Pergamon

Bioorganic & Medicinal Chemistry 7 (1999) 83–91

BIOORGANIC & MEDICINAL CHEMISTRY

Designed Four-helix Bundle Catalysts—the Engineering of Reactive Sites for Hydrolysis and Transesterification Reactions of *p*-Nitrophenyl Esters

Lars Baltzer,* Kerstin S. Broo, Helena Nilsson and Jonas Nilsson

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

Received 7 May 1998; accepted 21 September 1998

Abstract—Four-helix bundle proteins have been designed that catalyze the hydrolysis and transesterification reactions of *p*-nitrophenyl esters by a cooperative nucleophilic and general acid mechanism. The catalysts consist of two 42-residue peptides that fold into helix-loop-helix motifs and dimerise. They have previously been shown to recognize anionic and hydrophobic substrates and to follow saturation kinetics. The catalytic entity is a $HisH^+$–His pair in a helical segment spaced *i*, *i*+4, which can be supplemented by arginines and lysines in the adjacent helix. The binding residues have now been optimized for the catalysis of mono-*p*-nitrophenyl fumarate hydrolysis and found to vary with the location of the site. The catalytic efficiency of the $HisH^+$–His site in helix II in positions 30 and 34 is enhanced by the introduction of arginine and or lysine residues in positions 11 and 15, but not in 8 and 11 or in 15 and 19. The most efficient catalyst using this site, $JNIIR_{11}K_{15}$, catalyses the reaction with a second-order rate constant of $0.134\,M^{-1}\,s^{-1}$ in aqueous solution at pH 5.1 and 290 K. The second-order rate constant is larger than those of the corresponding sites with 'longer' and 'shorter' binding residues. Similar experiments have shown that the efficiency and selectivity of catalysts based on a $HisH^+$-11–His-15 site in helix I are enhanced the most by the introduction of Lys-30 and Arg-34.

Introduction

De novo designed folded polypeptides are templates of sufficient size and complexity to accommodate the incorporation of amino acid residues in large varieties of combinations and geometries without disrupting the motif. Concepts and principles in enzymology and strategies in the design of enzyme-like catalysts can thus be tested in sophisticated model systems. Provided that the tertiary fold can be predicted and determined, the naturally occurring amino acids, nonnatural amino acids, and cofactors can be organized, in principle, at will in three-dimensional space to provide reactive sites for the catalysis of large numbers of chemical reactions. The ultimate goal is to develop the understanding of how to engineer tailor-made catalysts with the efficiency and selectivity of natural enzymes for reactions that do not occur in nature. Designed coiled-coils,[1,2] four-helix bundle proteins,[3–5] and a ββα motif[6] have now been reported that fold into well-defined tertiary structures and the engineering of reactive sites on the surface of folded polypeptides that approach 100 residues in length is therefore due.

To date, the de novo design of folded polypeptide catalysts have focused on reactions along the enamine pathway,[7] upon catalysts that enhance the effective concentration in bimolecular ligation reactions[8] and on histidine catalysis of acyl transfer reactions of *p*-nitrophenyl esters.[9–11] These polypeptides all have some enzyme-like properties and they are efficient biocatalysts that follow saturation kinetics. They show rate enhancements over those of the uncatalyzed reactions of three orders of magnitude or more and they are engineered in helical motifs.

Oxaldie 1 and 2 are 14 residue peptides that fold into helical conformations and aggregate to form bundle-like structures.[7] The decarboxylation of oxaloacetate is catalyzed with second-order rate constants that are almost three orders of magnitude larger than that of the butylamine catalyzed reaction. An important catalytic feature, in addition to substrate and transition state binding, is the low pK_a of the N-terminal amino group that functions as the nucleophile in the formation of the enamine intermediate and k_{cat} and K_M are $0.0055\,s^{-1}$ and 8.7 mM, respectively.

The catalysis of the ligation reaction between the unprotonated N-terminal amino group of a helical peptide and the C-terminal thioester of a second helical peptide, was accomplished by a 33-residue peptide that

Key words: De novo design; catalysis; four-helix bundle; ester hydrolysis; reactive site.

*Corresponding author. Tel.: +46-31-772-2894; fax: +46-31-772-3840; e-mail: lars.baltzer@oc.chalmers.se

0968-0896/99/$ - see front matter
PII: S0968-0896(98)00218-1

folds into an amphiphilic helix.[8] The catalyst's function is primarily to increase the effective concentration of the reactants on the surface of the folded motif and the rate enhancements over background are larger than 4×10^3, mainly due to substrate binding in the μM range.

A different approach was taken in the catalysis of acyl-transfer reactions by KO-42 (Fig. 1), a histidine-based helix-loop-helix motif with 42 residues that dimerizes to form a four-helix bundle.[9–11] KO-42 catalyzes the hydrolysis and transesterification reactions of *p*-nitrophenyl esters in a two-step reaction (Fig. 2). In the first and rate limiting step the unprotonated form of the histidine side chain reacts with the ester to form an acyl intermediate under the release of *p*-nitrophenol. In the second step, the acyl intermediate reacts with the most potent nucleophile in solution to form the reaction products. The measured second-order rate constant at pH 4.1 and 290 K for the hydrolysis of mono-*p*-nitrophenyl fumarate was 1140 times larger than that of the 4-methyl imidazole catalyzed reaction. We have recently elucidated the functions of the reactive residues[10] and demonstrated substrate recognition and saturation kinetics.[11] We now wish to report that the reactive sites that can be engineered in the helix-loop-helix motif are highly specific and suitable for the rational design of catalysts with tailor-made specificities.

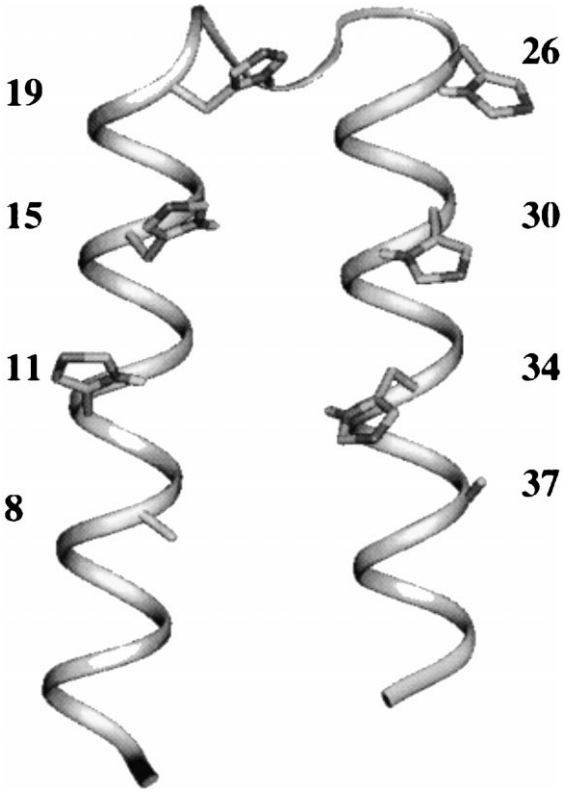


Figure 1. The modeled structure of KO-42. Only the histidine side chains and those of alanines 8 and 37 are shown as these are the positions relevant to the optimization of the reactive sites. KO-42 is a symmetric dimer, but only the monomer is shown for reasons of clarity of presentation.

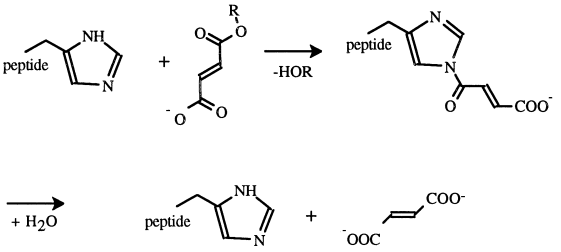


Figure 2. The reaction mechanism of histidine catalyzed hydrolysis of mono-*p*-nitrophenyl fumarate. The first and rate-limiting step is the formation of the acyl intermediate under the release of *p*-nitrophenol. In the second step the acyl intermediate reacts with the most efficient nucleophile to form the reaction products. In aqueous solution the reaction leads to hydrolysis, whereas in dilute trifluoroethanol the corresponding ester is formed.

Four-helix Bundle Histidine-based Catalysts

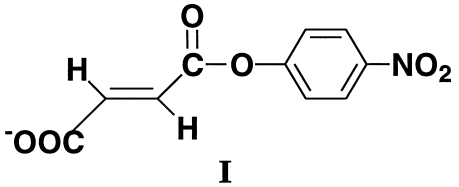
The structure and reactivity of the histidine-based four-helix bundle catalyst KO-42

The solution structure of KO-42 was determined by NMR and CD spectroscopy and equilibrium sedimentation ultracentrifugation.[9] The mean residue ellipticity at 222 nm was –24000 deg cm^2 $dmol^{-1}$ which corresponds to a helical content of more than 60%. The ^{1}H NMR spectrum was assigned in 2, 4, and 6% vol trifluoroethanol-d_3 (TFE-d_3) in 90% vol H_2O/10% vol D_2O at pH 5.8 and 323 K and the helical segments were identified from the medium-range NOEs αH–NH, i, $i+3$ and i, $i+4$, and from the αH chemical shift deviations from their random coil values. The helical segments extended from residues 3–18 and from residues 27–41, and the formation of the hairpin motif was verified by the identification of long-range NOE connectivities between the aromatic protons of Phe-35 and Phe-38 and the methyl groups of residues that form the hydrophobic core of the folded four-helix bundle motif. KO-42 was found by ultracentrifugation to form a dimer in the μM concentration range.

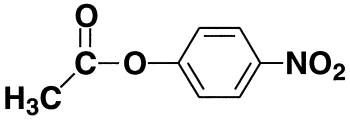
The reactivity of histidine residues is due to the reactivity of the imidazole side chain that is an efficient nucleophile at neutral pH and that, in addition, can function as a general-acid and a general-base catalyst. The protonated form of imidazole is a good proton donor and the unprotonated form is a good proton acceptor in hydrogen bonding. The versatility of the histidine side chain was exploited in KO-42 that was designed to have three histidine residues in each helix, His-11, His-15 and His-19 in helix I and His-26, His-30 and His-34 in helix II. KO-42 catalyzes the hydrolysis of mono-*p*-nitrophenyl fumarate (**I**) with a second-order rate constant of 0.10 M^{-1} s^{-1} in aqueous solution at pH 4.1 and 290 K, which is 1140 times larger than that of the 4-methyl imidazole catalyzed reaction. In 10% vol TFE the corresponding rate constant ratio for the formation of mono-trifluoroethyl fumarate is 620.

The pK_a values of KO-42 were determined by ^{1}H NMR spectroscopy and found highly perturbed; they were 6.9, 5.4, 7.0, 7.2, 5.3, and 5.2 for His-11, His-15, His-19, His-26, His-30, and His-34, respectively. The pH

dependence of the second-order rate constant was deceptively simple since an expression that describes the dissociation of a monoprotonic acid with a pK_a of 5.1 could be fitted to the experimental data in spite of the fact that the six histidine residues give rise to 64 different titrating species. At high pH, the measured second-order rate constant is very close to the sum of the calculated values for six histidines with the pK_a values of those of KO-42. The calculations were based on the Bronsted relation, the second-order rate constant of the 4-MeIm catalyzed reaction ($1.3\,M^{-1}\,s^{-1}$), the pK_a of 4-MeIm (7.95) and a Bronsted coefficient β of 0.8, which is the reported value for imidazole catalyzed hydrolysis of *p*-nitrophenyl acetate (**II**). In the pH independent region, KO-42 therefore functions as six non-cooperative His residues. At low pH, however, the pH dependence of the rate constants and the measured pH values rule out the possibility that the unprotonated peptide is the reactive species.



I



II

The kinetic solvent isotope effect at pH 4.7 was determined and found to be 2.0, which shows that there is isotopic fractionation in the transition state of the catalyzed reaction, suggesting that the reaction mechanism is cooperative nucleophilic and general-acid catalysis.

The minimal reactive sites of KO-42

KO-42 contains six histidine residues, three of which have depressed pK_a values, and it is therefore likely that several reaction pathways are catalyzed by independent reactive sites. In order to identify the reactive entities of KO-42 several peptides were synthesized that each contain two or three histidines in the same positions as in KO-42, so that each combination could be studied independently[10] (Fig. 3). The three histidines of helix I and the three histidines of helix II of KO-42 were incorporated in the peptides MN-42 and JN-42 and the second-order rate constants of the MN-42 and JN-42 catalyzed reactions of **I** at pH 5.1, $0.027\,M^{-1}\,s^{-1}$ and $0.065\,M^{-1}\,s^{-1}$, were found not to add up to that of the KO-42 catalyzed reaction, $0.31\,M^{-1}\,s^{-1}$. It was concluded that there is interhelical cooperativity in KO-42 and that protonated His residues in helix II of KO-42 provide transition state binding in the reactions where the His residues of helix I are the nucleophiles and general-acid catalysts, and vice versa. Support for this conclusion came from the observation that MNRR, a 42 residue peptide with the sequence of MN-42 except that Arg-30 and Arg-34 were incorporated in helix II, Figure 4, catalyzed the hydrolysis of **I** with a second-order rate constant of $0.080\,M^{-1}\,s^{-1}$, a rate constant that is approximately three times as large as that of the MN-42 catalyzed reaction.

Further dissection of the three residue sites led to the peptides JNI, JNII, JNIII, MNI, and MNII, where all the possible neighboring two-residue sites of KO-42 were engineered into helix-loop-helix motifs, i.e., JNI contains His-26–His-30, JNII contains His-30–His-34, MNI contains His-11–His-15 and MNII contains His-15–His-19.[10] JNIII with His-26–His-34 was incorporated in the series to be a reference peptide because the His residues are too far apart in the sequence to be capable of cooperative catalysis in ester hydrolysis. The second-order rate constants at pH 5.1 of JNI ($0.010\,M^{-1}\,s^{-1}$) and of JNII ($0.055\,M^{-1}\,s^{-1}$) add up to that of JN-42 ($0.065\,M^{-1}\,s^{-1}$) and those of MNI ($0.011\,M^{-1}\,s^{-1}$) and MNII ($0.008\,M^{-1}\,s^{-1}$) are close to that of MN-42 ($0.027\,M^{-1}\,s^{-1}$) and they are all larger than that of the 4-MeIm catalyzed reaction. The two-residue sites are therefore independent entities capable of cooperative catalysis of hydrolysis of *p*-nitrophenyl esters. The two-residue sites with histidines in positions i, $i+4$, are consequently the minimal cooperative reactive sites in helical segments and the cooperativity was also observed for peptide catalysts with the $HisH^+$–His configuration i, $i+3$.[12]

The reactivity difference between JNI and JNII was found to be due mainly to differences in pK_a values of the histidines and it is compatible with a model where the His with the highest number in the sequence is the nucleophile whereas the His with the lower number in the sequence in its protonated form is the proton donor. The conclusions were based on the calculation of the reactivities of the catalysts from their pK_a values, using the Bronsted equation and Bronsted coefficients β of 0.8 and α of 0.65.

Substrate recognition by folded polypeptide catalysts based on the $HisH^+$–His pair supplemented by binding residues

Because the minimal reactive sites were found to be the two-residue sites in helical segments the possibility of supplementing these entities with residues capable of further transition state binding was explored by introducing arginines, ornithines and lysines in binding positions.[11] A small peptide library was constructed by introducing all possible combinations of Arg, Gln and Lys in positions 30 and 34 in helix II of MN-42 (Fig. 4), and the reactivities of these peptides towards **I** and **II** were determined in aqueous solution at pH 5.1 and 290 K. The introduction of positively charged proton donors in helix II increased the second-order rate constants in all cases, but to different degrees, in the catalysis of hydrolysis of the two substrates. The largest

1

Ac-N-A-A-D-Nle-E-A-A-I-K-**H**-L-A-E-**H**-Nle-A-A-**H**-

20 23

-G-P-V-D-

42

NH_2-G-A-R-A-F-A-E-F-**H**-K-A-L-**H**-E-A-Nle-**H**-A-A-

KO-42

peptide	8	11	15	19	26	30	34	37	k_2 (I) M^{-1} s^{-1}	k_2 (II) M^{-1} s^{-1}
KO-42[a]	A	**H**	**H**	**H**	**H**	**H**	**H**	A	0.310	0.290
MN-42[b]	A	**H**	**H**	**H**	Q	Q	A	A	0.027	0.030
JN-42[b]	A	A	Q	K	**H**	**H**	**H**	A	0.065	-
MNKK[c]	A	**H**	**H**	**H**	Q	K	K	A	0.068	0.086
MNRR[b]	A	**H**	**H**	**H**	Q	R	R	A	0.080	0.049
MNKR[c]	A	**H**	**H**	**H**	Q	K	R	A	0.135	0.075
MNRK[c]	A	**H**	**H**	**H**	Q	R	K	A	0.110	0.106
MNRQ[c]	A	**H**	**H**	**H**	Q	R	Q	A	0.058	0.073
MNQR[c]	A	**H**	**H**	**H**	Q	Q	R	A	0.066	0.056
MNOrnR[c]	A	**H**	**H**	**H**	Q	Orn	R	A	0.092	0.081
MNKhR	A	**H**	**H**	**H**	Q	K	hR	A	0.090	0.059
MNI[b]	A	**H**	**H**	Q	Q	Q	A	A	0.011	-
MNII[b]	A	A	**H**	**H**	Q	Q	A	A	0.008	-
$MNIR_{26}R_{30}$	A	**H**	**H**	Q	R	R	A	A	0.035	-
$MNIR_{30}R_{34}$	A	**H**	**H**	Q	Q	R	R	A	0.028	-
$MNIR_{34}R_{37}$	A	**H**	**H**	Q	Q	Q	R	R	0.027	-
JNI[b]	A	A	Q	K	**H**	**H**	A	A	0.010	-
JNII[b]	A	A	Q	K	Q	**H**	**H**	A	0.054	0.048
JNIII[b]	A	A	Q	K	**H**	Q	**H**	A	0.007	-
$JNIIhR_{11}K_{15}$	A	hR	K	K	**H**	Q	**H**	A	0.087	0.053
$JNIIR_{11}R_{15}$[b]	A	R	R	K	Q	**H**	**H**	A	0.105	0.067
$JNIIR_{11}K_{15}$	A	R	K	K	Q	**H**	**H**	A	0.134	0.088
$JNIIR_{11}Orn_{15}$	A	R	Orn	K	Q	**H**	**H**	A	0.110	0.055
$JNIIK_{11}R_{15}$	A	K	R	K	Q	**H**	**H**	A	0.123	0.049
$JNIIK_{11}hR_{15}$	A	K	hR	K	Q	**H**	**H**	A	0.095	0.048
$JNIIOrn_{11}R_{15}$	A	Orn	R	K	Q	**H**	**H**	A	0.099	0.051
$JNIIR_{15}R_{19}$	A	A	R	R	Q	**H**	**H**	A	0.049	-
$JNIIK_8R_{11}$[c]	K	R	Q	K	Q	**H**	**H**	A	0.065	0.039
$JNIIR_8K_{11}$[c]	R	K	Q	K	Q	**H**	**H**	A	0.055	0.018

Figure 3. The amino acid sequence of KO-42 and the reactive site configurations of the peptide catalysts given in the one-letter code, together with the second-order rate constants for the peptide catalyzed reactions in aqueous solution at pH 5.1 and 290 K. Superscripts refer to [a] ref. 9, [b] ref. 10 and [c] ref. 11.

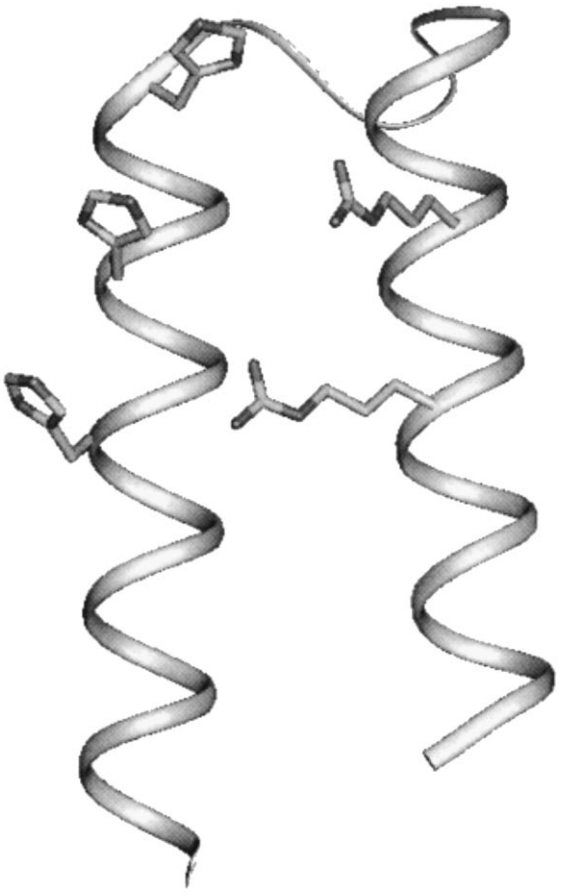


Figure 4. The structure of MNRR showing His-11, His-15 and His-19 as well as Arg-30 and Arg-34 in the positions of the binding residues of the peptides in the MN library. In MNI the histidine residues are His-11 and His-15, and in MNII the histidine residues are His-15 and His-19.

catalytic efficiency in the hydrolysis of anionic substrates was observed for the helix-loop-helix motif MNKR that has the sequence of MN-42 except that Lys-30 and Arg-34 have been introduced into helix II. The differences in the second-order rate constants for the hydrolysis of **I** ($0.135\,M^{-1}\,s^{-1}$) and of **II** ($0.075\,M^{-1}\,s^{-1}$) demonstrate a high degree of recognition of the fumaryl carboxylate anion. If Lys-30 is replaced by Ornithine-30 no fumaryl recognition is observed as the second-order rate constants for the hydrolysis of the two substrates at pH 5.1 and 290 K in aqueous solution are the same within experimental error.

Recognition of hydrophobic residues, too, was observed in the hydrolysis of a series of fatty acid esters, where the second-order rate constant increased with the length of the aliphatic chain, but hydrophobic recognition was only observed for the peptides that had the $HisH^+$–His site in helix II, the JN series of peptides. Apparently, the orientation of the helical segments differ so that the proximity of the reactive histidines to the hydrophobic core are not the same for the two helices. The rate enhancements due to the introduction of the aliphatic residues amounted to a factor of 14 in the JNIII catalyzed hydrolysis of *p*-nitrophenyl valerate ($k_2 = 0.097\,M^{-1}\,s^{-1}$) and **I** ($k_2 = 0.007\,M^{-1}\,s^{-1}$). However, it has not so far been possible to exploit simultaneous binding of a carboxylate anion and a hydrophobic residue and demonstrate synergistic rate enhancements.

A modest amount of chiral recognition was also observed in the hydrolysis of the *p*-nitrophenyl esters of D- and L-norleucine, where the D-form was hydrolyzed a factor of two faster than the L-form by peptides in the MN-42 based library. Chiral discrimination was observed when the MN motif was supplemented by arginines and lysines in positions 30 and 34 of helix II, whereas no discrimination was obtained in the absence of binding residues in helix II.

Substrate binding and saturation kinetics

The introduced flanking binding residues that enhance the reactivity of the peptide also binds the substrate and MNKR shows saturation kinetics with a k_{cat}/K_M of $0.17\,M^{-1}\,s^{-1}$. The dissociation constant K_M of the MNKR catalyzed reaction is 1 mM, which is clearly within the range of natural enzymes although k_{cat} is too small to provide overall enzyme-like efficiency. In order to optimize the peptides for the catalysis of chemical transformations and, in particular, for achieving cooperative transition-state binding of several functional groups of the substrate, the structure of the reactive sites must be known so that groups can be introduced in positions that favor the most efficient reaction pathway. Because it is very difficult to estimate by NMR spectroscopy the favored rotamers of the freely rotating side chains on the surface of the folded motif, as well as the 'twist' of the helical segments along the helical axis, we have designed a strategy where the optimum geometries of the side chains are probed by measuring the reactivities of the peptides as a function of the structure and the positions of the side chains on the surface of the folded four-helix bundle. We now wish to report on the optimum geometries of the reactive sites in helix-loop-helix motifs designed to catalyze acyl-transfer reactions of *p*-nitrophenyl esters.

Results and Discussion

Design and structure of catalytic helix-loop-helix motifs

The design of the four-helix bundle motif has been described in detail, previously.[9,13] In short, the amino acids were chosen based on their propensity for secondary structure formation and their capacity for stabilizing the folded helical segments by capping, salt bridge formation and stabilization of the helix dipole. Hydrophobic residues were incorporated so that in the folded state the helical segments would be amphiphilic and form a hairpin motif that could dimerise to form an antiparallel four-helix bundle. The de novo design of

folded polypeptides and proteins has been reviewed by Bryson et al.[14] On the surface of the folded motif, residues were incorporated that formed reactive sites with histidine residues flanked by side chains capable of transition state binding (Fig. 1).

The structures of the template peptides SA-42 and KO-42 were studied extensively by NMR and CD spectroscopy and by equilibrium sedimentation ultracentrifugation[9,13,15] and they were found to fold into the designed four-helix bundle motifs. All polypeptides reported here deviate from the sequence of KO-42 by four amino acid residues or less and the measured mean residue ellipticities are all in the range from -20000 to -25000 deg cm^2 dmol^{-1} in aqueous solution at room temperature and pH 5. We therefore conclude that they fold into the designed helix-loop-helix dimer motifs because the mean residue ellipticity is a good probe of dimerization as the unfolded monomers often have low helical contents.

Design of reactive sites that probe the geometry of the transition state

The HisH$^+$–His sites, supplemented by binding residues, can discriminate between anionic and uncharged substrate side chains in the hydrolysis and transesterification reactions of *p*-nitrophenyl esters.[11] The modest observed levels of discrimination do, however, suggest that several reaction pathways are used to form the reaction products and that bond distances and angles are not optimal. We therefore wish to optimize the geometry of the reactive site of the four-helix bundle motif so that the substrates and the transition states can be bound by complementary amino acid side chains in unstrained conformations.

Understanding the structure of the reactive site is thus a prerequisite for the design of selective and efficient catalysts but the dominant orientations of the reactive residues are difficult to determine. Kinetic measurements, however, provide a way of measuring whether a modification of the amino acid composition leads to improved transition state binding. We have now designed and synthesized a number of four-helix bundle catalysts with the objective of determining what the optimum positions and side chains are for the amino acids in reactive sites that catalyze the hydrolysis of *p*-nitrophenyl esters.

The most efficient HisH$^+$–His site in the helix-loop-helix motif is that of JNII, that has histidine residues with depressed pK_a values in positions 30 and 34, and catalyzes the hydrolysis of **I** with a second-order rate constant of 0.055 M^{-1} s^{-1}. It has previously been shown that the introduction of Arg-11 and Arg-15 enhances the second-order rate constant by approximately a factor of two ($k_2 = 0.105$ M^{-1} s^{-1}).[10] In order to determine whether better transition state stabilization could be obtained with binding residues in positions 8 and 11, or in 15 and 19 JNII was synthesized with arginines introduced pairwise in those positions. Positions 8, 11, 15, and 19 are next to the residues in the hydrophobic core and are therefore those that are the most likely to interact with the bound substrate or the transition state. Similar design principles for the MN series of peptides led to the synthesis of three 42-residue peptides with the HisH$^+$–His site in positions 11 and 15, and arginine residues in positions 26 and 30, 30 and 34 and 34 and 37. In addition, the lengths and functional groups of the side chains were varied to optimize bonding distances in the transition state. The length of the side chain that gives rise to the largest rate constant provides an estimate of the interresidue distances within the reactive site, although the bound conformation of the side chain will be an important factor too. To that end, JNII(His-30, His-34) was synthesized with combinations of arginine, homoarginine, lysine, and ornithine residues in positions 11 and 15 (Fig. 3).

The reactive site of JNII in the catalyzed hydrolysis of *p*-nitrophenyl esters

The second-order rate constant for JNII catalyzed hydrolysis of **I** is 0.055 M^{-1} s^{-1} and introduction of Arg-11 and Arg-15 increases the rate constant to 0.105 M^{-1} s^{-1}. The second-order rate constants therefore demonstrate that the incorporated residues provide extra transition state stabilization. The second-order rate constants for the corresponding reaction with peptides that contain Arg-8–Lys-11 and Lys-8–Arg-11 are 0.055 M^{-1} s^{-1} and 0.066 M^{-1} s^{-1} which shows that arginine residues in positions 8 and 11 do not bind the anionic fumaryl residue or the developing oxyanion, or bind very weakly. The second-order rate constant for the reaction catalyzed by JNII supplemented with Arg-15 and Arg-19 also does not provide any rate enhancement over that of the JNII catalyzed reaction as k_2 is 0.049 M^{-1} s^{-1}. The reactive site based on the HisH$^+$–His pair in positions 30 and 34 can therefore only be further stabilized by residues in positions 11 and 15 and the reactive site is thus geometrically well defined.

The reactive site geometries of MNI and MNII in the catalyzed hydrolyses of *p*-nitrophenyl esters

MN-42 catalyses the hydrolysis of **I** with a second-order rate constant of 0.027 M^{-1} s^{-1}. The corresponding rate constants for the MNI (His-11, His-15) and MNII (His-15, His-19) catalyzed reactions are 0.011 M^{-1} s^{-1} and 0.008 M^{-1} s^{-1}, and there is thus no dominant two-residue site in MN-42. MNRR (His-11, His-15, His-19, Arg-30, Arg-34) is more reactive than MN-42, k_2 is 0.080 M^{-1} s^{-1}, showing that one or both of the sites of helix I are improved by adjacent arginines. The second-order rate constant of MNIR$_{30}$R$_{34}$ was found to be 0.028 M^{-1} s^{-1}, which is slightly less than half of that of the MNRR catalyzed reaction, and the reactivity of MNRR can therefore be concluded to be due to two sites of approximately equal reactivity (Fig. 4). The second-order rate constant for the MNIR$_{26}$R$_{30}$ catalyzed reaction was 0.035 M^{-1} s^{-1} and that of the MNIR$_{34}$R$_{37}$ catalyzed reaction was 0.027 M^{-1} s^{-1}. The transition state binding by arginine residues in helix II is therefore less precise. From an inspection of the modeled

structure of MNRR (Fig. 4) it is very unlikely that Arg-26 can bind residues in an acyl intermediate at the side chain of His-15, suggesting that only Arg-30 binds in the $MNIR_{26}R_{30}$ catalyzed reaction. Since the magnitudes of the rate enhancements for the $MNIR_{26}R_{30}$, the $MNIR_{30}R_{34}$ and the $MNIR_{34}R_{37}$ catalyzed reactions are approximately equal, and roughly equal to that of half the rate enhancement of MNRR it seems likely that only one arginine binds in all of these peptides. The second-order rate constants of the peptide catalyzed reactions are therefore compatible with a model where Arg-30 may bind the anionic transition state of the reaction catalyzed by $HisH^+$-15–His-19, and Arg-34 may bind the transition state of the reaction catalyzed by $HisH^+$-11–His-15, but where Arg-30 is not in a position to bind in the transition state of the MNI catalyzed reaction and Arg-34 does not bind in the MNII catalyzed one. Unfortunately, there is also no extra binding energy provided by the flanking arginines in positions 26 and 37 in the MNI and MNII catalyzed reactions.

Reactivity as a function of the length of the side chains of the binding residues

It was shown previously that the reactivity of the MN-42 based motifs depended strongly on the nature of the side chain of the binding residues.[11] Two arginines bind less well than one arginine and one lysine, and the same holds for two lysines. The best binding occurs for unlike amino acids, perhaps due to repulsion of residues of equal charge and length. The largest rate constant for the peptides in the MN series was obtained by introducing Lys-30–Arg-34 ($k_2 = 0.135\,M^{-1}\,s^{-1}$), whereas Arg-30 and Lys-34 were slightly less reactive ($k_2 = 0.110\,M^{-1}\,s^{-1}$). In order to probe whether the rate constants of the MN series could be enhanced further Arg was replaced by homo arginine (hR) in MNKhR, because homoarginine has a longer side chain that can bind remote substrates better. The second-order rate constant has now been measured and it is $0.090\,M^{-1}\,s^{-1}$, which is lower than that of MNKR. It appears that in the catalysis of acyl-transfer reactions of mono-*p*-nitrophenyl fumarate the optimum binding distance for a residue in the 34 position in the best possible conformation is obtained if an arginine residue is used. MNKR is therefore the most efficient catalyst in the MN series of peptides and the rate enhancement accomplished by the introduction of Lys-30 and Arg-34 can be estimated for each of the two-residue sites. According to the arguments discussed above, the reactivity of MNIKR, i.e., the MN-42 motif with His-11, His-15, Lys-30 and Arg-34, should be approximately half of that of MNKR which catalyzes the hydrolysis of **I** with a second order rate constant of $0.135\,M^{-1}\,s^{-1}$. It should therefore be approximately $0.07\,M^{-1}\,s^{-1}$ which corresponds to a reactivity enhancement of more than a factor of six, and a $\Delta\Delta G$ of approximately 1.1 kcal/mol, over that of the MNI catalyzed reaction. The observed stabilization is approximately a factor of two larger than what is obtained from a single salt bridge,[16] suggesting that both Lys-30 and Arg-34 bind in the MNKR catalyzed reaction.

The same experiments were carried out for the JNII series of peptides where it was shown that residues in positions 8 and 19 contribute little to transition state binding in the catalysis of acyl-transfer reactions. JNII was therefore supplemented with homoarginines, arginines, lysines, and ornithines in positions 11 and 15; the second-order rate constants are given in Figure 3. Systematic variation of the residues shows interesting optima. Arg-11 has been combined with Arg-15, Lys-15 and Orn-15 and the maximum reactivity is observed for Lys-15 suggesting that Lys is the optimum residue in position 15 for the hydrolysis of **I**. When Lys-11 is supplemented by homoarginine-15 and Arg-15 the reactivity increases as the residue in position 15 is 'shortened', in agreement with the hypothesis that Lys-15 may provide the maximum transition state binding. Only a single catalyst with an even shorter side chain in position 11, that of Orn-11, has been studied. $JNIIOrn_{11}R_{15}$ shows a rate constant that is $0.099\,M^{-1}s^{-1}$ and a rate enhancement that is expected to arise from the substitution of Arg-15 by Lys-15 would bring its reactivity to within the range of that with Arg-11 Lys-15. These results suggest that Lys should be the optimum residue in position 11, too. The series of peptides that are based on Arg-15 do indeed show a maximum for Lys-11, whereas the catalysts with Arg-11 or Orn-11 are less reactive. Also, the two Lys-15 based catalysts containing Arg-11 or hArg-11 are more reactive the shorter the side chain in position 11 in good agreement with the conclusion that Lys-11 provides optimum stabilization.

The binding distances appear to be such that the naturally occurring residues Arg and Lys provide the best transition state stabilization possible as longer and shorter side chains give rise to less efficient catalysts. The Lys-11 Lys-15 combination will perhaps prove to be the most efficient although the lesson from earlier studies is that the combination of two identical residues is less reactive than non-identical ones.

The magnitude of the second-order rate constant for the $JNIIR_{11}K_{15}$ catalyzed reaction is the largest obtained for a single well-defined reactive site based on the reactivity of the $HisH^+$–His pair. The rate enhancement over that of the JNII catalyzed reaction approaches a factor of three. It is, however, for the catalyst $JNIIK_{11}R_{15}$ that has a slightly lower reactivity, that the most pronounced discrimination between **I** and **II** is observed, a factor of 2.5.

The rate constant maxima for both the JN and the MN series of peptides occur for the same type of binding residues, Arg and Lys, although in different order. In the MN series Lys has the lower number in the sequence whereas in the JN series Arg has the lower number. In the MN series the Lys residue is therefore located opposite to the nucleophilic histidine, whereas in the JN series it is opposite to the protonated one, suggesting that the reactive sites may bind the transition states in different geometries, something that may turn out to have interesting stereochemical consequences. The fact that Arg and Lys can bind cooperatively in the

hydrolysis of **I** and **II**, and that they bind better than residues with 'shorter' and 'longer' side chains shows that the dimensions of the template is sufficient for the engineering of versatile catalysts. The reactive sites that can be engineered on the surface of four-helix bundle motifs can clearly accommodate substrates and reagents with sizes that are common in organic synthesis.

Substrate recognition

The parent peptide catalyst, KO-42, showed little discrimination between **I** and **II**, probably because several reaction pathways with different stereochemical configurations were available. The MN series of catalysts with arginine and lysine residues in helix II was shown previously to be able to discriminate between the carboxylate anion of **I** and the nonpolar methyl group of **II**. The discrimination between substrates is even more pronounced as the transition state binding is improved in the catalysts presented here. The rate constant ratio $k_2(\mathbf{I})/k_2(\mathbf{II})$ is 2.5 for the catalyst $JNIIK_{11}R_{15}$, 2.0 for $JNIIK_{11}hR_{15}$ and 1.5 for MNKhR which reacts via two independent sites, each with a rate constant of approximately $0.05\,M^{-1}\,s^{-1}$. The least reactive peptides in the MN series show essentially no discrimination at all[11] and there is an observable trend that increased reactivity leads to increased selectivity, probably because the introduced binding residues bind specific groups in the substrate. There seems to be some justification for the expectation that as the efficiency of catalysts are improved by the introduction of residues that bind the substrate and the transition state, selectivity will improve.

Conclusions

The optimum amino acid composition for reactive sites in designed four-helix bundle proteins that catalyze the hydrolysis and transesterification reactions of *p*-nitrophenyl esters has been elucidated in peptides derived from KO-42. The reactivity of the $HisH^+$–His site of JNII can only be improved by the introduction of residues in the positions 11 and 15 and the optimum reactivity in the hydrolysis of mono-*p*-nitrophenyl fumarate was obtained by introducing Arg-11 and Lys-15. Longer and shorter side chains were shown to lead to less reactive catalysts. In the MN series of catalysts, that are intrinsically less reactive than the JN series, the binding by flanking residues was less precise although the rate enhancements were larger than in the JNII derived catalysts. The less well-defined geometry of the reactive sites of MN-42 suggests that JNII may provide the more pronounced discrimination between alternative transition state structures. It cannot be assumed that the binding residues discussed here are in unstrained conformations for the catalysis of **I** since the template is a crude one in terms of optimizing bond distances and angles, even for the shallow potentials of non-covalent bonds. Other substrates will require other reactive site geometries and it is likely that some will fit very well within the structural constraints of the four-helix bundle motif. The rate enhancements obtained in such reactions will provide good measures of what the catalytic limits are in reactions catalyzed by surface catalysts. Nevertheless, it has been demonstrated that the optimization of the reactive site geometry is possible in a rational way. Such optimization strategies may prove to be of general use in the design of folded polypeptide catalysts.

Experimental

Peptide synthesis, purification, and identification. The peptides were synthesized using a PerSeptive Biosystems Pioneer automated peptide synthesizer using a standard Fmoc chemistry protocol. The carboxy terminals were amidated upon cleavage from the resin by using a Fmoc-PAL-PEG-PS polymer (PerSeptive Biosystems) and the amino terminals were capped with acetic acid anhydride. The peptides were cleaved from the polymer and deprotected with TFA (9 mL), anisole (200 μL), ethanedithiol (300 μL), and thioanisole (500 μL) for 2 h at room temperature. After diethyl ether precipitation and lyophilization, the peptides were purified by reversed-phase HPLC on a semipreparative C-8 Kromasil, 7-μ column. The peptides were eluted isocratically using a solvent with 36–39% vol isopropanol in 0.1% TFA, flow rates of 5 mL min^{-1} and UV detection at 229 nm. The purity of each peptide was checked by reversed-phase analytical HPLC. The identities of the peptides were determined by electrospray mass spectrometry (VG Analytical, ZabSpec). The obtained molecular weights were within 1 au from the calculated and no high molecular weight impurities could be detected. The purity of each peptide was estimated to be more than 95% from HPLC and ES-MS.

CD spectroscopy. CD spectra were recorded on a Jasco J-720 spectropolarimeter, routinely calibrated with (+)-camphor-10-sulfonic acid. CD spectra were measured at room temperature in the wavelength interval 260 to 200 nm in 1-mm cuvettes. The samples were prepared in buffer solution and diluted by pipetting to the desired concentrations. The peptide concentrations of the stock solutions were determined by quantitative amino acid analysis.

Kinetic measurements. The kinetic experiments were carried out using Varian Cary 1 or Cary 4 spectrophotometers equipped with Varian temperature controllers by following the absorbance at 320 nm (*p*-nitrophenol) or 405 nm (*p*-nitrophenolate) as a function of time. A stock solution of peptide in 100 mM sodium acetate buffer was prepared; the pH was adjusted, if necessary, and the solution was centrifuged before diluting by pipetting to the desired concentrations. The concentrations of the peptide stock solutions were determined by quantitative amino acid analysis. In a typical kinetic experiment, 270 μL of peptide solution (0.2–0.5 mM) was temperature equilibrated in a 1-mm quartz cuvette, and 5 μL of substrate solution (7.4 mM) was added. The substrate mono-*p*-nitrophenyl fumarate was dissolved in 50:50 acetonitrile:buffer and

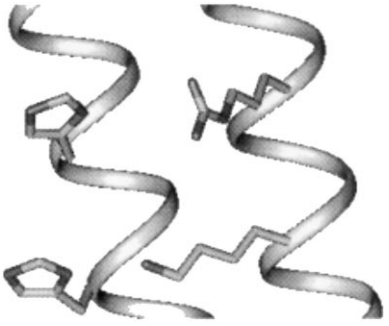


Figure 5. The structure of JNII (Arg-11, Lys-15) the most efficient single-site catalyst demonstrating the geometric relationship between the two His residues and the only positions that can provide transition state stabilization in the hydrolysis and transesterification of mono-*p*-nitrophenyl fumarate.

p-nitrophenyl acetate was dissolved in 100% acetonitrile. The reactions were followed for more than two half-lives and the pseudo-first order rate constants were determined by fitting a single exponential function to the experimental data using Igor Pro software (Wavemetrics Inc.). The reported rate constants are the results from linear regression analysis of the experimentally measured pseudo-first-order rate constants as a function of three or more peptide concentrations. The errors in the reported second order rate constants are due to errors in the linear regression analysis and the quantitative amino acid analysis. The error limits for the second order rate constants are probably $\pm 10\%$ (Fig. 5).

Acknowledgements

The authors are indebted to the Swedish Natural Science Research Council and Carl Tryggers Stiftelse for financial support.

References

1. Betz, S. F.; DeGrado, W. F. *Biochemistry* **1996**, *35*, 6955.
2. Hodges, R. S. *Biochem. Cell Biol*. **1996**, *74*, 133.
3. Dolphin, G. T.; Brive, L.; Johansson, G.; Baltzer, L. *J. Am. Chem. Soc*. **1996**, *118*, 11297.
4. Brive, L.; Dolphin, G. T.; Baltzer, L. *J. Am. Chem. Soc*. **1997**, *119*, 8598.
5. Hill, R. B.; DeGrado, W. F. *J. Am. Chem. Soc*. **1998**, *117*, 1138.
6. Struthers, M. D.; Cheng, R. P.; Impeirali, B. *Science* **1996**, *271*, 342.
7. Johnsson, K.; Allemann, R. K.; Widmer, H.; Benner, S. A. *Nature* **1993**, *365*, 530.
8. Severin, K.; Lee, H. D.; Kennan, A. J.; Ghadiri, M. R. *Nature* **1997**, *389*, 706.
9. Broo, K. S.; Brive, L.; Ahlberg, P.; Baltzer, L. *J. Am. Chem. Soc*. **1997**, *119*, 11362.
10. Broo, K. S.; Nilsson, H.; Nilsson, J.; Flodberg, A.; Baltzer, L. *J. Am. Chem. Soc*. **1998**, *120*, 4063.
11. Broo, K. S.; Nilsson, H.; Nilsson, J.; Baltzer, L. *J. Am. Chem. Soc*. **1998**, *120*, 10287.
12. Broo, K. S.; Baltzer, L., unpublished results.
13. Olofsson, S.; Johansson, G.; Baltzer, L. *J. Chem. Soc., Perkin Trans. 2*, **1995**, 2047.
14. Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neill, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935.
15. Olofsson, S.; Baltzer, L. *Folding Des*. **1996**, *1*, 347.
16. Lyu, P. C.; Gans, P. J.; Kallenbach, N. R. *J. Mol. Biol*. **1992**, *223*, 343.