Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The endoproteinase furin contains two essential Ca^{2+} ions stabilizing its N-terminus and the unique S1 specificity pocket

Manuel E. Than,[a]* Stefan Henrich,[a]‡ Gleb P. Bourenkov,[b] Hans D. Bartunik,[b] Robert Huber[a] and Wolfram Bode[a]

[a]Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18A, 82152 Martinsried, Germany, and [b]Max-Planck-Arbeitsgruppen für Strukturelle Molekularbiologie (MPG-ASMB), Gruppe Proteindynamik, DESY, Notkestrasse 85, 22603 Hamburg, Germany

‡ Current address: EML Research gGmbH, Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg, Germany.

Correspondence e-mail: than@biochem.mpg.de

Received 1 November 2004
Accepted 24 January 2005

The mammalian prohormone/proprotein convertase (PC) furin is responsible for the maturation of a great variety of homeostatic but also many pathogenic proteins within the secretory pathway and the endosomal pathway and at the cell surface. Similar to other members of the PC family, furin requires calcium for catalytic activity. In a previous paper, the structural association of the catalytic and the P-domain of furin was shown and data were presented indicating two or three calcium-binding sites. The exact number and the three-dimensional localization of the essential calcium sites within furin have now been determined by collecting X-ray diffraction data on either side of the Ca *K* absorption edge and by calculating a novel type of double difference map from these anomalous scattering data. Two calcium ions were unambiguously identified: the purely structural Ca-1 also conserved in the bacterial digestive subtilisins and the Ca-2 site specific to PCs and essential for the formation of the P1 specificity-determining S1-binding pocket. In addition, these anomalous diffraction data show that no tightly bound K^+ sites exist in furin.

1. Introduction

Many secreted proteins, hormones, enzymes and neuropeptides are initially synthesized as inactive precursors or proproteins which need to be activated *via* controlled proteolytic cleavage by the prohormone/proprotein convertases (PCs). Seven members of the mammalian PC family (furin, PC1/3, PC2, PACE4, PC4, PC5/6 and PC7/8/LPC/SPC7) have been identified and found to be homologous to the yeast endoproteinase Kex2/kexin (for reviews, see Steiner, 1998; Khatib *et al.*, 2002; Thomas, 2002).

Furin (also called SPC1/PACE; EC 3.4.21.75; MEROPS clan SB, family S8), the best studied member of the mammalian PC family, is a type I transmembrane protein that is expressed in all tissues examined (Thomas, 2002). It is initially translated as a 794-residue pre-proenzyme (human; 793 residues in mouse) which is autocatalytically cleaved and activated in a Ca^{2+}- and pH-dependent manner (Leduc *et al.*, 1992; Anderson *et al.*, 1997) during its migration through the secretory pathway. Once activated, it cycles between the trans-Golgi network (TGN), the cell surface and endosomal compartments, where it processes a diverse collection of membrane-bound and soluble protein precursors of growth factors, receptors, plasma proteins, Alzheimer-related secretases and cancer-associated extracellular matrix proteinases (reviewed in Nakayama, 1997; Molloy *et al.*, 1999; Khatib *et al.*, 2002; Rockwell *et al.*, 2002). In addition to these physiologically essential substrates, furin also activates many bacterial toxin precursors (including precursors to *Pseuodomonas aeruginosa*, diptheria and anthrax toxins (Molloy *et al.*, 1999) and processes viral-envelope

glycoproteins [*e.g.* those from pathogenic Ebola strains (Volchkov *et al.*, 1998), HIV-1 gp160 (Hallenberger *et al.*, 1992) or the avian influenza virus haemagglutinin (Nakayama, 1997; Rockwell *et al.*, 2002)]. In addition to membrane-bound forms, naturally truncated shed furin species have been described (reviewed in Thomas, 2002).

Similar to the other PCs or kexin, mature furin has an absolute requirement for Ca^{2+} ions (Molloy *et al.*, 1992) for activity *in vitro* and *in vivo*. Furthermore, the effect of high potassium concentrations on the activity of furin and kexin has been investigated (Rockwell & Fuller, 2002). The three-dimensional structures of decanoyl-Arg-Val-Lys-Arg-chloromethylketone-inhibited mouse furin (Henrich *et al.*, 2003) and of its yeast homologue kexin in complex with an Ala-Lys-Arg boronic acid inhibitor (Holyoak *et al.*, 2003) or an Ac-Arg-Glu-Lys-Arg boronic acid inhibitor (Holyoak *et al.*, 2004) have recently been solved. Two and three Ca^{2+}-binding sites were described for these two enzymes, respectively. However, these Ca^{2+} sites were identified solely on the basis of the electron-density level, the coordination sphere of the putatively bound ions and on the basis of homology to the related bacterial subtilisins. Though often used, this method gives no absolute certainty about the identity of sites of high electron density. In addition, the thorough investigation of metal-binding sites in various natively occurring (*e.g.* Bode *et al.*, 1987; Teplyakov *et al.*, 1990; Gallagher *et al.*, 1995; Almog *et al.*, 2003) and engineered (Bryan, 2000) subtilisins resulted in quite different ion assignments for some of the sites, with a maximum of five calcium-binding sites identified in the highly resolved X-ray structure of 'calcium-loaded' *Bacillus sphaericus* sphericase (Almog *et al.*, 2003). Very recently, the number and spatial location of described Ca^{2+}-binding sites in subtilisin-like serine proteases has been expanded by the crystal structure of KP-43 (Nonaka *et al.*, 2004), which in addition to a C-domain resembling to some degree the P-domain of the PCs exhibits three novel Ca^{2+}-binding sites, which correspond to none of the ion-binding sites described for the subtilisins or the PCs.

Here, we show for the first time the definitive three-dimensional localization of bound Ca^{2+} ions for furin using anomalous difference data in the vicinity of the *K* absorption edge of calcium. The coordination spheres of the two calcium ions identified in furin are compared with kexin and with the bacterial subtilisins, suggesting different functions for these two bound metals.

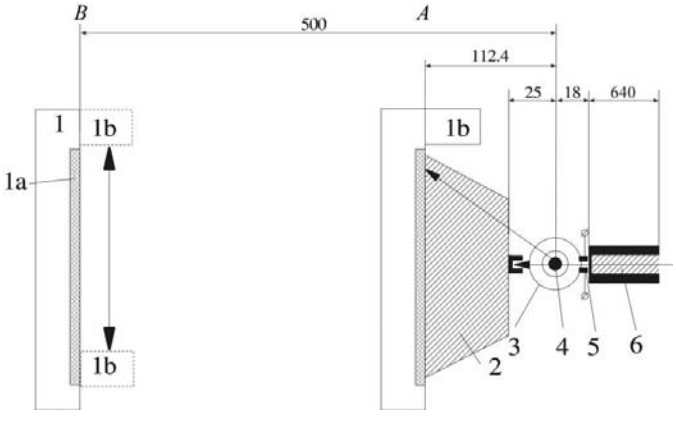

Figure 1
Schematic representation of the experimental setup of BW6/DORIS for long-wavelength measurements, showing the MAR 345 detector (1) with storage phosphor (1a) and its readhead (1b), the helium buffer (2) with Mylar windows (2 × 5 µm), the cryostream (MSC Xstream, operated at 100 K with N_2) (3), the sample position (4), the intensity monitor (2 mm long ionization chamber) (5) and the collimator (helium-filled, sealed with 5 µm Mylar) (6). The rotation exposures were taken at position *A*. Between the exposures, the detector was translated to the position *B* for scanning, which involves vertical translation of the readhead.

2. Materials and methods

2.1. Protein preparation and crystallization

The soluble ectodomain of mouse furin comprising the catalytic and P-domains was expressed from dihydrofolate reductase-amplified CHO cells (Cameron *et al.*, 2000) and purified and crystallized as described in Henrich *et al.* (2003). For this study, hexagonal crystals were grown in sitting drops by mixing equal volumes of protein (10 mg ml^{-1} in 20 m*M* HEPES–NaOH, 5 m*M* $CaCl_2$ pH 7.4) and precipitant solution [1.0 *M* $(NH_4)_2SO_4$, 0.4 *M* Li_2SO_4, 0.1 *M* sodium citrate, 6% 1,5-pentanediol pH 6.0] and equilibration against precipitant solution by the vapour-diffusion method. The crystals reached maximal dimensions of 200 × 200 × 100 µm and contained eight molecules per asymmetric unit. The crystals were transferred into perfluoropolyether, mounted in a cryoloop, transformed by exposing this loop to an air stream of 81% relative humidity (Kiefersauer *et al.*, 2000) and flash-frozen.

2.2. Beamline setup for long-wavelength data collection

Data collection was performed at the MPG-GBF beamline BW6/DORIS at Deutsches Elektronensynchrotron (DESY, Hamburg, Germany). This beamline was designed for X-ray fluorescence and diffraction measurements over a broad range of X-ray wavelengths extending from about 0.6 to 3.1 Å. Components that are essential for the soft X-ray regime include thin (beryllium and carbon) windows, an additional mirror for the suppression of harmonics and a helium-filled collimator system (for defining the aperture and divergence in the incident beam) that extends a short distance from the sample. The experiments described here were carried out using an Si(111) double-crystal monochromator which supresses the second-order harmonics. At nominal (first-order) wavelengths above ∼2.0 Å, the third- and fourth-order harmonic contaminations would be strongly noticeable in the diffraction pattern. The short (20 cm) quartz mirror which may be inserted into the X-ray beam between the double-focusing toroidal mirror and the sample position was adjusted to a glancing angle of 5 mrad, at which the third- and fourth-order components in the incident beam were suppressed. The quartz mirror provided >90% reflectivity and did not affect the focusing geometry.

Table 1
Data-collection statistics.

Values in parentheses are for the last shell.

Wavelength (Å)	3.15	2.99
Unit-cell parameters (Å, °)	$a = b = 135.5$, $c = 472.3$, $\alpha = \beta = 90$, $\gamma = 120$	
Space group	$P6_5$	
Resolution (Å)	50–3.9 (4.2–3.9)	
No. of reflections		
Measured	138969	140212
Unique (anomalous)	84127	82401
Completeness (%)	95.2 (92.6)	93.3 (95.7)
$R_{\rm sym}$†	0.15 (0.22)	0.095 (0.125)
$\langle I\rangle/\langle\sigma(I)\rangle$	4.8 (3.0)	7.3 (5.1)

† $R_{\rm sym} = \sum_h \sum_i |I_{i,h} - \langle I_h\rangle| / \sum_h \sum_i I_{i,h}$.

2.3. Data collection and processing

Data collection was performed on frozen crystals using a modified MAR 345 imaging-plate detector (MAR Research, Norderstedt, Germany) as depicted in Fig. 1. The front cover of the detector was removed and measurements were carried out in the dark. Both the collimator and the space between the crystal and the detector were filled with helium to minimize air absorption (for recent similar developments see, for example, Yang *et al.*, 2003; Polentarutti *et al.*, 2004). For each wavelength (λ = 3.15 and 2.99 Å), one angular segment of 55° was collected in frames of 0.5° with exposure times of 7 and 3 min, respectively. The data were processed using the *HKL* package (Otwinowski & Minor, 1997). The errors resulting from absorption were partly corrected by a local scaling procedure, which (similar to Matthews & Czerwinsky, 1975) defines an individual scale factor for every reflection by scaling the neighbouring reflections within a sphere (comprising at least 50 observed reflections) to a reference data set. The data collected previously (Henrich *et al.*, 2003) from the same crystal to 2.7 Å resolution at a wavelength of 1.05 Å were used as the reference. The data-processing statistics (after local scaling) are given in Table 1.

2.4. Determination of calcium sites and analysis

Model phases were calculated from a partially refined furin model obtained for the hexagonal crystal form ($R_{\rm cryst}$ = 22.2%, $R_{\rm free}$ = 25.0% for data in the resolution range 36–2.7 Å; Henrich *et al.*, 2003). Except for incomplete refinement and missing sugars, this model is essentially identical to, but of lower quality than, the triclinic model. The protein S^{δ} and S^{γ} atoms, as well as the Ca^{2+} ions and all solvent atoms of this structure, were excluded from phase calculations. Anomalous difference Fourier maps were calculated with *FFT* (Collaborative Computational Project, Number 4, 1994) for each data set and averaged with *MAPROT* (Collaborative Computational Project, Number 4, 1994) using the eightfold non-crystallographic symmetry inherent to the hexagonal crystals. Calcium-specific 'difference DANO maps' were calculated according to

$$\rho_{xyz} = \tfrac{1}{V}\sum_{hkl}[(F^{+}_{hkl,2.99} - F^{-}_{hkl,2.99}) - (F^{+}_{hkl,3.15} - F^{-}_{hkl,3.15})] \times \exp[-2\pi i(hx + ky + lz) + i(\alpha_{hkl} - \tfrac{\pi}{2})] \quad (1)$$

using *SFTOOLS* and *FFT* (Collaborative Computational Project, Number 4, 1994) and averaged similarly. All electron-density maps and structural comparisons were analyzed in *MAIN* (Turk, 1992) using this partially refined furin model and the following PDB-deposited structures: 1cse (subtilisin Carlsberg; Bode *et al.*, 1987), 1thm (thermitase; Teplyakov *et al.*, 1990), 1sup (subtilisin bpn′; Gallagher *et al.*, 1995), 1ea7 (*B. spaericus* extracelllular serine protease with five Ca^{2+} sites; Almog *et al.*, 2003), 1ot5 (kexin; Holyoak *et al.*, 2003) and 1r64 (kexin soaked in potassium malonate; Holyoak *et al.*, 2004). All structural pictures were created using using *MOLSCRIPT* (Kraulis, 1991), *BOBSCRIPT* (Esnouf, 1999) and *RASTER3D* (Merritt & Bacon, 1997).

3. Results

3.1. Data collection with soft X-rays at BW6

Measurements in the vicinity of the *K* absorption edge of calcium (3.07 Å) are close to or beyond the physical limit of most beamlines typically used for X-ray crystallography [for recent discussions and reviews on the use of soft(er) X-rays in protein crystallography refer to, for example, Carpentier *et al.*, 2000; Weiss *et al.*, 2001; Dauter, 2002; Helliwell, 2004]. This situation was further complicated by the large *c* axis of the weakly scattering hexagonal furin crystals. Nevertheless, useful signals could be obtained at BW6, which has specific equipment for the soft X-ray regime (for details, see §2.2 and Fig. 1). The change between the regular and long-wavelength setup can be performed at BW6 within an hour. The measurements of the long-wavelength and the reference short-wavelength data are possible within one and the same beamtime shift, without remounting the sample.

As estimated from the ionization-current measurements, the beamline delivers a flux density of 5×10^9 and 10^9 photon s^{-1} mm^{-2} at wavelengths of 2.99 and 3.15 Å, respectively. In contrast to previous experiments on BW6 in the long-wavelength range with trypsin crystals (Hauschild *et al.*, 2000), it was not possible to observe interpretable diffraction patterns at $\lambda \geq 3$ Å with the weakly diffracting furin crystals using the MAR CCD detector and exposure times up to 20 min per degree of rotation. This was mostly owing to absorption by iron impurities in the beryllium window of the CCD detector. A significant improvement in the detection efficiency has been achieved by using the MAR 345 imaging-plate detector after removal of its carbon front cover and by partly eliminating the air absorption of the diffracted beam by introducing a helium buffer that extended immediately towards the surface of the detector. This allowed a fivefold reduction in dose to yield a similar signal-to-noise ratio in the highest resolution shell. Using the setup depicted schematically in Fig. 1 and exposure times of 7 and 3 min per 0.5° of rotation at λ = 3.15 and 2.99 Å, respectively (Ca *K* absorption edge $\pm$ 0.08 Å, $\simeq$ 100 eV), weak diffraction

patterns were observed. The wavelengths were chosen to be close to but sufficiently distant from the exact energetic location of the edge to safely allow for the shift in the exact energetic location of the edge between pure elements and protein-bound ions. The precise choice of the wavelengths very close to the absorption edge was not required for the present experiment, but might be determined from a fluorescence scan in future applications of this method that use white lines. At $\lambda = 3.15$ Å, the averaged intensity, integrated over the peak area (4×4 pixels), was 350 units[1] over the background of 4100 units (integrated over the same area for all data between 50 and 3.9 Å) and 140/1800 units in the last resolution shell (4.2–3.9 Å). For $\lambda = 2.99$ Å, the corresponding ratios were 980/8800 (all data) and 440/4500 (last shell). The relatively poor accuracy in the data collected at $\lambda = 3.15$ Å (Table 1) is because of the low flux density at this wavelength. In both data sets, the signal-to-noise level was close to that expected from counting statistics. The estimates of the strength of the raw anomalous signal, calculated by averaging $|F^+ - F^-|/\sigma(|F^+ - F^-|)$ values in the resolution shells, significantly exceeded unity only at low resolution (1.5 in the 50–8 Å shell) for the $\lambda = 2.99$ Å data set and was below 1.2 in all resolution shells for $\lambda = 3.15$ Å. The fraction of $|F^+ - F^-|/\sigma(|F^+ - F^-|) > 3$ was below 1% in both data sets.

3.2. Determination of bound Ca^{2+} and other metal ions

The first data set was collected from the hexagonal furin crystals (Henrich *et al.*, 2003) at a wavelength of 2.99 Å in order to observe anomalous signals for calcium and also for lighter elements such as potassium or sulfur (Fig. 2). In fact, several well defined peaks could be observed in the corresponding anomalous difference Fourier map (Fig. 3*a*) above the absolute value of the highest dubious peak at 7σ. Except for two (at 37σ and 30σ, respectively), all these peaks (between 10σ and 15σ) could be interpreted by the anomalous scattering contribution of S atoms. The 'difference DANO map' calculated according to (1) using anomalous differences from both data sets (*i.e.* $\lambda = 2.99$ and 3.15 Å) showed two strong peaks at 23σ and 17σ as the only signals above the absolute value of the first noise peak at 7σ (Fig. 3*b*). These peaks corresponded to the two non-sulfur sites seen in the anomalous difference map using only the 2.99 Å data.

Since all contributions to the anomalous signal from elements other than calcium are practically eliminated by (1), these two sites definitely correspond to bound calcium ions. Furthermore, the data show that no other light metal ion heavier than sulfur (*e.g.* no K^+) is bound to the hexagonal furin crystals at any significant level. Since the anomalous signal of sulfur is easily visible in Fig. 3(*a*), an anomalous contribution from any reasonably occupied K^+ site would have resulted in an easily recognizable signal. However, all non-sulfur sites of the anomalous difference Fourier map calculated from the 2.99 Å data are calcium, excluding the presence of further tightly bound elements heavier than sulfur in the hexagonal crystals.

[1] Digital units output by MAR 345 software on a Cartesian image.

3.3. Structure and coordination sphere of the two Ca^{2+} sites

The Ca-1 site is situated towards the northeast of the catalytic domain (if viewed in the standard orientation with the substrate/inhibitor proceeding from left to right through the horizontally oriented active-site cleft; Fig. 3*b*) and shows a fairly regular pentagonal bipyramidal (sevenfold) coordination of the central Ca^{2+} ion with O atoms at distances of 2.2–2.6 Å (Fig. 4*a*). The five equatorial coordination sites are made up by the side-chain O atoms of Asp162 (OD1 and OD2), Asn208, Asp115 and the main-chain carbonyl O atom of Gly212, while the two apical sites are occupied by the main-chain carbonyl O atoms of Val205 and Val210. All main-chain carbonyls belong to the Cα2–Cβ4 loop, which is well conserved within the PCs (Henrich *et al.*, 2005) and shows a highly similar conformation in the subtilisins. Considering the charges of the Ca^{2+} ion and the two acidic side chains involved, an overall net charge of ± 0 can be assigned to this site.

The Ca-2 site is located below the S1-specificity pocket and shows a more complex coordination of its calcium ion. The first coordination sphere is created by the side-chain O atom of Glu331 (OE1 and OE2), Asp258 and Asp301 as well as three internal water molecules, again creating a pentagonal bipyramidal coordination geometry (Fig. 4*b*). These residues are part of an extended hydrogen-bonding network including additional polar main- and side-chain atoms, additional water molecules and even the guanidino moiety of the bound P1 arginine side chain. Considering the three acidic side chains directly contacting the Ca^{2+} and the nearby Asp306, which is also located at the bottom of the S1 pocket, an overall net charge of -2 can be calculated. This site seems to harbour an extremely high concentration of negative charges, which are apparently neutralized to some degree by the tightly bound Ca^{2+} ion. Interestingly, the residues forming this Ca-2 site and

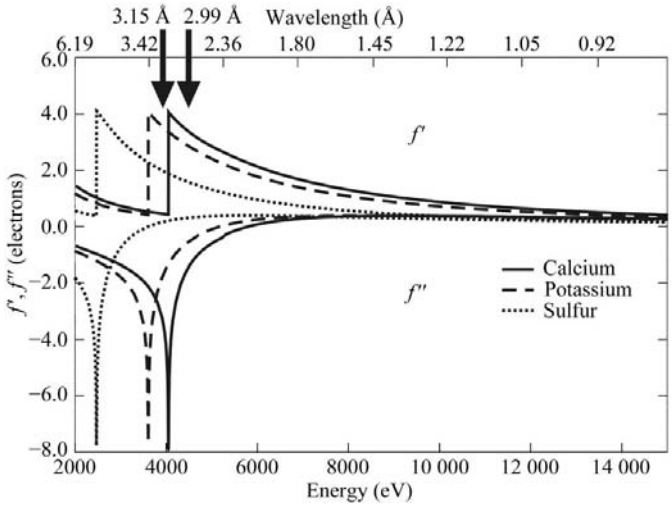

Figure 2
Anomalous scattering terms f' and f'' of calcium (solid line), potassium (dashed line) and sulfur (dotted line) as function of the X-ray energy and wavelength. The theoretical values were calculated using the website http://www.bmsc.washington.edu/scatter/. The energetic location of the two wavelengths used is indicated.

the bottom of the S1 pocket also participate in part in the interface to the PC-specific P-domain and are not conserved in the bacterial subtilisins.

4. Discussion

X-ray crystallography in the wavelength range 2–3 Å is complicated mainly by two physical limitations. The strong increase in absorption in the beamline and detector windows, along the air path and in the sample reduces the X-ray intensity incident on the detector. Secondly, the quantum-detection efficiency of typically used detectors (CCD and image plate) rapidly decreases (Tate *et al.*, 1995; Amemiya, 1995). Nevertheless, weak but clearly interpretable data could be obtained even for the weakly diffracting furin crystals, emphasizing the suitability of BW6 for measurements at the Ca *K* absorption edge. The absorption of low-energy X-rays was minimized by the beamline design (see §2.2) and by replacing the air paths inside the collimator and from the crystal to the detector with helium. Futhermore, local scaling to a reference data set collected at a significantly lower wavelength proved to be essential to correct for the strong effects of X-ray absorption by the crystal. This scaling procedure employing a common reference also ensured that both data sets were on the same scale for the calculation of the difference DANO map (see below) without the need for any additional scaling procedure.

The identification of sites of high electron density during the interpretation of diffraction data is a frequently occurring problem in protein crystallography. Such sites are often assigned based on the presence of potential ligands in the protein or crystallization buffers. If the presence of several different ions or ligands is suspected or if ions and/or ligands might have been carried along through the protein purification, other methods have to be applied. High-resolution diffraction data and the consideration of coordination geometry and of metal–ligand distances often allow a reasonable distinction between certain well known metal ions (Bode *et al.*, 1987; Yang *et al.*, 2002; Harding, 2004). The bond-valence method provides a generalized numerical approach for identifying metal ions, but requires a resolution of better than 1.5 Å for the X-ray data (Müller *et al.*, 2003). In addition, *B*-factor and/or occupancy refinements of the individual atoms in question can help to distinguish between different possible ligands. However, this latter technique might be of limited use if the atoms in question have very similar atomic numbers, if the real occupancy is unclear and/or if the diffraction data are not well resolved. The widespread use of tunable synchrotron radiation has given rise to several studies where optimized anomalous scattering data for the element in question have been used to calculate anomalous difference Fourier maps, showing the spatial location of the anomalously scattering atoms and facilitating several element distinctions (see, for example, Einspahr *et al.*, 1985; Tereshko *et al.*, 2001; Weiss *et al.*, 2001, 2002; DeLa Barre & Brunger, 2003; Steinbacher *et al.*, 2004; Ferreira *et al.*, 2004). The first application of such a method permitted distinction between Mn^{2+} and Ca^{2+} ions in pea lectin crystals (Einspahr *et al.*, 1985). In general, the contrast between two different elements *A* and *B* is enhanced as function of the ratio $(f''_A/f_{oA})/(f''_B/f_{oB})$, where $f'' = f''(\lambda)$ is the imaginary part and f_o is the wavelength-independent real part of the atomic scattering factors. Assuming X-ray wavelengths below the *K* absorption edge of all elements in question, this contrast is enhanced twofold to threefold for the distinction between heavier (*e.g.* Zn^{2+}, Mn^{2+}) and lighter (*e.g.* Ca^{2+}, K^+) ions and tenfold again for the distinction between the lighter ions and water. This contrast remains constant over a broad wavelength range and does not generally permit the distinction between elements that have similar atomic numbers (such as K^+ and Ca^{2+}). For such distinctions and for a direct assignment of the nature of completely unknown ions, the exact energetic location of the absorption edge of the element in question has to be considered for data collection. Even though the ratio between imaginary and real parts of the form factor of a given element (f''_A/f_{oA}) is largest at the maximum of the X-ray absorption curve, it also shows a significant anomalous scattering contribution (see above) at much lower wavelength (higher energy, see also Fig. 2). Therefore, the use of anomalous differences that are measured at one wavelength only might result in wrong assignments and only the comparison of the anomalous scattering contributions determined at either side of the absorption edge and measured in close vicinity to the edge can unequivocally identify a specific element (Dobbek *et al.*, 1999, 2002). The present application followed such an approach. The 'difference DANO map' [as given for the Ca *K* edge in (1)] was calculated from anomalous data recorded on either side of the Ca *K* edge and resulted directly in a spatially resolved and element specific electron-density map. This application was complicated by the energetic location of this edge in the soft X-ray regime. In general, the essentially same approach can be followed for any element that has an X-ray absorption edge that can be accessed for diffraction data collection.

The 'difference DANO map' calculated according to (1) clearly showed two calcium sites located within each of the eight furin molecules in the asymmetric unit of the hexagonal crystals used in this study. Owing to the relatively high concentration of (a maximum of) 1.4 *M* SO_4^{2-} ions in the final equilibrated mother liquor, some of the 5 m*M* Ca^{2+} ions that were added together with the protein solution should have been precipitated. This raises the question whether all physiologically relevant Ca-binding sites were still occupied in the crystals. Based on the solubility constant of $CaSO_4$ in water (2.45×10^{-5} mol^2 l^{-2}), the concentration of free Ca^{2+} was at least 17 μ*M* in the crystalline state. Presumably, the Ca^{2+} binding is strongly enhanced in the crystalline state and in the presence of the inhibitor as seen previously for the Ca^{2+}-binding sites that (under comparable experimental conditions) were fully occupied in trypsin (Bode & Schwager, 1975). Therefore, all Ca^{2+}-binding sites essential for the activation of furin (at 1–3 m*M* Ca^{2+}; Molloy *et al.*, 1992) should at least be partially occupied under these conditions. All significant peaks in the anomalous difference Fourier map that did not correspond to these two Ca^{2+} sites could be interpreted as occupied

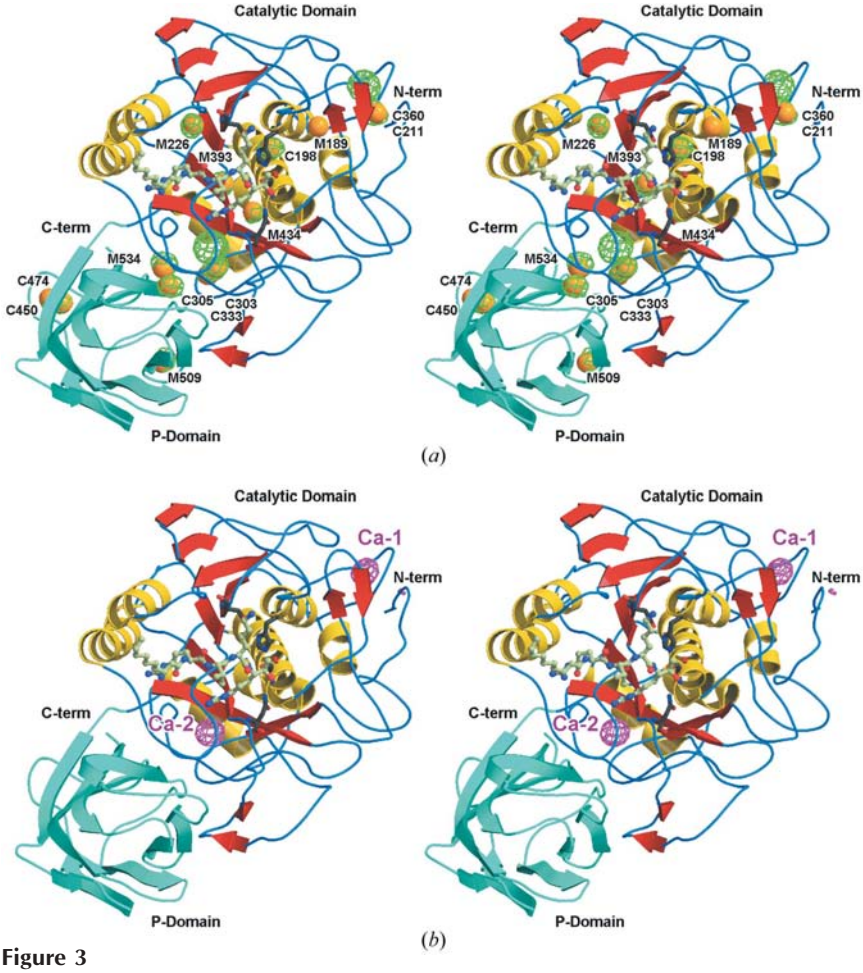

Figure 3
Stereo ribbon plot of soluble furin together with (*a*) the anomalous difference Fourier density map determined at 2.99 Å (contoured at 6σ, green) and all S atoms (yellow balls) and (*b*) the Ca-specific 'difference DANO map' (contoured at 6σ, purple). For both panels, the catalytic domain is shown as yellow helices, red β-strands and dark blue irregular structures, while the P-domain is depicted entirely in light blue. The active-site residues (dark grey) and the dec-RVKR-cmk inhibitor residues (light-coloured ball-and-stick model) are shown with all non-H atoms. The methionine and cysteine residues of furin as well as the two bound Ca^{2+} ions are labelled in (*a*) and (*b*), respectively.

by sulfur, precluding the presence of any tightly bound K^+ ions.

This latter statement might be challenged by the fact that no K^+ was added to the buffer used for crystallization. However, during expression all required constituents were present in the medium and most chemicals used for the purification contain micromolar traces of this metal ion, so that at least very tightly bound K^+ ions should have been visible. Therefore, it can be concluded that only the two identified Ca^{2+} ions and no K^+ ions exist as essential high-affinity metal-binding sites within the proprotein convertase furin. Further, low-affinity or surface-located ion-binding sites might be present as found in kexin, where a maximum of three Ca^{2+}- and four K^+-binding sites have been described (Holyoak *et al.*, 2004) in crystals that were grown in the presence of higher SO_4^{2-} and hence lower resulting Ca^{2+} concentrations and were soaked in high concentrations of potassium malonate prior to flash-freezing. In fact, a third site of somewhat unusual electron density [corresponding to the suggested Ca-401 binding site in subtilisin Carlsberg (Bode *et al.*, 1987) and the Ca-3 site of kexin (Holyoak *et al.*, 2003)] had been observed in the fully refined $2F_{obs} - F_{calc}$ electron-density map of furin (Henrich *et al.*, 2003). However, its coordination sphere did not allow for a clear assignment at that time. Based on the current data, no Ca^{2+} or K^+ is bound at this site. Hence, a tightly bound water molecule (average *B* factor of 5.2 Å^2 for the eight molecules per asymmetric unit of the triclinic crystals) or alternatively an Na^+ ion (which cannot unequivocally be distinquished from water with the available diffraction data) should be bound at this site, both of which are equally consistent with the refined $2F_{obs} - F_{calc}$ electron-density map.

Various binding sites for metal ions described for furin, kexin and selected bacterial digestive subtilisins are compared in Fig. 5. The Ca-1 site, which is located close to the N-terminus of the catalytic domain, is conserved among all published structures. In subtilisin, its deletion by mutations still results in active protein, which however exhibits significantly lower stability (Strausberg *et al.*, 1995). In furin and in the other PCs, this Ca^{2+} seems by analogy to subtilisin to be responsible for structural stabilization. In fact, Asp115 is tightly bound to Ca-1, which apparently results in the tight attachment of the N-terminal residues to the rest of the catalytic domain. Modelling of an inactive pro-furin chimera (Henrich *et al.*, 2005) shows that the amino acids preceding Gln120 must deviate from their conformation in mature furin to allow binding of the Arg-Ala-Lys-Arg108 'primary activation cleavage segment' into the specificity subsites of the active-site cleft. The requirement for calcium during maturation of furin under physiological conditions (Anderson *et al.*, 1997) might be explained by the necessity of forming this Ca-1 site during the final folding steps of the mature proteinase after the first activating cleavage.

Of all the compared subtilases, the Ca-2 site is unique to furin and kexin. Based on the conservation of the liganding residues in all PCs (data not shown) and the striking differences from the bacterial subtilisins, which precluded the correct modelling of this site based on subtilisin structures, the presence of the Ca-2 binding site seems to be one of the key differences between the highly specific PCs and the digestive subtilisins. This site is surrounded by a remarkably high concentration of negatively charged amino-acid side chains, resulting in a delicate balance between the electrostatic repulsion between all the negative charges and the attraction to the positively charged Ca-2 and the arginine side chain at

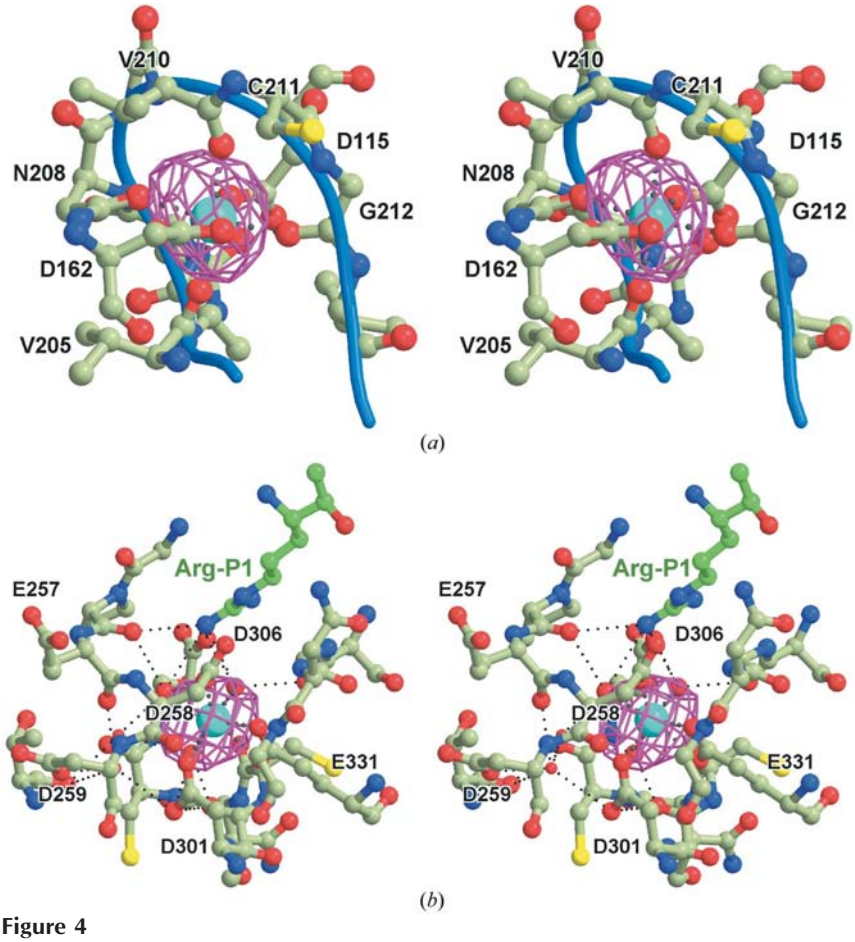

Figure 4
Details of calcium coordination in furin. The two defined calcium ions (cyan balls) are shown together with the coordinating residues in ball-and-stick representation (grey, carbon; blue, nitrogen; red, oxygen) and the Ca-specific 'difference DANO map' (contoured at 6σ, purple) for the Ca-1 (*a*) and the Ca-2 (*b*) sites. Contacts to the first coordination sphere are indicated with grey dotted lines. In (*a*), the Cα2–Cβ4 loop is also shown as a blue backbone worm. In (*b*), hydrogen bonds of the extended network including the P1 aginine (green) and several water molecules (red balls) are indicated by black dotted lines. Some key coordinating residues are labelled for orientation.

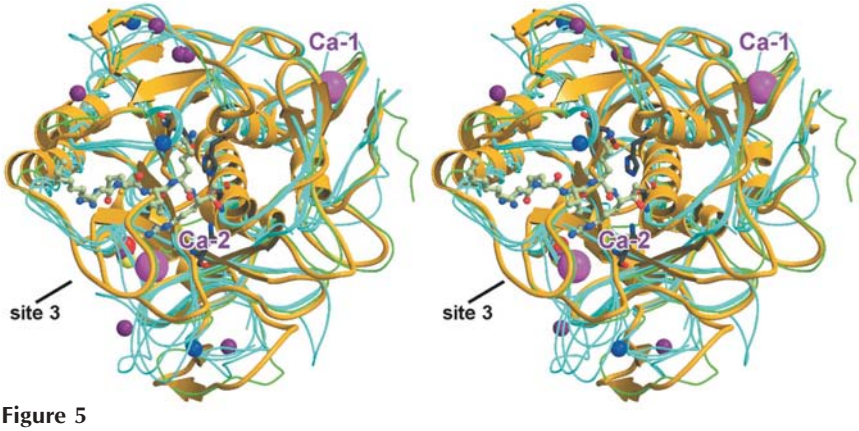

Figure 5
Metal ion-binding sites in furin and various homologues. The catalytic domain of furin is shown entirely in gold in the same ribbon representation as in Fig. 3. The two determined Ca^{2+} sites and the water molecule of furin corresponding to the third Ca^{2+} site in kexin are depicted as large pink and small red balls, respectively, and labelled. The bacterial digestive subtilases (thermitase, subtilisin bpn′, subtilisin Carlsberg, *B. spaericus* extracelllular serine protease with five Ca^{2+} sites) and kexin (two structures) are shown as cyan and green C^{α} worms, respectively (for PDB codes and references, see §2). The bound calcium, potassium and sodium ions described for the other structures are indicated as small balls in dark purple, dark blue and yellow, respectively.

P1. Hence, the Ca-2 located directly below the S1 specificity pocket might be required to stabilize the high negative electrostatic potential at this site, which is essential to confer the PCs with their high degree of substrate specificity.

We thank I. Lindberg for the furin preprations used for crystallization and for carefully reading this manuscript. This work was supported by DFG grants SFB596 (to WB and MET) and TH 862/1-1 to MET.

References

Almog, O., González, A., Klein, D., Greenblatt, H. M., Braun, S. & Shoham, G. (2003). *J. Mol. Biol.* **332**, 1071–1082.
Amemiya, Y. (1995). *J. Synchrotron Rad.* **2**, 13–21.
Anderson, E. D., VanSlyke, J. K., Thullin, C. D., Jean, F. & Thomas, G. (1997). *EMBO J.* **16**, 1508–1518.
Bode, W., Papamokos, E. & Musil, D. (1987). *Eur. J. Biochem.* **166**, 673–692.
Bode, W. & Schwager, P. (1975). *J. Mol. Biol.*, **98**, 693–717.
Bryan, P. N. (2000). *Biochim. Biophys. Acta*, **1543**, 203–222.
Cameron, A., Appel, J., Houghten, R. A. & Lindberg, I. (2000). *J. Biol. Chem.* **275**, 36741–36749.
Carpentier, P., Berthet-Colominas, C., Capitan, M., Chesne, M.-L., Fanchon, E., Lequien, S., Stuhrmann, H., Thiaudiere, D., Vicat, J., Zielinski, P. & Kahn, R. (2000). *Cell. Mol. Biol.* **46**, 915–935.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z. (2002). *Curr. Opin. Struct. Biol.* **12**, 674–678.
DeLa Barre, B. & Brunger, A. T. (2003). *Nature Struct. Biol.* **10**, 856–863.
Dobbek, H., Gremer, L., Kiefersauer, R., Huber, R. & Meyer, O. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 15971–15976.
Dobbek, H., Gremer, L., Meyer, O. & Huber, R. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8884–8889.
Einspahr, H., Suguna, K., Suddath, F. L., Ellis, G., Helliwell, J. R. & Papiz, M. Z. (1985). *Acta Cryst.* B**41**, 336–341.
Esnouf, R. M. (1999). *Acta Cryst.* D**55**, 938–940.
Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J. & Iwata, S. (2004). *Science*, **303**, 1831–1838.
Gallagher, T., Gilliland, G., Wang, L. & Bryan, P. (1995). *Structure*, **3**, 907–914.
Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H. D. & Garten, W. (1992). *Nature (London)*, **360**, 358–361.
Harding, M. M. (2004). *Acta Cryst.* D**60**, 849–859.
Hauschild, J., Bourenkov, G. P. & Bartunik, H. D. (2000). *Acta Cryst.* A**56**(*Suppl.*), s232.
Helliwell, J. R. (2004). *J. Synchrotron Rad.* **11**, 1–3.

Henrich, S., Cameron, A., Bourenkov, G. P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W. & Than, M. E. (2003). *Nature Struct. Biol.* **10**, 520–526.
Henrich, S., Lindberg, I., Bode, W. & Than, M. E. (2005). *J. Mol. Biol.* **345**, 211–227.
Holyoak, T., Kettner, C. A., Petsko, G. A., Fuller, R. S. & Ringe, D. (2004). *Biochemistry*, **43**, 2412–2421.
Holyoak, T., Wison, M. A., Fenn, T. D., Kettner, C. A., Petsko, G. A., Fuller, R. S. & Ringe, D. (2003). *Biochemistry*, **42**, 6709–6718.
Khatib, A. M., Siegfried, G., Chretien, M., Metrakos, P. & Seidah, N. G. (2002). *Am. J. Pathol.* **160**, 1921–35.
Kiefersauer, R., Than, M. E., Dobbek, H., Gremer, L., Melero, M., Strobl, S., Dias, J. M., Soulimane, T. & Huber, R. (2000). *J. Appl. Cryst.* **33**, 1223–1230.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Leduc, R., Molloy, S., Thorne, B. A. & Thomas, G. (1992). *J. Biol. Chem.* **267**, 14303–14308.
Matthews, B. W. & Czerwinsky, E. W. (1975). *Acta Cryst.* A**31**, 480–497.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Molloy, S. S., Anderson, E. D., Jean, F. & Thomas, G. (1999). *Trends Cell Biol.* **9**, 28–35.
Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R. & Thomas, G. (1992). *J. Biol. Chem.* **267**, 16369–16402.
Müller, P., Kopke, S. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 32–37.
Nakayama, K. (1997). *Biochem. J.* **327**, 625–635.
Nonaka, T., Fujihashi, M., Kita, A., Saeki, K., Ito, S., Horikoshi, K. & Miki, K. (2004). *J. Biol. Chem.* **279**, 47344–47351.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Polentarutti, M., Glazer, R. & Carugo, K. D. (2004). *J. Appl. Cryst.* **37**, 319–324.
Rockwell, N. C., Krysan, D. J., Komiyama, T. & Fuller, R. S. (2002). *Chem. Rev.* **102**, 4525–4548.
Rockwell, N. C. & Fuller, R. S. (2002). *J. Biol. Chem.* **277**, 17531–17537.
Strausberg, S. L., Alexander, P. A., Gallagher, D. T., Gilliland, G. L., Barnett, B. L. & Bryan, P. N. (1995). *Biotechnology*, **13**, 669–673.
Steinbacher, S., Schiffmann, S., Bacher, A. & Fischer, M. (2004). *Acta Cryst.* D**60**, 1338–1340.
Steiner, D. F. (1998). *Curr. Opin. Chem. Biol.* **2**, 31–39.
Tate, M. W., Eikenberry, E. F., Barna, S. L., Wall, M. E., Lowrance, J. L & Gruner, S. M. (1995). *J. Appl. Cryst.* **28**, 196–205.
Teplyakov, A. V., Kuranova, I. P., Harutyunyan, E. H., Vainshtein, B. K., Frommel, C., Hohne, W. E. & Wilson, K. S. (1990). *J. Mol. Biol.* **214**, 261–279.
Tereshko, V., Wilds, C. J., Minasov, G., Prakash, T. P., Maier, M. A., Howard, A., Wawrzak, Z., Manoharan, M. & Egli, M. (2001). *Nucleic Acid. Res.* **29**, 1208–1215.
Thomas, G. (2002). *Nature Rev. Mol. Cell Biol.* **3**, 753–766.
Turk, D. (1992). PhD thesis. Technische Universität München.
Volchkov, V. E., Feldmann, H., Volchkova, V. A. & Klenk, H. D. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5762–5767.
Weiss, M. S., Panjikar, S., Nowak, E. & Tucker, P. A. (2002). *Acta Cryst.* D**58**, 1407–1412.
Weiss, M. S., Sicker, T., Djinovic Carugo, K. & Hilgenfeld, R. (2001). *Acta Cryst.* D**57**, 689–695.
Yang, C., Pflugrath, J. W., Courville, D. A., Stence, C. N. & Ferrara, J. D. (2003). *Acta Cryst.* D**59**, 1943–1957.
Yang, W., Lee, H.-W., Hellinga, H. & Yang, J. S. (2002). *Proteins*, **47**, 344–356.