

*Arolas et al.*

## Expression and purification of integral membrane metallopeptidase HtpX

Joan L. Arolas<sup>1,\*</sup>, Raquel García-Castellanos<sup>1</sup>, Theodoros Goulas<sup>1</sup>, Yoshinori Akiyama<sup>2</sup>  
& F. Xavier Gomis-Rüth<sup>1,\*</sup>

<sup>1</sup>Proteolysis Lab, Department of Structural Biology, Molecular Biology Institute of Barcelona, CSIC; Barcelona Science Park; E-08028 Barcelona (Spain)

<sup>2</sup>Institute for Virus Research, Kyoto University; Kyoto 606-8507 (Japan)

\*Correspondence: [jlacri@ibmb.csic.es](mailto:jlacri@ibmb.csic.es) or [xgrcri@ibmb.csic.es](mailto:xgrcri@ibmb.csic.es); phone: (+34) 93 402 0186 or (+34) 93 402 0187; fax: (+34) 93 403 4979

The authors state they have no competing financial interest

## Highlights

- The integral membrane metallopeptidase HtpX was overexpressed in *E. coli*.
- A catalytically ablated HtpX mutant was designed to avoid self-cleavage.
- The protein was extracted from membranes and purified in octyl- $\beta$ -D-glucoside.
- The protein was purified by cobalt-affinity, anion-exchange and size-exclusion.
- The overall system renders milligram amounts of HtpX and fosters structural studies.

## Abstract

Little is known about the catalytic mechanism of integral membrane (IM) peptidases. HtpX is an IM metallopeptidase that plays a central role in protein quality control by preventing the accumulation of misfolded proteins in the membrane. Here we report the recombinant overexpression and purification of a catalytically ablated form of HtpX from *Escherichia coli*. Several *E. coli* strains, expression vectors, detergents, and purification strategies were tested to achieve maximum yields of pure and well-folded protein. HtpX was successfully overexpressed in *E. coli* BL21(DE3) cells using a pET-derived vector attaching a C-terminal His<sub>8</sub>-tag, extracted from the membranes using octyl- $\beta$ -D-glucoside, and purified to homogeneity in the presence of this detergent in three consecutive steps: cobalt-affinity, anion-exchange, and size-exclusion chromatography. The production of HtpX in milligram amounts paves the way for structural studies, which will be essential to understand the catalytic mechanism of this IM peptidase and related family members.

*Keywords:* metallopeptidase; integral membrane protein; protein quality control; *Escherichia coli*; overexpression

## Introduction

The study of membrane proteins constitutes a major frontier in biochemistry. Genome analyses in prokaryotes and eukaryotes reveal that 20-30% of the encoded proteins are integral membrane (IM) proteins and another 10-20% are membrane-associated [1]. IM proteins are crucial in cellular communication and transport, among many other processes, and they are commonly related to disease, which makes them relevant as pharmaceutical targets [2]. However, the functional and structural study of these proteins faces enormous challenges owing to low endogenous expression levels and technical hurdles related to their recombinant expression and purification. Accordingly, in recent years much effort has been made to find expression systems that yield sufficient amounts of native-like protein and detergents/lipids that mimic the native membrane environment [3, 4]. Nevertheless, IM proteins are still dramatically under-represented in the three-dimensional structural database of the Protein Data Bank (PDB; see <http://www.pdb.org/> and [5]).

IM proteins include a large number of peptidases of different catalytic types. Among them are IM metallopeptidases (MPs) from families M48, M50 and M56 of the MEROPS database (<http://merops.sanger.ac.uk>), which contain a characteristic metal-binding motif, HEXXH (*X* stands for any residue; [6-8]). These IMMPs participate in a variety of physiological mechanisms that require proteolysis, and they target transcription factors (site-2 proteases), nuclear prelamin A/mating pheromone  $\alpha$ -factor (FACE1/Ste24p), and signal molecules that trigger bacterial resistance (BlaR1 and MecR1 from family M56), among others [7]. The biological relevance of IMMPs fosters the need to acquire atomic structural information, mainly by means of X-ray crystallography, in order to be able to modulate the activity of these enzymes. However, the only IMMP structures published to date are those of site-2 protease from *Methanocaldococcus jannaschii* (family M50) and human/yeast FACE1/Ste24p (family M48) [9-11]. Another member of this latter family is HtpX, which was initially described as a heat-shock, Cpx-controlled protein [12, 13]. HtpX is widespread in bacteria and archaea, and it participates in the proteolytic quality control of membrane proteins together with a two-pass transmembrane ATP-dependent MP, FtsH [14-16]. Disruption mutations of HtpX and FtsH revealed that they are essential for *Escherichia coli* growth and can compensate for each other when one of them is inactivated [13].

HtpX was previously expressed in *E. coli* and purified under denaturing conditions owing to its strong self-cleavage. After refolding, the protein proved to be an active zinc-dependent MP [16]. Here we have developed a bacterial overexpression system and a purification protocol

*Arolas et al.*

based on the extraction of soluble protein from membranes. The system renders milligram amounts of a well-folded HtpX variant with the catalytic glutamic acid of the HEXXH motif mutated to alanine.

## Materials and methods

### Generation of constructs

The gene coding for HtpX from *E. coli* (293 residues, 32 kDa; UniProt [UP] sequence database access code P23894) was previously cloned into a derivative of vector pTYE007 [13]. This vector confers resistance towards ampicillin and attaches a 4-kDa C-terminal tag comprising a factor Xa peptidase recognition site and a hexahistidine (His<sub>6</sub>)-tag among other residues [16]. This construct gave rise to a 328-residue form of HtpX (36 kDa; hereafter referred to as HtpX-4kDa-tag; see Table 1). A single-residue point mutant (glutamic acid-to-alanine at position 140; E<sup>140</sup>A) of HtpX-4kDa-tag was generated using the QuikChange Site-Directed Mutagenesis kit (Stratagene) according to the manufacturer's instructions. The gene coding for E<sup>140</sup>A HtpX plus a 2-kDa C-terminal tag comprising a factor Xa peptidase recognition site and an octahistidine (His<sub>8</sub>)-tag was subsequently cloned into a modified pET28 (Novagen) vector that confers resistance towards kanamycin. This construct gave rise to a 313-residue form of E<sup>140</sup>A HtpX (34 kDa; hereafter referred to as HtpX-2kDa-tag).

This latter construct was also cloned into modified pETM30 and pET28 vectors comprising an N-terminal His<sub>6</sub>-tag followed by a fusion protein and a tobacco etch virus (TEV) peptidase cleavage site (Table 1). The fusion proteins tested were maltose-binding protein (MBP; ~43 kDa), Ztag (Z-domain of *Staphylococcal* protein A; ~10 kDa), GB1 (B1 immunoglobulin binding domain of *Streptococcal* protein G; ~9 kDa), thioredoxin (Trx; ~12 kDa), NusA (~59 kDa), glutathione S-transferase (GST; ~27 kDa), and Mystic (membrane-integrating protein from *Bacillus subtilis*; ~13 kDa). A construct without fusion protein between the N-terminal His<sub>6</sub>-tag and the TEV peptidase cleavage site was prepared as reference.

The gene coding for E<sup>140</sup>A HtpX with only a C-terminal His<sub>8</sub>-tag was further cloned into another modified pET28 vector, which confers resistance towards kanamycin and does not attach extra residues (301 residues; 33 kDa; hereafter referred to as HtpX-His<sub>8</sub>). HtpX orthologs from *Aquifex aeolicus* (302 residues, UP access code O67798), *Listeria monocytogenes* (304 residues, UP access code C1L1N4), *Thermoanaerobacter tengcongensis* (299 residues, UP access code Q8R936), *Pyrococcus furiosus* (289 residues, UP access code Q8U1S0), *Methanopyrus kandleri* (318 residues, UP access code Q8TYX0), and *Methanosarcina acetivorans* (294 residues, UP access code Q8TP15) were cloned into the same modified pET28 vector attaching a C-terminal His<sub>6</sub>-tag. All constructs were verified by DNA sequencing.

### Protein expression assays

Both wild-type and E<sup>140</sup>A HtpX-4kDa-tag were overexpressed in *E. coli* AD16 [16] or C43(DE3) cells, which were grown at 37°C in Luria Bertani (LB) or Terrific Broth (TB)

medium supplemented with 100 µg/ml ampicillin. E<sup>140</sup>A HtpX-2kDa-tag with different N-terminal fusion tags (see above) and the HtpX orthologs were overexpressed in *E. coli* AD16, C43(DE3), or BL21(DE3) cells grown at 37°C in LB or TB medium supplemented with 30 µg/ml kanamycin. E<sup>140</sup>A HtpX-2kDa-tag without any N-terminal fusion tag and E<sup>140</sup>A HtpX-His<sub>8</sub> were overexpressed in *E. coli* BL21(DE3) cells grown at 37°C in TB medium supplemented with 30 µg/ml kanamycin. Expression trials were carried out in 50-ml tubes under different growth and induction conditions (i.e. inducing expression at an OD<sub>600</sub> of 0.5-1 with 0.2-1 mM isopropyl-β-D-thiogalactopyranoside [IPTG] and incubated either for 5h at 37°C or overnight at 18°C). Expression experiments were carried out by growing cultures in TB medium at 37°C in 2-liter Erlenmeyer flasks (6 liters in total) or in a 5-liter fermentor (Infors AG), inducing expression at an OD<sub>600</sub> of 1.5 or 6.0, respectively, with 0.2 mM IPTG, and incubating overnight at 18°C. Further expression experiments were performed using a 50-liter pilot-plant Biostat-U fermentor (Braun Biotech) at the Fermentation Facility (Department of Microbiology, Faculty of Biology) of the University of Barcelona (Spain) by growing cultures with TB medium at 37°C, inducing expression at an OD<sub>600</sub> of 6.0 with 0.2 mM IPTG, and incubating overnight at 18°C. Cells were harvested by centrifugation at 7,000xg for 30 min at 4°C and stored at -80°C. Cultures grown in the 50-liter fermentor typically reached a final OD<sub>600</sub> of 12 and were divided into five aliquots of approximately 200 g of cells each.

#### *Protein purification of constructs*

Wild-type and E<sup>140</sup>A HtpX-4kDa-tag were purified as follows. The cell pellet was thawed in ice, washed twice in 1x TBS buffer, and resuspended in the same buffer plus EDTA-free protease inhibitor cocktail tablets and DNase I (both Roche Diagnostics). Cells were lysed using a cell disrupter (Constant Systems Ltd.) at a pressure of 2.4 kbar, and the cell debris was removed by centrifugation at 9,000xg for 30 min at 4°C. Membranes were collected by ultracentrifugation at 150,000xg for 1 h at 4°C in a Beckman Optima L-90K using a 50.2 Ti rotor (Beckman) and 26.3-ml polycarbonate bottles with cap assembly (Beckman). Membranes were subsequently homogenized using a glass Potter and solubilized under gentle stirring by overnight incubation at 4°C with 50 mM Tris-HCl, 500 mM NaCl, 10 mM 2-mercaptoethanol, pH 8.0, containing 1% Igepal CA-630 (octylphenoxypolyethoxyethanol, previously known as Nonidet P-40; purchased from Sigma) and EDTA-free protease inhibitor cocktail tablets. Membrane debris was further removed by ultracentrifugation as described above and the resulting supernatant was filtered (0.22 µm pore size; Millipore). The sample was incubated under gentle stirring for 5 h at 4°C with nickel-nitrilotriacetic acid resin (Invitrogen) previously equilibrated with 50 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, 10 mM 2-mercaptoethanol, pH 8.0, containing 0.085% Igepal CA-630. The resin was collected in a 25-ml disposable Bio-Rad column, washed extensively, and the His<sub>6</sub>-tagged protein eluted with the same buffer plus 300 mM imidazole. The sample

was desalted and detergent-exchanged using a PD-10 column (GE Healthcare) previously equilibrated with 20 mM Tris-HCl, 2 mM 2-dithiothreitol, pH 8.0, containing either 0.05% dodecyl- $\beta$ -D-maltoside (DDM), 0.2% decyl- $\beta$ -D-maltoside (DM), or 1% octyl- $\beta$ -D-glucoside (OG) (all of Anagrade purity [ $>99\%$ ] from Anatrace). The critical micelle concentration (CMC) of DDM, DM, and OG in water is  $\sim 0.0087\%$ ,  $\sim 0.087\%$ , and  $\sim 0.53\%$ , respectively. The protein variants were further purified by anion-exchange chromatography using a Mono Q 5/50 GL column (bed volume 1 ml; GE Healthcare) with a 0-500 mM NaCl linear gradient in 15 column volumes, and by size-exclusion chromatography on a Superdex 75 10/300 GL column (bed volume 24 ml; GE Healthcare) previously equilibrated with 20 mM Tris-HCl, 150 mM NaCl, pH 7.5, containing 0.05% DDM, 0.2% DM, or 1% OG. The two latter chromatographic steps were conducted using an ÄKTA Purifier System (GE Healthcare) at 4°C.

Purification of E<sup>140</sup>A HtpX-2kDa-tag fused with different N-terminal tags and the six HtpX orthologs was performed as follows. The cell pellet was thawed in ice, resuspended in 1x TBS buffer, and lysed in an ice-bath using a digital sonifier (Branson) in two cycles of brief (10 s) sonication (set at 10% amplitude). The cell debris was subsequently removed by centrifugation at 13,000xg for 10 min at 4°C, and the supernatant was analyzed by 15% Tricine-SDS-PAGE and Western blotting using a nitrocellulose membrane (GE Healthcare) and a His-probe Antibody (H-3) HRP (Santa Cruz Biotechnology). Purification of E<sup>140</sup>A HtpX-2kDa-tag without any N-terminal fusion tag and E<sup>140</sup>A HtpX-His<sub>8</sub> was carried out initially as described above for HtpX-4kDa-tag. Membranes were next solubilized under gentle stirring by overnight incubation at 4°C with 50 mM Tris-HCl, 500 mM NaCl, pH 8.0, containing 2% OG and EDTA-free protease inhibitor cocktail tablets. After cell debris removal by ultracentrifugation (see above), the supernatant was filtered, and the sample was incubated under gentle stirring for 5 h at 4°C with TALON metal-affinity resin (charged with cobalt instead of nickel; Clontech) previously equilibrated with 50 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, pH 8.0, containing 1% OG. The resin was collected in a 25-ml disposable column, washed extensively, and the His<sub>8</sub>-tagged protein eluted with the same buffer plus 300 mM imidazole. The sample was desalted using a PD-10 column (GE Healthcare) previously equilibrated with 20 mM Tris-HCl, pH 8.5, containing 1% OG or using dialysis against this same buffer. The protein was further purified by anion-exchange chromatography using a Mono Q PC 1.6/5 column (bed volume 0.1 ml; GE Healthcare) with a one-step elution with 1 ml of the same buffer plus 500 mM NaCl, and by size-exclusion chromatography on a Superdex 75 10/300 GL column previously equilibrated with 20 mM Tris-HCl, 150 mM NaCl, pH 7.5, containing 1% OG. Protein identity and purity were assessed by 15% Tricine-SDS-PAGE stained with Coomassie blue. Ultrafiltration steps were performed with Vivaspin 4 and Vivaspin 500 filter devices of 50-kDa cut-off (Sartorius Stedim Biotech). Approximate protein concentration was determined by measuring A<sub>280</sub> in a spectrophotometer (NanoDrop) and the calculated absorption coefficient,  $E_{0.1\%} = 0.65$ . N-

terminal sequencing through Edman degradation and peptide-mass fingerprinting of tryptic protein digests were carried out at the Proteomics and Genomics Facility of Centro de Investigaciones Biológicas (Madrid, Spain).

*Circular dichroism assays*

Samples for far-UV circular dichroism spectroscopy were prepared at a final protein concentration of 0.3 mg/ml in 20 mM Tris-HCl, 150 mM NaCl, pH 7.5, containing 1% OG. Measurements were carried out in a Jasco J-815 spectrometer at 20-90°C by using a 1-mm path length cell.

*Crystallization trials*

Crystallization assays were carried out at the IBMB/IRB Crystallography Platform (Barcelona, Spain) by the sitting-drop vapor diffusion method using 96x2-well MRC plates (Innovadyne). A TECAN Freedom EVO robot was used to prepare reservoir solutions, and a Cartesian Microsys 4000 XL robot (Genomic Solutions) was used for nanodrop dispensing. Crystallization plates were stored in Bruker steady-temperature crystal farms at 4°C and 20°C. Crystals appeared at 20°C from equivolometric drops containing protein solution (~5 mg/ml) and 0.1 M glycine, 0.1 M lithium sulfate monohydrate, 30% (v/v) polyethylene glycol 400, pH 9.3 as reservoir solution. Crystals were cryo-protected with 20% (v/v) glycerol. Diffraction experiments were performed at 100K from liquid-N<sub>2</sub> flash-cryo-cooled crystals (Oxford Cryosystems 700 series cryostream) on a MarCCD detector (microfocus beam line ID23-2) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) within the Block Allocation Group “BAG Barcelona”.



## Results and discussion

### *Design of a stable HtpX variant*

The prototypic IMMP HtpX from *E. coli* is a 293-residue protein that is predicted to comprise four transmembrane segments, two within the first 55 residues and the other two within residues 150-215 (Fig. 1A). However, the localization of the two latter segments remains controversial [13]. The position of the characteristic HEXXH zinc-binding motif reveals that the catalytic moiety is most likely to be found on the cytosolic side of the membrane. This is in agreement with studies reporting that HtpX cleaves only the cytoplasmic regions of membrane protein SecY *in vivo* [16] and contrasts with Oma1, another family-M48 IMMP involved in maintaining the integrity of the mitochondrial inner membrane, which cleaves protein substrates on both sides of the membrane [17]. According to the zinc-binding motif, H<sup>139</sup> and H<sup>143</sup> most likely coordinate the catalytic zinc ion in HtpX. This is consistent with the crystal structure of a soluble, artificial fragment of an HtpX ortholog from *Vibrio parahaemolyticus* comprising residues 57-157 (PDB entry 3CQB; deposited by a Structural Genomics Consortium but not published). Despite showing the equivalent histidines at positions compatible with catalytic zinc binding, the structure exhibited an overall non-functional active site due to the lack of the third residue that is required for proper zinc coordination in competent MPs. This residue, probably a glutamic acid by analogy with FACE1/Ste24p (PDB entries 2YPT/4IL3; [10, 11]), is predicted to be located within a “glutamate helix” [6, 18] spanning residues 220-230, which is found after the fourth transmembrane helix as inferred from homology modeling and secondary structure predictions (see Fig. 1B and [7]). Consistently, mutation of the only glutamic acid found within this helix to glutamine (E<sup>222</sup>Q) abolished the complementation activity of HtpX [13].

Previously, wild-type HtpX had been reported to undergo rapid self-cleavage after homologous recombinant overexpression during cell disruption and/or membrane solubilization with detergent [16]. The protein was therefore purified under denaturing conditions and refolded in the presence of a metal chelator [16]. When supplemented with zinc ion, the enzyme was reported to be catalytically active against casein and the membrane protein SecY but it also showed self-cleavage around position Leu<sup>260</sup>. The authors produced a single-residue point mutant, H<sup>139</sup>F, which disrupted the active site by preventing zinc coordination, resulting in an inactive enzyme that did not undergo self-cleavage nor show activity against protein substrates [16].

By analogy with other zinc-dependent MPs, E<sup>140</sup> of the zinc-binding motif may function as a general base and acid during catalysis by aligning and activating the catalytic water molecule that attacks the substrate [18, 19]. Glutamic acid-to-alanine mutation is a common strategy used in the field of MP research to generate stable protein variants that are catalytically ablated but

maintain the correct active-site architecture and zinc coordination, thus constituting valid models for structural studies. This strategy has been successfully used to determine the structure of numerous MP zymogenic forms, which usually show self-activation at the high protein concentrations required for crystallization [20-22]. Accordingly, we constructed HtpX mutant E<sup>140</sup>A to prevent self-cleavage and thus yield a stable form for structural studies.

#### *Screening for the best expression and purification system for HtpX*

HtpX was overexpressed and extracted from the membranes in detergent instead of using denaturing agents and refolding from insoluble inclusion bodies. To check whether this approach rendered well-folded protein, wild-type HtpX with a 4-kDa C-terminal tag (HtpX-4kDa-tag), which was the construct used in the previous report [16], was overexpressed and purified. The protein obtained after the purification protocol underwent strong self-cleavage, indicating that it was catalytically active and hence most likely in a (at least near) native-like conformation (see Fig. 2A). We subsequently overexpressed and purified E<sup>140</sup>A HtpX-4kDa-tag, which proved to be stable and did not show self-cleavage under the conditions tested (see “Materials and methods” and below). Both wild-type and E<sup>140</sup>A HtpX-4kDa-tag were overexpressed in *E. coli* AD16 or C43(DE3) cells harboring a variant of vector pTYE007, which attached 35 extra C-terminal residues, including a His<sub>6</sub>-tag for metal-affinity purification. To reduce the number of additional residues, as these could interfere with crystallization experiments, a construct with a shorter C-terminal tag (35→20 extra residues) was designed and cloned into a modified pET28 vector (Table 1). It included a longer His-tag (eight instead of six residues), which had been reported to improve binding to the metal-affinity resin in other cases [23]. This construct (E<sup>140</sup>A HtpX-2kDa-tag) was next cloned into modified pETM30 and pET28 expression vectors comprising a collection of N-terminal fusion proteins removable by TEV peptidase. A construct without fusion protein (i.e. containing only an N-terminal His<sub>6</sub>-tag) was also cloned as reference (see Table 1).

Expression trials of all these constructs were carried out in three *E. coli* strains using different growth and induction conditions (see “Materials and methods”) and monitored by Western blotting using a His-tag antibody as probe. This screening revealed that (i) there was expression in C43(DE3) and BL21(DE3) cells but not in AD16 cells (data not shown)—interestingly, the protein had been previously overexpressed in this latter *E. coli* strain harboring a derivative of vector pTYE007 [16]—; (ii) best expression yields were obtained in BL21(DE3) cells induced overnight at 18°C with low IPTG concentrations (data not shown); and (iii) the attachment of Ztag, GB1, Trx, GST, and Mystic as fusion proteins increased expression levels of HtpX, with the latter displaying about 5-fold higher expression when compared to the reference construct without fusion protein (Fig. 2B). Production using these fusion proteins was, however, ruled out because this improvement in protein expression was not expected to compensate for the poor

cleavage of the fusion protein by TEV peptidase in the presence of detergent [24, 25]. We further cloned and screened the expression of six HtpX orthologs from other bacteria and archaea, a strategy that had proved useful in the field of membrane proteins [9, 26]. However, none of the orthologs was detected by Western blotting using the His-tag antibody under the expression conditions tested (see “Materials and methods” and Fig. 2C), and therefore all subsequent studies were performed with the *E. coli* ortholog.

Accordingly, E<sup>140</sup>A HtpX-2kDa-tag without fusion protein was overexpressed as described above and successfully purified in three consecutive steps: nickel-affinity, anion-exchange, and size-exclusion chromatography. Even though the protein was initially extracted from the membranes using the non-ionic non-denaturing detergent Igepal CA-630, this detergent was subsequently replaced by other detergents more suitable for structural studies. In particular, HtpX proved to be soluble and stable in the presence of DDM, DM, and OG. DDM and DM are common choices due to their gentle nature [27, 28]. However, they form relatively large micelles that were concentrated together with the protein when employing a 50-kDa cut-off filter device (a 100-kDa filter did not retain the protein during concentration), and the high detergent concentration resulted in strong phase separation in preliminary crystallization trials. Therefore, we focused on OG, which assembles into more compact micelles (~23 kDa) and is in general more successful in the crystallization of small and medium-sized proteins such as HtpX (32 kDa) [27, 28].

#### *Large-scale production of homogeneous HtpX*

Although, in general, the detergent used for solubilization does not need to be the same as the one used for purification and crystallization, it is advisable to ensure maximum homogeneity of the final sample [29]. To produce homogeneous and soluble HtpX, OG was used to extract the protein initially from the membranes at 2% and thereafter maintained at 1%. Extraction yields were similar to those obtained with Igepal CA-630 (data not shown). In addition, a new construct with only a C-terminal His<sub>8</sub>-tag was designed (i.e. lacking the 12 extra residues of E<sup>140</sup>A HtpX-2kDa-tag) and cloned into the pET28-modified vector. This gave rise to E<sup>140</sup>A HtpX-His<sub>8</sub>, which showed essentially the same binding to the metal-affinity resin. The protein was subsequently purified by anion-exchange chromatography, which, however, did not yield pure protein mainly due to the co-elution of proteins such as expression host cytochromes (giving a yellowish color to the sample). These contaminating proteins bound in a non-specific manner to the agarose resin of the nickel-affinity chromatography step (Fig. 2D). To overcome this problem, the protein was purified using cobalt-affinity resin instead, which rendered protein with a high degree of purity (Fig. 2E). E<sup>140</sup>A HtpX-His<sub>8</sub> was further loaded onto a precision anion-exchange column (with a bed volume of 0.1 ml), which contributed to reducing the sample volume without affecting the concentration of detergent. The purity and homogeneity of

the protein was assessed by size-exclusion chromatography, which revealed a single sharp, Gaussian peak, thus suggesting that the protein was well folded and stable (Fig. 2F). This peak migrated as a ~62-kDa molecule as inferred from the protein standards previously loaded onto the same column. This would correspond to HtpX-His<sub>8</sub> (33 kDa) in complex with detergent molecules that expand its hydrodynamic radius [30].

Overexpression of E<sup>140</sup>A HtpX-His<sub>8</sub> following the aforementioned approach was scaled up in a 50-liter fermentor to produce milligram amounts of protein. Growth and induction conditions were essentially the same as those used for expression experiments in Erlenmeyer flasks or in a 5-liter fermentor (see “Materials and methods”). The cell pellet was subsequently processed using the aforementioned three-step purification protocol, which yielded ~5 mg of pure protein from 50 liters of culture. Circular dichroism assays of the purified protein confirmed a well-folded conformation as the spectrum showed the typical shape (minima at 208 and 222 nm) of a mostly  $\alpha$ -helical protein that was lost after thermal denaturation (see Fig. 2G). Finally, initial crystallization attempts using 1% OG and concentrating the protein in a 50-kDa filter rendered protein crystals that diffracted up to ~18Å (Fig. 2H), which provided a starting point for upcoming experiments. Overall, our newly developed production and purification system opens the door for the crystallization and structure determination of HtpX, which may uncover the physiological function and regulation of this IMMP.

## **Acknowledgments**

We thank Tibisay Guevara, Mar López-Pelegri n, and Javier M endez (University of Barcelona) for excellent technical assistance and Hartmut Luecke (University of California, Irvine) and Robin Rycroft for helpful suggestions. The study was supported in part by grants from European, Spanish, and Catalan agencies (FP7-HEALTH-2010-261460 “Gums&Joints”; FP7-PEOPLE-2011-ITN-290246 “RAPID”; FP7-HEALTH-2012-306029-2 “TRIGGER”; BFU2012-32862; CSD2006-00015; Fundaci n “La Marat n de TV3” 2009-100732; 2009SGR1036), by a “JAE” research contract from CSIC (co-funded by FSE) and a short-term EMBO fellowship (to J.L.A.), and by a “Juan de la Cierva” research contract from the Spanish Ministry (to R.G-C.).

## References

- [1] E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* 7 (1998) 1029-1038.
- [2] J. Drews, Drug discovery: a historical perspective. *Science* 287 (2000) 1960-1964.
- [3] R.M. Bill, P.J. Henderson, S. Iwata, E.R. Kunji, H. Michel, R. Neutze, S. Newstead, B. Poolman, C.G. Tate, H. Vogel, Overcoming barriers to membrane protein structure determination. *Nat Biotechnol* 29 (2011) 335-340.
- [4] S. Wagner, M.L. Bader, D. Drew, J.W. de Gier, Rationalizing membrane protein overexpression. *Trends Biotechnol* 24 (2006) 364-371.
- [5] D. Kozma, I. Simon, G.E. Tusnady, PDBTM: Protein Data Bank of transmembrane proteins after 8 years. *Nucleic Acids Res* 41 (2013) D524-529.
- [6] N. Cerda-Costa, F.X. Gomis-Ruth, Architecture and function of metallopeptidase catalytic domains. *Protein Sci* 23 (2014) 123-144.
- [7] M. Lopez-Pelegrin, N. Cerda-Costa, F. Martinez-Jimenez, A. Cintas-Pedrola, A. Canals, J.R. Peinado, M.A. Marti-Renom, C. Lopez-Otin, J.L. Arolas, F.X. Gomis-Ruth, A novel family of soluble minimal scaffolds provides structural insight into the catalytic domains of integral membrane metallopeptidases. *J Biol Chem* 288 (2013) 21279-21294.
- [8] N.D. Rawlings, A.J. Barrett, A. Bateman, MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 40 (2012) D343-350.
- [9] L. Feng, H. Yan, Z. Wu, N. Yan, Z. Wang, P.D. Jeffrey, Y. Shi, Structure of a site-2 protease family intramembrane metalloprotease. *Science* 318 (2007) 1608-1612.
- [10] E.E. Pryor, Jr., P.S. Horanyi, K.M. Clark, N. Fedoriw, S.M. Connelly, M. Koszelak-Rosenblum, G. Zhu, M.G. Malkowski, M.C. Wiener, M.E. Dumont, Structure of the integral membrane protein CAAX protease Ste24p. *Science* 339 (2013) 1600-1604.
- [11] A. Quigley, Y.Y. Dong, A.C. Pike, L. Dong, L. Shrestha, G. Berridge, P.J. Stansfeld, M.S. Sansom, A.M. Edwards, C. Bountra, F. von Delft, A.N. Bullock, N.A. Burgess-Brown, E.P. Carpenter, The structural basis of ZMPSTE24-dependent laminopathies. *Science* 339 (2013) 1604-1607.
- [12] D. Kornitzer, D. Teff, S. Altuvia, A.B. Oppenheim, Isolation, characterization, and sequence of an *Escherichia coli* heat shock gene, *htpX*. *J Bacteriol* 173 (1991) 2944-2953.
- [13] N. Shimohata, S. Chiba, N. Saikawa, K. Ito, Y. Akiyama, The Cpx stress response system of *Escherichia coli* senses plasma membrane proteins and controls HtpX, a membrane protease with a cytosolic active site. *Genes Cells* 7 (2002) 653-662.
- [14] Y. Akiyama, Quality control of cytoplasmic membrane proteins in *Escherichia coli*. *J Biochem* 146 (2009) 449-454.

- [15] S. Langklotz, U. Baumann, F. Narberhaus, Structure and function of the bacterial AAA protease FtsH. *Biochim Biophys Acta* 1823 (2012) 40-48.
- [16] M. Sakoh, K. Ito, Y. Akiyama, Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from *Escherichia coli*. *J Biol Chem* 280 (2005) 33305-33310.
- [17] M. Kaser, M. Kambacheld, B. Kisters-Woike, T. Langer, Oma1, a novel membrane-bound metallopeptidase in mitochondria with activities overlapping with the m-AAA protease. *J Biol Chem* 278 (2003) 46414-46423.
- [18] F.X. Gomis-Ruth, Structure and mechanism of metalloproteases. *Crit Rev Biochem Mol Biol* 43 (2008) 319-345.
- [19] B.W. Matthews, Structural basis of the action of thermolysin and related zinc peptidases. *Acc Chem Res* 21 (1988) 333-340.
- [20] I.V. Demidyuk, T.Y. Gromova, K.M. Polyakov, W.R. Melik-Adamyanyan, I.P. Kuranova, S.V. Kostrov, Crystal structure of the protealysin precursor: insights into propeptide function. *J Biol Chem* 285 (2010) 2003-2013.
- [21] T. Goulas, J.L. Arolas, F.X. Gomis-Ruth, Structure, function and latency regulation of a bacterial enterotoxin potentially derived from a mammalian adamalysin/ADAM xenolog. *Proc Natl Acad Sci U S A* 108 (2011) 1856-1861.
- [22] T. Guevara, I. Yiallouris, R. Kappelhoff, S. Bissdorf, W. Stocker, F.X. Gomis-Ruth, Proenzyme structure and activation of astacin metallopeptidase. *J Biol Chem* 285 (2010) 13958-13965.
- [23] A.K. Mohanty, M.C. Wiener, Membrane protein expression and production: effects of polyhistidine tag length and position. *Protein Expr Purif* 33 (2004) 311-325.
- [24] A.K. Mohanty, C.R. Simmons, M.C. Wiener, Inhibition of tobacco etch virus protease activity by detergents. *Protein Expr Purif* 27 (2003) 109-114.
- [25] J.M. Vergis, M.C. Wiener, The variable detergent sensitivity of proteases that are utilized for recombinant protein affinity tag removal. *Protein Expr Purif* 78 (2011) 139-142.
- [26] I. Manolaridis, K. Kulkarni, R.B. Dodd, S. Ogasawara, Z. Zhang, G. Bineva, N. O'Reilly, S.J. Hanrahan, A.J. Thompson, N. Cronin, S. Iwata, D. Barford, Mechanism of farnesylated CAAX protein processing by the intramembrane protease Rce1. *Nature* (2013).
- [27] F. Junge, B. Schneider, S. Reckel, D. Schwarz, V. Dotsch, F. Bernhard, Large-scale production of functional membrane proteins. *Cell Mol Life Sci* 65 (2008) 1729-1755.
- [28] G.G. Prive, Detergents for the stabilization and crystallization of membrane proteins. *Methods* 41 (2007) 388-397.
- [29] Z.E. Newby, J.D. O'Connell, 3rd, F. Gruswitz, F.A. Hays, W.E. Harries, I.M. Harwood, J.D. Ho, J.K. Lee, D.F. Savage, L.J. Miercke, R.M. Stroud, A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat Protoc* 4 (2009) 619-637.

- [30] R.M. Garavito, S. Ferguson-Miller, Detergents as tools in membrane biochemistry. *J Biol Chem* 276 (2001) 32403-32406.
- [31] A. Bernsel, H. Viklund, A. Hennerdal, A. Elofsson, TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res* 37 (2009) W465-468.
- [32] C. Cole, J.D. Barber, G.J. Barton, The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* 36 (2008) W197-201.

**Tables**

Table 1. Constructs used in this work

Name of construct *	Size of construct	Sequence
HtpX-4kDa-tag	328 residues; ~36 kDa	HtpX-G-S-P-G-L-Q-E-F-I-E-G-R-↓-H-H-H-H-H-H-I-D-E-E-Q-K-L-I-S-E-E-D-L-L-R-K-R
HtpX-2kDa-tag	313 residues; ~34 kDa	HtpX-G-S-P-G-L-Q-E-F-I-E-G-R-↓-H-H-H-H-H-H-H-H
His <sub>6</sub> -HtpX	~34 kDa	M-G-S-S-H-H-H-H-H-H-S-S-G-E-N-L-Y-F-Q-↓-G-P-HtpX
His <sub>6</sub> -MBP-HtpX	~75 kDa	M-K-H-H-H-H-H-H-P-MBP-S-G-S-G-S-G-S-E-N-L-Y-F-Q-↓-G-A-HtpX
His <sub>6</sub> -Ztag-HtpX	~42 kDa	M-K-H-H-H-H-H-H-P-Ztag-G-S-G-S-G-S-E-N-L-Y-F-Q-↓-G-A-HtpX
His <sub>6</sub> -GB1-HtpX	~41 kDa	M-K-H-H-H-H-H-H-P-GB1-G-S-G-S-G-S-E-N-L-Y-F-Q-↓-G-A-HtpX
His <sub>6</sub> -Trx-HtpX	~46 kDa	M-K-H-H-H-H-H-H-P-Trx-G-S-G-S-G-S-E-N-L-Y-F-Q-↓-G-A-HtpX
His <sub>6</sub> -NusA-HtpX	~91 kDa	M-K-H-H-H-H-H-H-P-NusA-G-S-G-S-G-S-E-N-L-Y-F-Q-↓-G-A-HtpX
His <sub>6</sub> -GST-HtpX	~56 kDa	M-K-H-H-H-H-H-H-N-T-S-S-N-S-GST-T-S-G-S-G-G-G-G-S-M-S-E-N-L-Y-F-Q-↓-G-A-HtpX
His <sub>6</sub> -Mistic-HtpX	~48 kDa	M-G-S-S-H-H-H-H-H-H-N-T-S-S-N-S-Mistic-G-S-G-S-E-N-L-Y-F-Q-↓-G-P-HtpX
HtpX-His <sub>8</sub>	301 residues; ~33 kDa	HtpX-H-H-H-H-H-H-H-H

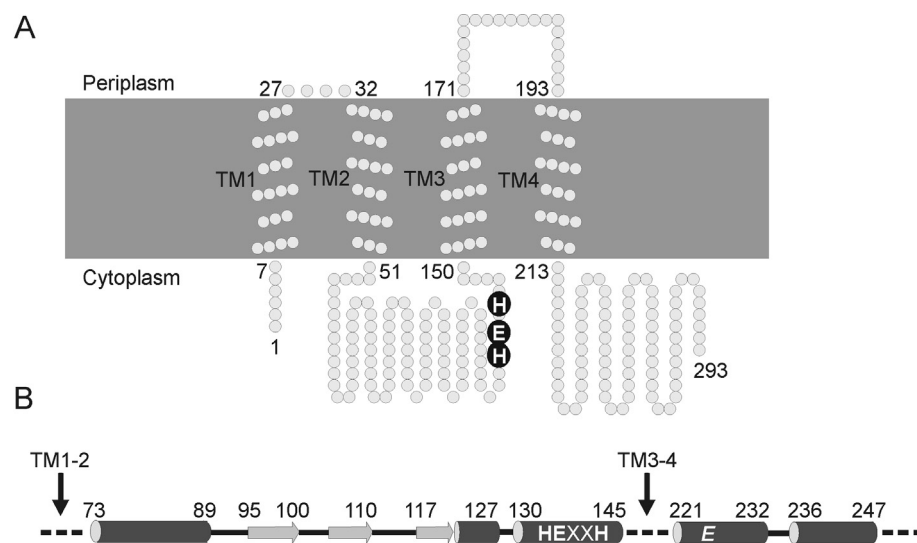
\* Constructs displayed from top to bottom by order of use. The arrow shows a factor Xa or TEV peptidase cleavage site.



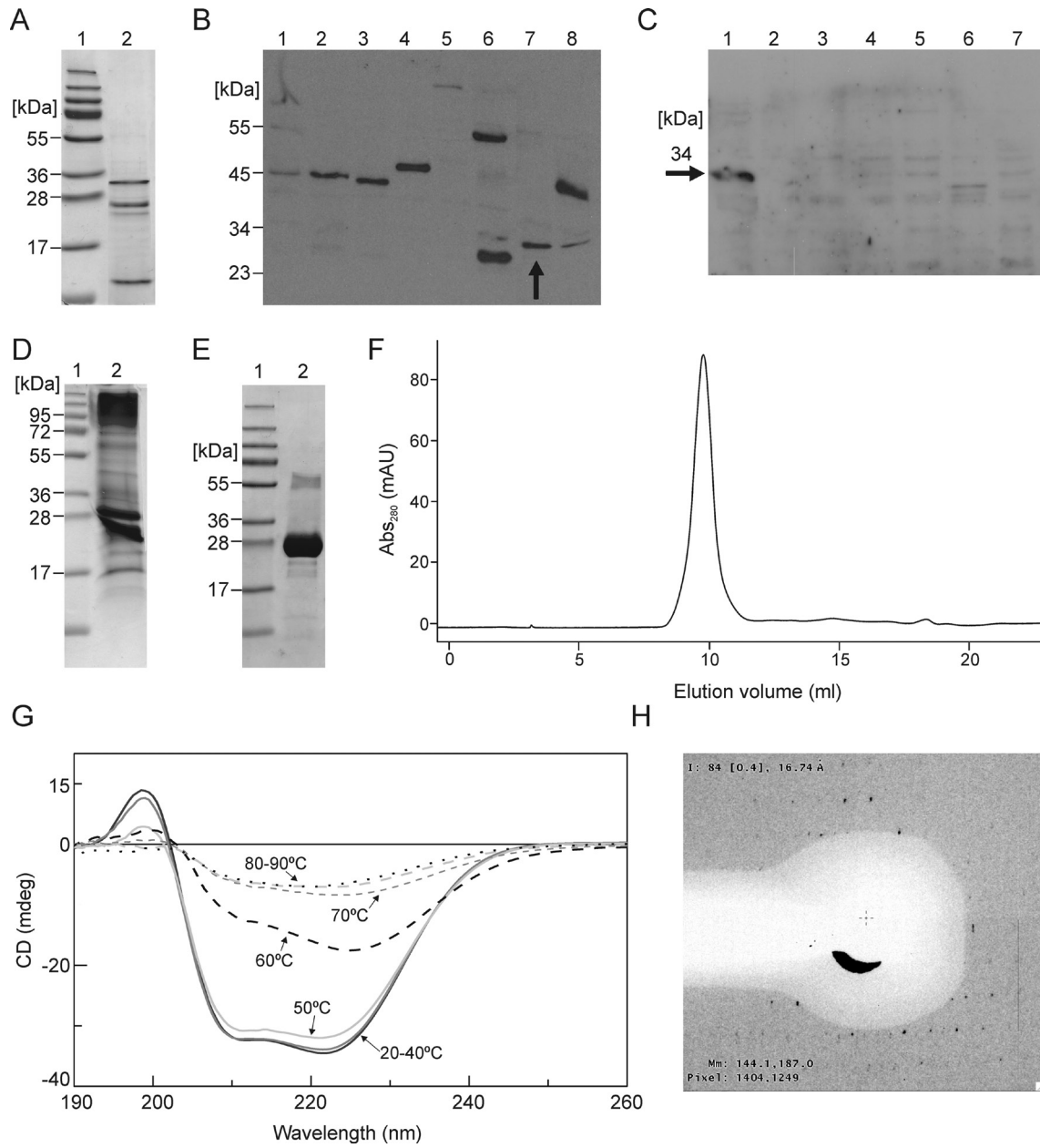
## Figure legends

Fig. 1. Topological architecture of HtpX. (A) The transmembrane topology of *Escherichia coli* HtpX (UP access code P23894) was obtained using the TOPCONS program [31] and represented using the TOPO2 server (<http://www.sacs.ucsf.edu/TOPO2/>). The N- and C-terminal residues as well as the flanking transmembrane (TM) residues and TM segments are labeled. The residues of the characteristic HEXXH zinc-binding motif are shown in black circles. (B) Suggested regular secondary structure elements of *E. coli* HtpX, based on the partial crystal structure of an artificial fragment of an orthologous protein (residues 57-157; PDB entry 3CQB) and bioinformatics predictions with Jpred3 [32], are indicated with rods (helices) and arrows (strands). The residues of the zinc-binding motif and the third putative zinc ion ligand are shown in regular and italic letters, respectively.

Fig. 2. Overexpression and purification of HtpX mutant E<sup>140</sup>A. (A) SDS-PAGE of wild-type HtpX-4kDa-tag after purification by nickel-affinity and anion-exchange chromatography (lane 2). PageRuler Plus Prestained Protein Ladder (Fermentas) was used as molecular mass standard (lane 1). Bands lower than HtpX-4kDa-tag correspond to self-cleavage products as assessed by peptide mass fingerprint analysis. (B) Western blotting of E<sup>140</sup>A HtpX-2kDa-tag constructs with an N-terminal His<sub>6</sub>-tag followed by (lanes 1-6 and 8) MBP, Ztag, GB1, Trx, NusA, GST, or Mystic fusion proteins and a TEV peptidase cleavage site. E<sup>140</sup>A HtpX-2kDa-tag without fusion protein is indicated by an arrow in lane 7. (C) Western blotting of HtpX orthologs from (lanes 1-7) *E. coli* (indicated by an arrow), *P. furiosus*, *A. aeolicus*, *M. kandleri*, *M. acetivorans*, *T. tengcongensis*, and *L. monocytogenes*. SDS-PAGE of E<sup>140</sup>A HtpX-His<sub>8</sub> after purification by nickel-affinity (D) or cobalt-affinity (E) chromatography. Protein ladder as in (A). (F) Size-exclusion chromatography of E<sup>140</sup>A HtpX-His<sub>8</sub> after purification by cobalt-affinity and anion-exchange chromatography in the presence of 1% OG. The protein elutes at 9.80 ml, which corresponds to ~62 kDa. The column was calibrated using blue dextran (void volume; 7.62 ml), conalbumin (75 kDa; 9.65 ml), ovalbumin (44 kDa; 10.30 ml), and cytochrome C (12 kDa; 13.37 ml). (G) Circular dichroism spectra of E<sup>140</sup>A HtpX-His<sub>8</sub> collected at temperatures ranging from 20 to 90°C. (H) Close-up view of an X-ray diffraction pattern from a crystal obtained in preliminary crystallization trials.



**FIG. 1.**



**FIG. 2.**