research papers

Following directed evolution with crystallography: structural changes observed in changing the substrate specificity of dienelactone hydrolase

**Hye-Kyung Kim, Jian-Wei Liu,
Paul D. Carr and David L. Ollis***

Research School of Chemistry, Australian
National University, Building 35, Science Road,
Canberra, ACT 0200, Australia

Correspondence e-mail: ollis@rsc.anu.edu.au

The enzyme dienelactone hydrolase (DLH) has undergone directed evolution to produce a series of mutant proteins that have enhanced activity towards the non-physiological substrates α-naphthyl acetate and p-nitrophenyl acetate. In terms of steady-state kinetics, the mutations caused a drop in the K_m for the hydrolysis reaction with these two substrates. For the best mutant, there was a 5.6-fold increase in k_{cat}/K_m for the hydrolysis of α-naphthyl acetate and a 3.6-fold increase was observed for p-nitrophenyl acetate. For α-naphthyl acetate the pre-steady-state kinetics revealed that the rate constant for the formation of the covalent intermediate had increased. The mutations responsible for the rate enhancements map to the active site. The structures of the starting and mutated proteins revealed small changes in the protein owing to the mutations, while the structures of the same proteins with an inhibitor co-crystallized in the active site indicated that the mutations caused significant changes in the way the mutated proteins recognized the substrates. Within the active site of the mutant proteins, the inhibitor was rotated by about 180° with respect to the orientation found in the starting enzyme. This rotation of the inhibitor caused the displacement of a large section of a loop on one side of the active site. Residues that could stabilize the transition state for the reaction were identified.

1. Introduction

The utility of directed evolution for generating mutant enzymes with improved or new activities is well known (Arnold & Volkov, 1999; Dalby, 2003). However, directed evolution also has the potential to provide a great deal of information concerning how enzymes function. It can provide a series of mutant proteins that have a range of kinetic properties; by looking at the structures of these mutant proteins with inhibitors in the active site, it should be possible to determine which features are important for activity. We have taken this approach in probing the catalytic mechanism of the enzyme dienelactone hydrolase (DLH) when it acts on non-physiological substrates.

DLH is the third enzyme of the halocatechol branch of the β-ketoadipate pathway and catalyses the hydrolysis of both (E) and (Z) dienelactone (4-carboxymethylene-but-2-ene-4-olide) to maleyl acetate (2-oxo-but-1,3-diene-1,4-dicarbox-ylate; Schmidt & Knackmuss, 1980; Fig. 1). The enzyme was first identified in *Pseudomonas* sp. B13 with its gene encoded on a plasmid. DLH is a 25.5 kDa monomer that belongs to the α/β-hydrolase fold class of enzymes (Ollis *et al.*, 1992). It has a

catalytic triad consisting of a nucleophilic cysteine (Cys123), a histidine (His202) and an aspartic acid (Asp171) (Pathak *et al.*, 1988). The refined structure of DLH shows Cys123 to be discretely disordered, with two conformations related by a rotation about the $C^{\alpha}-C^{\beta}$ bond (Pathak & Ollis, 1990). In one of these conformations the nucleophile exists as a thiolate and is pointing away from the active site. It is thought that the nucleophile can be shifted to an active conformation upon substrate binding through a charge-relay system that involves Arg206 and Glu36 (Cheah, Austin *et al.*, 1993; Beveridge & Ollis, 1995). The catalytic nucleophile Cys123 has been mutated to a serine (C123S), in which the side chain is in the active conformation with its oxygen hydrogen bonded to His202. Consistent with the mechanism proposed for wild-type DLH, the C123S mutant has very little activity towards dienelactone (Cheah, Ashley *et al.*, 1993), but is capable of interconverting the *E* and *Z* forms of dienelactone (Walker *et al.*, 2000). The C123S mutant also has a low level of activity towards *p*-nitrophenyl acetate (Fig. 1) and other simple esters (Pathak *et al.*, 1991). The initial object of the present work was to use directed evolution to enhance the ability of DLH to degrade simple esters.

In the present study, we developed a procedure to evolve DLH to increase α-naphthyl acetate activity. The mutants we obtained also had enhanced activity towards *p*-nitrophenyl acetate. We examined the structures of the mutant enzymes, noting the structural differences. We also obtained the structures of several key mutants in the presence of a covalently bound inhibitor, phenylmethylsulfonyl fluoride (PMSF; Fig. 1), which mimics the transition state of the reaction in question (Robinson *et al.*, 2000). We compared inhibitor binding in the wild-type and mutant proteins with enhanced activities.

2. Methods and materials

All chemicals were purchased from Sigma or Aldrich. Molecular-biology reagents and enzymes were purchased from Roche. Primers were obtained from Genework. Qiagen DNA purification kits were used for all DNA purification.

The *Escherichia coli* strain DH5α [*sup*E44 Δ*lac*U169 (*φ*80 *lac*ZΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1] (Hanahan, 1983) was used for all aspects of the work described. Cells were grown at 310 K. Cell lines were maintained on LB agar plates supplemented with 100 µg ml^{-1} ampicillin.

2.1. Preparation of a DLH mutant pool

A pool of 39 genes with one or more active-site mutations was used as the starting point for directed evolution. These mutations were generated by site-directed mutagenesis as part of an unrelated study (Jeffries, 1998; Walker, 2001). The genes used in the initial round of shuffling contained the following mutations: Q35G, E36A, E36D, E36G, E36I, I37D, I37E, I37H, R81A, R81K, W88A, W88H, C123S, R206A and R206K.

2.2. DNA shuffling

The coding sequences of DLH mutants were amplified by PCR from a mixture of plasmid DNA from 39 DLH mutants. The DLH mutant coding sequences were used for DNA shuffling essentially as described by Stemmer (1994). Primers DLH-1 (5′-CTT TAA GAA GGA GAT ATA CAT ATG-3′) and DLH-2 (5′-GAG CTC GAA TTC TTA TTA TG-3′) were used to amplify the shuffled DLH coding sequences. The shuffled DLH was cloned into a constitutive expression vector pCY76 (*par*$^{+}$, *bla*$^{+}$, *lacZpo, T7φ10^{tir+}*) (Yang *et al.*, 2003), then used to transform competent cells by electroporation. The transformed cells were plated on to LB plates supplemented with 100 µg ml^{-1} ampicillin.
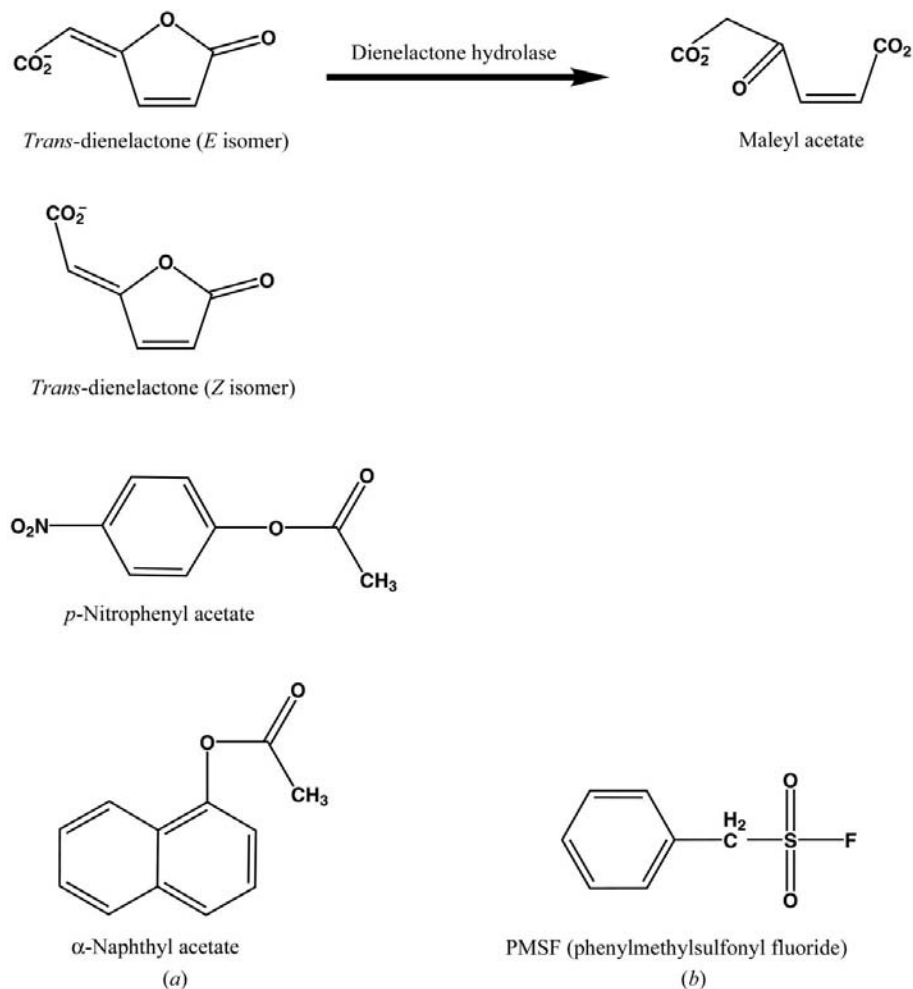

Figure 1
(*a*) The substrates and (*b*) an inhibitor of DLH.

research papers

2.3. Screening

After incubation at 310 K overnight, the colonies on each plate were blotted onto 3MM filter paper (Whatman, England). The filter papers with colonies were soaked in a solution containing 0.3 mM α-naphthyl acetate, 0.5 mM Fast Garnet GBC salt, 20 mM HEPES pH 7.0, 1 mM EDTA pH 8.0 and 10% acetonitrile. A total of 20 000 colonies were screened and 106 'darker' colonies were identified. The darker colonies were those that stained significantly faster than the others. These colonies were picked and grown in 96-well plates. The overnight cultures were then assayed for their ability to hydrolyse α-naphthyl acetate. This was achieved by adding a 100 µl aliquot of each culture to the corresponding wells of a 96-well plate that contained 100 µl reaction mixture. The reaction mixture consisted of 0.6 mM α-naphthyl acetate, 1 mM Fast Garnet GBC salt, 20 mM HEPES pH 7.0, 1 mM EDTA pH 8.0 and 2% acetonitrile. The reaction with α-naphthyl acetate was monitored with a Labsystem Multiskan UV–Vis spectrophotometer at 540 nm. Activities were determined from the slope of the line of best fit through ten data points in each case. Clones exhibiting the highest activities were chosen for the next round of directed evolution. Four rounds of directed evolution were performed. In each round, a library of approximately 20 000 colonies was screened and between 20 and 100 of the best clones were chosen for next round of directed evolution. The DLH mutants were sequenced using an ABI 3730 DNA sequencer.

2.4. Expression and purification

DH5α cells bearing DLH were grown in LB supplemented with 100 µg ml^{-1} ampicillin at 310 K overnight before cells were harvested and resuspended in a buffer consisting of 20 mM HEPES, 1 mM EDTA pH 7.0. Cells were lysed with a French press. The soluble fraction was loaded onto a DEAE Fractogel column. DLH was eluted between 50 and 100 mM NaCl and gave a prominent peak on the chromatogram. SDS–PAGE analysis confirmed that this peak contained the partially purified proteins. The peak fractions were pooled and loaded onto a Sephadex G75 column. The purified protein was concentrated using an Amicon Centricon 10 centrifugal filter and the concentration was determined with a UV spectrophotometer using the extinction coefficient $\varepsilon_{280\,nm}$ = 32 430 M^{-1} cm^{-1} for DLH and the mutants. The extinction coefficient was calculated using the method of Gill & von Hippel (1989) as implemented in the *ProtParam* program found on the ExPASy Proteomics Server.

2.5. Crystallization

Crystals of the native and mutant DLH proteins were formed by vapour diffusion of hanging drops containing 4 µl protein at a concentration of 11 mg ml^{-1}, 4 µl 20 mM sodium citrate buffer pH 6.5 and 1.2 M ammonium sulfate. Co-crystals of the proteins with the inhibitor PMSF were formed by vapour diffusion of sitting drops containing 4 µl protein at a concentration of 11 mg ml^{-1} and 12 µl 20 mM sodium citrate buffer pH 6.5, 1.2 mM ammonium sulfate and 2 µl 100 mM PMSF dissolved in 2-propanol. Crystals were grown in a cold room until their size reached 0.1 $\times$ 0.1 $\times$ 0.5 mm on average.

2.6. Data collection and structure determination

Crystals were transferred into a cryobuffer overnight prior to data collection. The cryobuffer was the same as the crystallization buffer with the glycerol concentration increased to 35%. Data were collected from crystals that were flash-cooled to 100 K in the nitrogen stream of an Oxford Cryostream (Oxford Cryosystems, UK). The X-ray source was a Rigaku RU-200 rotating-anode X-ray generator producing Cu $K\alpha$ radiation with power set at 59 kV and 84 mA. Data were recorded on a Rigaku R-AXIS IIC imaging-plate detector. For the mutants G2-06 and G3-16, data were collected from the synchrotron source at the BioCARS beamline under the same cryobuffer and cooling regime. All data were processed and scaled using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

Refinement was carried out using the program *CNS_SOLVE* (Brünger *et al.*, 1998). A test set consisting of 5% of the data was excluded from the refinement calculations for cross-validation purposes. Cycles of torsion-angle dynamics simulated annealing, minimization and individual temperature-factor refinement were interspersed with manual building and automatic solvent placement. These were performed using standard *CNS_SOLVE* refinement protocols.

An initial model was used without water molecules and with a modified residue at position 123 from cysteine to serine (PDB code 1din). *SIGMAA* (Read & Moult, 1992) difference electron-density maps ($2mF_o - DF_c$, $mF_o - DF_c$) were calculated to check the course of the refinement and manual adjustments were progressively made to the model. Omit electron-density maps were also calculated for selected residues to check that side chains had been modelled in the correct positions. Omit maps were also used to build the loop comprising residues 165–185 in mutants crystallized with the bound inhibitor PMSF. The program *O* (Jones *et al.*, 1991) was used for manual rebuilding. Only the PMS portion of PMSF is required for incorporation into the structure, because the F atom is cleaved from the molecule during the binding reaction. The S atom of the inhibitor is covalently linked to the OG atom of Ser123, creating a new residue, Pms123 (Robinson *et al.*, 2000). Pms123 from PDB entry 1ggv was used as the starting position of the PMS moiety. The stereochemistry was checked using the programs *PROCHECK* (Laskowski *et al.*, 1993) and *WHATCHECK* (Hooft *et al.*, 1996).

The structures of the mutant proteins were overlaid on DLH-C1 using the least-squares algorithm of Kabsch (1976) to minimize the differences in the position of the main-chain atoms of 233 residues. The schematic diagram of the active sites of DLH and mutants shown in Fig. 4 was drawn using the *ChemDraw Ultra* program (CambridgeSoft Co., UK).

2.7. Kinetic assays

The kinetics of hydrolysis of substrates by DLH and the mutant forms of DLH were obtained using a Shimadzu

UV-2101PC absorption spectrophotometer equipped with a thermostated cell compartment at 298 K (Shimadzu, Japan). The kinetic constants for the two substrates (α-naphthyl acetate, p-nitrophenyl acetate) were determined by varying the concentrations of the substrate with a constant protein concentration. The substrates were dissolved in acetonitrile and used at less than 1%(v/v) in the final assay buffer. The assay buffer contained 20 mM HEPES pH 7.0, 1 mM EDTA. The rate of hydrolysis of α-naphthyl acetate was measured by monitoring the molar quantity of α-naphthol released from the hydrolysis by the enzyme at 236.7 nm after zeroing the protein absorbance at this wavelength. Absolute initial reaction velocities (μM min^{-1}) were calculated using an extinction coefficient $\varepsilon_{236.7\,nm} = 24\,000\,M^{-1}\,cm^{-1}$ (Mastropaolo & Yourno, 1981). The rate of hydrolysis of p-nitrophenyl acetate was measured by monitoring the molar quantity of p-nitrophenolate at 400 nm and the absolute initial reaction velocities (μM min^{-1}) were calculated using an extinction coefficient $\varepsilon_{400\,nm} = 8740\,M^{-1}\,cm^{-1}$ (Pathak $et\ al.$, 1991). The kinetic constants (V_{max} and K_m) were obtained by fitting the data to the equation (Michaelis & Menten, 1913)

$$V = V_{max}S/(K_m + S), \tag{1}$$

where V is the initial velocity, V_{max} is the maximum velocity, K_m is the apparent dissociation constant and S is the substrate concentration. k_{cat}, a first-order rate constant, is calculated according to the equation $k_{cat} = V_{max}/E$, where E is the protein concentration used in the assay.

In the presence of an excess of either substrate, the DLH mutants exhibited biphasic kinetics, with the initial phase much faster than the second. This situation occurs when an enzyme-bound intermediate accumulates during the reaction and can be described by

$$E + S \overset{K_s}{\longleftrightarrow} ES \overset{k_2}{\longrightarrow} EI + P_1 \overset{k_3}{\longrightarrow} E + P. \tag{2}$$
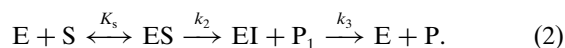
In the first stage, the enzyme and the substrate combine to give an enzyme–substrate complex (ES). This is assumed to be rapid and reversible with no chemical changes taking place. K_s is the dissociation constant of the enzyme–substrate complex. The substrate reacts to form stoichiometric amounts of the enzyme-bound intermediate (EI) and product (P_1) with a rate constant of k_2. In the second stage, the intermediate breaks down to form the free enzyme (E) and release the product (P) with a rate constant of k_3. At high substrate concentration, DLH displays initial burst-phase kinetics consistent with the formation of an enzyme intermediate. Burst kinetics of P $versus$ time allowed the rate constants k_2 and k_3 to be calculated using the following expression (Fersht, 1999),

$$\text{Burst} = [E]_0\left(\frac{k_2}{k_2 + k_3}\right)^2, \quad \text{Slope} = \frac{[E]_0 k_2 k_3}{k_2 + k_3}. \tag{3}$$

Using these estimated rate constants, it was possible to calculate k_{cat}^* and K_s values from

$$k_{cat}^* = \frac{k_2 k_3}{k_2 + k_3}, \quad K_s = \frac{K_m(k_2 + k_3)}{k_3} \tag{4}$$

Table 1
Sequences of the DLH clones and mutants and their activity enhancement compared with wild-type DLH (C1).

Clone	Mutations	Activity† (%)
C1‡§	C123S	100
C2	E36D	104
C3	E36D R81A C123S R206A	114
C4‡	E36N C123S	140
C5‡	E36A C123S	—
C6‡§	C123S R206A	65
C7	E36D C123S F173A	—
G2-06‡§	E36D C123S	183
G2-07	E36D C123S A229V	187
G2-13	E36D C123S R206T A229V	122
G2-16	E36D C123S A134T A229V	164
G2-28	E36D R45Q C123S	172
G3-12	E36D C123S G211D A229V	172
G3-16‡§	E36D C123S A134S S208G A229V K234R	256
G4-03	E36D C123S G211D K234N	205
G4-07	E36D R45Q C123S A205D A229V	148
G4-09	E36D C123S A205D A229V	222
G4-11	E36D C123S A205D A229V	224
G4-12	E36D C123S A205D A229V	230
G4-14	E36D C123S A205D	217
G4-113‡§	E36D R105H C123S G211D K234N	233

† Activity (%) = activity (mutant)/activity (C1) in a whole cell hydrolysis ratio measurement of α-naphthyl acetate. Experiments were repeated six times with less than 5% error. ‡ Kinetic properties of this mutant were measured and are reported in Table 2. § Structure presented in Table 3.

(Bolon & Mayo, 2001). Calculated k_{cat}^* values are compared with those obtained from the kinetic analysis determined from UV absorbance in Table 2.

3. Results

3.1. Generating mutants

Directed evolution was used to generate a series of DLH mutants that had enhanced hydrolytic activity towards α-naphthyl acetate. Table 1 summarizes the results of four rounds of evolution. The labels given to mutant proteins in this table will be used to identify mutant proteins in the subsequent discussion. In the first round of evolution, a pool of 39 DLH genes containing various combinations of active-site mutations were shuffled. The mutations in these genes are given in §2. Only two of these mutations found their way into the second and subsequent rounds of evolution. The C123S mutation was in 33 of the initial 39 genes and was present in all genes used for shuffling in the second and subsequent rounds of evolution. For this reason, the DLH-C1 protein (with the C123S mutation; Table 1) was used in place of wild-type DLH (wtDLH) for comparisons with other mutants. The E36D mutation was present in only one of the 39 genes that were shuffled in the first round, but was in all genes selected at the end of round 1. The effect of this mutation alone on the wild-type protein was tested in the mutant protein designated DLH-C2 and found to cause only a small increase in activity in the absence of the C123S mutation (Table 1). The addition of E36D to the C123S mutation causes a significant increase in activity (DLH-G2-06; Table 1). For this reason, the DLH-G2-

Table 2
Kinetic data.

(*a*) α-Naphthyl acetate.

Mutant	K_m (×10 μM)	k_{cat} (×10^{-4} s^{-1})	V_{max} (×10^{-2} μM s^{-1})	k_{cat}/K_m† (×10 M^{-1} s^{-1})	k_2 (×10^{-3} s^{-1})	k_3 (×10^{-4} s^{-1})	k_{cat}^*‡ (×10^{-4} s^{-1})	K_s§ (×10 μM)
DLH-C1	7.4 ± 0.8	63.4 ± 1.3	5.1 ± 0.2	8.6 (1.0) ± 0.7	17.1 ± 0.8	76.4 ± 0.4	52.8 ± 1.0	23.8 ± 1.8
G4-113	3.2 ± 0.3	86.2 ± 0.7	6.2 ± 0.1	27.2 (3.2) ± 2.3	68.2 ± 8.7	89.8 ± 0.5	79.3 ± 1.5	27.2 ± 0.4
G2-06	2.3 ± 0.3	71.2 ± 0.8	5.4 ± 0.1	31.2 (3.6) ± 3.0	96.5 ± 21.7	71.1 ± 0.4	66.3 ± 1.3	33.2 ± 3.2
G3-16	2.0 ± 0.2	98.1 ± 0.7	7.1 ± 0.1	48.3 (5.6) ± 5.0	128.2 ± 27.5	100.3 ± 0.5	93.0 ± 1.9	28.0 ± 2.3
DLH-C4	5.1 ± 0.5	61.0 ± 1.4	4.7 ± 0.2	11.9 (1.4) ± 1.0	22.3 ± 1.4	62.9 ± 0.3	49.1 ± 0.9	23.4 ± 1.4
DLH-C5	—	—	—	—	—	—	—	—
DLH-C6	16.8 ± 0.6	95.4 ± 0.4	7.4 ± 0.1	5.7 (0.66) ± 0.2	14.6 ± 0.4	160.1 ± 1.0	76.5 ± 1.2	32.2 ± 0.9

(*b*) *p*-Nitrophenyl acetate.

Mutant	K_m (×10 μM)	k_{cat} (×10^{-4} s^{-1})	V_{max} (×10^{-2} μM s^{-1})	k_{cat}/K_m† (×10 M^{-1} s^{-1})	k_2 (×10^{-3} s^{-1})	k_3 (×10^{-4} s^{-1})	k_{cat}^*‡ (×10^{-4} s^{-1})	K_s§ (×10 μM)
DLH-C1	2.1 ± 0.2	107.7 ± 1.5	7.2 ± 0.2	52.0 (1.0) ± 4.3	100.6 ± 14.3	116.9 ± 0.6	104.7 ± 2.0	19.9 ± 0.5
G4-113	0.8 ± 0.2	123.0 ± 1.9	8.9 ± 0.2	149.5 (2.9) ± 30.5	75.8 ± 7.0	141.8 ± 0.7	119.5 ± 2.3	5.2 ± 0.8
G2-06	0.8 ± 0.1	97.9 ± 1.7	7.4 ± 0.2	130.9 (2.5) ± 19.4	72.4 ± 8.4	104.0 ± 0.5	91.0 ± 1.7	6.0 ± 0.4
G3-16	0.7 ± 0.1	126.8 ± 0.8	9.3 ± 0.2	185.8 (3.6) ± 24.3	63.6 ± 5. 0	142.3 ± 0.8	116.3 ± 2.2	3.7 ± 0.3

† Values in parentheses in are scaled to DLH-C1 (1.0). ‡ k_{cat}^* is a calculated value from k_2 and k_3 using the equation $k_{cat}^* = k_2 k_3/(k_2 + k_3)$. § K_s is calculated value from K_m, k_2 and k_3 using the equation $K_s = K_m(k_2 + k_3)/k_3$.

06 protein was selected for detailed kinetic and structural characterization. From Fig. 2 it is evident that there were a number of mutations (R45Q and A229V) that did not give rise to a large increase in activity when combined individually with the E36D and C123S mutations. There were also a number of mutants with four changes that exhibited an activity similar to that of DLH-G2-06. This observation suggested that changes such as A134T, R206T and G211D had a minor effect, if any, on activity. The best mutant, DLH-G3-16, was selected for further characterization: it had six changes with one, S208G, thought to be responsible for a significant increase in activity owing to its close proximity to the active site.

Of the mutations in the initial pool, it was thought that the mutations R206A and R81A would open up the active site as well as remove charged groups from within the active site and that this should facilitate the hydrolysis of neutral esters considerably larger than the physiological substrate. DLH-C3 contained the R206A and R81A as well as the E36D and C123S mutations and was found to have an activity slightly higher than DLH-C1 but well below the activity of DLH-G2-06 (Table 1), suggesting that the removal of these two arginine side chains had a detrimental effect on activity. Directed evolution also produced DLH-G2-13 which contained C123S, E36D, R206T and A229V, which had a level of activity that was below that of DLH-G2-06, suggesting that the nature of the residue at position 206 was important for activity.
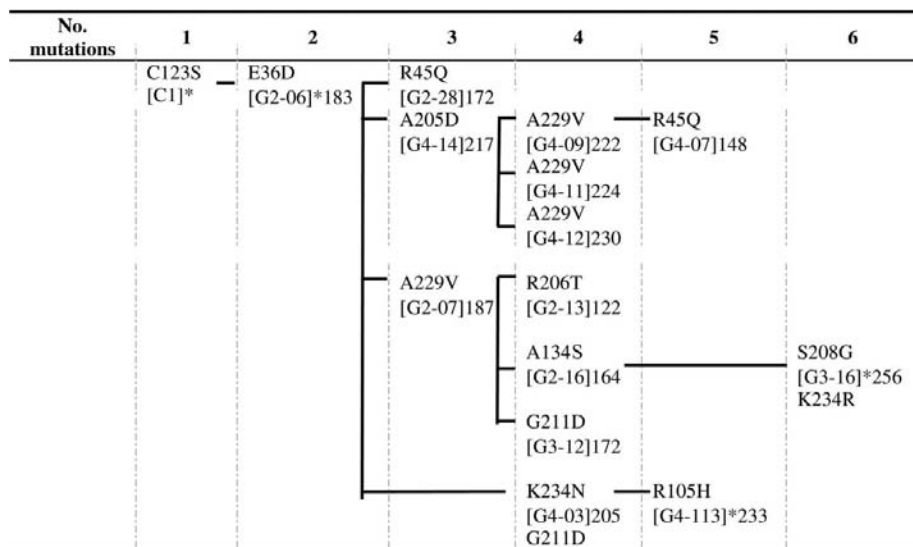
3.2. Catalytic properties of evolved DLH proteins

A number of mutant forms of DLH were selected for kinetic characterization with respect to the hydrolysis of α-naphthyl acetate and *p*-nitrophenyl acetate; the resulting K_m and k_{cat} values are given in Table 2. α-Naphthyl acetate was used for selection in the directed evolution and it is not surprising that better results were obtained with it rather than with *p*-nitrophenyl acetate. For both substrates, the most noticeable effect on the kinetic parameters was a drop in K_m. For α-naphthyl acetate there was a small increase in the k_{cat} values for some mutants, while for *p*-nitrophenyl

Figure 2
Evolutionary tree of DLH. Clone name and activity are also shown. * indicates the mutants selected for the kinetic and structural studies.

Table 3
Summary of X-ray diffraction data and refinement statistics.

Values in parentheses are for the outer data shell. DLH and the mutants form in space group $P2_12_12_1$. All data were collected in-house ($\lambda = 1.5418$ Å), except for those for G2-06 and G3-16 which were collected at a synchrotron ($\lambda = 1.2915$ Å).

	DLH-C1	DLH-C6	G2-06	G3-16	G4-113	DLH-C1–PMS	G2-06–PMS	G3-16–PMSF
X-ray diffraction data								
Unit-cell parameters								
a (Å)	48.2	48.4	48.4	48.5	48.5	48.5	47.6	47.6
b (Å)	71.3	70.6	70.9	70.4	70.8	70.9	70.4	70.7
c (Å)	76.9	77.3	76.9	77.2	77.4	77.5	77.6	77.6
$\alpha = \beta = \gamma$ (°)	90	90	90	90	90	90	90	90
Resolution (Å)	30–1.7	15–1.7	34–1.5	36–1.4	34–1.8	25–1.9	20–1.7	15–1.7
	(1.70–1.76)	(1.70–1.76)	(1.50–1.55)	(1.40–1.45)	(1.80–1.86)	(1.90–1.97)	(1.70–1.76)	(1.70–1.76)
Observed reflections	151210	230539	257291	252026	272040	162672	314687	353135
Unique reflections	28216	29795	43132	52804	25363	21641	29337	29298
Completeness (%)	94.4 (92.0)	96.7 (98.2)	98.2 (99.0)	97.4 (81.8)	97.3 (95.2)	98.1 (96.2)	100 (99.9)	99.4 (94.5)
R_{merge}† (%)	9.5 (25.5)	4.2 (20.2)	4.9 (19.5)	4.1 (13.4)	3.5 (21.5)	3.3 (16.7)	4.9 (24.5)	4.9 (24.1)
Average $I/\sigma(I)$	19.4	21.6	18.8	27.7	16.9	19.4	18.7	22.6
Refinement statistics								
Resolution (Å)	30–1.7	15–1.7	34–1.5	36–1.4	34–1.8	25–1.9	20–1.7	15–1.7
No. of reflections (working)	26847	28347	41037	50221	24088	20566	27895	27860
No. of reflections (test)	1369 [4.9%]	1448 [4.9%]	2095 [4.9%]	2583 [4.9%]	1275 [5.0%]	1075 [5.0%]	1442 [4.9%]	1438 [4.9%]
R_{cryst}/R_{free}‡ (%)	0.165/0.202	0.174/0.195	0.171/0.198	0.167/0.181	0.173/0.196	0.188/0.212	0.189/0.216	0.187/0.211
No. of atoms/molecules/ions (excluding H atoms)								
Protein atoms	1777	1771	1776	1777	1779	1787	1786	1787
Water molecules	199	184	185	199	158	114	118	124
SO_3^{2-} ions	3	1	1	1	0	2	2	2
Glycerol molecules	1	2	4	5	3	2	1	1
PMSF molecules	0	0	0	0	0	1	1	1
R.m.s. deviations from ideal geometry								
Bond lengths (Å)	0.025	0.023	0.026	0.024	0.034	0.025	0.025	0.026
Bond angles (°)	2.125	2.009	2.249	2.189	2.599	1.880	2.429	2.448
Dihedrals (°)	24.437	24.626	24.813	24.775	24.770	24.586	24.902	24.637
Impropers (°)	1.528	1.4253	1.572	1.692	1.914	1.941	2.127	2.102
Average B factors (Å^2)								
Mean	13.48	16.34	14.60	9.58	17.84	19.75	19.81	17.74
Minimum	7.5	9.06	7.95	4.9	10.55	9.72	10.52	9.98
Maximum	41.41	45.27	40.96	30.34	37.22	43.47	46.8	41.62
Ramanchandran plot (%)								
Most favoured	93.8	92.8	93.3	93.3	92.8	93.3	91.2	90.7
Additional allowed	5.2	6.7	6.2	6.2	6.7	6.2	7.7	8.3
Generously allowed	0.5	0.0	0.0	0.0	0.0	0.0	0.5	0.5
Disallowed§	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

† $R_{merge} = \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections. ‡ $R_{cryst} = (|F_{obs}| - |F_{calc}|)/|F_{obs}|$; R_{free} was calculated in the same manner as R_{cryst} using a test set of data not used in the refinement process. § Nucleophile in catalytic triad.

acetate the k_{cat} values were all very similar to those of DLH-C1. By comparing the K_m values for DLH-C1 and DLH-G2-06, it appears that the introduction of the E36D mutation is responsible for the drop in K_m with both substrates, as well as a modest increase in k_{cat} with α-naphthyl acetate. The introduction of other mutations is responsible for smaller falls in K_m and slight increases in k_{cat}. For the best mutant, DLH-G3-16, with α-naphthyl acetate the K_m is 3.7-fold lower than that of DLH-C1 and k_{cat} is 1.6-fold higher, so that the specific constant, k_{cat}/K_m, is significantly higher (5.6-fold). This same mutant also gave the best increases with p-nitrophenyl acetate, with k_{cat}/K_m 3.6-fold higher than DLH-C1.

Given that the E36D caused a significant drop in K_m, it was decided to further investigate the nature of the change caused by the mutation. Two different mutants (E36N and E36A) were made to see if the charge of the residue or the length of the side chain were important for the rate enhancements. The kinetic properties of DLH-C4 (with the E36N mutation) show

that the k_{cat}/K_m value is only 1.4-fold higher than the wild type, with a only slightly better K_m than C123S, clearly much higher than the K_m for the DLH-G2-06 protein with only the C123S and E36D mutations. This result indicates that it is the negative charge on Glu36 that is necessary for the rate enhancement in DLH-G2-06. The results for DLH-C5 with the E36A mutation were more difficult to interpret. This mutant did not have significant activity, suggesting that it is the size of the side chain that is important for its activity. However, it also expressed very poorly and we cannot rule out the possibility that the E36A mutation causes the protein to become unstable.

The best mutant, DLH-G3-16, contained S208G. In wtDLH, Ser208 makes a stabilizing hydrogen bond to the side chain of Arg206, so that it can form an ion pair with Glu36 (Cheah, Austin *et al.*, 1993). It was thought that removal of the Arg206 side chain would give bulky substrates better access to the active site. The DLH-C6 mutant was generated by site-

directed mutagenesis and tested for activity with α-naphthyl acetate. From the data in Table 2, it is apparent that the removal of the Arg206 side chain caused a large increase in K_m and a modest increase in k_{cat}, with k_{cat}/K_m down to 66% of the value for the DLH-C1 mutant, indicating that the side chain of Arg206 is necessary for the lower K_m of the mutant enzymes.

The data given in Table 2 show that the mutants exhibit different kinetic properties with respect to the two substrates. In the hydrolysis of α-naphthyl acetate, the DLH mutants shows as much as a 7.5-fold increase in k_2 compared with the DLH-C1 mutant, while the increase in k_3 is more modest. In this case, it is the formation of the intermediate that has been enhanced by the mutation. For p-nitrophenyl acetate, the values for k_2 are below that of DLH-C1, while the values for k_3 are slightly higher than those of the DLH-C1 mutant. In this case, the rate enhancement appears to arise from a reduction in K_s, the initial binding event.

3.3. The structures of DLH mutants with and without PMS

X-ray crystallography was used to determine the structural changes arising from the introduction of mutations. It was thought that the structures of mutant proteins would shed light on how they might affect enzyme activity. The mutants selected consisted of DLH-C1, DLH-C6, DLH-G2-06, DLH-
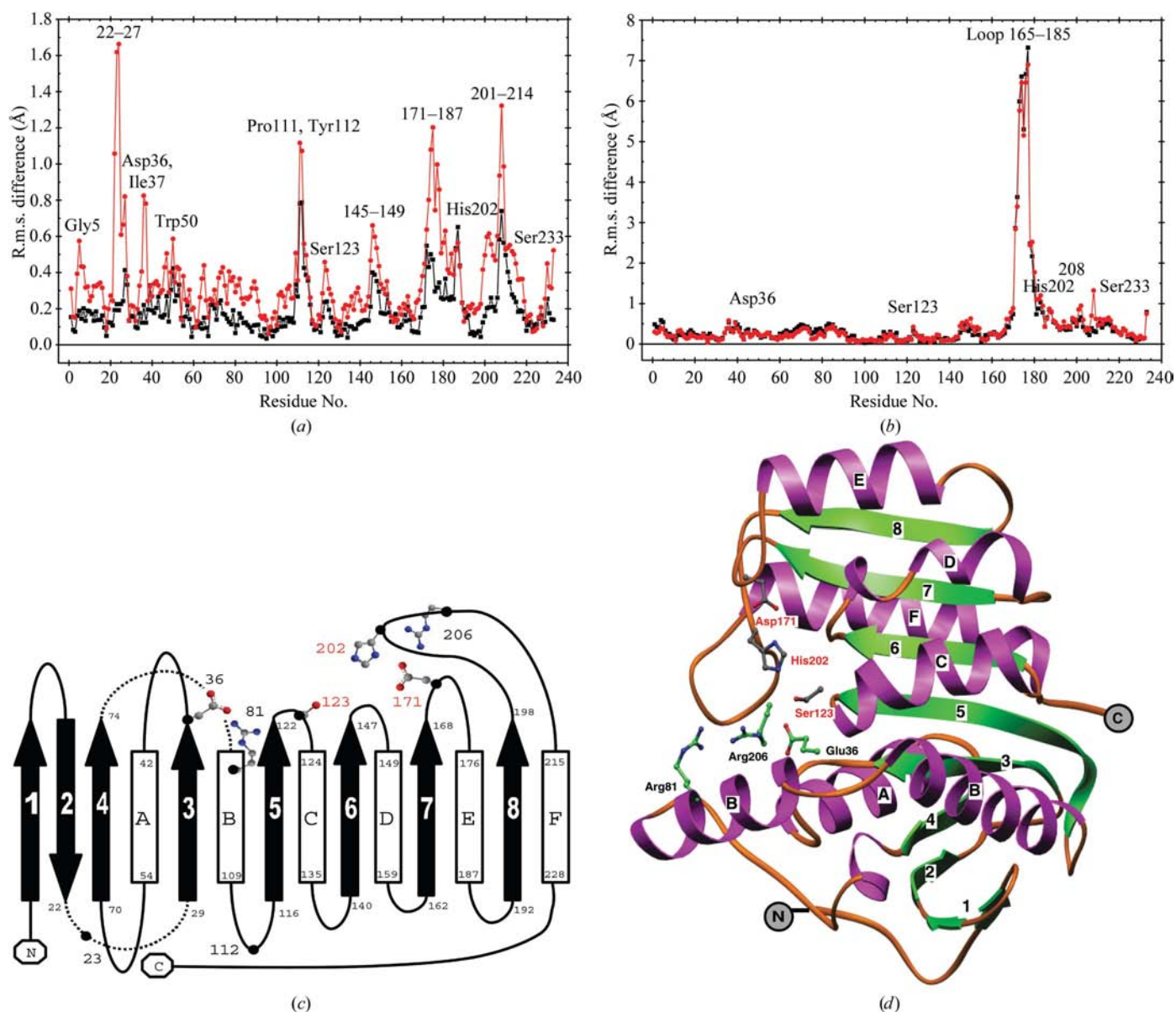

Figure 3
(a) R.m.s. distance difference plot showing the distance between main-chain atoms in DLH-C1 and the corresponding atoms in DLH-G2-06 (black squares) and DLH-G3-16 (red circles). (b) R.m.s. distance difference plot showing the distance between main-chain atoms in DLH-C1-PMS and the corresponding atoms in DLH-G2-06-PMS (black squares) and DLH-G3-16-PMS (red circles). (c) Two-dimensional schematic diagram of DLH showing secondary structure, active-site residues and the residues that are shifted in mutant proteins. The catalytic triad is shown in red. (d) Three-dimensional schematic diagram of DLH showing secondary structure, active-site residues and the residues that are shifted in mutant proteins. The catalytic triad is shown in red.

G3-16 and DLH-G4-113. DLH-G2-06 was selected to monitor the effect of the E36D mutation, while DLH-G3-16 was selected because it was the most active and contained the S208G mutation that was thought to be important for rate enhancement. DLH-G4-113 lacked the S208G mutation but had similar activity to DLH-G2-06 with few other mutations and was selected for this reason. The structure of DLH-C1 was chosen as the starting point for comparison. Although the structure of this protein was known, it was re-determined for the present work since the methods used to prepare crystals, collect data and refine the structures differed to those used

previously. However, the r.m.s. distance between the corresponding main-chain atoms obtained in the two studies was less than 0.5 Å, indicating that the structures are essentially the same. The structures were determined for the proteins with and without the inhibitor PMSF. All structures were determined with near-complete data sets which extended to at least 1.9 Å (Table 3). The final R factors and the refinement statistics are given in Table 3, which shows that the structures were well determined. The final structures were checked for stereochemical correctness with *PROCHECK* (Laskowski *et al.*, 1993) and *WHATCHECK* (Hooft *et al.*, 1996). A Ramachandran plot (Ramachandran & Sasisekharan, 1968) showed that over 91% of the non-glycine and non-proline residues were found in the most favoured regions of the plot, with between 5 and 8% in the additional allowed regions. In about half the mutants Tyr145 was assigned to the generously allowed region. As in all α/β-hydrolase enzymes (Ollis *et al.*, 1992), the nucleophile (Ser123) is at the apex of a sharp turn that is the only residue in a disallowed region. The average values of the B factors for all the main chains ranged from 11.8 to 18.6 Å^2, while those for the side chains ranged from 15.3 to 23.0 Å^2. As would be expected, the regions exhibiting high B factors are the more flexible parts of the molecule such as loop regions or surface residues.
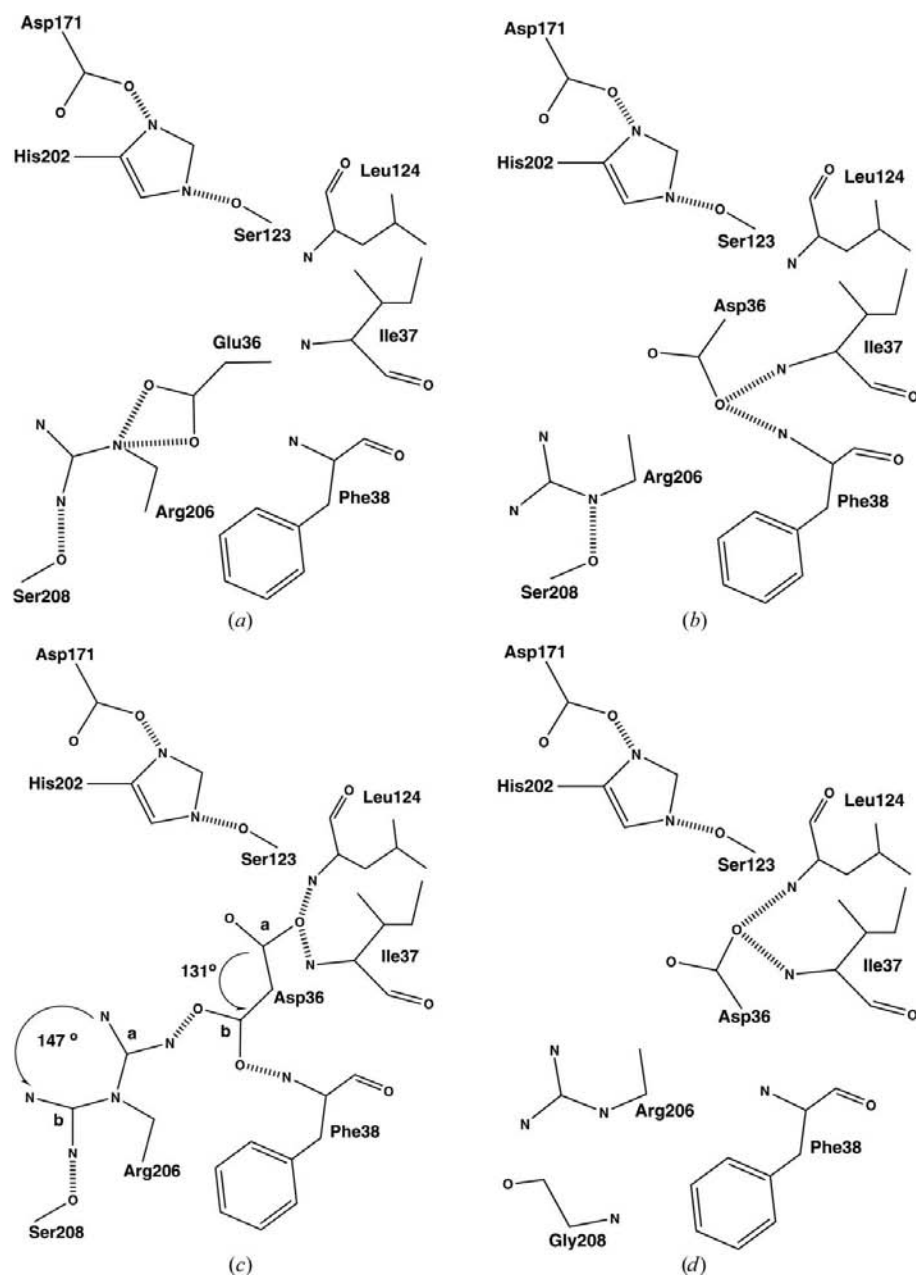
3.4. The structural changes for the protein in the absence of inhibitor

The overall structure of DLH-C1 and the mutant proteins are essentially the same in the absence of inhibitor. Fig. 3(a) shows the r.m.s. distance between the main-chain atoms of mutant proteins and the corresponding atoms of the DLH-C1 molecule. The difference between the DLH-C1 and DLH-G2-06 structures shows the effects of the E36D mutation. As can be seen from the plot, the nucleophile at the end of strand 5 is displaced by the E36D change. The loops at the end of strands 6, 7 and 8 are also displaced (Fig. 3a). In addition to the displacements described above, the comparison of DLH-C1 and DLH-G3-16 shows that loops distant from the active site undergo displacement during the evolutionary process. Changes at the C-terminus of the DLH-G3-16 mole-

Figure 4
Schematic diagrams showing the active sites of (*a*) DLH-C1 with only the C123S mutation, (*b*) DLH-G2-06 with the C123S and E36D mutations and (*c*) DLH-G4-113 with the C1232S, E36D, R105H, G211D and K234N mutations. Arrows indicate the relative orientations of discreetly disordered side chains Asp36 and Arg206. (*d*) DLH-G3-16 with the C123S, E36D, A134S, S208G, V229V and K234R mutations.

cule could explain shifts in other residues, especially residues around 23 and 111 (Figs. 3a and 3c). Both residues 23 and 111 are close to the C-terminus of the protein.

The active site of DLH has undergone significant change during the course of directed evolution. As mentioned above, the active-site nucleophile and the loops at the ends of strands 6, 7 and 8 are displaced owing to the E36D mutation (Fig. 3a). The loops at the end of strands 7 and 8 contain the catalytic triad residues; specifically, Asp171 is in the loop following strand 7 and His202 is at the end of strand 8 (Figs. 3c and 3d). The displacements caused by the E36D change occur in such a way that the integrity of the catalytic triad is preserved; that is, the hydrogen bonds linking the triad residues are retained. Other links have been lost and the relative positions of other residues have been altered by the E36D change. The residues that are displaced in the mutant proteins are those involved in substrate binding and activation of the nucleophile in the native protein, in particular Arg206 and Glu36. The structural changes associated with the introduction of mutations are depicted in Fig. 4. In DLH-C1 there is a salt link connecting the side chains of Glu36 and Arg206 that is modified in DLH-G2-06 owing to the E36D mutation. In DLH-G2-06 the orientation of the side chains of residue 36 has changed slightly so that it can still interact with Arg206. In the DLH-C1 and DLH-G2-06 proteins, there is also a hydrogen bond between Ser208 and Arg206 that stabilizes the side chain of Arg206. However, in DLH-G3-16 this link is absent because of the S208G change, so that Arg206 moves away from Asp36 and allows the side chain of this residue to rotate about its $C^{\alpha}-C^{\beta}$ bond by 147° and point towards Ser123. The negatively charged carboxyl group of Asp36 in the best mutant, DLH-G3-16, is stabilized by forming hydrogen bonds with the main-chain N atoms of Ile37 and Leu124 that forms part of the oxyanion hole in the DLH-C1 protein. In DLH-G3-16, the distance between Asp36 (OD1 and OD2) and Arg206 (NE) is 9.3 and 7.2 Å and the distance between Asp36 (OD1 and OD2) and Ser123 (OG) is ∼4 Å (3.7 and 4.1 Å). In the DLH-C1 mutant the distance between Glu36 (OE1 and OE2) and Arg206 (NE) is ∼3 Å (3.2 and 2.9 Å) and the distances between Glu36 (OE1 and OE2) and Ser123 (OG) are 7.0 and 5.8 Å. Interestingly, in the DLH-G4-113 protein the side chains of both Asp36 and Arg206 are discretely disordered; there are two conformations related by a rotation of 131° about the $C^{\alpha}-C^{\beta}$ bond for Asp36 and by a rotation of 147° about the NE—CD bond for Arg206. Asp36a is located in close proximity to Ser123 (∼3.5 Å) and Asp36a OD1 forms a hydrogen bond with Ile37 N (2.4 Å) and is also close to Ile124 N (3.5 Å). Asp36b faces towards Arg206a and Asp36b OD2 forms a hydrogen bond with Arg206a NH1 (3.0 Å). Arg206b NH1 forms a strong hydrogen bond with Ser208 OG. The structural data suggest that the conformations of the Asp36 side chain are in equilibrium and that this equilibrium is influenced by the proximity of Arg206. B-factor refinement was used to determine the occupancies of the two conformers. Temperature factors of 12.6 Å^2 for both conformers were obtained for occupancies of 0.45 (a) and 0.55 (b) in Asp36 and a temperature factor of 20.7 Å^2 for occupancies of

0.40 (a) and 0.60 (b) in Arg206, indicating that side chains of both Asp36 and Arg206 occupy both positions with approximately the same occupancy.

It was noted in the previous section that DLH-C6 with the C123S and R206A mutations had a low level of activity (Table 2) owing to a large increase in K_m. The structure of this protein revealed that the removal of the arginine side chain had little effect on the overall structure. In this mutant, the Glu36 side chain was in the same position as in DLH-C1 and did not undergo a rotation as observed in proteins with an aspartic acid at position 36. Even with the side chain of Arg206 removed, the side chain of Glu36 is too long to allow it to make a stabilizing link with the backbone of Ile37. In DLH-G2-13 with C123S, E36D and R206T mutations, one would expect that the change R206T would have a similar effect on the position of Asp36 as the S208G mutation. However, the activity of this protein is much lower than DLH-G3-16, which has Arg206 along with the S208G mutation. Although the side chain of Arg206 moves owing to the introduction of the S208G mutation, it has a beneficial effect on activity.

3.5. Inhibitor (PMSF) binding

To understand how mutations might affect the mode of substrate binding, the structures of DLH-C1, DLH-G2-06 and DLH-G3-16 were determined as a complex with PMS. The proteins all reacted with PMSF to form a complex in which Ser123 is covalently linked to the sulfur of the PMS group. The structure of the DLH-C1-PMS complex has been reported previously (Robinson et al., 2000), but the resolution of the structure was lower than that reported here and the crystal form differed from that used in the present study. Nevertheless, the overall structure of the complex as well as the position and orientation of the PMS group is the same as described in the previous report (Robinson et al., 2000).

The differences between DLH-C1 and its complex with PMS are relatively small and have, with one exception, been reported previously. The exception involved His202 of the catalytic triad. The positions of the catalytic triad residues in the complexes with PMS differed from those observed in the uninhibited protein. In the previous report of the C123S-PMS structure (Robinson et al., 2000) the imidazole ring of His202 was disordered, while in the present study it was found in good electron density. Compared with the uninhibited protein, the imidazole ring was rotated by 140° about its $C^{\alpha}-C^{\beta}$ bond and was stabilized in this position by a hydrogen bond with the Ser203 OG. The hydrogen bond between His202 and the third member of the catalytic triad, Asp171, was lost.

There were two significant differences between the PMS complexes with DLH-C1 and those mutants containing E36D. Firstly, the orientation of the PMS group in the DLH-C1 differed from that observed in other complexes containing the E36D mutation. Secondly, in mutant proteins containing E36D, the PMS caused the displacement of the loop following strand 7. Fig. 3(b) shows the magnitude of the displacements in this loop. The two orientations of the PMS and the displacement of the loop following strand 7 are illustrated in Fig. 5.

In DLH-C1-PMS, a collection of mainly hydrophobic residues surrounds the PMS. Within 4 Å of the PMS atoms are Phe38, Tyr85, Trp88, Phe173 and Ile37, along with the charged residues Asp36, Arg81 and Arg206. These same residues also encapsulate the PMS ring in complexes formed with the other mutants. In DLH-C1-PMS one of the O atoms of PMS, OD1, is close to Trp88, while OD2 forms stabilizing hydrogen bonds with Ile37 N (3.4 Å) that forms part of the oxyanion hole. These interactions are lost in the mutant proteins, primarily because of the E36D change. In these proteins, the side chain of Asp36 occupies the oxyanion hole and the PMS is rotated by about 180° so that OD1 of the PMS is close to His202, forming a hydrogen bond with His202 ND1 (2.9 Å for G3-16-PMS and 3.1 Å for G2-06-PMS). OD2 is close to Val147 and also stabilized by forming a hydrogen bond with a nearby water molecule [W11 (2.9 Å) for G3-16-PMS and W11 (3.0 Å) for G2-06-PMS]. A water molecule, W11 in G3-16-PMS and W11 in G2-06-PMS, is found hydrogen bonded to OD1 of the PMS moiety, Ser203 OG and Asp36 OD2. This water molecule could stabilize the oxyanion of the transition state of the acylation reaction. Alternatively, His202 could also carry out this function with a slight rotation of its side chain.

Apart from a change in the orientation of PMS, the mutant proteins also show a large shift in the position of a loop (Figs. 3b and 5). The loop consists of residues 165–185 and links strand 7 and helix E. It includes Asp171 that forms part of the catalytic triad and Phe173 that forms one side of the cavity for the physiological substrate. The C$^\alpha$ of Asp171 moves approximately 3 Å, while the side-chain atoms move between 7 and 9 Å. In its new position Asp171 can form a hydrogen bond with the imidazole of His172, which also base-stacks with the side chain of Phe173. In the mutant proteins, PMS causes Phe173 to be displaced by about 7 Å so that it sits on the periphery of the active site. Apart from these large changes, there are a number of smaller changes in the active site. For example, in the DLH-G3-16 mutant Arg206 is moved away from a close contact distance with the ring owing to the loss of hydrogen-bonding stabilization brought about by the S208G mutation.

The structures of the PMS complexes containing the E36D mutation revealed that the inhibitor and therefore presumably the substrate bind to the protein by displacing residues around Phe173. Attempts were made to engineer a substrate-binding site into mutant proteins so that the substrate would not need to displace the loop containing Phe173. These attempts resulted in the production of DLH-C7 with the F173A mutation. Although the protein appeared to be stable, this mutant had a very low activity (Table 1).

4. Discussion

Directed evolution was used to generate a number of mutant forms of DLH that had enhanced activity towards non-physiological substrates. The method used for selecting mutants has given results that can be explained, to some extent, by the mechanism employed by the enzyme. Like the serine proteases, the reactions catalyzed by DLH proceed *via*

the formation of a covalent intermediate. As can be seen from Table 2, the rate of formation of the intermediate is much greater than the rate at which it is broken down. The rate-limiting step for the overall reaction is the second step, the breakdown of the intermediate. If directed evolution is to produce mutants that have an increase in the overall rate at which substrate is broken down, then one would expect these mutants to show faster breakdown of the intermediate. This was not observed. In the case of α-naphthyl acetate, improvements in the rate of intermediate formation are observed, suggesting that the selection process was involved in this step. The selection process must therefore reflect the concentration of α-naphthol liberated during the formation of the intermediate. The initial screen used utilized clones that produced high levels of DLH with substrate levels much greater than that of the enzyme. The rapid formation of the intermediate produced sufficient α-naphthol to be readily detected. This is consistent with the selection process; clones were selected if they quickly produced a change in colour.

It has been possible to use structure analysis to understand the molecular basis for the rate enhancements produced by directed evolution. The initial round of evolution selected two mutations, C123S and E36D, from a library of active-site mutations. The first mutation was expected, while the second was not. Both mutations left the protein with an active catalytic triad. In the case of the E36D mutation, the close proximity of the nucleophile caused a shift in the position of the triad residues, but the necessary hydrogen bonds were retained between the residues. The E36D mutation also has a number of other effects. A salt link between Glu36 and Arg206 was lost when residue 36 became an aspartic acid. In subsequent rounds of evolution, a second mutation was found to have a positive effect on hydrolysis. This mutation, S208G, resulted in the loss of a hydrogen bond that stabilized Arg206, so that the side chain of Arg206 could move away from residue 36. The movement of Arg206 allowed the side chain of Asp36 to rotate around and form a hydrogen bond with what has been referred to as the 'oxyanion hole' of DLH. As noted in the last section, the side chain of Asp36 in the uninhibited enzyme appears to exist as an equilibrium between two conformations. The effect of the S208G mutation is to cause the side chain of Arg206 to move and this causes a shift in the equilibrium between the two conformations of Asp36.

The effects of mutations on the catalytic efficiency of the mutant proteins were probed by structural analysis of the complexes with PMS, which was intended to mimic the transition state of a simple ester with a large leaving group. In DLH-C1 (C123S mutation only) the sulfate oxygen OD1 of PMS forms a hydrogen bond with the backbone nitrogen of Ile37, mimicking the oxyanion in the oxyanion hole. The second oxygen, OD2, would take the place of a methyl group of the substrate and was found pointing towards Trp88. The imidazole of the His202 was displaced by C1 of the PMS moiety. In the hydrolysis of a peptide bond by a serine protease, the histidine would be required to protonate the leaving group, but in the case of substrates such as *p*-nitrophenyl acetate and α-naphthyl acetate, the pK_a of the leaving

groups is low so that protonation is not required. Along with the displacement of His202, there were relatively minor displacements of residues around Phe173, as shown in Fig. 3(a). The effects of the E36D mutation on the complex with PMS are dramatic; it caused the PMS to bind in a different orientation. The side chain of Asp36 blocks the oxyanion hole and forces the substrate to assume an orientation that can be described as a rotation of approximately 180° about the long axis of the substrate. In the structure of the proteins with the E36D mutation, it appears as if the sulfate O atoms of PMS displaces the imidazole of His202. In the uninhibited protein, the imidazole of His202 must be hydrogen bonded to the catalytic serine, but as this latter residue attacks the substrate to form a tetrahedral intermediate, the side chain of the histidine can be displaced. In its final location, the imidazole of His202 could form a hydrogen bond with sulfate oxygen OD1 of the PMS moiety. This oxygen also forms a hydrogen bond with a water molecule that is held in place by a hydrogen bond with Asp36 and the side chain of Ser203. In the hydrolysis of the substrate, this water molecule and the imidazole of His202 could stabilize the oxyanion as represented by sulfate oxygen OD1. The sulfate oxygen OD2 would represent the methyl group of acetate of
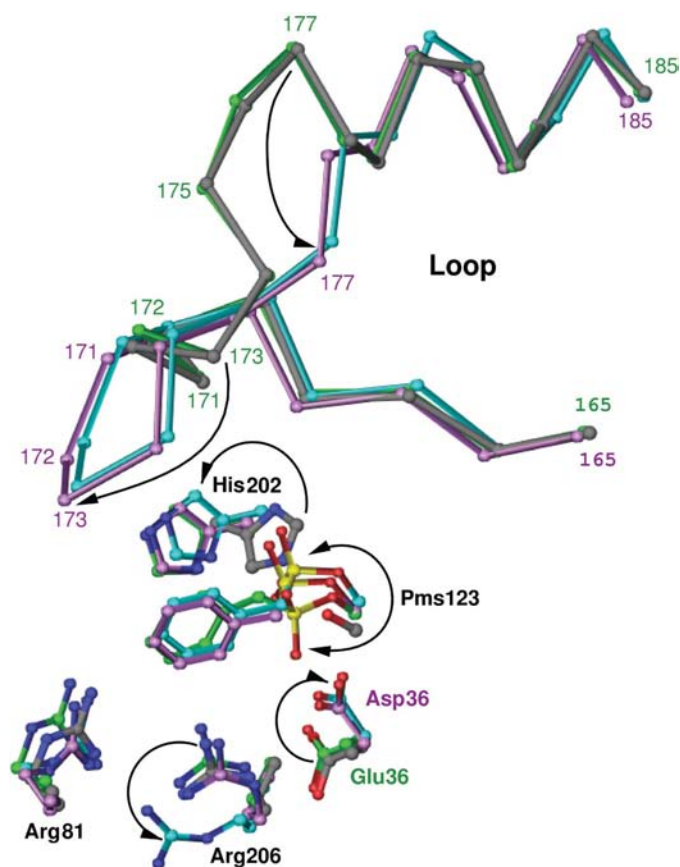
the substrate. In its present location, OD2 would make contact with side chain of Phe173 causing it to be displaced, as described in §3. On the breakdown of the tetrahedral intermediate, the acyl intermediate would probably occupy space initially occupied by Phe173. This intermediate could be broken down by base attack catalyzed by the imidazole of His202 or by the acid group of Asp36.

During the evolution of DLH in this study, the initial attack of the nucleophile required that the elements of the catalytic triad remain intact. The introduction of an aspartic acid at position 36 allowed the enzyme to bind non-physiological substrates in such a way that initial attack could proceed, but that once the transition-state intermediate had been formed, the active site could be deformed and the histidine used for purposes quite different to those carried out by the corresponding histidine found in serine proteases. Why has the enzyme evolved to bind the substrate in a significantly different manner? Two comments can be made. Firstly, the K_m values for the mutant enzymes are significantly lower than the starting protein, indicating that the substrate binds with a greater affinity. By moving His202, the substrate creates a cavity that can better accommodate its bulk. Secondly, the mutant proteins must be better able to stabilize the transition state during the formation of the intermediate. This can be achieved by hydrogen bonds that form with His202 as well as a water molecule bound through E36D. These residues are capable of stabilizing the anionic transition state; in essence, they are forming an 'oxyanion hole'.

The results obtained in this work were not predicted beforehand. We could not foresee which mutants would be selected, nor could we predict the effects of mutations that were obtained. Furthermore, attempts to improve enzyme performance by making site-specific mutants were not successful. However, directed evolution has yielded new enzymes with a novel variation on the reaction mechanism. The substrates used in this study are relatively easy to hydrolyze so that unnecessary steps in the reaction process, such as the protonation of the leaving group, could be bypassed. The selection process appears to have depended upon the first step of the reaction and no improvement in the slower steps in the reaction pathway could be detected. Although only a modest increase in the reaction rate has been obtained, it has been possible to arrive at a plausible explanation for this enhancement by using structure analysis.

Figure 5
Diagram of the refined structures showing the active site and the loop region (residues 165–185, C^α only) of DLH-C1 (grey bonds) DLH-C1-PMS (green bonds), G2-06-PMS (lavender bonds) and G3-16-PMS (cyan bonds). The structures of G2-06-PMS and G3-16-PMS were overlaid onto DLH-C1-PMS using all residues with the program *LSQKAB* (Kabsch, 1976).

References

Arnold, F. H. & Volkov, A. A. (1999). *Curr. Opin. Chem. Biol.* **3**, 54–59.

Beveridge, A. J. & Ollis, D. L. (1995). *Protein Eng.* **8**, 135–142.

Bolon, D. N. & Mayo, S. L. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 14274–14279.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cheah, E., Ashley, G. W., Gary, J. & Ollis, D. (1993). *Proteins*, **16**, 64–78.

Cheah, E., Austin, C., Ashley, G. W. & Ollis, D. (1993). *Protein Eng.* **6**, 575–583.

Dalby, P. A. (2003). *Curr. Opin. Struct. Biol.* **13**, 500–505.

Fersht, A. (1999). *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, pp. 143–146. New York: W. H. Freeman.

Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.

Hanahan, D. (1983). *J. Mol. Biol.* **166**, 557–580.

Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.

Jeffries, C. M. (1998). BSc (Hon) thesis. The Australian National University, Australia.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Mastropaolo, W. & Yourno, J. (1981). *Anal. Biochem.* **115**, 188–193.

Michaelis, L. & Menten, M. L. (1913). *Biochem. Z.* **49**, 333–369.

Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. & Goldman, A. (1992). *Protein Eng.* **5**, 197–211.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pathak, D., Ashley, G. & Ollis, D. (1991). *Proteins*, **9**, 267–279.

Pathak, D., Ngai, K. L. & Ollis, D. (1988). *J. Mol. Biol.* **204**, 435–445.

Pathak, D. & Ollis, D. (1990). *J. Mol. Biol.* **214**, 497–525.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.

Read, R. J. & Moult, J. (1992). *Acta Cryst.* A**48**, 104–113.

Robinson, A., Edwards, K. J., Carr, P. D., Barton, J. D., Ewart, G. D. & Ollis, D. L. (2000). *Acta Cryst.* D**56**, 1376–1384.

Schmidt, E. & Knackmuss, H. J. (1980). *Biochem. J.* **192**, 339–347.

Stemmer, W. P. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 10747–10751.

Walker, I. (2001). PhD thesis. The Australian National University, Australia.

Walker, I., Easton, C. J. & Ollis, D. L. (2000). *Chem. Commun.*, pp. 671–672.

Yang, H., Carr, P. D., McLoughlin, S. Y., Liu, J. W., Horne, I., Qiu, X., Jeffries, C. M., Russell, R. J., Oakeshott, J. G. & Ollis, D. L. (2003). *Protein Eng.* **16**, 135–145.