A designed folded polypeptide model system that catalyses the decarboxylation of oxaloacetate

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

Organic Chemistry, Department of Chemistry, Göteborg University, SE-412 96 Göteborg, Sweden. E-mail: Lars.Baltzer@oc.chalmers.se

Received (in Cambridge) 7th May 1998, Accepted 19th August 1998

NP-42, a 42-residue polypeptide that folds into a hairpin helix-loop-helix motif and dimerises to form a four-helix bundle, has been designed to catalyse the decarboxylation of oxaloacetate. The residues of the reactive site are Arg-10, Lys-11, Arg-15 and Arg-34. The lysine residue reacts with the carbonyl carbon of the substrate to form the imine intermediate that is decarboxylated to form the pyruvate product. The second-order rate constant of the NP-42 catalysed reaction in aqueous solution at pH 7.0 and 298 K determined by ^{1}H NMR spectroscopy is 0.015 M^{-1} s^{-1}, which is approximately a factor of 10 larger than that of the butylamine catalysed reaction. The study of the reaction by ^{1}H NMR spectroscopy permits the direct observation of reactants and products and the problem of determining extinction coefficients is thus avoided. The corresponding second-order rate constant determined by UV spectroscopy is 0.010 M^{-1} s^{-1}. The NP-42 catalysed reaction has been shown not to follow saturation kinetics and the reaction follows pseudo-first-order kinetics over a range of substrate concentrations from 2.5 to 30 mM. NP-42 is thus a well-defined model system for the further development of efficient catalysts capable of substrate recognition.

Introduction

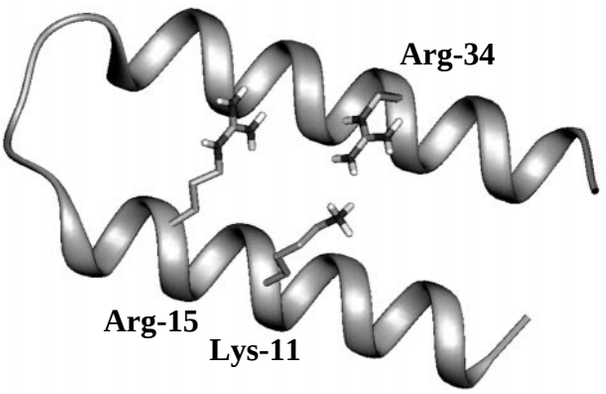
The study of enzymatic reactions is driven not only by the chemists' desire to understand their functions but also by the expectation that the discovered reaction mechanisms may be exploited in the design of novel catalysts with tailored specificities. The approaches are many and range from the design and synthesis of small organic mimics to the engineering of catalytic antibodies and the reengineering of natural enzymes. The ultimate measure of our understanding of enzymatic catalysis is our ability to repeat natural activities in man-made mimics and the efficiencies and selectivities of model systems therefore provide a measure of the level of understanding we can claim to have about how enzymes work.

A large number of enzyme mimics have now been reported but progress in this field has been slow. The reasons for this can only be guessed at but likely problems are the rigidity of the molecular framework, difficulties in synthesis, problems with solubility when there is a change in charge along the reaction pathway and the difficulties in determining the structures of *e.g.* catalyst–substrate complexes. We recently reported on the design and synthesis of helix-loop-helix motifs that dimerise in aqueous solution to form four-helix bundles.[1–3] Folded polypeptides form somewhat flexible tertiary structures, are easily synthesized, are water soluble and are readily subjected to structure determination by, for example, NMR spectroscopy. They appear to be ideal templates for the construction of enzyme mimics, provided that the functionality can be introduced in a controlled way.

Through the site-selective functionalisation reaction[4,5] a NAD+/NADH cofactor mimic was recently incorporated into a folded polypeptide[6] to form a covalently linked peptide–cofactor complex designed to reduce carbonyl groups. In a further development of the functionalisation strategy a pyridoxal phosphate cofactor was introduced into a folded four-helix bundle motif where the site-selectivity was controlled by non-covalent interactions between the phosphate group and an arginine residue in the reactive size.[7] We now wish to report the exploitation of only naturally occurring amino acid residues for the study of the decarboxylation of oxaloacetate to form pyruvic acid and carbon dioxide. The decarboxylation reaction has been studied extensively in organic model systems[8–10] and follows a reaction mechanism where a primary amine attacks the carbonyl carbon of the substrate to form an imine that is in a rapid equilibrium with an enamine. The imine form is decarboxylated because the protonated nitrogen can act as an electron sink and stabilise the high energy transition state. The reactivity is to a large degree dependent on the pK_a of the primary amine[11] and the depression of the pK_a is therefore a key feature in the design of the catalyst. Extra transition state binding by flanking groups is expected to enhance the reactivity even further.

The decarboxylation of oxaloacetate was previously reported to be catalysed by two fourteen-residue peptides designed by Johnssen *et al.*, oxaldie 1 and oxaldie 2, with rate enhancements over that of the butylamine catalysed reaction of approximately three orders of magnitude.[12] A key component of the function of the catalyst was the low pK_a, 7.2, of the amino terminal of oxaldie 1. Once the amino terminal group was blocked by acetylation to form oxaldie 2 the reactivity was decreased by a factor of three. However, considerable transition state binding was also provided by the flanking lysine residues since oxaldie 2 is also a powerful catalyst, although the pK_a of the reactive lysine is larger than 9. The main reason for our interest in the decarboxylation of oxaloacetate is because it is an ideal model system from which to engineer polypeptide catalysts that catalyse the aldol condensation reaction.[13] In spite of their impressive rate enhancements the Johnssen catalysts are unsuitable for that purpose because their tertiary structure is poorly defined. They aggregate to form helical bundles with relatively low helical contents and the relative positions of the amino acid side chains therefore vary substantially over time. Some uncertainty also prevails about the nature of the intermediates and their extinction coefficients.

We have developed a model system for the study of the oxaloacetate decarboxylation reaction based on a structurally more stable motif, the four-helix bundle. We have studied the kinetics by NMR and UV spectroscopy to avoid any uncertainties about the nature of the reacting species and we have engineered a catalyst that is susceptible to structural determination



1 11 15 19
Ac-N-A-A-D-Nle-E-H-A-I-R-**K**-L-A-E-**R**-Nle-A-A-G-

-G-P-V-D-
20 23

NH_2-G-A-R-A-F-A-E-F-**R**-R-A-L-R-E-A-Nle-Q-A-A-
42 34 30 24

NP-42

Fig. 1 The modelled structure of NP-42 showing only the side chains of the residues of the reactive site. Only the monomer is shown for clarity of presentation. The amino acid sequence of NP-42 is presented in the one-letter code where A is alanine, D is aspartic acid, E is glutamic acid, F is phenylalanine, G is glycine, H is histidine, I is isoleucine, K is lysine, L is leucine, N is asparagine, P is proline, Q is glutamine, R is arginine, V is valine and Nle is norleucine. Reactive site residues are shown in bold.

and that catalyses the reaction more efficiently, by approximately a factor of ten, than butylamine. The decarboxylase mimic provides the necessary platform for the further development of the model system towards the efficiency and selectivity of an enzyme-like catalyst.

Results and discussion

The design and synthesis of NP-42, a 42 residue polypeptide that folds into a hairpin helix-loop-helix motif and dimerises to form a four-helix bundle

The amino acid sequence of NP-42, Fig. 1, was chosen based on the propensities for secondary structure formation of the amino acid residues. Further stabilisation of the folded motif was provided by the introduction of residues capable of capping, salt bridge formation and stabilisation of the helix dipole. The amino acid sequence is similar to that of SA-42, which has been shown by extensive investigations by NMR and CD spectroscopy and equilibrium sedimentation ultracentrifugation to fold into an antiparallel hairpin helix-loop-helix dimer.[1,2] The driving force for the formation of the four-helix bundle motif is the hydrophobic interactions between amphiphilic helices. The design of folded polypeptides and proteins has been reviewed by Bryson *et al.*[14]

The structure of the folded motif is conveniently studied by CD spectroscopy since the mean residue elipticity at 222 nm, θ_{222}, provides a measure of its helical content. The monomeric form of the peptide is usually less structured with a low helical content and the state of aggregation is therefore readily disclosed by the magnitude of θ_{222}. The mean residue ellipticity of NP-42 is -22300 ± 1000 deg cm^2 dmol^{-1} in aqueous solution at pH 7.0 and room temperature and NP-42 most likely folds into a helix-loop-helix dimer, as does the template peptide SA-42.[1] The concentration dependence of θ_{222} is shown in Fig. 2 and θ_{222} is approximately independent of concentration in the range from 0.2 to 1 mM. Below 0.2 mM the helical content rapidly decreases and a dimer dissociation constant of roughly 10 μM

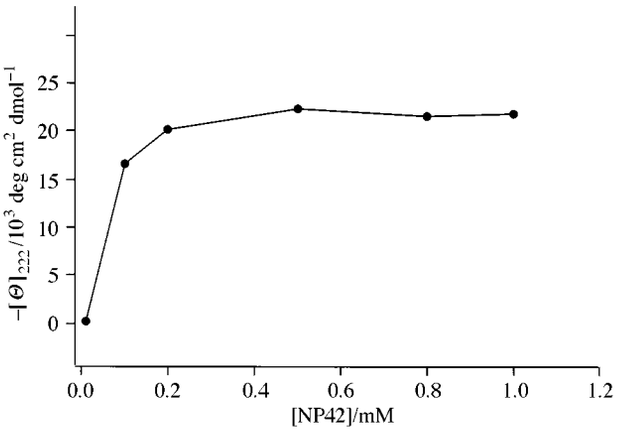


Fig. 2 The concentration dependence of the mean residue ellipticity at 222 nm, θ_{222}, of NP-42 in aqueous solution at pH 7.0 and room temperature.

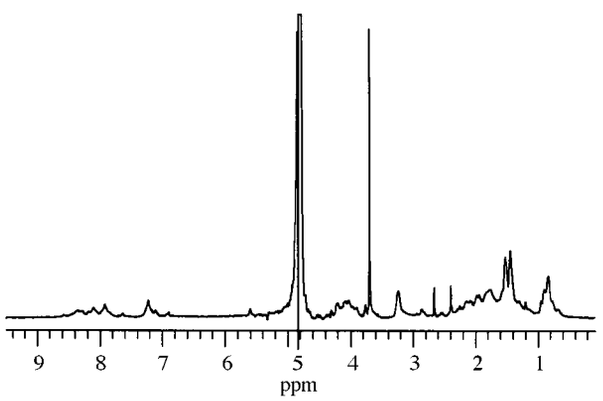


Fig. 3 The ^{1}H NMR spectrum of 0.2 mM NP-42 and 5 mM oxaloacetate in 90% H_2O–10% D_2O at 298 K and pH 7.0. The oxaloacetate methylene resonance at 3.71 ppm and the pyruvate methyl resonance at 2.41 ppm are clearly visible.

can be estimated from the peptide concentration at which θ_{222} has half of its maximum values. The helical content is independent of pH in the range from 4 to 8.

The reactive site of NP-42 contains a single lysine residue, Lys-11, flanked by Arg-10, Arg-15 and Arg-34. From the studies of the site-specific depression of histidine pK_a values in helical segments[15] it is probable that the flanking residues affect the dissociation constants by as much as 0.5 pK_a units, but the direct measurement of a pK_a in that range is fraught with uncertainty because the structure of the motif, and consequently the pK_a of the lysine residue, is changed as the positively charged residues are deprotonated. Similar effects were observed in the case of oxaldie 2 where the 'middle' lysine had a pK_a that was depressed by about a pK_a unit by four flanking lysine residues.[12] The flanking arginine residues were also included to bind the carboxylate anions in the peptide-bound intermediates. The amino acid sequence of NP-42 and the modelled folded structure of the polypeptide is shown in Fig. 1.

The peptide catalysed decarboxylation of oxaloacetate

The NP-42 catalysed decarboxylation of oxaloacetate was studied by ^{1}H NMR and UV spectroscopy. In the NMR spectrum of NP-42 and oxaloacetate in 90% H_2O–10% D_2O at pH 7.0 and 298 K under conditions of a 25-fold excess of substrate over catalyst, Fig. 3, the oxaloacetate methylene protons and the pyruvate methyl group are clearly visible at 3.71 and 2.41 ppm, respectively. The resonances were assigned from the spectra of the pure compounds in the same solvents in separate experiments. The direct observation of the reactants and products in a reaction with several intermediates simplifies the

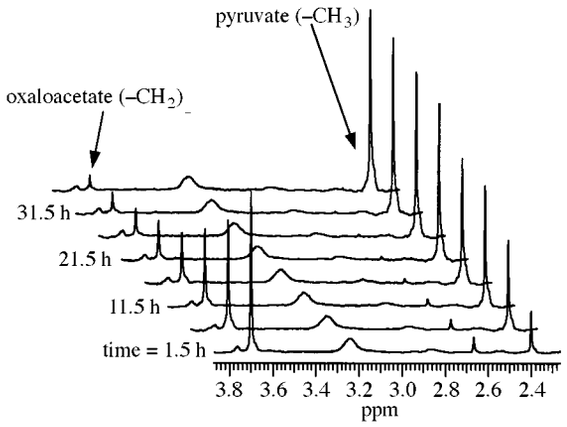


Fig. 4 The NP-42 catalysed decarboxylation reaction of oxaloacetate studied by NMR spectroscopy. The exponential decay of the oxaloacetate resonance (3.71 ppm) and the exponential growth of the pyruvate resonance (2.41 ppm) are clearly visible.

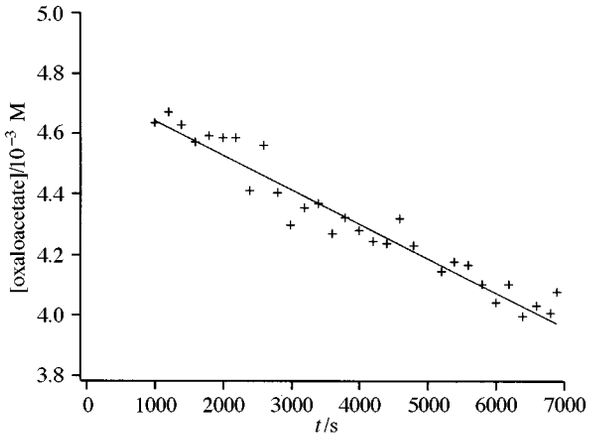


Fig. 5 Sample plot of concentration *versus* time in initial rate determination of NP-42 catalysed decarboxylation of oxaloacetate. The intensity to concentration transformations were calculated from the intensity extrapolated to time zero and the total amount of added substrate.

kinetic treatment considerably, and eliminates the laborious search for extinction coefficients of short-lived species.

The decarboxylation reaction was followed for several half-lives, Fig. 4, and it was shown that the sums of the integrals of the oxaloacetate and pyruvate resonances, compensated for the difference on the numbers of protons, stayed constant to within 95% over more than 3 half-lives. The effect of isotopic fractionation is very small under the reaction conditions. There is thus no appreciable formation of byproducts in the reaction and it can be monitored by following either the decrease of the oxaloacetate intensity or the increase in the pyruvate intensity. An attempt to detect intermediates directly by NMR spectroscopy was unsuccessful.

A single exponential function can be fitted to the experimental results under conditions of excess substrate over peptide to show that the reaction is pseudo-first-order with regards to oxaloacetate, and follows the rate law of eqn. (1), where OAA is

$$-\mathrm{d[OAA]}/\mathrm{d}t = (k_2[\text{NP-42}] + k_0)[\mathrm{OAA}] \qquad (1)$$

oxaloacetate, k_2 is the second-order rate constant for the NP-42 catalysed reaction and k_0 is the first-order rate constant for the spontaneous background decarboxylation. Having established the applicable rate law the rate constants may be determined by initial rate experiments, by dividing the initial slope of the plot of oxaloacetate intensity *versus* time by the total concentration of oxaloacetate, Fig. 5. The study of kinetics by NMR techniques have traditionally suffered from the uncertainty in the determination of intensities and integrals, but with recent developments in NMR technology the precision in the measurements are now sufficient for the accurate determination of rate constants, particularly under conditions of excess substrate. The determinations of the rate constants by UV spectroscopy followed the procedure of Johnsson *et al.*, using the reported extinction coefficient at 285 nm of 0.36 $\mathrm{mM^{-1}\ cm^{-1}}$, which is the measured difference between the extinction coefficient of oxaloacetate and that of pyruvate.[16]

The initial rate measurements in the NP-42 catalysed reaction over a substrate concentration range from 2.5 to 30 mM and a peptide concentration of 0.2 mM showed no trace of saturation kinetics, and the reaction was therefore treated as a pseudo-first-order reaction regardless of peptide or substrate concentration. The second-order rate constants were therefore determined in the usual way from the slope of the plot of the pseudo-first-order rate constants as a function of peptide concentration. The second-order rate constants determined by ^{1}H NMR and UV spectroscopy were 0.015 and 0.010 $\mathrm{M^{-1}\ s^{-1}}$, respectively, in reasonable agreement considering that they were determined by two different spectroscopic techniques and that the extinction coefficient is somewhat uncertain. The second-order rate constant for the butylamine catalysed reaction was reported previously, and it is 0.0011 $\mathrm{M^{-1}\ s^{-1}}$. The second-order rate constant of the NP-42 catalysed reaction is therefore approximately a factor of ten larger than that of the butylamine catalysed one.

The catalytic efficiency of NP-42

The efficiency of the oxaldie 1 catalysed reaction is to some degree due to the low pK_a of the amino terminal of the peptide, but mainly to the binding energy provided by the flanking lysine residues in the transition state of the reaction. The pK_a of the lysine of NP-42 has not been measured but it is unlikely that it is severely depressed. We have previously reported on extensive studies of histidine pK_a values in folded helix–loop–helix dimers and the maximum pK_a depression obtained on the surface of the motif was 1.2 pK_a units.[15] A similar effect was also reported for the 'middle' lysine of oxaldie 2, that was surrounded by four lysine residues.[12] No pK_a depression of the lysine residue beyond at most one pK_a unit is therefore probable in NP-42. The main factor so far is therefore probably the binding energy, although modest, that is provided by flanking arginine residues. The introduction of *e.g.* artificial amino acid residues with low pK_a values is expected to enhance the reactivity of the catalyst substantially beyond the order of magnitude that has already been accomplished. However, the fact that no saturation kinetics are observed shows that substrate binding is not efficient. To increase the catalytic efficiency, further development towards catalysts capable of substrate and transition state binding is now under way by systematic variation of the residues in the reactive site.

Experimental

Peptide synthesis, purification and identification

The peptide NP-42 was synthesised on a Fmoc-Gly-PEG-PS polymer (PerSeptive Biosystems) using a PerSeptive Biosystems Pioneer automated peptide synthesiser and a standard Fmoc chemistry protocol. The peptide was cleaved from the polymer and deprotected with TFA (10 ml), anisole (222 μl), ethanedithiol (333 μl) and thioanisole (555 μl) for two hours at room temperature. After diethyl ether precipitation and lyophilisation, the peptide was purified by reversed-phase HPLC on a semi-preparative C-8 Kromasil, 7 μ column. It was eluted isocratically using a solvent with 38% propan-2-ol in 0.1% TFA, a flow rate of 10 ml $\mathrm{min^{-1}}$ and UV detection at 229 nm. The peptide was identified by electrospray mass spectrometry (VG Analytical, ZabSpec). The obtained molecular weight was

within 1 au from the calculated and no high molecular weight impurities could be detected. The purity was estimated to be more than 95% from ES-MS.

NMR and CD spectroscopy

The ^{1}H NMR spectra were recorded at 500 MHz using a Varian Inova 500 MHz NMR spectrometer. The peptide concentrations were estimated from weighing, assuming a water content of 25%, and subsequently determined by quantitative amino acid analysis. The spectra were recorded in aqueous phosphate buffer in H_2O–D_2O 90 : 10 and the pH (uncorrected) was adjusted by the addition of dilute DCl or NaOD. CD spectra were recorded on a Jasco J-270 spectropolarimeter, routinely calibrated with (+)-camphor-10-sulfonic acid. CD spectra were measured at room temperature in the wavelength interval 280 to 190 nm in a 1 mm cuvette. 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. The pH dependence of the mean residue ellipticities were determined without buffer.

Kinetic measurements

The kinetic experiments were carried out using a Varian Cary 5 spectrophotometer equipped with a Varian temperature controller by following the absorbance at 285 nm as a function of time. A stock solution of peptide in buffer was prepared, the pH was adjusted if necessary and the solution was centrifuged before diluting by pipetting to the desired concentrations. The buffer that was used was 50 mM Bis–Tris, and the ionic strength was adjusted to 0.15 M by the addition of NaCl. The concentrations of the peptide stock solutions were determined by quantitative amino acid analysis. In a typical kinetic experiment, 300 μl of a 0.2 mM peptide solution was temperature equilibrated in a 1 mm quartz cuvette, and 50 μl of substrate solution was added. The reaction was followed for approximately four hours and the initial rate was determined by fitting the equation for a straight line to the experimental data using Igor Pro software (Wavemetrics Inc.) and an extinction coefficient of 0.36 mM^{-1} s^{-1}. The rate constant for the background reaction was determined in independent experiments and the average value after three measurements was 1.2×10^{-5} s^{-1}. The pseudo-first order rate constants were calculated from the initial rates by division with the total concentration of oxaloacetate, and the second-order rate constants were calculated by subtracting the rate constant for the background reaction and dividing by the peptide concentration. The reported second-order rate constant is the average of five runs. The rate constants determined by NMR spectroscopy were determined from the plots of intensities *versus* time. The linewidths did not vary over time and it proved to be more convenient to measure intensities rather than integrals. The spectra were processed with matched filters to maximise signal-to-noise, and with zero filling to improve the intensity measurement. The reported second-order rate constant was calculated from the plots of the measured pseudo-first-order rate constants as a function of three peptide concentrations. The errors in the reported second-order rate constant is due to errors in the linear regression analysis and the quantitative amino acid analysis. The data processing of the kinetic results were processed in a manner analogous to that of the UV spectroscopic measurements.

Acknowledgements

We are indebted to the Swedish Natural Science Research Council and Carl Tryggers Stiftelse for financial support. We thank The Swedish NMR Centre at Göteborg University for access to a 500 MHz NMR spectrometer and Charlotta Damberg for expert help with the measurements.

References

1 S. Olofsson, G. Johansson and L. Baltzer, *J. Chem. Soc.*, *Perkin Trans. 2*, 1995, 2047.
2 S. Olofsson and L. Baltzer, *Folding Des.*, 1996, **1**, 347.
3 K. S. Broo, L. Brive, P. Ahlberg and L. Baltzer, *J. Am. Chem. Soc.*, 1997, **119**, 11362.
4 L. Baltzer, A.-C. Lundh, K. Broo, S. Olofsson and P. Ahlberg, *J. Chem. Soc.*, *Perkin Trans. 2*, 1996, 1671.
5 K. Broo, L. Brive, A.-C. Lundh, P. Ahlberg and L. Baltzer, *J. Am. Chem. Soc.*, 1996, **118**, 8172.
6 M. Kjellstrand, K. Broo, L. Andersson, C. Farre, Å. Nilsson and L. Baltzer, *J. Chem. Soc.*, *Perkin Trans. 2*, 1997, 2745.
7 M. Allert, M. Kjellstrand, K. Broo, Å. Nilsson and L. Baltzer, *Chem. Commun.*, 1998, 1547.
8 K. J. Pedersen, *Acta Chem. Scand.*, 1954, **8**, 710.
9 R. W. Hay, *Aust. J. Chem.*, 1965, **18**, 337.
10 W. J. Spetznagel and I. M. Klotz, *J. Am. Chem. Soc.*, 1976, **98**, 8199.
11 D. L. Leussing and N. V. Raghavan, *J. Am. Chem. Soc.*, 1980, **102**, 5635.
12 K. Johnsson, R. K. Allemann, H. Widmer and S. A. Benner, *Nature*, 1993, **365**, 530.
13 C. F. Barbas III, A. Heine, G. Zhong, T. Hoffmann, S. Gramatikova, R. Björnestedt, B. List, J. Anderson, E. A. Stura, I. A. Wilson and R. A. Lerner, *Science*, 1997, **278**, 2085.
14 J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neill and W. F. DeGrado, *Science*, 1995, **270**, 935.
15 K. S. Broo, L. Brive, R. S. Sott and L. Baltzer, *Folding Des.*, 1998, **3**, 303.
16 K. Johnsson, Ph.D. Thesis, ETH, Zürich, 1992.

Paper 8/03453B