# Setting the Stage for New Catalytic Functions in Designed Proteins—Exploring the Imine Pathway in the Efficient Decarboxylation of Oxaloacetate by an Arg–Lys Site in a Four-Helix Bundle Protein Scaffold

**Malin Allert[a] and Lars Baltzer*[b]**

**Abstract:** Fourteen 42-residue polypeptides have been designed to identify reactive sites for the catalysis of the decarboxylation of oxaloacetate, a chemical transformation that proceeds through the formation of an imine intermediate. The sequences fold into helix–loop–helix motifs and dimerise to four-helix bundles. The catalytically active lysine residues were incorporated in several surface exposed positions, but also in positions characterised by hydrophobic properties to reduce their p$K_a$ values. The molecular environments of the Lys residues were systematically varied, to find which residues were able to stabilise and bind the imine intermediate in the decarboxylation reaction. A two-residue Arg–Lys site formed the main component of the reactive site of the helix–loop–helix dimer Decarb-K34_R33, which obeyed saturation kinetics in catalysing the reaction with a $k_{cat}/K_M$ of $0.59\,M^{-1}s^{-1}$. The rate constant measured was nearly three orders of magnitude larger than the second-order rate constant of the butylamine-catalysed reaction ($0.0011\,M^{-1}s^{-1}$), and four orders of magnitude larger than the pseudo first-order rate constant of the uncatalysed reaction ($1.3 \times 10^{-5}\,s^{-1}$). The sequence of Decarb-K34_R33 contained only a single lysine residue. It was flanked by an arginine in the preceding position in the sequence. A flanking Arg residue provided more efficient catalysis than a flanking Lys or Gln residue. Arginines in flanking positions in the helix, in positions four residues before or after the Lys in the sequence, are not as important in catalysis as the Arg of the Arg–Lys pair. The effect of p$K_a$ on the catalytic efficiency of the Lys residue in the decarboxylation reaction is well known. The identification of the role of the flanking Arg residue in catalysing decarboxylation, its optimal position, and the importance of conformational stability reported here sets the stage for developing a number of catalytic systems that depend on the formation of imine intermediates, but that lead to different reaction products.

**Keywords:** decarboxylation · enzyme catalysis · imines · peptides · protein models

## Introduction

The quest for the fundamental principles of enzyme catalysis is a critical test of our understanding of the interplay between structure, dynamics, molecular interactions and molecular recognition in proteins. An understanding at the atomic level of how substrates, intermediates and transition states are recognised and bound by the catalyst, while the products are released, and how the catalytic residues control the chemical transformations remains elusive. While the study of native enzymes has provided considerable insights into the mechanisms of catalysis, a comprehensive set of rules for the construction of new enzymes has not emerged, probably because of the structural and functional complexity of naturally occurring proteins. In contrast, in de novo designed proteins,[1–13] catalytic residues can be introduced according to reaction mechanistic principles, to probe the relationship between structure and function in well-defined sites that have no hereditary biological function. In spite of their simplicity in comparison with native proteins, de novo designed sequences can serve as scaffolds of sufficient complexity for the introduction of multiple functions in variable geometries and are thus excellent model systems for the study of biocatalysis. A number of de novo designed proteins have now been reported, together with their three-dimensional structures and a variety of designed protein scaffolds have therefore become available.[4, 14–21]

The introduction of catalytic sites into folded polypeptide model systems have shown that catalytic efficiencies of three to four orders of magnitude are achievable in constructs that

[a] M. Allert
Department of Chemistry, Organic Chemistry, Göteborg University
41296 Göteborg (Sweden)

[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.wiley-vch.de/home/chemistry/ or from the author.

  0947-6539/02/0811-2549 $ 17.50+.50/0 

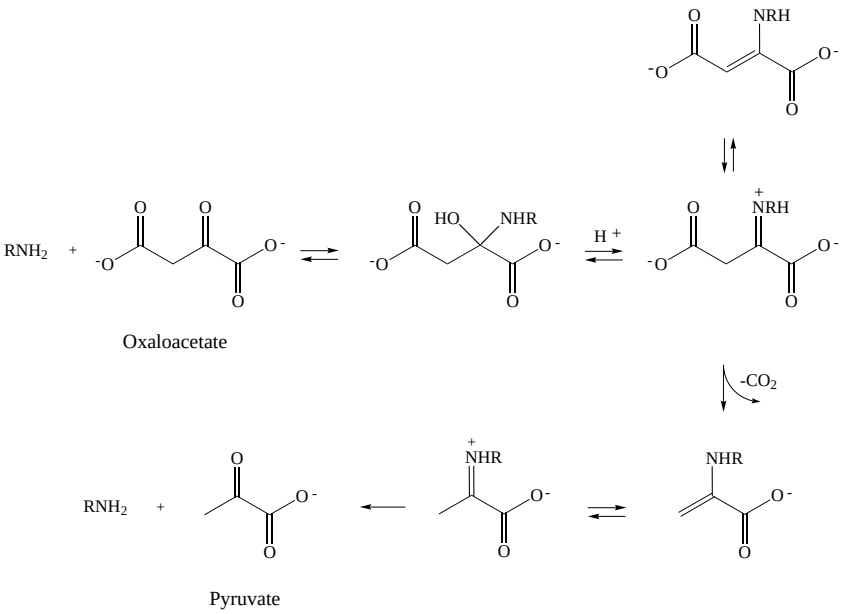
are very simple relative to native enzymes.[22–36] A landmark achievement in peptide design was that of Johnsson et al. who reported a 14-residue peptide (oxaldie 1) that folded into a helical structure and catalysed the decarboxylation of oxaloacetate with a $k_{cat}/K_M$ of $0.47\,M^{-1}s^{-1}$ in aqueous solution at pH 7, more than two orders of magnitude more efficiently than butylamine.[22] The catalysis of the self-replication of 33-residue peptides, with rate constants ($k_{cat}/K_M$) more than three orders of magnitude larger than the second-order rate constants of the corresponding uncatalysed reactions, established that it was possible to catalyse bimolecular reactions by binding both reactants to the surface of a folded polypeptide.[23, 25, 29, 30] We have previously reported helix–loop–helix dimers that follow saturation kinetics in the catalysis of the hydrolysis and transesterification of nitrophenyl esters with rate enhancements that were more than three orders of magnitude greater than those of the imidazole-catalysed reactions.[26, 31] Their catalytic efficiency was based mainly on cooperative nucleophilic and general acid catalysis by unprotonated histidine residues flanked by protonated histidines, and supplemented by Arg and Lys residues that gave a modest degree of substrate selectivity. Catalysts have also been designed that make use of cofactors introduced into model protein scaffolds in efforts to mimic the transamination reaction of amino acid biosynthesis.[37]

The reports of successful designs in which saturation kinetics and large rate enhancements were achieved in designed polypeptides and proteins suggest that we now know how to construct binding sites for small molecules and peptides with mM affinities or better. The catalytic sites reported so far lag behind their natural role models considerably in terms of their capacity for efficient chemical transformations, probably because the number of residues and geometries that can be introduced and optimised in reactive sites in designed proteins is too small. Nevertheless, the observation of efficient catalytic sites in simple protein scaffolds suggests that these may be considered to be primitive precursors of enzyme-like catalysts that can be supplemented by groups capable of native-like recognition and binding as our understanding of protein structural refinement and folding advances.

In many enzymes the nature of the chemical reaction to be catalysed is controlled by a small number of residues, whereas other groups in the active site are mainly involved in non-covalent binding of intermediates and transition states that depends more on the character of the substrate than on the chemical transformation. For example, the efficiency of the His–HisH$^+$ site in a helix–loop–helix motif was refined by the incorporation of groups capable of substrate and transition state binding in a rational way, but the reaction mechanism remained unaltered.[31] Although this is a somewhat superficial view of enzyme catalysis, it has proven to be of some value in catalyst design, and the identification of simple sites that determine and control the pathways of chemical transformations are therefore central to the development of new catalysts. An important group of reactions that are of great fundamental and practical value are those based on the formation of imines from nucleophilic lysine residues and carbonyl groups. Imines are key intermediates in racemisation, transamination, decarboxylation, Michael addition and aldol condensation reactions, and control of imine formation opens up a whole range of chemistries to catalyst design. The search for the basic catalytic unit required for efficient imine formation is therefore of great value.

The introduction of a lysine residue into the binding site of an antibody to depress its p$K_a$ value and to provide a chiral environment proved to be a powerful strategy for the catalysis of a variety of aldol and decarboxylation reactions.[38–41]

The decarboxylation of oxaloacetate can be catalysed by a pathway that proceeds through the formation of an imine from a primary amine and the keto group of the substrate (Scheme 1).[22, 42–44] The 14-residue sequences oxaldie 1 and oxaldie 2 (RHN-L-A-K-L-L-K-A-L-A-K-L-L-K-K-CONH in which R corresponds to a hydrogen for oxaldie 1 and an acetyl group for oxaldie 2) that catalysed the reaction by almost three orders of magnitude[22] contained five lysine residues, and the formation of imines was demonstrated by trapping with $NaBH_3CN$, although the catalytically active lysines were never identified. In the folded state the lysine residues of oxaldie 1 and oxaldie 2 are aligned on one face of the helix. The CD spectra of the oxaldie sequences were highly concentration dependent, and under the experimental conditions the helical contents were low. Oxaldie 2 was more ordered than oxaldie 1 owing to the acetylation of the N-terminus, and the secondary structures were in both cases




Scheme 1. The reaction mechanism for the decarboxylation of oxaloacetate catalysed by a primary amine.

  0947-6539/02/0811-2550 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 11

stabilised by aggregation. To determine which residues and configurations formed the most reactive sites for the decarboxylation of oxaloacetate, and to understand better the molecular basis for catalysis, we have designed fourteen 42-residue helix–loop–helix forming sequences in which the molecular environment of the lysine residues in the folded state has been systematically varied. In addition to providing a better understanding of the mechanism of the decarboxylation reaction, the purpose of the design was also to probe the factors that control the formation of imines, which is of general interest in the development of new catalysts for other reactions along the imine pathway. We now wish to report that we have identified a simple two-residue Lys–Arg site in a conformationally stable helix that is an efficient catalytic unit for the transformation of oxaloacetate. When supplemented by flanking residues capable of binding and recognition of other imines, several catalytic systems can be envisioned that are capable of efficiently and selectively transforming carbonyl compounds into a vast number of different chemical structures. This two-residue site is thus a key building block in the design of new catalysts along the imine pathway that can be supplemented by residues capable of recognition and further refined in terms of substrate selectivity.

## Results

**Design and characterisation of the four-helix bundle motif**: Fourteen 42-residue sequences were designed to form helix–loop–helix motifs and dimerise into four-helix bundle proteins, and lysine residues were incorporated in positions in which their capacity for catalysing the decarboxylation of oxaloacetate could be determined. The polypeptides were synthesised by solid-phase techniques using an automated peptide synthesiser and the Fmoc protecting group strategy. The peptides were purified using reversed-phase HPLC and identified by electrospray mass spectrometry. The design of the antiparallel helix–loop–helix dimer motif is conveniently discussed in terms of the heptad repeat pattern (a-b-c-d-e-f-g)$_n$ (Figure 1) in which residues in the a- and d-positions form the hydrophobic core, those in the g- and c-positions form the exposed surface of the dimer, and the residues in the b- and e-positions are at the dimer interface and control dimerisation. The design of the peptide sequences reported here (Figure 2) was based on the sequence of SA-42[45, 48] and the sequences are similar to those of the peptides KO-42,[26] LA-42b[46] and KA-I[47] that have been described in detail, previously. These sequences were designed to fold into two amphiphilic helical segments connected by a four-residue loop, and to dimerise in an antiparallel mode to form a four-helix bundle. The a- and d-residues that form the hydrophobic core of the folded polypeptides were identical among the parent sequences, as were the b- and e-residues that control the dimerisation of the helix–loop–helix monomers. The $\alpha$-amino isobutyric acid (Aib) residues used in SA-42 were replaced by alanines in later designs. The reactive sites based on g- and c-residues were modified in each sequence according to the purpose for which the sequence was designed. Their solution structures were studied extensively by NMR and CD spectroscopy and by analytical ultracentrifugation (SA-42 and KO-42) and found to fold according to prediction. High-resolution NMR structures could not be obtained as the dimers have partly disordered hydrophobic cores and are in fast exchange on the NMR timescale, resulting in time-averaged NOEs that cannot be used in structure calculations. Nevertheless, the helical regions and the relative positions of the helices were well defined, according to kinetic measurements of catalysis for which the introduction of arginine and lysine residues in helix I increased the reactivity of the cooperative HisH$^+$–His site in helix II towards esters.[31, 49]

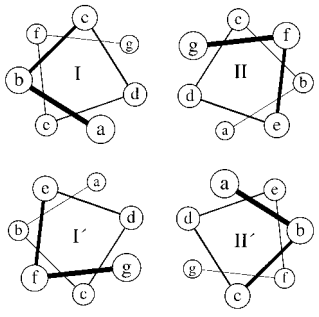



Figure 1. Schematic representation of the heptad repeat pattern used in the design of antiparallel helix–loop–helix dimers, in which the residues in the a- and d-positions form the hydrophobic core; those in the g- and c-positions form the exposed surface positions of the motif, and the residues in the b- and e-positions are localised at the dimer interface.

The sequences reported here were designed according to the same principles, and the hydrophobic core residues remained essentially the same as those of the peptides described above, with the exception of sequences in which single lysine residues were deliberately introduced into the core to depress their p$K_a$ values and to enhance their reactivities, and with the further exception that norleucines were replaced by leucine residues in some sequences. All sequences were approximately 80% homologous to the template sequence SA-42, and, if replacements of Lys by Arg are disregarded, the homology is even larger. We assumed that as a result of the similarity with sequences that have been subjected to extensive structural characterisation, the sequences were likely to fold into the same motif.

Under these conditions we considered the mean residue ellipticity at 222 nm to be a good probe of structure and, in particular, of dimer formation. A large negative value of the mean residue ellipticity is strong evidence in favour of helix–loop–helix dimer formation, as the corresponding value of the monomer is typically very low. This is evident from the concentration dependence of the CD spectrum, in which at low peptide concentrations the mean residue ellipticities commonly have low values as a result of dissociation of the dimer.[45, 48] At peptide concentrations in the mM range, further aggregation is possible, as has been observed previously from the line broadening in the $^1$H NMR spectrum of a solution of 5 mM SA-42.[45, 48] The CD spectra of all sequences (Table 1) were recorded at the 0.2 mM concentration used for kinetic measurements and showed the two minima at 208 and 222 nm that are typical of helical proteins. The sequences that have low helical content are invariably

| Peptide | Sequence |
|---|---|
| Heptad repeat pattern | -g-a-b-c-d-e-f-g-a-b-c-d-e-f-g-a-b-c-d- -a-b-c-d-e-f-g-a-b-c-d-e-f-g-a-b-c-d-e- |
| Decarb-K5 | Ac-N-A-A-D-K-E-H-A-I-R-R-L-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-L-A-R-R-F-E-*H*-F-A-R-A-G-COOH |
| Decarb-K8 | Ac-N-A-A-D-X-E-H-K-I-R-R-L-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-L-A-R-*H*-F-E-R-F-A-R-A-G-COOH |
| NP42 (K11) | Ac-N-A-A-D-X-E-H-A-I-R-K-L-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-L-A-R-R-F-E-A-F-A-R-A-G-COOH |
| Decarb-K12 | Ac-N-A-A-D-X-E-H-A-I-R-R-K-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-L-A-R-R-F-E-*H*-F-A-R-A-G-COOH |
| Decarb-K16 | Ac-N-A-A-D-X-E-H-A-I-R-R-L-A-E-R-K-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-L-A-R-R-F-E-*H*-F-A-R-A-G-COOH |
| Decarb-K19 | Ac-N-A-A-D-X-E-H-A-I-R-R-L-A-E-R-X-A-A-K-G-P-V-D-A-A-*H*-X-A-E-R-L-A-R-R-F-E-A-F-A-R-A-G-COOH |
| Decarb-K26 | Ac-N-A-A-D-X-E-H-A-I-R-R-L-A-E-R-X-A-A-G-G-P-V-D-A-A-K-X-A-E-R-L-A-R-R-F-E-*H*-F-A-R-A-G-COOH |
| Decarb-K27 | Ac-N-A-A-D-X-E-H-A-I-R-R-L-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-K-A-E-R-L-A-R-R-F-E-*H*-F-A-R-A-G-COOH |
| Decarb-K31 | Ac-N-A-A-D-X-E-H-A-I-R-R-L-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-K-A-R-R-F-E-*H*-F-A-R-A-G-COOH |
| Decarb-K34_R33 | Ac-N-A-A-D-X-E-H-*H*-I-R-R-L-A-E-R-X-A-A-G-G-P-V-D-A-A-Q-X-A-E-R-L-A-R-K-F-E-R-F-A-R-A-G-COOH |
| Decarb-K34_K33 | Ac-N-A-A-D-L-E-*R*-*H*-I-Q-R-L-A-Q-R-L-A-A-R-G-P-V-D-A-A-Q-L-A-E-R-L-A-K-K-F-E-*E*-*Y*-A-R-A-G-CONH$_2$ |
| Decarb-K34_Q33 | Ac-N-A-A-D-L-E-R-H-I-Q-R-L-A-Q-R-L-A-A-R-G-P-V-D-A-A-Q-L-A-E-R-L-A-*Q*-K-F-E-E-Y-A-R-A-G-CONH$_2$ |
| Decarb-1 | Ac-N-A-A-D-L-E-K-K-I-Q-R-L-A-Q-R-L-A-A-Q-G-P-V-D-A-A-Q-L-A-E-R-L-A-K-K-F-E-R-F-A-R-A-G-CONH$_2$ |
| Decarb-2 | Ac-N-A-A-D-L-E-A-A-I-Q-R-L-A-K-K-L-A-A-R-G-P-V-D-A-A-Q-L-A-K-K-L-A-R-E-F-E-R-F-A-R-A-G-CONH$_2$ |
| Decarb-K | Ac-N-A-A-D-L-E-K-A-I-Q-K-L-A-Q-K-L-A-R-K-G-P-V-E-A-A-K-L-A-E-K-L-A-K-K-F-E-K-F-A-R-A-G-CONH$_2$ |
| Position in sequence | 1 5 10 15 20 25 30 35 40 |

Figure 2. The sequences of the designed 42-residue peptides in the one-letter code, in which X is used for norleucine. The amino acid residues referred to in italics are those that have been changed in the sequence relative to that of NP-42.[50] At the top of the Figure, the heptad repeat pattern is given and at the bottom the relative positions of the amino acids in the sequences are shown.

Table 1. The mean residue ellipticity at 222 nm of the polypeptide catalysts at 0.2 mM concentration in aqueous solution.

| Peptide | $\theta_{222}$ [deg cm$^2$ dmol$^{-1}$][a] |
|---|---|
| Decarb-K5 | −10300 |
| Decarb-K8 | −20500 |
| NP-42 (K11) | −22500 |
| Decarb-K12 | −8100 |
| Decarb-K16 | −10000 |
| Decarb-K19 | −18400 |
| Decarb-K26 | −20500 |
| Decarb-K27 | −10900 |
| Decarb-K31 | −8200 |
| Decarb-K34_R33 | −23000 |
| Decarb-K34_K33 | −20800 |
| Decarb-K34_Q33 | −18400 |
| Decarb-1 | −14800 |
| Decarb-2 | −19700 |
| Decarb-K | −15300 |

[a] The estimated experimental accuracy is ±500 deg cm$^2$ dmol$^{-1}$.

those in which the lysines have been incorporated into a- and d-positions and are part of the hydrophobic core.

The mean residue ellipticities at 222 nm ranged from −8000 to −23000 deg cm$^2$ dmol$^{-1}$ at pH 7.0 and room temperature. The low helical content of some of the peptides may be a result of dissociation of the dimers to form monomers, and since the lowest helical content was recorded for those peptides (Decarb-K12, Decarb-K16, Decarb-K27 and Decarb-K31) in which charged lysine residues were incorporated into a- or d-positions in the hydrophobic core, the most likely explanation is that the sequence modifications led to decreased stability of the folded motif. The negative value of the mean residue ellipticity of Decarb-K34_R33 at 0.2 mM concentration was large, demonstrating a well developed structure, but decreased at low μM concentrations, in agreement with a model in which the peptide is involved in a monomer–dimer equilibrium (Figure 3a).[45, 48] The pH dependence of the mean residue ellipticity of Decarb-K34_R33 indicated that it was fully folded at pH 7, but largely unstructured at low pH (Figure 3b). The most efficient catalysts show the largest helical contents, and the least efficient catalysts have relatively little helical structure, with some notable exceptions such as NP-42 (K11), Decarb-K26 and Decarb-K34_Q33, that have high helical contents but only modest activities. The conformational stability of the catalysts seems to be an important factor in catalysing the decarboxylation reaction.

Detailed NMR analyses of the solution structures of the peptides presented here have not been carried out because of the large sequence homologies with the parent polypeptides, but the one-dimensional $^1$H NMR spectra are also informative. The folded polypeptides are biomacromolecules in chemical exchange between several conformers at a rate that is intermediate on the NMR timescale. Unlike the resonances of native proteins and small molecules, those of partially disordered proteins are broadened, and the chemical shift dispersion is intermediate between proteins with well-defined



0947-6539/02/0811-2552 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 11

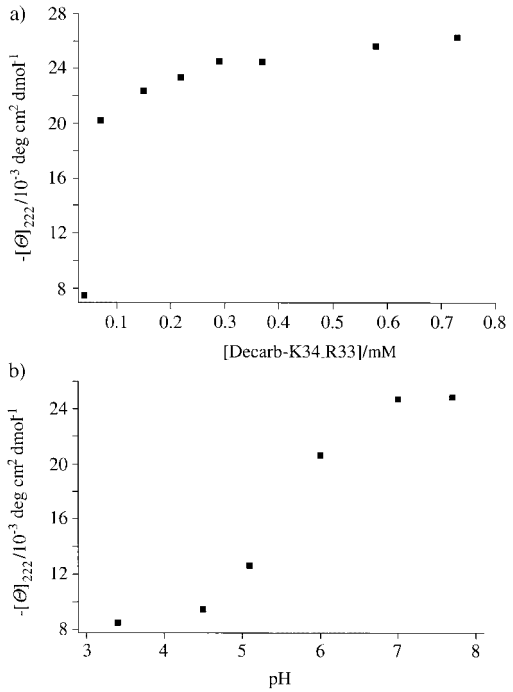



Figure 3. a) The concentration dependence of $\Theta_{222}$, the mean residue ellipticity at 222 nm, of Decarb-K34_R33 in 50 mM Bis-Tris buffer solution containing 0.15 M NaCl at pH 7.0 and room temperature; b) pH dependence of $\Theta_{222}$, the mean residue ellipticity at 222 nm, of 0.4 mM Decarb-K34_R33 in $H_2O$ at room temperature.

tertiary structures and those of small molecules. The $^1$H NMR spectrum of Decarb-K34_R33 shows considerable line broadening and poor shift dispersion suggesting that the sequence folds into a helix–loop–helix dimer under the experimental conditions (Figure 4).

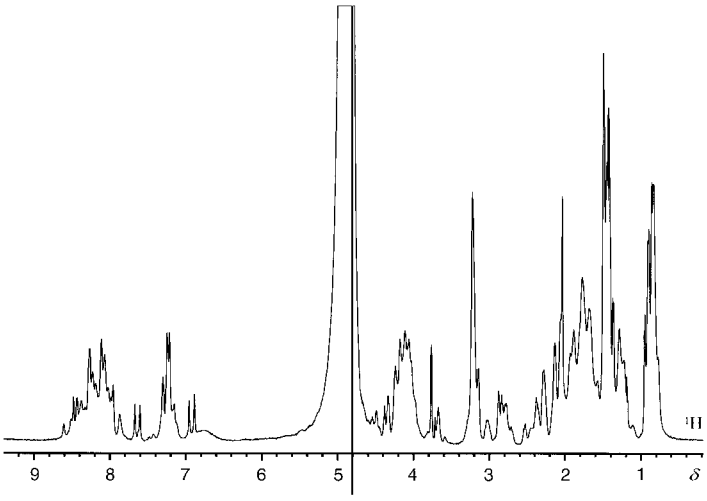



Figure 4. The $^1$H NMR spectrum of Decarb-K34_R33 in $H_2O/D_2O$ (90:10) at pH 5.4 and 298 K.

In addition, the temperature dependence of the $^1$H NMR spectrum of Decarb-K34_R33 was recorded in the interval 278–318 K, and the line broadening as well as the shift dispersion decreased with temperature. The folded polypeptide was therefore in exchange between several conformers at an intermediate rate (see Supporting Information). The similarity of the $^1$H NMR spectra of the polypeptides reported here suggests that their folds are similar.

**Design of reactive sites for decarboxylation**: Lysine residues are required for imine formation, and nine sequences were designed with a single lysine residue in a position that was systematically varied (Figures 2, 5) to probe the reactivity of lysines in various helical positions and to determine the effect of flanking residues. Previously, the sequence NP-42 was designed according to the same principle.[50] In the sequences Decarb-K5, Decarb-K12, Decarb-K16, Decarb-K19, Decarb-K27, Decarb-K31 and Decarb-K34_R33, lysine residues were

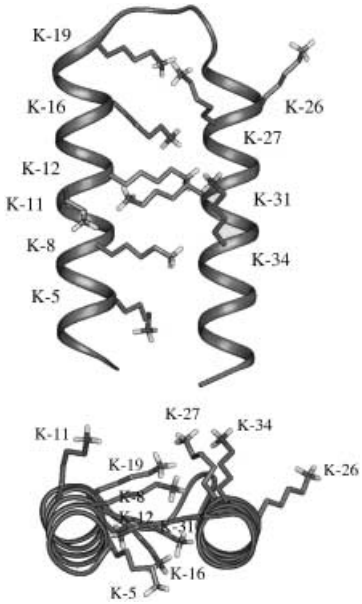



Figure 5. Schematic representation of the helix–loop–helix motif showing the relative positions of the ten single lysine residues of the single-lysine sequences.

introduced in a- or d-positions. It has been shown previously that the reactivities of Lys and His residues in a- and d-positions are enhanced as a result of p$K_a$ depression induced by the hydrophobic environment.[32, 47] Flanking residues in all of these sequences were mainly other hydrophobic groups, except in the sequences Decarb-K12, Decarb-K16, Decarb-K31 and Decarb-K34_R33 in which an Arg residue was incorporated in the position preceding the Lys in the sequence, that is, in a g- or c-position with the lysine in an a-or d-position. Position 34 is a d-position, and the basic residues placed in that position in closely similar sequences have been shown to have depressed p$K_a$ values; His-34 in the peptide JNIII[31] has a p$K_a$ of 5.4, and Lys-34 in the peptide KA-I[47] has a p$K_a$ of 9.3, both approximately 1.2 p$K_a$ units below their random coil values. In the sequences Decarb-K8, NP-42 and Decarb-K26, in which the lysines were positioned in the solvent exposed g- and c-positions, the Lys residues were flanked by arginines in neighbouring g- and c-positions.

Four sequences were designed where lysines were incorporated in pairwise configurations in the sequence to pro-

vide comparisons with those where the imine was flanked by residues in neighbouring g- and e-positions configurations on the face of a helix. In Decarb-1, four lysines were incorporated in positions 7 and 8, and in positions 33 and 34; position 34 is known to depress the p$K_a$ values of charged residues. In Decarb-K34_K33, Lys-7 and Lys-8 were replaced by Arg-7 and His-8, since this combination of lysine residues appears to destabilise the four-helix bundle motif. The mean residue ellipticity of Decarb-1 was only two thirds of that of Decarb-K34_K33 (Table 1). In Decarb-2, four lysines were introduced in pairwise configurations in positions 14, 15, 29 and 30, that is, in positions that were not expected to induce p$K_a$ depression in lysine side chains. In this sequence, Lys-15 and Lys-30 were expected to be close in space in the folded state and provide the opportunity for an interhelical reactive site, by analogy with His – HisH$^+$ sites that were able to span both helices.[35] In Decarb-2, the pairwise lysine residues were flanked by arginines in g- and c-positions on the solvent exposed face of the helix, and the helical content was approximately equal to that of Decarb-K34_K33 (Table 1). Decarb-1, Decarb-K34_K33 and Decarb-K all have lysine pairs in positions 33 and 34, but different flanking groups. In Decarb-1, two Arg residues were incorporated into positions 30 and 37. In Decarb-K34_K33, an Arg was introduced into position 30, but a glutamic acid was incorporated into position 37. Decarb-K mimics some of the properties of oxaldie 1 for comparison between scaffolds of different conformational stability, since lysines 7, 11, 15 and 19 are lined up in a configuration that is similar to that of the lysine residues in oxaldie 1. They form a lysine-rich face in the folded helical form, as do lysines 26, 30, 33 and 37 in helix II. In Decarb-K, two lysine residues flank the lysine pair in positions 33 and 34. Decarb-K was expected to have a more ordered structure than the oxaldie sequences as a result of the more stable structure of helix – loop – helix dimer. Decarb-K34_Q33 is a reference peptide with a Lys in position 34 and a Gln in position 33 to compare directly the reactivity of the Lys – Lys and the Lys – Arg pairs with a single Lys in the same position flanked by a neutral hydrogen bond donor.

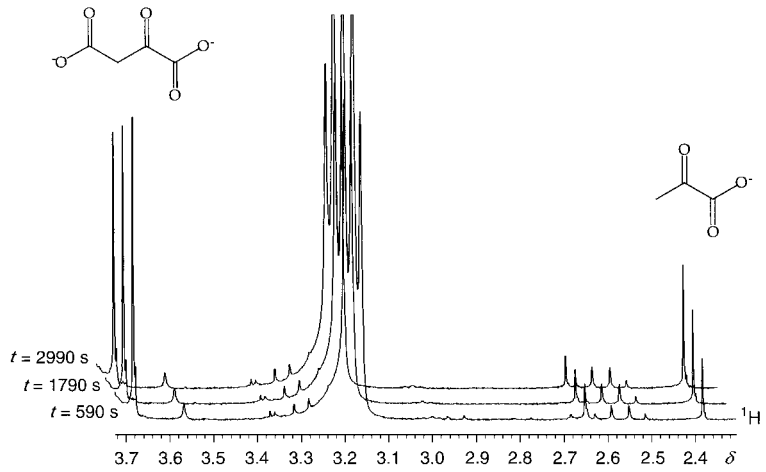



Figure 6. Part of the $^1$H NMR spectrum of Decarb-K34_R33 under reaction conditions showing the growth of pyruvate at 2.38 ppm and the disappearance of oxaloacetate at 3.68 ppm with time. Chemical shifts are given relative to TSP.

**Kinetic measurements**: The reactivities of the designed helix – loop – helix dimers were screened by measuring the initial rates of pyruvate formation by $^1$H NMR spectroscopy as a function of time at 298 K and 20 mM concentration of oxaloacetate at a catalyst concentration of 0.2 mM (Figure 6). The measured rate constants, after subtraction of background decarboxylation rates, followed by division by the substrate and peptide concentrations, correspond to the second-order rate constants for reactions that do not follow saturation kinetics. When saturation kinetics are obeyed, $k_{cat}/(K_M+[S])$ is obtained instead, in which [S] is the concentration of substrate (Table 2).

Table 2. Rate constants of peptide-catalysed decarboxylation obtained from measurements of initial rates. Reported values are initial rates divided by peptide and substrate concentrations after subtraction of rate constants of back-ground decarboxylation.

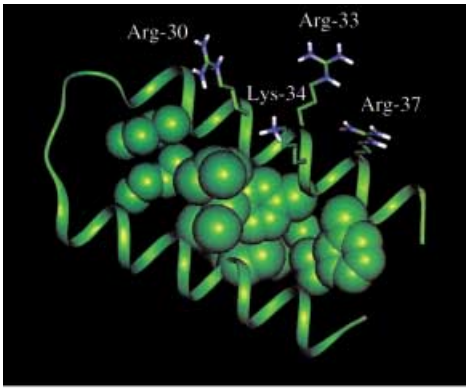
| Peptide | Rate constant [$M^{-1}s^{-1}$] |
|---|---|
| Decarb-K5 | 0.036 |
| Decarb-K8 | 0.15 |
| NP-42 (K11) | 0.015 |
| Decarb-K12 | 0.071 |
| Decarb-K16 | 0.11 |
| Decarb-K19 | 0.14 |
| Decarb-K26 | 0.058 |
| Decarb-K27 | 0.034 |
| Decarb-K31 | 0.050 |
| Decarb-K34_R33 | 0.22 |

The second-order rate constants are directly comparable and the identification of the most reactive catalyst follows directly from their relative magnitudes. An efficient catalyst that follows saturation kinetics with very low values of $k_{cat}$ and $K_M$, but a high value of $k_{cat}/K_M$, may be saturated with substrate at 20 mM concentration and may, in principle, avoid detection under these conditions. For that reason, we have also screened the library by UV spectroscopy by following the decrease in absorbance at 285 nm as described by Johnsson et al.[22, 53] The measurements are not accurate enough for rate constant measurements, since at 285 nm at least two species absorb, one of which increases in concentration and one of which decreases with time. They are also inaccurate because the first phase is a rapid increase in absorbance owing to imine formation that makes the initial slope of the decrease in absorbance difficult to estimate. These difficulties are clearly manifested in the spread of initial rate values observed upon attempts to determine the steady-state parameters by UV spectroscopy. Nevertheless, it is a good technique for identifying catalysts with low dissociation constants because the

  0947-6539/02/0811-2554 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 11

absorbance at 285 nm reflects the amount of imine formed at the side chain of the catalytically active Lys residue. Initial rate constants determined by $^1$H NMR spectroscopy were obtained that approached 0.15 $M^{-1}s^{-1}$ and the values of $k_{cat}$ and $K_M$ were estimated from the Michaelis–Menten equation. Since the concentration of substrate ([S]) is 20 mM, an analysis of the possible values of $k_{cat}$ and $K_M$ that satisfy the equation indicated that in order for the catalyst to have a higher value of $k_{cat}/K_M$ than Decarb-K34_R33, $K_M$ would have to be less than approximately 5 mM, that is, four times lower than Decarb-K34_R33. No evidence was obtained from UV spectroscopy that any catalyst had a distinctly lower dissociation constant than Decarb-K34_R33. Consequently, it is unlikely that efficient catalysts remained undetected in the screening process.

In the $^1$H NMR screening of single-lysine peptides, the initial rates were determined and compared, and the two most efficient peptides (Decarb-K34_R33 and Decarb-K8) were selected and subjected to steady-state kinetic analysis (Figure 7).




**Decarb-K34.R33**

Ac-N-A-A-D-Nle-E-H-H-I-R-R-L-A-E-R-Nle-A-A-G-
-G-P-V-D-
COOH-G-A-R-A-F-R-E-F-K-R-A-L-R-E-A-Nle-Q-A-A-

Figure 7. The modelled structure of Decarb-K34_R33 showing the side chains of the residues in the catalytic site and in neighbouring positions. Hydrophobic core residues are also presented since they control folding and p$K_a$ depression. Only the monomer is shown for the sake of clarity.

Steady-state kinetics were also recorded for all peptides with Lys–Lys pairs and for Decarb-K34_Q33 (Table 3). The steady-state parameters were determined by fitting an expression describing the dependence of the initial rate ($v$) on the concentration of substrate. Three polypeptides were found to follow saturation kinetics with large rate enhancements relative to the second-order rate constant ($k_2$) for the butylamine-catalysed reaction, 0.0011 $M^{-1}s^{-1}$ (Figure 8).

The value of the steady-state rate constant $k_{cat}/K_M$ of the Decarb-K34_R33-catalysed reaction divided by $k_2$ for the butylamine-catalysed reaction, approaches three orders of magnitude. However, the second-order rate constants of the catalysts that do not follow saturation kinetics are more than two orders of magnitude larger than that of the butylamine-catalysed reaction. The most efficient catalysts have the most developed structures, whereas several of the less efficient catalysts have low or high helical contents under the reaction conditions.

Table 3. Steady-state kinetic parameters of peptide-catalysed decarboxylation of oxaloacetate in aqueous solution at pH 7.0 and 298 K.

| Peptide | $k_{cat}$ [$s^{-1}$] | $K_M$ [mM] | $k_{cat}/K_M$ [$M^{-1}s^{-1}$] | $k_2$ [$M^{-1}s^{-1}$] |
|---|---|---|---|---|
| Decarb-K34_R33 | 0.013 | 22 | 0.59 | |
| Decarb-K34_K33 | 0.013 | 52 | 0.25 | |
| Decarb-K34_Q33 | | | | 0.079 |
| Decarb-1 | | | | 0.18 |
| Decarb-2 | | | | 0.14 |
| Decarb-K | 0.013 | 23 | 0.57 | |
| Butylamine[a] | | | | 0.0011 |
| Spontaneous[a] | 0.000013 | | | |

[a] Data from Johnsson et al.[22]

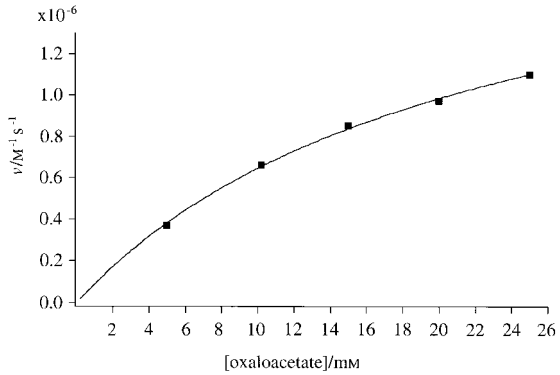



Figure 8. A plot of initial rate ($v$) for the formation of pyruvate catalysed by Decarb-K34_R33 at a peptide concentration of 0.2 mM versus different concentrations of oxaloacetate (5–25 mM). The solid line represents the best-fit of the Michaelis–Menten equation to the experimental data giving a $k_{cat}$ of 0.013 $s^{-1}$ and a $K_M$ of 22 mM.

**Kinetics at low peptide concentrations**: To probe the role of conformational stability, the reaction rate was determined in aqueous solution at low concentration of polypeptide in which there is considerable dissociation of dimer to form the less ordered monomer. The plot of the initial rate of the Decarb-K34_K33-catalysed reaction versus peptide concentration indicates that the reactivity decreases with peptide concentration, and is therefore dependent on structure (see Supporting Information).

## Discussion

The design of new enzymes that catalyse chemical reactions not catalysed by nature remains the ultimate goal in de novo catalyst design. Achievements, so far, include the demonstration of folded polypeptides capable of substrate recognition and binding,[22–25, 28–36, 49] non-covalent binding of both reactants in bimolecular reactions,[23–25, 27–30] cooperative nucleophilic and general acid catalysis[26, 31–36] and rate enhancements in turnover reactions that surpass three orders of magnitude.[22–36] Our understanding of biomolecular interactions has reached a level where binding sites for small molecules and peptides can be constructed with affinities that are compara-

ble to those of native enzymes. The organisation of reactive residues in proteins is considerably more difficult as revealed by the modest values of $k_{cat}$ reported for de novo designed catalysts, and the problem of refinement of structure and function is perhaps best handled by the tools of molecular biology, in strategies that mimic the evolution of competent enzymes from primitive precursors. The key question is what stage in development the primitive catalysts must reach before they can be subjected to evolution in the test tube, and whether de novo design can bridge the gap between reaction mechanistic concepts and new enzymes. Catalytic efficiencies in model catalysts have approached and surpassed three orders of magnitude, in reactive sites based on only a few amino acid residues, and in polypeptide scaffolds of surprising simplicity. The results suggest that much larger rate enhancements can be obtained by the further introduction into such sites of groups capable of substrate, intermediate and transition state binding, preferably with differential affinity. The availability of more sophisticated protein scaffolds that allow the introduction of several residues beyond those that are involved in the making and breaking of bonds, at optimal distances, should therefore make it possible to bring new catalysts that originate from de novo design into the range of efficiencies and selectivities of the least efficient, but competent enzymes that are found in nature. The refinements of designed catalysts at that level of performance by molecular biological techniques are thus likely to succeed. The search for new chemical reactions to be introduced into manmade proteins is therefore timely and one of the most promising pathways that can be explored is that which follows the pathway of imine formation.

Primary amines and carbonyl groups reversibly form imines in aqueous solution with release of water, with equilibrium constants that disfavour imine formation because of the high concentration of solvent. The resulting low concentrations make it prohibitively difficult to carry out chemical transformations between imines and other reactants unless they are bound in an active site, or, unless the concentration of imine is increased by interactions with stabilising groups that drive the equilibrium towards completion. To open up the imine pathway to polypeptide catalysis and subsequently to enzyme design, a protein scaffold will have to be engineered that drives the equilibrium between the amine and the carbonyl compound towards the imine to counteract the unfavourable mass balance effect. The imine forming residues in proteins are the side chains of lysine residues which, owing to their high p$K_a$ values in solvent exposed positions, are not very reactive at neutral pH.[51] Catalysts capable of lysine p$K_a$ depression and binding and stabilisation of the imine intermediate are thus a prerequisite for efficient catalysis. The oxaloacetate decarboxylation, which is a reaction that proceeds via an imine intermediate, is an excellent model reaction for identifying sequences and conformations that favour imine formation. The decarboxylation reaction has therefore now been studied in fifteen designed helix – loop – helix motifs in which the position of the lysine has been systematically varied in terms of its position and the configurations of flanking residues that will optimise the reactivity of the catalyst.

Based upon reaction mechanistic principles, lysines flanked by positively charged Arg or Lys residues were expected to catalyse the decarboxylation of oxaloacetate more efficiently than unperturbed primary amines, as a result of p$K_a$ depression that arises from charge – charge repulsion. In addition, stabilisation of the negatively charged intermediate was expected owing to charge – charge attraction.[52] These expectations were born out as all polypeptide catalysts were more efficient than butylamine in catalysing oxaloacetate decarboxylation (Tables 2 and 3). The introduction of lysine residues in positions that according to the heptad repeat pattern are hydrophobic, was expected to depress p$K_a$ values even more than flanking arginines and to enhance the catalytic capacity further. The rate constants obtained from the screening experiments were unfortunately unimpressive for most catalysts in which Lys residues were incorporated into hydrophobic a- or d-positions. Decarb-K5, Decarb-K12, Decarb-K27 and Decarb-K31 have rate constants in the range $0.036 - 0.071\,M^{-1}s^{-1}$, corresponding to 30- to 70-fold rate enhancements over that of the butylamine-catalysed reaction. The observed rate enhancements were probably a result of a combination of factors, mainly p$K_a$ depression and imine stabilisation. The helical contents of the polypeptides were low, and a reason for the poor reactivity may be that the structures were partially unfolded and that the effect on p$K_a$ was much less than anticipated as there are no charge – charge repulsions to depress p$K_a$ values in the unfolded peptide. Two single-lysine catalysts (Decarb-K16 and Decarb-K19) had considerably higher reactivity with rate constants that were factors of 100 and 140-times larger than that of the butylamine-catalysed reaction. This is particularly interesting in the case of Decarb-K16 that has a poorly developed helical structure with a mean residue ellipticity under the reaction conditions of only $-10000\ deg\,cm^2\,dmol^{-1}$. In a more ordered structure, the reactivity might have been much higher. However, none of them were as reactive as Decarb-K34_R33 (Figure 7) in which Lys-34 is in a d-position that gives rise to p$K_a$ values that are depressed by more than a p$K_a$ unit relative to solvent exposed lysine side chains.[47] In spite of the fact that Lys-34 was located in the hydrophobic core, Decarb-K34_R33 had a high helical content with a mean residue ellipticity of $-23000\ deg\,cm^2\,dmol^{-1}$. In addition to being partially embedded in the hydrophobic core, Lys-34 was flanked by three Arg residues to further lower its p$K_a$ and to provide an opportunity for imine stabilisation by binding of the carboxylate substituents. A low p$K_a$ value of the primary amine is clearly an important catalytic factor in the decarboxylation of oxaloacetate, and thus in imine formation. A reactive primary amine is especially efficient in combination with flanking positively charged residues that can stabilise the imine intermediate. However, conformational stability is also central to efficient catalysis, as there appears to be a rough correlation between helical content and catalytic efficiency.

The screening of single lysine catalysts made it possible to design a set of more complex polypeptides in the search for the optimal reactive sites, and to compare the effects of flanking groups on the reactivity of lysines in the decarboxylation reaction (Tables 2 and 3). Three peptides were efficient catalysts and followed saturation kinetics; the $k_{cat}$/



0947-6539/02/0811-2556 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 11

$K_M$ of Decarb-K34_R33, Decarb-K34_K33 and Decarb-K was 0.59, 0.25 and 0.57 $M^{-1}s^{-1}$, respectively. It was particularly gratifying to find that the "simplest" site—belonging to Decarb-K34_R33—with a single lysine residue flanked by arginines was the most efficient in catalysing the decarboxylation reaction. Oxaldie 1 catalysed the reaction with a $k_{cat}/K_M$ of 0.47 $M^{-1}s^{-1}$,[22] but this was to a large degree a result of the reactivity of the free amino terminus with a p$K_a$ of 7.2, and when the amino terminus was acylated to form the catalyst oxaldie 2, $k_{cat}/K_M$ dropped to 0.15 $M^{-1}s^{-1}$, which is nevertheless an impressive rate constant for a small peptide. There are five lysines in the sequences of these catalysts, and the reactive site of Decarb-K34_R33 is almost four times as efficient as oxaldie 2, although there is only a single lysine residue in the sequence. Decarb-K34_K33 is less efficient than Decarb-K34_R33, although the sequences are very similar and they both have a lysine in position 34. The main difference seems to be that Lys-34 is flanked by Arg-33 in the latter, whereas it is flanked by Lys-33 in the former (Figure 9).

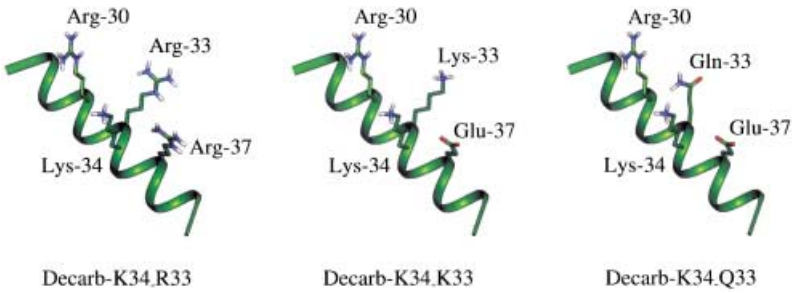



Figure 9. Modelled peptide segments showing the different environments surrounding Lys-34 in the peptides Decarb-K34_R33, Decarb-K34_K33 and Decarb-K34_Q33.

This observation makes it likely that the reactivity, to a large degree, resides in the interplay between the residues in positions 33 and 34. The replacement of Arg by Gln led to the sequence Decarb-K34_Q33 that was almost an order of magnitude less reactive than Decarb-K34_R33. The Decarb-K34_Q33-catalysed reaction does not follow saturation kinetics and the removal of Arg and Lys appears to diminish the capacity of the catalyst to interact with the imine intermediate. $K_M$ of Decarb-K34_Q33 is not observable; in Decarb-K34_K33 it is 52 mM and in Decarb-K34_R33 it is 22 mM. The formation of the Michaelis–Menten complex in this case is most likely the formation of the imine from the lysine side chain and the substrate, and the reason that they obeyed saturation kinetics was perhaps that the imine was stabilised by groups that flanked the catalytically active Lys residue (Figure 10).

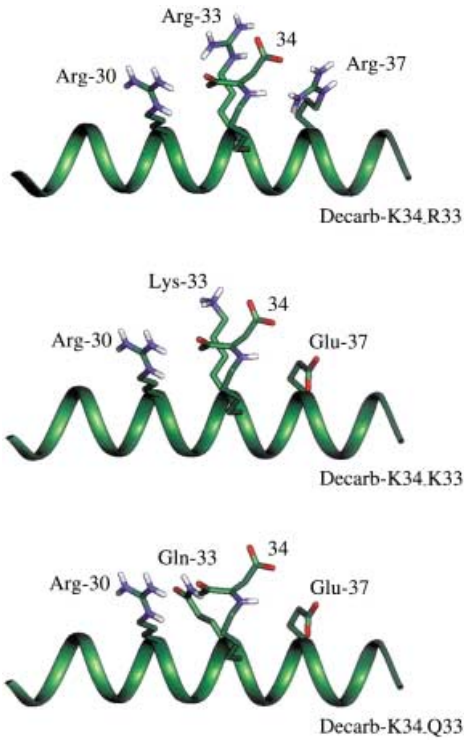



Figure 10. Modelled peptide segments showing the different environments surrounding the imine formed at the side chain of Lys-34 in the peptides Decarb-K34_R33, Decarb-K34_K33 and Decarb-K34_Q33.

In the case of Decarb-K34_R33, the flanking groups were arginines in positions 30, 33 and 37; in the case of Decarb-K34_K33, Lys-34 was flanked by Arg-30, Lys-33 and Glu-37. The results suggest an important role of a positively charged residue in position 33, since the two-residue sites Lys-33–Lys-34 and Arg-33–Lys-34 were present in all of the most efficient catalysts, and because the replacement of Lys-33 by Gln-33 reduced the catalytic efficiency by a factor of three, and because the replacement of Arg-33 by Gln-33 reduced the reactivity by a factor of 7.5. We conclude that the catalytic efficiency of Lys-34, enhanced by p$K_a$ depression, was strongly increased by the presence of Arg-33 to form an efficient two-residue site for imine formation. Arg-33 is clearly a better stabiliser of the imine intermediate than Lys-33. The modelled structure of a helical segment in which an imine has been formed by the reaction between a lysine side chain and oxaloacetate is shown in Figure 10, in which the flanking residue is Lys, Gln or Arg, to illustrate the interactions that stabilise the intermediate. The contribution to the overall catalysis attributable to p$K_a$ depression can be estimated from the Brønsted equation.[54] Although the Brønsted coefficient ($\beta$) for the imine-catalysed decarboxylation of oxaloacetate is not known, it can be estimated from the second-order rate constants of the glycine amide- and butylamine-catalysed reactions at pH 7, which are 0.011 and 0.0011 $M^{-1}s^{-1}$, respectively.[53] From the p$K_a$ value of glycine amide (7.93) and butylamine (10.6), a very approximate value of $\beta$ of 0.6 can be estimated. A calculated value of the second-order rate constant for the reaction catalysed by an amine with a p$K_a$ of 9.3 is $3.6 \times 10^{-3}$ $M^{-1}s^{-1}$ at pH 7 using a Brønsted coefficient of 0.6, whereas the observed rate constant $k_{cat}/K_M$ was 0.59 $M^{-1}s^{-1}$. The overall rate enhancement of the Decarb-K34_R33-catalysed reaction as a result of p$K_a$ depression was thus only a factor of 3.2, if $k_{cat}/K_M$ is compared to the second-

order rate constant. The dominant contribution arose from the catalytic efficiency of the Arg – Lys pair.

Decarb-K mimics the catalyst oxaldie 2[22] and included the same active site as Decarb-K34_R33, except that the arginines were replaced by lysines, and that many more lysines were incorporated in the sequence, all of which contributed to the observed overall reactivity towards oxaloacetate. The observed rate constant in the oxaldie 2-catalysed reaction was the sum of those of several competing catalytic lysine residues, all of which must be lower than the observed one. The rate constant $k_{cat}$ was 0.013 $s^{-1}$ for all the sequences that followed saturation kinetics, that is, Decarb-K, Decarb-K34_R33 and Decarb-K34_K33. The binding of the intermediate was stronger in Decarb-K34_R33 and Decarb-K, although in the latter several binding constants probably contributed, as there were many lysine residues in the sequence. The binding of the imine intermediate by Decarb-K34_R33 must be due to interactions with Arg-30, Arg-33 and/or Arg-37. Arg-30 was present in several sequences in which Lys-34 was also present and appeared to be inefficient or have a very small contribution. Arg-37 was also not likely to contribute substantially to imine binding. It was present in the sequence Decarb-1 without enhancing the reactivity of the Lys-33 – Lys-34 pair, although the helical content of Decarb-1 was low and the comparison is possibly unfair at 0.2 mM concentration. To further probe the role of the helical content of Decarb-1 in catalysis, the reaction was also carried out at 0.5 mM concentration at which the mean residue ellipticity was $-18000$ deg $cm^2$ $dmol^{-1}$ rather than $-14800$ deg $cm^2$ $dmol^{-1}$ that was the measured value at 0.2 mM concentration. The reactivity of Decarb-1 was the same at the two concentrations, within the experimental error, and therefore Arg-37 does not play an important role in supplementing the Arg – Lys pair. Arg-37 may well contribute to a lesser extent by non-specific charge – charge interactions with the intermediate, but Arg-33 is clearly the most important of the flanking residues, and the Arg-33 – Lys-34 pair is the key component of the catalytic machinery.

Decarb-K34_R33, Decarb-K34_K33, Decarb-K34_Q33 and Decarb-2 all had high helical contents, whereas those of Decarb-1 and Decarb-K were low with mean residue ellipticities of approximately $-15000$ deg $cm^2$ $dmol^{-1}$. Again, there is a correlation between catalytic efficiency and conformational freedom as Decarb-1, which has a reactive site that is essentially identical to that of Decarb-K34_K33, has only modest catalytic efficiency and does not follow saturation kinetics. The peptide catalysts oxaldie 1 and oxaldie 2 were only helical to a limited extent, and the introduction into a four-helix bundle scaffold of lysine residues in a configuration that mimics those of the oxaldie catalysts enhanced the reactivities. Decarb-K was 20% more efficient than oxaldie 1 and almost four-fold more efficient than oxaldie 2 which is a more fair comparison, since Decarb-K had an acetylated amino-terminus. The reason that Decarb-K was as efficient as Decarb-K34_R33 may be that it reacts through many reactive sites, and that their reactivities add up to that of Decarb-K34_R33. The fact that Decarb-K34_K33 and Decarb-K34_R33 are both efficient catalysts that follow saturation kinetics support the conclusion that the active sites are improved by conformational stability. The design of a reactive site with a single Lys residue that is more efficient than sequences with several lysine residues therefore amounts to an important development in catalyst design. It also provides the opportunity to design catalysts for comparable but more complex substrates by the incorporation of further binding residues that enhance the capacity for substrate discrimination and intermediate and transition state stabilisation.

While the Arg – Lys site is very efficient in the decarboxylation of oxaloacetate, obviously a flanking Arg specifically binds imines formed from negatively charged carbonyl containing substrates. Nevertheless, the stage has been set for the exploration of reactions along the imine pathway. Positively charged substrates are most likely going to be recognised by reactive sites that contain lysine residues and flanking negatively charged aspartates or glutamates. Hydrophobic substrates may be recognised by residues with hydrophobic side chains and hydrophobic pockets. A number of reactions that follow the imine pathway are therefore now open to catalyst design and their further refinement into mature biocatalysts in more complex protein scaffolds may take us one step further towards the goal of tailor-made enzymes for reactions not catalysed by nature.

## Conclusion

The progress in protein and catalyst design to date, has set the stage for the further refinement of manmade biocatalysts, in terms of controlling the function of individual amino acid residues and expanding the chemical repertoire. It is fair to say that we now know how to bind small molecules and peptides by non-covalent forces, and how to control $pK_a$ values of ionisable residues to bring about efficient catalysis. The key to efficient catalysis is to introduce cooperativity between residues that are only poorly reactive by themselves. To that end, the Arg – Lys two-residue site has been shown to be an efficient and cooperative catalyst for oxaloacetate decarboxylation, a reaction along the imine pathway. Catalysts for other reactions along this pathway may now be developed. The identification of the Arg – Lys site reported here therefore sets the stage for the further development of catalysts for imine-mediated chemical reactions.

## Experimental Section

**Peptide synthesis, purification and identification**: The peptides were synthesised on an automated peptide synthesiser (Pioneer, Applied Biosystems) using standard Fmoc protection strategies on a 0.1 mmol scale. The Fmoc-protected amino acids were activated in situ by using TBTU [*O*-(7-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate] (0.5 M in DMF) and DIPEA (*N*,*N*-diisopropylethylamine) (1.0 M in DMF). The Fmoc protecting group was removed by treatment with piperidine (20% *v*/*v* in DMF). For side chain protection, *tert*-butoxy (O*t*Bu) was used for Asp and Glu, trityl (Trt) for His, Asn and Gln, *tert*-butyl (*t*Bu) for Tyr, *tert*-butoxycarbonyl (Boc) for Lys and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) for Arg. Standard coupling times were 60 minutes, except for Arg and difficult parts in the sequence for which 90 minutes were used. At the end of the synthesis, the

  0947-6539/02/0811-2558 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 11

amino-terminus was capped by acetic anhydride (0.3 M in DMF). The ten single-lysine peptides were synthesised on an Fmoc-Gly-PEG-PS polymer (Applied Biosystems) with a substitution level of 0.16–0.23 mmol $g^{-1}$, and the free carboxylic acid was obtained upon cleavage from the resin. The carboxy termini of the remaining peptides were amidated upon cleavage from the resin using an Fmoc-PAL-PEG-PS polymer (Applied Biosystems) with a substitution level of 0.18–0.22 mmol $g^{-1}$. The peptide was cleaved from the resin and deprotected by treatment with a mixture containing trifluoroacetic acid (TFA)/ethanedithiol/$H_2O$/triisopropylsilane (94:2.5:2.5:1, 20 mL $g^{-1}$ resin), for two hours at room temperature. After filtration and concentration of the mixture, the peptides were precipitated by the addition of cold diethyl ether, centrifuged and lyophilised. The peptides were purified by reversed-phase HPLC on a semi-preparative C-8 column (HICHROM), eluated isocratically with propan-2-ol (30–38 % *v*/*v* in TFA (0.1 % *v*/*v*) at a flow rate of 8 mL $min^{-1}$ under UV detection at 229 nm. The purity of the peptides were determined by reversed-phase analytical HPLC. No peaks other than the desired peptides were present in the chromatogram. The identities of the peptides were determined by electrospray mass spectrometry (ES-MS) on a magnetic sector instrument (VG ZabSpec). The obtained molecular weights were within 1 a.u. from the calculated weight and no high molecular weight impurities could be detected.

**Circular dichroism spectroscopy**: CD spectra were recorded on a spectropolarimeter (Jasco J-270), calibrated with (+)-camphor-10-sulfonic acid. The spectra were measured at room temperature in the interval 280–190 nm in 0.1, 0.5 or 1 mm cuvettes. The concentration of peptide was determined by quantitative amino acid analysis. Stock solutions of the peptides were used for better inter-experimental accuracy. Oxaloacetate was purchased from Sigma–Aldrich and its purity was checked by NMR spectroscopy. The stock solutions of peptides and oxaloacetate were prepared in Bis-Tris buffer solution (50 mM) that contained NaCl (0.15 M) and the pH was adjusted with NaOH and HCl. In the pH titration experiment, no buffer was used.

**NMR spectroscopy**: The $^1H$ NMR spectroscopic measurements of the temperature dependence of the structure of Decarb-K34_R33 in the temperature interval 278–318 K were performed using a 600 MHz spectrometer (Varian Inova). Decarb-K34_R33 was dissolved in $H_2O$/$D_2O$ (90:10) at pH 5.4 to a final concentration of 0.5 mM.

**Kinetic measurements followed by NMR spectroscopy**: The $^1H$ NMR spectroscopic measurements of reaction rates were performed using a 400 MHz spectrometer (Varian Unity). The decarboxylation reaction was studied in 90 % Bis-Tris propane (50 mM)/10 % $D_2O$, containing NaCl (150 mM) at pH 7.0 and 298 K, with peptide concentrations of 0.2 mM and oxaloacetate concentrations of 20 mM. The reaction rate was determined by following the decrease of the oxaloacetate peak at $\delta = 3.68$ and the increase of the pyruvate intensity at $\delta = 2.38$ with time. The chemical shifts were measured relative to TSP. In a typical kinetic experiment, 0.2 mM peptide solution was thermally equilibrated for at least 30 minutes before an aliquot of oxaloacetate was added from a freshly prepared solution. The measurements were started within 2 min after briefly shaking the NMR tube, and the reactions were followed for a least five hours. The initial rates were obtained from the first five percent of the reaction from the plots of integrals versus time of the growth of pyruvate. Pyruvate integrals were converted to pyruvate concentrations, under the assumption that the concentrations of other intermediates and by-products were negligible and that therefore the sum of the concentrations of oxaloacetate and pyruvate throughout the reaction was 20 mM as reported previously,[50] by dividing the integral of the pyruvate peak by the sum of the integrals of pyruvate and oxaloacetate. The concentration of pyruvate was plotted against time, and the initial rate was determined by fitting a straight line to the experimental results (see Supporting Information). For the screening experiments, oxaloacetate concentrations of 20 mM were used. For the peptides Decarb-K8, Decarb-K34_R33, Decarb-K34_K33, Decarb-K34_Q33, Decarb-1, Decarb-2 and Decarb-K the initial rates for a number of different concentrations of oxaloacetate (5–25 mM) were determined. The initial rates were plotted against the oxaloacetate concentrations, and for the peptides that followed saturation kinetics the equation $v = k_{cat}[E]_0[S]/(K_M+[S])$ was fitted to the experimental results.

**Kinetic measurements followed by UV**: The kinetic experiments were carried out with spectrophotometers (Varian Cary 1 or 5) equipped with temperature controllers (Varian). The reaction was followed by measuring the absorbance at 285 nm as a function of time.[53] Stock solutions of peptides and oxaloacetate were prepared in Bis-Tris buffer solution (50 mM) containing NaCl (150 mM) and the pH was adjusted when necessary. The concentrations of the peptide stock solutions were determined by quantitative amino acid analysis. In a typical kinetic experiment, 275 μL of solution was used in a 1 mm quartz cuvette. The peptide solution (0.2 mM) was thermally equilibrated for at least 30 minutes before adding an aliquot from a freshly prepared oxaloacetate solution. The measurements were started after briefly shaking the cuvette and reintroduction into the thermostated cell compartment, and the reactions were followed for twelve hours. The initial rates were determined by fitting the equation for a straight line to the experimental data using an extinction coefficient of 0.36 $mM^{-1} s^{-1}$.[53]

The initial rates for different Decarb-K34_K33 concentrations (5 μM–0.5 mM) were determined under the same reaction conditions as described above with an oxaloacetate concentration of 20 mM.

## Acknowledgement

We are indebted to the Swedish Natural Science Research Council and to Carl Tryggers Stiftelse for financial support. We thank the Swedish NMR Centre at Göteborg University for access to a 600 MHz NMR spectrometer and Charlotta Damberg for expert help with the instrument. We are indebted to Anders Holmberg for the synthesis of Decarb-2 and Decarb-K and to Gunnar Stenhagen for expert mass spectrometric assistance.

[1] M. Mutter, G. G. Tuchscherer, C. Miller, K.-H. Altmann, R. I. Carey, D. F. Wyss, A. M. Labhardt, J. E. Rivier, *J. Am. Chem. Soc.* **1992**, *114*, 1463–1470.
[2] J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Niel, W. F. DeGrado, *Science* **1995**, *270*, 935–941.
[3] S. Kamtekar, M. H. Hecht, *FASEB J.* **1995**, *9*, 1013–1022.
[4] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342–345.
[5] C. E. Schaftmeister, R. M.Stroud, *Curr. Opin. Biotechnol.* **1998**, *9*, 350–353.
[6] W. D. Kohn, R. S. Hodges, *TIBTECH* **1998**, *16*, 379–389.
[7] E. Lacroix, T. Kortemme, M. Lopez de la Paz, L. Serrano, *Curr. Opin. Struct. Biol.* **1999**, *9*, 487–493.
[8] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
[9] B. R. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745–754.
[10] L. Baltzer, H. Nilsson, J. Nilsson, *Chem. Rev.* **2001**, *101*, 3153–3163.
[11] M. S. Searle, *Perkin Trans. 2* **2001**, *7*, 1011–1020.
[12] L. Baltzer, J. Nilsson, *Curr. Opin. Biotechnol.* **2001**, *12*, 355–360.
[13] C. Smith, L. Regan, *Acc. Cem. Res.* **1997**, *30*, 152–161.
[14] B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado, D. Eisenberg, *Science* **1993**, *259*, 1288–1293.
[15] Y. Fezoui, D. L. Weaver, J. J. Osterhout, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3675–3679.
[16] E. Ilyina, V. Roongta, K. H Mayo, *Biochemistry* **1997**, *36*, 5245–5250.
[17] C. E. Schaftmeister, S. L. LaPorte, L. J. W. Miercke, R. M. Stroud, *Nature Struct. Biol.* **1997**, *4*, 1039–1046.
[18] B. R. Hill, W. F. DeGrado, *J. Am. Chem. Soc.* **1998**, *120*, 1138–1145.
[19] T. Kortemme, M. Ramirez-Alvarado, L. Serrano, *Science* **1998**, *281*, 253–256.
[20] J. J. Skalicky, R. B. Gibney, F. Rabanal, R. J. Bieber Urbauer, P. L. Dutton, A. J. Wand, *J. Am. Chem. Soc.* **1999**, *121*, 4941–4951.
[21] S. R. Griffiths-Jones, M. S. Searle, *J. Am. Chem. Soc.* **2000**, *122*, 8350–8356.
[22] K. Johnsson, R. K. Allemann, H. Widmer, S. A. Benner, *Nature* **1993**, *365*, 530–532.
[23] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525–528.
[24] S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *J. Am. Chem. Soc.* **1997**, *119*, 10559–10560.
[25] K. Severin, D. H. Lee, A. J. Kennan, M. R. Ghadiri, *Nature* **1997**, *389*, 706–709.

[26] K. S. Broo, L. Brive, P. Ahlberg, L. Baltzer, *J. Am. Chem. Soc.* **1997**, *119*, 11362 – 11372.
[27] S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *Nature* **1998**, *396*, 447 – 450.
[28] C. Micklatcher, J. Chmielewski, *Curr. Opin. Chem. Biol.* **1999**, *3*, 724 – 729.
[29] A. J. Kennan, V. Haridas, K. Severin, D. H. Lee, R. Ghadiri, *J. Am. Chem. Soc.* **2001**, *123*, 1797 – 1803.
[30] A. Saghatelian, Y. Yokobayashi, K. Soltani, M. R. Ghadiri, *Nature* **2001**, *409*, 797 – 801.
[31] K. S. Broo, H. Nilsson, J. Nilsson, L. Baltzer, *J. Am. Chem. Soc.* **1998**, *120*, 10287 – 10295.
[32] K. S. Broo, L. Brive, R. S. Sott, L. Baltzer, *Folding Design* **1998**, *3*, 303 – 312.
[33] L. Baltzer, K. S. Broo, *Biopolymers* **1998**, *47*, 31 – 40.
[34] K. S. Broo, H. Nilsson, J. Nilsson, A. Flodberg, L. Baltzer, *J. Am. Chem. Soc.* **1998**, *120*, 4063 – 4068.
[35] J. Nilsson, K. S. Broo, R. S. Sott, L. Baltzer, *Can. J. Chem.* **1999**, *77*, 990 – 996.
[36] J. Nilsson, L. Baltzer, *Chem. Eur. J.* **2000**, *6*, 2214 – 2220.
[37] M. A. Shogren-Knaak, B. Imperiali, *Bioorg. Med. Chem.* **1999**, *7*, 1993 – 2002.
[38] J. Wagner, R. A. Lerner, C. F. Barbas III, *Science* **1995**, *270*, 1797 – 1800.
[39] R. Björnestedt, G. Zhong, R. A. Lerner, C. F. Barbas III, *J. Am. Chem. Soc.* **1996**, *118*, 11720 – 11724.
[40] C. F. Barbas III, A. Heine, G. Zhong, T. Hoffmann, S. Gramatikova, R. Björnestedt, B. List, J. Andersson, E. A. Stura, I. A. Wilson, R. A. Lerner, *Science* **1997**, *278*, 2085 – 2092.
[41] T. Hoffmann, G. Zhong, B. List, D. Shabat, J. Anderson, S. Gramatikova, R. A. Lerner, C. F. Barbas III, *J. Am. Chem. Soc.* **1998**, *120*, 2768 – 2779.
[42] R. W. Hay, *Aust. J. Chem.* **1965**, *18*, 337 – 351.
[43] J. P Guthrie, F. Jordan, *J. Am. Chem. Soc.* **1972**, *94*, 9136 – 9141.
[44] D. L. Leussing, N. V. Raghavan, *J. Am. Chem. Soc.* **1980**, *102*, 5635 – 5643.
[45] S. Olofsson, G. Johansson, L. Baltzer, *J. Chem. Soc. Perkin Trans. 2* **1995**, 2047 – 2056.
[46] L. K. Andersson, G. T. Dolphin, J. Kihlberg, L. Baltzer, *J. Chem. Soc. Perkin Trans. 2* **2000**, 459 – 464.
[47] L. K. Andersson, M. Caspersson, L. Baltzer, unpublished results.
[48] S. Olofsson, L. Baltzer, *Folding Design* **1996**, 347 – 356.
[49] L. Baltzer, K. S. Broo, H. Nilsson, J. Nilsson, *Bioorg. Med. Chem.* **1999**, *7*, 83 – 91.
[50] M. Allert, M. Kjellstrand, K. Broo, Å. Nilsson, L. Baltzer, *J. Chem. Soc. Perkin Trans. 2* **1998**, 2271 – 2274.
[51] C. Tanford, *Adv. Protein Chem.* **1962**, *17*, 69 – 165.
[52] F. H. Westheimer, *Tetrahedron* **1995**, *51*, 3 – 20.
[53] K. Johnsson, Thesis, Eidgenössische Technische Hochschule, Zürich (Switzerland), **1992**.
[54] A. Fersht, *Structure and Mechanism* in *Protein Science*, W. H. Freeman, USA, **1999**, Chapter 2.

Received: October 11, 2001 [F3606]

  0947-6539/02/0811-2560 $ 17.50+.50/0 