA-42 in aqueous solution to calculate fraction folded. The measured free energy of unfolding for RA-42 was 8.5 kcal mol^{-1}.

nected by a short loop. The amino acid sequence of RA-42 is almost identical to that of SA-42 and RA-42 can therefore be assumed to form a similar structure.

The role of Orn-34 could therefore be probed by exploiting the properties of the solvent to distance Orn-34 from the vicinity of His-11 in 30 vol% TFE. If there were significant contributions to transition state binding in aqueous solution by Orn-34 they would be lost in 30 vol% TFE and the rate enhancements should be decreased. The rate enhancements were, however, increased and so transition state binding by Orn-34 seems to make an insignificant contribution to the catalysis.

In RA-42 His-11, Orn-15 and Orn-34 have replaced Ala-11, His-15 and Ala-34 in comparison with the sequence of SA-42. The mean residue ellipticity of SA-42[5] was −25 000 deg cm^2 $dmol^{-1}$ and that of RA-42[2] was −18 800 deg cm^2 $dmol^{-1}$ which amounts to a substantial reduction in helical content by the incorporation of the residues in the reactive site. In order to investigate further the effect of amino acid substitutions on structure the free energy of unfolding was measured for SA-42 and for RA-42. GuaHCl is commonly used to denature proteins[10] and the results are generally interpreted with the assumption of a two-state model so that an equilibrium constant can be calculated for each concentration of GuaHCl used, Fig. 4. The plateau regions are used to calculate the values for the folded and unfolded forms and each intermediate point is used to calculate the free energy of unfolding for that concentration of GuaHCl, which is then extrapolated to aqueous solution. The contribution to the free energy of unfolding of a lysine residue in a helix is only 0.12 kcal mol^{-1} (1 cal = 4.184 J) less than that of an alanine[11] and from the structural similarity between ornithine and lysine one would not expect this number to be substantially different for an ornithine residue. The expected difference in the free energy of unfolding of RA-42 relative to SA-42 based on this simple consideration would not be expected to exceed a few tenths of a kcal mol^{-1}. The measured free energy of unfolding of RA-42 was, however, 8.5 kcal mol^{-1} which is 4.3 kcal mol^{-1} lower than that of SA-42, Fig. 4.

The large destabilisation of the supersecondary structure cannot be due to the difference in helix propensity as that difference is small but it seems to be due to repulsive interactions between the side chains of the introduced ornithines and residues in their vicinity in the folded state. Orn-15 is placed in between His-11 and Lys-19 and Orn-34 is placed next to Lys-33 in the sequence although they will be close in the unfolded state as well. It appears that reactive sites consisting of positively charged amino acid residues may destabilise the supersecondary structure of the polypeptide.

Fig. 5 Chemical shift index of the αH of RA-42, *i.e.* the chemical shifts of RA-42 relative to random coil shifts *versus* position in the sequence as given by Wishart.[12] A negative deviation of the chemical shift is typical of amino acid residues in helix conformation. Amino acids are given in one-letter code, see Fig. 1. Aib has no αH and the amino acids at the amino terminal have not been assigned due to fast exchange of the amide protons.

The mechanism of destabilisation was however elusive and attempts to determine the effect on structure by NMR spectroscopy did not result in the identification of specific sites where the structure deviated from that in SA-42. The αH chemical shift index[12] which is an indicator of the extent of helical conformation of the polypeptide, was determined, Fig. 5, and the difference in comparison with that of SA-42[6] was found to be small.

The measured reaction rates, Table 1, suggest however, that the destabilisation of the helical structure does not reduce the reactivity of the polypeptide. From these results it is not possible to determine conclusively whether the decreased helical content and the decreased free energy of unfolding are manifested as partial unfolding of helical structures or as a shift in the equilibrium towards a more random coil conformation. The small effect on the αH chemical shifts at the helical ends and the lack of plateau at low concentrations of GuaHCl do however suggest that the effect is a result of a shift in equilibrium towards a larger fraction of coil conformation. The effect on reactivity is therefore probably only manifested as a small decrease in the concentration of active conformation as compared to a hypothetical undestabilised polypeptide. Since the mean residue ellipticity has only decreased by 20% in comparison with the template peptide SA-42 the effect is probably small. As a result, minor destabilising effects on polypeptide structure are not expected to have any substantial adverse effect on reactivity. The reactive site of RA-42, His-11 and Orn-15 has now been incorporated into a helix–loop–helix motif, GTD-43, which has a well-defined tertiary structure[13] and the second-order rate constants were virtually identical.[14]

In conclusion, we have shown that the observed increase in reactivity [k_2 (RA-42)/k_2 (4-MeIm)] in dilute TFE is mainly due to stronger transition state binding by Orn-15 and that Orn-34 does not participate appreciably in the reactive site. The rate enhancement in TFE solution is probably due to an electrostatic effect caused by the decreased polarity of the medium. The decreased helical content is parallelled by a decreased free energy of unfolding in comparison with SA-42. The decrease is larger than expected from the helix propensities of the involved amino acid residues which suggests electrostatic repulsion between neighbouring, positively charged, residues in helical structures. The reactivity of RA-42 is, however, not significantly reduced by the decrease in stability of the helix–loop–helix motif. The results obtained are of general interest and

pave the way for further development of designed polypeptide catalysts.

Experimental

The synthesis, purification and identification of RA-42 as well as the synthesis of mono-*p*-nitrophenyl fumarate have been described previously.[2] The kinetics were carried out by following the increase in absorbance of free *p*-nitrophenol at 320 nm at 290 K. The concentration of substrate was typically 0.1 mM and the concentration of RA-42 was 0.6–1 mM. The concentration of peptide was determined by amino acid analysis.

The CD spectra were recorded on a Jasco J 720 spectrometer, that was routinely calibrated with d-camphor sulfonic acid. The dependence of the mean residue ellipticity on the concentration of TFE was measured by diluting aliquots of peptide stock solution with varying amounts of TFE and buffer and measuring the mean residue ellipticity at 222 nm and room temp. No precautions were taken to control the temperature of the cuvette as the temperature dependence of the mean residue ellipticity has previously been shown to be small. The free energies of unfolding were measured by recording the mean residue ellipticity at constant concentration of peptide as a function of guanidinium hydrochloride (GuaHCl) concentration at pH 7. A constant volume of solution from a stock solution of peptide was mixed with varying volumes of GuaHCl solution and the total volume was made up by adding buffer solution.

^{1}H NMR spectra were recorded on a Varian 500 Unity NMR spectrometer at 323 K as described previously.[2] The assignment of the spectrum has been described previously.[2]

Acknowledgements

We thank Per Ahlberg for stimulating discussions. We gratefully acknowledge financial support from the Natural Science Research Council.

References

1 J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neill and W. F. DeGrado, *Science*, 1995, **270**, 935.
2 L. Baltzer, A.-C. Lundh, K. Broo, S. Olofsson and P. Ahlberg, *J. Chem. Soc., Perkin Trans. 2*, 1996, 1671.
3 K. Broo, L. Brive, A.-C. Lundh, P. Ahlberg and L. Baltzer, *J. Am. Chem. Soc.*, 1996, **118**, 8172.
4 J. W. Nelson and N. R. Kallenbach, *Proteins*, 1986, **1**, 211.
5 S. Olofsson, G. Johansson and L. Baltzer, *J. Chem. Soc. Perkin Trans. 2*, 1995, 2047.
6 S. Olofsson and L. Baltzer, *Folding Design*, 1996, **1**, 347.
7 W. C. Johnson, Jr., *Proteins*, 1990, **7**, 205.
8 A. Jasanoff and A. R. Fersht, *Biochemistry*, 1994, **33**, 2129.
9 K. Johnsson, R. K. Allemann, H. Widmer and S. A. Brenner, *Nature*, 1993, **365**, 530.
10 C. N. Pace, *Methods Enzymol.*, 1986, **131**, 266.
11 K. O'Neill and W. F. DeGrado, *Science*, 1990, **250**, 646.
12 D. S. Wishart, B. D. Sykes and F. M. Richards, *J. Am. Chem. Soc.*, 1996. in the press.
13 G. T. Dolphin, L. Brive and L. Baltzer, *J. Am. Chem. Soc.*, 1996, **118**, 11297.
14 G. T. Dolphin and L. Baltzer, unpublished results.

Paper 6/05908B
Received 27*th August* 1996
Accepted 10*th October* 1996