Mutagenesis and crystallographic studies of the catalytic residues of the papain family protease bleomycin hydrolase: new insights into active-site structure

Paul A. O'FARRELL¹ and Leemor JOSHUA-TOR²

W.M. Keck Structural Biology Laboratory, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, U.S.A.

Bleomycin hydrolase (BH) is a hexameric papain family cysteine protease which is involved in preparing peptides for antigen presentation and has been implicated in tumour cell resistance to bleomycin chemotherapy. Structures of active-site mutants of yeast BH yielded unexpected results. Replacement of the active-site asparagine with alanine, valine or leucine results in the destabilization of the histidine side chain, demonstrating unambiguously the role of the asparagine residue in correctly positioning the histidine for catalysis. Replacement of the histidine with alanine or leucine destabilizes the asparagine position, indicating a delicate arrangement of the active-site residues. In all of the mutants, the C-terminus of the protein, which lies in the active site, protrudes

further into the active site. All mutants were compromised in their catalytic activity. The structures also revealed the importance of a tightly bound water molecule which stabilizes a loop near the active site and which is conserved throughout the papain family. It is displaced in a number of the mutants, causing destabilization of this loop and a nearby loop, resulting in a large movement of the active-site cysteine. The results imply that this water molecule plays a key structural role in this family of enzymes.

Key words: bleomycin hydrolase, conserved water, crystal structure, cysteine protease, mutagenesis, papain.

INTRODUCTION

Bleomycin hydrolase (BH) is an evolutionarily conserved aminopeptidase that is involved in antigen presentation and bleomycin chemotherapy resistance, and has been implicated as a risk factor in Alzheimer's disease [1–4]. Recently, it has been shown to have a possible protective function against homocysteine toxicity [5]. The crystal structures of the yeast and human forms of BH have been determined, showing the protein to be a homohexamer with a central channel into which the six active sites open [6,7]. Each monomer consists of a protease core, augmented by two other domains that are involved in the oligomerization of the enzyme. The protease core itself has a structure typical for C1 proteases. The configuration of the active-site residues is identical with that of papain. The α -carbon atoms of the catalytic triad residues cysteine, histidine and asparagine, and of the glutamine which stabilizes the oxyanion intermediate, fall within 0.15 Å (1 Å = 0.1 nm) rmsd (root mean square deviation) of their papain counterparts (Figure 1). An unusual feature of the BH active site is the presence of the protein's C-terminus, which intrudes into the active-site cleft and occupies the S2–S4 substrate-binding subsites. This architectural feature converts the core papain-type endopeptidase into an aminopeptidase by restricting substrate entry to one residue before the point of cleavage [8]. Subsequently, a similar role to that of the BH C-terminus has been shown for the mini-chain of cathepsin H [9,10] and for the exclusion domain of cathepsin C [11].

Previously, we have mutated the active-site cysteine to alanine, in the yBH (yeast BH) [8], and to serine, in the hBH (human BH) [7]. These mutations abolished the enzymatic activity as expected, but also demonstrated the flexibility of the C-terminus of the pro-

tein. In both of these mutants, the C-terminus rotates around Gly-450 (451 in hBH) and extends further into the active site, placing the terminal alanine residue more-or-less in the S1 subsite, which would thus prevent substrate binding and hydrolysis. It appears that the presence of the cysteine is essential to prevent this movement. It should be noted that some movement of the C-terminus is necessary to allow the autocleavage reaction that removes one gene-encoded residue that is not found in the mature protein [6,8]. Other than this movement of the C-terminus into an intrusive conformation, no significant structural changes were seen in these mutants.

Initially, in an effort to produce a stable enzyme–substrate complex, we mutated the active-site histidine, asparagine and glutamine residues of BH. Because these mutations were expected to severely abridge the catalytic activity, we also deleted the gene-encoded C-terminal residue, which would be removed autocatalytically in the intact enzyme. In order to confirm an intact active site, the mutant proteins were crystallized and their structures analysed. Kinetic parameters were also determined for each of these mutants. To our surprise, these mutations resulted in dramatic alterations in the active-site structures. Probing these effects allowed a much clearer understanding of the role of the active-site residues as well as of an important water molecule in the active-site architecture of this important and ubiquitous family of cysteine proteases.

MATERIALS AND METHODS

Mutagenesis

Mutagenesis was carried out using the QuikChange[®] mutagenesis system from Stratagene. Mutagenic oligonucleotides were

Abbreviations used: ASU, asymmetric unit; BH, bleomycin hydrolase; hBH, human BH; yBH, yeast BH; DTT, dithiothreitol; E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; rmsd, root mean square deviation.

¹ Present address: Royal College of Surgeons in Ireland – Medical University of Bahrain, PO Box 15503, Adliya, Kingdom of Bahrain.

² To whom correspondence should be addressed (email leemor@cshl.edu).

Atomic co-ordinates and structure factors have been deposited in the PDB (Protein Data Bank) (accession codes are given for the following mutants: N392A: 2DZY; N392V: 2DZZ; N392L: 2E00; H369A: 2E01; H369L: 2E02; and Q67E: 2E03).

Figure 1 Catalytic sites of yBH and papain

yBH (PDB entry 1GCB) is shown in coil representation in green, with the side chains of the catalytic triad and Gln-67 shown in ball-and-stick representation, with nitrogen atoms in blue, oxygen in red and sulfur in yellow. Papain (PDB entry 1PPN [25]) is shown in coil representation in brown, with side chains also in brown. Residues shown are 65–75, 367–371 and 391–395 for yBH, and 17–27, 157–161 and 174–178 for papain. Residues are labelled near the C α atom in yBH/papain numbering. The alignment was based on the C α atoms of the catalytic triad and the glutamine residue.

manufactured by Operon or Sigma-GenoSys. Mutagenesis was carried out in two steps. First, the codon for the genetically encoded C-terminal residue (Lys-454), which would normally be removed autocatalytically, was mutated to a stop codon. The resultant construct was then used as a template for the second step in which the targeted active-site residue was changed. All mutant genes were fully sequenced on both strands to confirm the mutation and to ensure no spurious changes had occurred during the procedure.

Protein expression

Mutant yBH was produced as previously described for the human protein [7]. Briefly, the proteins were expressed as His₆-tagged constructs in *Escherichia coli* BL21. Cells were grown at 37 °C to mid-exponential phase and induced with 400 μ M IPTG (isopropyl β -D-thiogalactoside) for 5 h at room temperature (22 °C). The cells were harvested by centrifugation and stored frozen at –70 °C. After thawing, the cells were disrupted by sonication and the protein was purified by a combination of nickel-affinity, ion-exchange and size-exclusion chromatographies. The protein was concentrated by ultrafiltration to 25–50 mg/ml and stored at 4 °C.

Protease assays

Protease assays were carried out essentially as previously described [7], using L-arginine-7-amido-4-methylcoumarin (Sigma) as the fluorogenic substrate. In the assays, BH (10–1000 pmol) was incubated in reaction buffer [50 mM KH₂PO₄, 50 mM Na₂HPO₄, 2 mM EDTA, and 10 mM DTT (dithiothreitol), pH 7.5] at 30 °C. Each sample was contained within a final volume of 200 μ l in black-walled 96-well microtitre plates. In the standard assay, reactions were initiated through the addition of substrate to a final concentration of 0.125 mM and were monitored for a period of 3 h. Points from the linear portion of the reaction curve were used for analysis. For determination of kinetic constants a range of substrate concentrations from 0 to 0.125 mM was used. Initial rates were determined from the linear portion of the reaction curve and $V_{\max}$ and K_m were calculated by curve-fitting to the Michaelis–Menten equation. Active enzyme concentration was determined

by stoichiometric titration with E-64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; proteinase inhibitor] [12]. Because of the low activity of the mutant proteins, pre-incubation with E-64 was carried out overnight to ensure that the reaction went to completion. K_{cat} was determined from the enzyme concentration and $V_{\max}$.

Crystallization and data collection

Crystallization was carried out at 17 °C by the hanging-drop vapour-diffusion method. Drops were prepared by mixing 1 μ l of reservoir solution with 1 μ l of protein solution [12 mg/ml in 25 mM Tris/HCl, pH 8.5, 1 mM DTT and 10 % (w/v) glycerol]. BH mutants crystallized from 14–20 % (w/v) PEG [poly(ethylene glycol)] 4000 and 100 mM Tris/HCl (pH 8.5). All of the mutant proteins crystallized in space group P6₃22 with one polypeptide in the ASU (asymmetric unit) generating the biological hexamer through crystal symmetry.

Data were collected on beamline X26C at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory (BNL). Crystals were soaked in cryoprotectant solution [reservoir solution plus 30 % (w/v) ethylene glycol], mounted on loops and frozen by plunging directly into liquid nitrogen. Data were collected from frozen crystals at 95 K. The reflection intensities were integrated and scaled using the HKL suite of programs [13]. Each data set was collected from a single crystal.

Refinement

Initial phases were generated using the wild-type protein as model. Refinement was carried out with the program CNS [14] with a maximum-likelihood target function [15] and using a combination of simulated annealing, minimization and individual *B*-factor refinement. Model building was performed using the program O [16]. The electron density maps are essentially complete except where noted in the text for the 60s loop (residues 61–71). For all mutants, the 380s loop (residues 380–384) was poorly ordered and atomic positions could not be accurately determined, as was the case for this loop in the wild-type protein. In addition, a number of surface lysine and arginine side chains were poorly ordered. Data collection and refinement statistics are given in Table 1.

Structural alignments

The wild-type and mutant yBH structures were aligned using the C α atoms of residues 2–49, 101–135, 251–265, and 405–441. These were the residues that were least affected by the mutations. Alignments were carried out using the program LSQMAN [17].

Figures

Figures 1–4 were prepared with the programs RASTER3D and BOBSCRIPT [18,19].

RESULTS

The structures of six mutants of yBH (N392A, N392V, N392L, H369A, H369L and Q67E) were determined in the present study. The mutants crystallized in space group P6₃22 with one monomer in the ASU, and with the biological hexamer being generated by space group symmetry. All mutants formed the hexamer in solution, as demonstrated by size-exclusion chromatography. The overall structures of the mutant proteins are very similar to that of the wild-type, with an rmsd of all C α atoms of 0.5–0.8 Å, as calculated for the monomer. There are, however, some major

Table 1 Data collection and refinement statistics

Data were truncated to the resolution where $I/\sigma I$ is at least 2.0 and R_{sym} is better than 50 %.

	N392A	N392V	N392L	H369A	H369L	Q67E
Resolution (final shell)	50–2.57 (2.66–2.57)	50–2.15 (2.23–2.15)	50–2.0 (2.07–2.0)	50–1.73 (1.79–1.73)	50–2.2 (2.28–2.2)	50–2.13 (2.21–2.13)
Completeness %	99.9 (100)	88.2 (57.1)*	96.7 (77.6)	89.0 (59.2)*	86.8 (43.3)*	99.0 (95.2)
R_{sym}	14.1 (48.7)	10.1 (43.5)	10.7 (43.6)	5.8 (43.3)	18.3 (38.0)	9.6 (47.3)
Redundancy	9.1 (9.1)	17.2 (4.5)	11.1 (6.8)	9.5 (2.6)	10.0 (1.5)	10.4 (5.0)
Unique reflection	19700	29770	39821	56353	27215	34254
Total reflection	178643	512277	442314	533382	271832	356973
$I/\sigma I$	41.5 (13.9)	40.5 (3.0)	46.6 (6.1)	45.8 (4.1)	18.9 (3.1)	45.3 (4.3)
Space group	P6 ₃ 22	P6 ₃ 22	P6 ₃ 22	P6 ₃ 22	P6 ₃ 22	P6 ₃ 22
Cell a=b, c	151.0, 89.7	152.5, 89.4	150.9, 89.8	151.8, 89.3	152.6, 89.0	152.3, 89.3
$R_{\text{cryst}}/R_{\text{free}}$	19.8/24.8	19.0/23.2	21.0/24.5	18.9/22.4	22.7/26.0	20.6/25.6
Protein atoms	3677	3653	3689	3639	3593	3616
Water	187	258	259	321	122	174
Average B	26.8	33.2	28.7	27.6	34.1	31.5

* Completeness is low due to detector limitations, not diffraction limit of crystal.

Table 2 Summary of structural and catalytic effects of mutations

+, Observed; –, not observed; n.a., not applicable; n.d., not determined.

	Histidine position compromised	Cysteine position changed	Asparagine position changed	390s core water displaced	390s loop elevated B -factors	60s loop disordered	C-terminus intrusion	Relative $k_{\text{cat}}/K_{\text{m}}$ (mut/wt)
yN392A	+	–	n.a.	–	–	–	+	0.0183
yN392V	+	+	n.a.	+	+	+	+	0.0053
yN392L	+	–	n.a.	–	–	–	+	0.0052
yH369A	n.a.	+	+	+	–	+	+	0.0001
yH369L	n.a.	+	+	+	+	+	+	n.d.*
yQ67E	+	+	+	+	+	+	+	0.0004

* Preliminary results indicated a severe abridgement of catalytic activity.

structural changes at the active site, as discussed below. When the atoms of the most changed portions are omitted (see the Materials and methods section), the rmsd falls to 0.3–0.5 Å.

Asn-392

The active-site asparagine residue in C1 cysteine proteases is completely conserved [20], and has long been considered to be part of an active-site triad. However, mechanistic studies have shown that this class of enzyme functions with a catalytic dyad (cysteine and histidine), with the asparagine not being directly involved in the catalysis. The observation that the asparagine forms a hydrogen bond with the catalytic histidine led to the suggestion that its role is to correctly orient the histidine for the various steps of the hydrolysis reaction [21,22]. Mutagenesis experiments in papain have shown that replacing this residue with any other reduces the catalytic activity of the enzyme [23,24]. Those studies, however, did not include determination of the structures of the mutant enzymes, although thermal stability studies suggested that some structural instability was introduced by the mutations.

We have mutated the active-site Asn-392 of yBH to leucine, valine and alanine (N392L, N392V and N392A respectively) and determined the crystal structures of the mutant proteins. The replacement residues were chosen to test the effect of removing the hydrogen bond, and of progressively decreasing the length of the side chain. The enzyme activity of all of the mutants was drastically reduced (see Table 2). There was also a pronounced destabilization of the position of the catalytic His-369 in all cases.

Omit maps of the alanine and valine mutants clearly show density for two conformations of the imidazole side chain of His-369 (Figure 2A). Conformation 1 is similar to that observed in the wild-type protein, pointing more-or-less towards the mutated 392 position. However, conformation 2 involves a rotation about the C α –C β bond of approx. 78°, placing the histidine side chain in what is its most commonly observed rotamer. In the case of the N392L mutant, there is a somewhat lower contribution of conformation 2. It should also be noted that for conformation 1 of N392L, there is a rotation of the imidazole side chain around the C β –C γ bond of approx. 35° compared with the wild type. We have estimated the relative occupancies of conformations 1 and 2 for each mutant as 0.5/0.5, 0.45/0.55 and 0.6/0.4 for N392A, N392V and N392L respectively.

Unexpectedly, structural changes in the N392V mutant were not restricted to the histidine. In fact, an examination of the N392V mutant shows dramatic structural changes at and beyond the active site (Figure 2B). The catalytic cysteine rotates by approx. 137° with the sulfur moving by 4.3 Å and the C α by 0.9 Å. This results in the thiol of the catalytic Cys-73 pointing towards the 390s loop (residues 391–403) and places it in a position to interact with Ne2 of conformation 1 of the catalytic histidine, rather than N δ 1. The 390s loop becomes somewhat disordered. The position and conformation of the loop are much the same as in the wild-type protein; however, the B -factors increase considerably, and the electron density becomes significantly weaker (Figure 2C). In addition, a water molecule in the centre of this loop is displaced (Figure 2B). This water molecule (wat-512 in wild-type yBH, PDB entry 1GCB [6]), which makes a number of hydrogen bonds

Figure 2 Structural changes at the active site

(A) The position of the catalytic His-369 is compromised by the removal of Asn-392. Simulated-annealed omit maps were generated with the program CNS [14], omitting His-369 and a surrounding sphere of 2.5 Å. The maps are shown contoured at 3 σ around the final model for each protein. Two conformations are visible for the histidine side chain in each of the mutant proteins. Conformation 1 is the most similar to that of the wild-type protein and is shown in green. Conformation 2 is shown in blue. The occupancies for conformations 1/2 were estimated as 0.5/0.5, 0.45/0.55 and 0.6/0.4 for the N392A, N392V and N392L mutants respectively. (B) Comparison of wild-type and N392V mutant yBH. Wild-type yBH is shown in green and N392V in pink, in coil representation. Only residues 58–75, 367–371, 391–402 and 451–453 are shown. The side chains of residues 73, 369, 392 and 67 (not visible in N392V) are shown, with nitrogen in blue, oxygen in red and sulfur in yellow. Both conformations are shown for His-369 of N392V. The C-terminal oxygens and carboxyl carbon are also shown. Residues are labelled near the α -carbon. A dotted line represents residues 61–71 of N392V, which are not visible in the electron density. The core water molecule of the 390s loop in the wild-type protein is shown as a red sphere. Semi-transparent CPK (Corey–Pauling–Koltun) spheres show the steric clash between the wild-type water molecule and the C γ 2 atom of Val-392 in the N392V mutant. (C) Elevated *B*-factors in the 390s loop. The 390s loops for wild-type and mutant yBHs are shown in coil representation, coloured by *B*-factor from blue at or below 40 Å² to red at or above 70 Å². The loops of H369L, N392V and Q67E show highly elevated *B*-factors compared with the wild-type. The *B*-factors of the loops of the other mutants are only slightly elevated. The overall *B*-factors of all structures are comparable (Table 1). (D) Effect of H369A mutation on asparagine conformation and 390s loop. The view and representation are similar to Figure 1. The mutated histidine residue is labelled in *italic*. A dotted line represents residues 65–71, which are not visible in the electron density.

to backbone groups of the 390s loop, is conserved in all papain-type proteases examined and will be considered further in the Discussion section. Even more surprising is that the 60s loop (residues 61–71), which leads to the catalytic cysteine, becomes very disordered. Omit maps show only a few disconnected pockets of electron density along the chain's expected route and the chain cannot be traced. The compromised position of the active-site cysteine, described above, is also probably due to the disorder of

the 60s loop. Although Cys-73 does not interact directly with the loop, a stable 60s loop appears to be important for its proper positioning. For example, Arg-72, directly preceding Cys-73, interacts with Gln-67 and Ser-70, both of which are disordered in the N392V mutant. Removing these interactions is likely to affect the positioning of Cys-73 as well.

The consequences of the alanine and leucine mutations are different from those of the valine mutation. Although the histidine

position is compromised, the large movement of Cys-73 and the destabilization of the 60s and 390s loops seen in the N392V mutant do not occur. The *B*-factors of the 390s loop are similar to those of the wild-type (see Figure 2C) and the 60s loop is clearly defined.

His-369

His-369 is an essential residue, which forms the catalytic thiolate–imidazolium pair with Cys-73, and which hydrogen-bonds with Asn-392. To examine the structural effects of removing these interactions, we replaced the histidine with alanine; to examine these effects while maintaining some steric bulk at this position, we also replaced it with leucine. The crystal structures of both mutants were determined. As expected, in both mutants the catalytic activity was severely curtailed (Table 2).

In both cases, we see similar overall structural changes to those observed with the N392V mutant. The 60s loop becomes completely disordered, and the position of the catalytic cysteine is altered as before, now pointing in the direction of the 390s loop (Figure 2D). In H369A, the loss of the hydrogen bond with the cysteine allows the side chain of Asn-392 to swing away from its wild-type position. It falls into the centre of the 390s loop, displacing the central water molecule of that loop, and changing the hydrogen-bonding network. O δ 1 of the asparagine forms hydrogen bonds with the amide nitrogen of Trp-394, and with the amide of Ser-393. In addition, the carbonyl oxygen of Gly-395, which interacts with the water in the wild-type protein, now receives a hydrogen bond from the amide nitrogen of Ser-398 (Figure 2D). Although the geometry of the 390s loop is distorted in H369A, the structure is fairly stable, with *B*-factors comparable with those of the wild-type (Figure 2C). In H369L, there are two conformations for Asn-392. One points towards Met-371 where it could hydrogen-bond to the δ -sulfur atom, while the other is in a similar position to that found in H369A. In this mutant, the 390s loop is not very well ordered, and has elevated *B*-factors (Figure 2C).

Gln-67

Gln-67 is a completely conserved catalytic site residue, although it is not considered to be part of the so-called active-site triad. Its side chain forms part of the so-called 'oxyanion hole', whose role is to stabilize the anionic intermediate that is transiently formed during catalysis. Gln-67 interacts with the backbone carbonyl of Ser-70 in the 60s loop, and with the side chain of Ser-393 in the 390s loop. In the yBH C73S and hBH C73A mutants described in our earlier studies, the observed intrusion of the C-terminus into the active site allows its carboxylate oxygens to interact with Gln-67 [7,8]. We have replaced Gln-67 with glutamate to assess whether the presence of a negative charge at this position, effectively abolishing the oxyanion hole, would prevent the intrusion of the C-terminus. However, the crystal structure of the mutant shows that this is not the case. The C-terminus does intrude, making similar contacts to those described below for N392V and the His-369 mutants. We also see many of the same changes that are apparent in those mutants. Cys-73 is displaced, the 60s loop is disordered and *B*-factors increase noticeably for the 390s loop (Figure 3C). Asn-392 is seen in two conformations, one similar to the wild-type conformation, and the other similar to that observed in yBH H369A. The catalytic histidine itself also appears in two conformations. It seems that the negatively charged glutamate cannot be accommodated in the same manner as a glutamine at this position, and that this causes considerable disruption to the local structure. It is likely that this general disruption of the

structure is responsible for the large loss in enzymatic activity observed with this mutant (Table 2).

The C-terminus

In all of the structures examined in the present study the C-terminus intrudes into the active site, in a manner similar to that observed with mutants of the catalytic cysteine described in our earlier studies [7,8]. This intrusion brings the C α of the terminal Ala-453 approx. 3.2 Å deeper into the active site than in the wild-type protein. However, the mutants in the present study differ somewhat from the cysteine mutants described earlier, and they fall into two classes with regard to the C-terminus: those mutants in which the catalytic cysteine changes position (N392V, H369A, H369L and Q67E), and those where it does not (N392A and N392L). N392V will be described as an example of the first class as all of these mutants behave similarly. In N392V, the carboxylate carbon of Ala-453 is shifted by 1.35 Å compared with C73A (Figure 3A). Its terminal oxygens interact with the amide nitrogens of Ala-370 and Cys-73 (Figure 3D). In C73A, on the other hand, the terminal oxygens interact with the side chain of Gln-67, the backbone amide of C73 and N δ 1 of His-369 [8] (Figure 3C). This shift probably occurs because of the movement of Cys-73, which places its amide nitrogen in a different position, and because of the high mobility of Gln-67, which would destabilize an interaction with the C-terminus such as that observed with the C73A mutant.

In the second class, N392A and N392L, the catalytic cysteine does not change position. In these mutants, no electron density can be seen for the terminal alanine residue. It is clear, however, that the C-terminus is in the intrusive conformation. Density for the side chain of the penultimate residue, Leu-452, definitively places it in the same position that it occupies in C73A and in N392V, and the side chain of Met-367, which moves to accommodate the repositioned Leu-452 in C73A and N392V (see Figure 3), also moves in N392A and N392L. We considered the possibility that an autocleavage reaction in these mutants had removed the alanine in a manner similar to the deletion of the genetically encoded Lys-454 in the wild-type enzyme. However, MS of trypsin digests of N392A and N392L clearly show a peptide terminating in Ala-453, with no evidence for a truncated peptide with Leu-452 as the terminus (results not shown). The lack of electron density for Ala-453 in these mutants must therefore be due to a high degree of mobility for this residue. A steric clash with the thiol of Cys-73 does not allow Ala-453 of N392A or N392L to form the stabilizing interactions with Gln-67 and N δ 1 of His-369 that occur in C73A, which may explain the mobility of the terminal alanine. Such a clash does not occur in those mutants described above in which there is a large movement of the cysteine.

DISCUSSION

In the present study, we have demonstrated unambiguously the important role of the hydrogen bond with the active-site asparagine in maintaining the catalytic histidine in the correct position for catalysis. Although previous mutagenesis work in papain has shown the asparagine to be important for enzyme activity [23,24], to our knowledge the crystal structures presented here clearly demonstrate the instability of the histidine position in the absence of the stabilizing hydrogen bond for the first time. The fact that the alternative conformation of the histidine side chain is the most commonly observed rotamer suggests that the hydrogen bond is necessary to prevent the histidine from adopting what would otherwise be a thermodynamically more favourable conformation.

Figure 3 C-terminal of wild-type and mutant yBH

(A) The Figure shows an overlay in coil representation of wild-type, C73A and N392V C-terminal interactions in green, yellow and pink respectively. Residues represented are 65–76, 140–145, 366–374 and 449–453. The side chains of Gln-67 (not visible in N392V), Cys-73, His-369 (two conformations in N392V), Ala-453, Leu-452, Ala-451 and Met-367 are shown as bonds, with nitrogen, oxygen and sulfur in blue, red and yellow respectively. The terminal oxygens and carboxyl carbon are also shown. In (A, D), residues 65–71 of N392V, which are not visible in the electron density, are shown as a dashed pink coil. The overlay in (A) shows the intrusion of the C-terminus in the mutant proteins and the different positions adopted by the terminal alanine in C73S and N392V. The potential steric clash between Leu-452 of the mutant proteins and Met-367 of the wild-type protein is highlighted with a red circle. (B–D) The Figures show the interactions that stabilize the C-terminus in each structure. Hydrogen bonds are shown as dashed black lines. The structures were aligned using the C α atoms of residues 2–49, 101–135, 251–265, and 405–441.

The mutant proteins examined in the current study revealed the importance of a previously unnoted conserved water molecule in maintaining the integrity of the active-site structure (wat-512 in wild-type yBH, PDB entry 1GCB [6]). This water molecule is conserved in the papain fold (in papain, PDB entry 1PPN [25], it is wat-222). A survey of ten non-redundant papain-type proteases in the PDB with a resolution better than 2.0 Å shows this water to be present in all cases [PDB IDs: 1CS8, 1IWD, 1M6D, 1ME4, 1MEM, 1NQC, 1OOE, 1PPN, 1YAL and 2ACT]. In fact, it is ranked in the lowest 5% of water molecules with respect to *B*-factor in each of these structures, indicating that it is well ordered and most likely quite tightly bound. Thus it has all the characteristics of a 'structural' water molecule [26]: it is buried, without access to the bulk solvent, its hydrogen bonds are to backbone groups, it is conserved in other proteins of the same fold, and it has a low *B*-factor. In yBH, wat-512 is bounded by the 390s loop which contains the active-site Asn-392. It makes extensive hydrogen-bonding interactions that stabilize this loop, thus forming a structural core for this part of the molecule (Figure 4). The 390s loop itself interacts with and appears to position the 60s loop (residues 61–71). This is an important loop, as it leads directly to the catalytic cysteine at position 73 and contains Gln-67, which stabilizes the oxyanion intermediate. There is also a direct hydrogen bond from the γ -oxygen of Ser-393, in the 390s

loop, to O δ 1 of Gln-67. This stabilizing network of hydrogen bonding is conserved throughout the C1 family of peptidases.

The unexpected structural changes seen with some of the mutants can be explained when we consider the effect on this core water molecule. An intriguing observation is that minimal structural changes occurred when Asn-392 was mutated to alanine or leucine, while quite drastic changes were observed when this residue was changed to valine. When the structures are examined more closely, we note that in N392A and N392L, the core water molecule of the 390s loop is intact, while in N392V it is displaced. When the structures of N392V and the wild-type protein are overlaid, it can be seen that the C γ 2 of the substituted valine falls less than 2.3 Å from the water position, and thus the water is ejected (Figure 2B). This ejection only occurs with the β -branched valine. The shorter alanine side chain does not reach the water position, while the leucine side chain does not branch at the β position and is able to assume a conformation that does not impinge on the water. This ability of β -branched amino acids to block the approach of water to the protein backbone has been observed before. For example, in a mutagenesis study of lysozyme [27], replacement of a valine residue with any other (with the exception of β -branched isoleucine) allowed the entry of a new water molecule to a polar cavity within that enzyme. Other studies have used mutagenesis in an attempt to eject bound water

Figure 4 Hydrogen-bonding network centred on the core water molecule of the 390s loop

The 390s loop of wild-type yBH is shown in stereo representation, illustrating the hydrogen-bonding network supported by the core water molecule. Only residues 65–75, 367–371 and 391–402 are shown for clarity. The side chains of residues 73, 369, 392, 393 and 67 are shown as bonds with nitrogen in blue, oxygen in red, sulfur in yellow and carbon in grey. Residues are labelled near the α -carbon. Backbone nitrogens or oxygens involved in hydrogen bonds are shown in blue and red respectively. Hydrogen bonds are shown as black dashed lines. Residues 61–64 and the hydrogen bond between Glu-391 and Ser-61 are omitted for clarity. Co-ordinates are from PDB entry 1GCB.

molecules. For example, Vriend et al. [28] replaced an alanine with serine in order to eject a water molecule from the neutral protease of *Bacillus stearothermophilus*. In this case, the serine hydroxyl interacts with the former hydrogen-bonding partners of the ejected water molecule, maintaining the integrity of the structure.

It is evident that it is the loss of this core water molecule that results in the large structural changes we observe with N392V. As described above, wat-512 forms the core of the 390s loop, which it stabilizes with three hydrogen bonds. The methyl group that replaces it cannot hydrogen-bond and the 390s loop becomes unstable – the electron density becomes weaker and in some places discontinuous, and the *B*-factors rise. As the 60s loop is anchored to a large degree by four hydrogen bonds from the 390s loop (see Figure 4), it becomes disordered.

The interpretation that the loss of the core water molecule of the 390s loop is the cause of the dramatic structural changes observed in N392V is reinforced by the structural changes observed in the H369A and H369L mutants. In both of these cases, a similar destabilization of the 60s loop occurs: no electron density is seen for the loop and the chain cannot be traced.

As described above, the movement of the asparagine side chain that results from the loss of its hydrogen bond with the catalytic histidine causes the ejection of the core water molecule, which in turn results in a new hydrogen-bonding pattern that distorts the geometry of the 390s loop (Figure 2D). On examination, it can be seen that this changes the 390s loop's ability to interact with the 60s loop. The carbonyl oxygen of Ser-393 shifts by 1.9 Å, while the amide nitrogen of Gly-395 shifts by 1.4 Å, in both cases compromising the length and geometry of their hydrogen bonds to the 60s loop. The side chain of Glu-391, which interacts with Ser-61 in the wild-type protein, is in a very different position and now interacts with Thr-374. Presumably, it is the loss of these stabilizing interactions that causes the disorder of the 60s loop, and the resulting compromised position of the catalytic cysteine.

Thus we have uncovered the importance of a structurally conserved water molecule in the papain fold that maintains the architecture of the active site. Even though this water molecule lies 11.7 Å away, it is crucial for the positioning of the sulfur of the nucleophilic cysteine.

The preservation of this active-site structure throughout the papain family suggests that equivalent mutations to those described here for BH could result in similar dramatic structural changes in other C1 class proteases. That there is a complex network of structural interactions right at the active site is no surprise. However, here we have shown that this goes far beyond the immediate vicinity of the active site.

We thank Annie Heroux and Dieter K. Schnieder (BNL) for assistance at X26C (NSLS), Francois St. Pierre for assistance with crystallization and data collection, Michael P. Myers and Danny Perkowski (Cold Spring Harbor Laboratory) for MS and members of the Joshua-Tor laboratory for comments on this paper. This work was supported by NIH (National Institutes of Health) grant 5R01-CA711746 (to L. J.-T.). The NSLS is supported by the US Department of Energy, Division of Material Sciences and Division of Chemical Sciences.

REFERENCES

- Stoltze, L., Schirle, M., Schwarz, G., Schroter, C., Thompson, M. W., Hersh, L. B., Kalbacher, H., Stevanovic, S., Rammensee, H. G. and Schild, H. (2000) Two new proteases in the MHC class I processing pathway. *Nat. Immunol.* **1**, 413–418
- Nishimura, C., Nishimura, T., Tanaka, N. and Suzuki, H. (1987) Potentiation of the cytotoxicity of bleomycin against Ehrlich ascites carcinoma by bleomycin hydrolase inhibitors. *J. Antibiot.* **40**, 1794–1795
- Sebt, S. M., Jani, J. P., Mistry, J. S., Gorelik, E. and Lazo, J. (1991) Metabolic inactivation: a mechanism of human tumor resistance to bleomycin. *Cancer Res.* **51**, 227–232
- Montoya, S. E., Aston, C. E., DeKosky, S. T., Kamboh, M. I., Lazo, J. S. and Ferrell, R. (1998) Bleomycin hydrolase is associated with risk of sporadic Alzheimer's disease. *Nat. Genet.* **18**, 211–212
- Zimny, J., Sikora, M., Guranowski, A. and Jakubowski, H. (2006) Protective mechanisms against homocysteine toxicity: the role of bleomycin hydrolase. *J. Biol. Chem.* **281**, 22485–22492
- Joshua-Tor, L., Xu, E. H., Johnston, S. A. and Rees, D. C. (1995) Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. *Science* **269**, 945–950
- O'Farrell, P. A., Gonzalez, F., Zheng, W., Johnston, S. A. and Joshua-Tor, L. (1999) Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease. *Structure* **7**, 619–627
- Zheng, W., Johnston, S. A. and Joshua-Tor, L. (1998) The unusual active site of Gal6/bleomycin hydrolase can act as a carboxypeptidase, aminopeptidase, and peptide ligase. *Cell* **93**, 103–109
- Dodt, J. and Reichwein, J. (2003) Human cathepsin H: deletion of the mini-chain switches substrate specificity from aminopeptidase to endopeptidase. *Biol. Chem.* **384**, 1327–1332
- Guncar, G., Podobnik, M., Pungercar, J., Strukelj, B., Turk, V. and Turk, D. (1998) Crystal structure of porcine cathepsin H determined at 2.1 Å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure* **6**, 51–61
- Turk, D., Janjic, V., Stern, I., Podobnik, M., Lamba, D., Dahl, S. W., Lauritzen, C., Pedersen, J., Turk, V. and Turk, B. (2001) Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J.* **20**, 6570–6582
- Barrett, A. J. and Kirschke, H. (1981) Cathepsin B, cathepsin H, and cathepsin L. *Methods Enzymol.* **80**, 535–561
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
- 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. et al. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921

- 15 Pannu, N. S. and Read, R. J. (1996) Improved structure refinement through maximum likelihood. *Acta Crystallogr. Sect. A Found. Crystallogr.* **52**, 659–668
- 16 Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. Sect. A Found. Crystallogr.* **47**, 110–119
- 17 Kleywegt, G. J. and Jones, T. A. (1994) A super position. *ESF/CCP4 Newsletter*, **31**, 9–14
- 18 Esnouf, R. M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 938–940
- 19 Merrit, E. A. and Bacon, D. J. (1997) Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524
- 20 Berti, P. J. and Storer, A. C. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J. Mol. Biol.* **246**, 273–283
- 21 Drenth, J., Kalk, K. H. and Swen, H. M. (1976) Binding of chloromethyl ketone substrate analogues to crystalline papain. *Biochemistry* **15**, 3731–3738
- 22 Rullmann, J. A., Bellido, M. N. and van Duijnen, P. T. (1989) The active site of papain. All-atom study of interactions with protein matrix and solvent. *J. Mol. Biol.* **206**, 101–118
- 23 Vernet, T., Chatellier, J., Tessier, D. C. and Thomas, D. Y. (1993) Expression of functional papain precursor in *Saccharomyces cerevisiae*: rapid screening of mutants. *Protein Eng.* **6**, 213–219
- 24 Vernet, T., Tessier, D. C., Chatellier, J., Plouffe, C., Lee, T. S., Thomas, D. Y., Storer, A. C. and Menard, R. (1995) Structural and functional roles of asparagine 175 in the cysteine protease papain. *J. Biol. Chem.* **270**, 16645–16652
- 25 Pickersgill, R. W., Harris, G. W. and Garman, E. (1992) Structure of monoclinic papain at 1.60 Å resolution. *Acta Crystallogr. Sect. B Struct. Sci.* **48**, 59–67
- 26 Rashin, A. A., Iofin, M. and Honig, B. (1986) Binding of buried structural water increases the flexibility of proteins. *Biochemistry* **25**, 3619–3625
- 27 Xu, J. B. W., Quillin, M. L., Baldwin, E. P. and Matthews, B. W. (2001) Structural and thermodynamic analysis of the binding of solvent at internal sites in T4 lysozyme. *Protein Sci.* **10**, 1067–1078
- 28 Vriend, G., Berendsen, H., van der Zee, J., van den Burg, B., Venema, G. and Eijssink, V. (1991) Stabilization of the neutral protease of *Bacillus stearothermophilus* by removal of a buried water molecule. *Protein Eng.* **4**, 941–945

Received 29 April 2006/20 September 2006; accepted 28 September 2006

Published as BJ Immediate Publication 28 September 2006, doi:10.1042/BJ20060641