
John W. Little
Department of Biochemistry and Molecular Biophysics,
University of Arizona,
Tucson, AZ 85721, USA
Email: jlittle@email.arizona.edu

Handbook of Proteolytic Enzymes 2nd Edn
ISBN 0-12-079610-4
Copyright © 2004 Elsevier Ltd
All rights of reproduction in any form reserved

616. UmuD and UmuD' proteins

Databases
MEROPS name: UmuD protein
MEROPS classification: clan SF, family S24, peptidase S24.003
Species distribution: More than one superkingdom
Sequence known from: Escherichia coli, IncJ plasmid R391, Morganella morgani, plasmid pKM101, plasmid Rts1, Salmonella enterica, Salmonella typhimurium, Serratia marcescens, Shigella flexneri, Vibrio cholerae
Tertiary structure: Available

Name and History
The Escherichia coli umuC locus was identified in the late 1970s in genetic screens for strains that were UV-nonmutable (Kato & Shinoura, 1977). In their manuscript, Kato and Shinoura reported the identification of three discrete loci that they called umuA, umuB and umuC. Mapping studies indicated that umuA and umuB mutants were likely to reside in lexA and recA respectively, while umuC appeared to be a novel locus involved in damage-induced mutagenesis (Kato & Shinoura, 1977). Shortly thereafter, in an independent study, Steinborn similarly isolated mutants of E. coli that were also nonmutable after exposure to UV light, which he called uvm (for UV-mutagenesis) (Steinborn, 1978). The use of the uvm name was, however, dropped in the early 1980s when it was realized that the locus is allelic with umuC.

Cloning and sequencing of the umu locus revealed that instead of encoding a single gene, it in fact consists of a two-gene operon that is regulated by the transcriptional repressor LexA (Bagg et al., 1981; Shinagawa et al., 1983; Elledge & Walker, 1983; Kitagawa et al., 1985; Perry et al., 1985). The larger of the two genes was called umuC, while the smaller gene was called umuD. The umuD gene is located upstream of umuC with the TGA stop codon of umuD and the ATG start codon of umuC overlapping by one base pair (Kitagawa et al., 1985; Perry et al., 1985). Perry et al. also noticed that the smaller umuD gene encodes a protein with similarity to the C-terminal domain of LexA, including conserved cleavage site and active-site residues, suggesting that UmuD may undergo post-translational processing (Perry et al., 1985). Indeed, in 1988, Shinagawa et al. and Burckhardt et al. demonstrated that the 15 kDa UmuD protein undergoes both RecA-mediated cleavage and autocatalysis at high pH to generate peptides of ~3 kDa and ~12 kDa (Burckhardt et al.,
Activity and Specificity

The ability of E. coli UmuD (Burchhardt et al., 1988; Shinagawa et al., 1988) and several of its orthologs, including Salmonella typhimurium UmuD (Woodgate et al., 1991; McDonald et al., 1998b), the R-plasmid-encoded MucA (Shiba et al., 1990; Hauser et al., 1992) or RumA (Kulaeva et al., 1995) proteins to undergo post-translational cleavage in vitro and in vivo has been reported. In all cases, processing is greatly stimulated by RecA protein. In vivo, moderately efficient cleavage occurs in cells that have been exposed to cellular DNA damage and in which RecA is believed to be in a so-called ‘activated’ filamentous state bound to single-stranded DNA. Constitutive cleavage occurs in E. coli strains expressing mutant RecAs that are in the activated state in the absence of exogenous DNA damage (Shinagawa et al., 1988; Woodgate & Ennis, 1991; Ennis et al., 1995; Konola et al., 1998).

In vitro, the RecA-mediated reaction requires single-stranded DNA and magnesium for efficient nucleoprotein filament formation. In the absence of RecA, E. coli UmuD, MucA and RumA all undergo autoproteolysis at alkaline pH. The rate of autoproteolysis varies considerably between the orthologs, with the fastest being MucA (t1/2 at pH 10 \(-75\) min) (Hauser et al., 1992) and the slowest being E. coli UmuD (t1/2 at pH 10 >10 h) (Burchhardt et al., 1988; Hauser et al., 1992). Both of which are considerably slower than that of E. coli LexA (t1/2 at pH 10 \(-8\) min) (Little, 1984; Hauser et al., 1992).

The UmuD-like proteins exist as dimers in solution (Woodgate et al., 1989; Battista et al., 1990) and cleavage is believed to occur via an intermolecular reaction in which the substrate tail of one protomer is cleaved in the active site of a dimer-mate (McDonald et al., 1998a, 1999). While both proteins form homodimers in solution, when UmuD2 and UmuD′2 are mixed together in vitro, they preferentially associate to form UmuD/UmuD′ heterodimers (Battista et al., 1990).

Structural Chemistry

The E. coli UmuD protein is 139 amino acids in length and has a calculated molecular mass of 15 063 Da and a pI of 4.5. The 115 residue UmuD′ protein has a molecular mass of 12 285 Da and also has a pI of 4.5. The crystal structure of the UmuD′2 protein was solved at 2.5 Å resolution, revealing an extended N-terminal tail and a globular C-terminal catalytic and dimerization domain containing a mostly beta protein fold (Peat et al., 1996a,b) (Figure 616.1A). NMR spectroscopy has been used to map the UmuD′ dimerization interface (Ferentz et al., 1997); to solve the UmuD′ solution structure (Ferentz et al., 2001) (Figure 616.1B); and to propose a structure for the UmuD/UmuD′ heterodimer (Ferentz et al., 2001; Sutton et al., 2002).

UmuD utilizes a serine–lysine dyad mechanism. Site-directed mutagenesis (Nohmi et al., 1988) and structural studies (Peat et al., 1996a) are consistent with Ser60 serving as the nucleophile with Lys97 acting as the general base.
as the nucleophile and Lys97 the general base. A structural
alignment of UmuD′ with the acyl-enzyme of signal peptidase
(another member of the clan SF) suggested an ori-
centation for the cleavage site in the UmuD-binding site.
These studies also revealed that the nucleophilic Ser60
hydroxyl of UmuD attacks the scissile bond (located between
Cys24Gly25 of UmuD) from the si-face rather than the re-
face as seen in most serine proteases (Paetzel & Strynadka,
1999) (Figure 616.2).

NMR analysis of the UmuD/UmuD′ dimer (Ferentz et al.,
2001) and crystallographic analysis of the analogous protein
LexA (Luo et al., 2001) are consistent with the proposed
substrate orientation. The crystal structure of LexA with
its bound cleavage site suggests that the main-chain amide
hydrogens from Ser60 and Asp59 in UmuD could serve as
the oxyanion hole (Luo et al., 2001) (Figure 616.2B). Crystal
structures are now available for four members of the clan SF:
UmuD′, λCI repressor, LexA repressor, and signal peptidase
(Peat et al., 1996a; Bell et al., 2000; Luo et al., 2001; Paetzel
et al., 1998). A superposition of their active sites reveals
that the Nζ of the lysine general base is coordinated by three
hydrogen bonds. In the case of UmuD′, the neutral ε-amino
group of Lys97 (the deprotonated state is a requirement for it
serving as the general base) would have two hydrogen bond
acceptors (Val96 O and Thr95 Oγ) and one hydrogen bond
donor (Ser60 OγH). It is proposed that the pKa of the ε-amino

group of Lys97 is depressed by its burial upon binding of the
cleavage site in an energetically unfavorable position (Luo
et al., 2001) (Figure 616.2).

The NMR solution structure suggests that the UmuD′
dimer is structurally dynamic and that Ser60 and Lys97 are
not within hydrogen bonding distance in solution (Ferentz
et al., 2001) (Figure 616.1B). In contrast, all crystal struc-
tures of the clan SF proteases so far have shown the nucle-
ophilic serine and general-base lysine to be within hydrogen
bonding distance. Ferentz and colleges suggest that the crys-
tal packing forces may, therefore, result in the stabilization
of the catalytically competent conformation and that an inter-
action with a RecA nucleoprotein filament in vivo facilitates
the cleavable conformation (Ferentz et al., 2001).

Preparation

The UmuD protein was initially overexpressed in E. coli
from a temperature-inducible λPl promoter (Burckhardt
et al., 1988). Homodimeric UmuD′ can be purified from
the same UmuD overproducing strain if the cells are also
exposed to the DNA-damaging agent mitomycin C, so as to
promote in vivo conversion of UmuD to UmuD′ (Woodgate
et al., 1989). Both UmuD and recombinant UmuD′ are now
routinely expressed from an IPTG-inducible T7 promoter
(Frank et al., 1993; Ferentz et al., 1997). Purification is
relatively simple and involves ammonium sulfate precipita-
tion, ion-exchange and gel-filtration chromatography. Under
these conditions, up to 10 mg of highly purified UmuD or
UmuD′ protein can be isolated from 1 liter of an induced
E. coli culture.

Biological Aspects

Since their discovery, the Umu proteins have been hypoth-
esized to participate in damage-induced mutagenesis. For

Figure 616.2 (A) The UmuD active site. The side chains of residues in the S1 and S3 binding pockets are shown in ball-and-stick. The crystal structure of UmuD′ shows that Ser60 Oγ and Lys97 Nε are within hydrogen bonding distance (Peat et al., 1996a). The atomic coordinates 1UUM (pdb code) were used to produce this figure. (B) A schematic of the possible interactions between the UmuD cleavage site region in the binding site of its dimer mate. The residues involved in forming the S1 and S3 binding sites are indicated. The cleavage site residues are in parenthesis. Potential hydrogen bonding interactions between the extended cleavage site region and the β strands that line each side of the binding sites are shown. The main-chain amide hydrogens of Ser60 and Asp59 would make up the oxyanion hole.
many years it was believed that the Umu proteins somehow modified the cell’s main replicase, so that it would traverse otherwise replication-blocking lesions. However, in the past few years, the E. coli UmuC protein has been shown to possess intrinsic DNA polymerase activity (Woodgate et al., 1989; Frank et al., 1996; Tang et al., 1998). This large family of lesion-bypassing DNA polymerases is found in all three kingdoms of life, yet interestingly, UmuD-like orthologs have only been identified in gram-negative bacteria, their self-transmissible R-plasmids or bacteriophages. Even more intriguing is the fact that the ‘founding members’ of the recently described Y-family of DNA polymerases (Ohmori et al., 2001). In vitro studies suggest that UmuD stimulates the catalytic activity of the UmuC protein (Reuven et al., 1999), to which it is normally complexed (Woodgate et al., 1989; Bruck et al., 1996; Tang et al., 1998).

### Further Reading

A structural analysis and comparison of the Ser/Lys protease has recently been performed (Paetzel et al., 2002). For a general discussion on serine-lysin proteases see Paetzel & Dalbey (1997). Recent reviews on UmuD and its role in the SOS response include those of Goodman & Woodgate (2000), Sutton et al. (2000, 2002) and Gonzalez & Woodgate (2002). There have been numerous reviews on the Y-family DNA polymerases, among which are Woodgate (1999), Goodman & Tippen (2000), Friedberg et al. (2000), Livneh (2001) and references therein.

### References


---

**Mark Paetzel**
Simon Fraser University,
Department of Molecular Biology and Biochemistry
South Science Building
8888 University Drive Burnaby,
British Columbia,
Canada V5A 1S6
Email: mpaetzel@sfu.ca

**Roger Woodgate**
Section on DNA Replication,
Repair and Mutagenesis,
Building 6, Room 1A13,
National Institute of Child Health and
Human Development,
National Institutes of Health,
9000 Rockville Pike,
Bethesda, MD 20892-2725, USA
Email: woodgate@helix.nih.gov

---

**Handbook of Proteolytic Enzymes 2nd Edn**
ISBN 0-12-079610-4
Copyright © 2004 Elsevier Ltd
All rights of reproduction in any form reserved

---

**617. Signal peptidase I**

**Databanks**

**MEROPS name:** Signal peptidase I

**MEROPS classification:** clan SF, family S26A, peptidase S26.001

**IUBMB:** EC 3.4.21.89

**CAS registry:** 65979-36-4

**Species distribution:** Eubacteria

**Sequence known from:** Agrobacterium tumefaciens, Aquifex aeolicus, Azotobacter vinelandii, Bordetella pertussis, Bradyrhizobium japonicum, Brucella melitensis, Brucella suis, Buchnera aphidicola, Buchnera sp. APS, Escherichia coli, Haemophilus influenzae, Neisseria meningitidis, Pasteurella multocida, Pseudomonas aeruginosa, Pseudomonas fluorescens, Rathia solanacearum, Rhodobacter capsulatus, Salmonella enterica, Salmonella typhimurium, Shewanella oneidensis, Shigella flexneri, Sinorhizobium meliloti, Treponema pallidum, Vibrio cholerae, Wigglesworthia brevipalpis, Xanthomonas axonopodis, Xanthomonas campestris, Xylella fastidiosa, Yersinia pestis

**Tertiary structure:** Available