FEBS Letters

FEBS PRESS science publishing by scientists

Co-expression with chaperones can affect protein 3D structure as exemplified by loss-of-function variants of human prolidase

Elżbieta Wątor*, Maria Rutkiewicz[§], Manfred S. Weiss and Piotr Wilk*

Macromolecular Crystallography, Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin, Germany

Correspondence
P. Wilk, Malopolska Centre of Biotechnology, Jagiellonian University, Gronostajowa 7a Str., Krakow, 30-387, Poland
Tel: +48 12 664-61-06
E-mail: wilk.piotr@uj.edu.pl

Present addresses
*Malopolska Centre of Biotechnology, Jagiellonian University, ul. Gronostajowa 7a, Krakow, 30-387, Poland
[§]Macromolecular Structure and Interaction, Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, Berlin, 13125, Germany

(Received 6 May 2020, revised 19 June 2020, accepted 23 June 2020, available online 14 July 2020)

doi:10.1002/1873-3468.13877

Edited by Peter Brzezinski

Prolidase catalyzes the cleavage of dipeptides containing proline on their C terminus. The reduction in prolidase activity is the cause of a rare disease named 'Prolidase Deficiency'. Local structural disorder was indicated as one of the causes for diminished prolidase activity. Previous studies showed that heat shock proteins can partially recover prolidase activity *in vivo*. To analyze this mechanism of enzymatic activity rescue, we compared the crystal structures of selected prolidase mutants expressed in the absence and in the presence of chaperones. Our results confirm that protein chaperones facilitate the formation of more ordered structures by their substrate protein. These results also suggest that the protein expression system needs to be considered as an important parameter in structural studies.

Databases
The reported crystal structures and their associated structure factor amplitudes were deposited in the Protein Data Bank under the accession codes 6SRE, 6SRF, and 6SRG, respectively.

Keywords: chaperones; loss of function; mutation; prolidase deficiency; protein misfolding; rare diseases

Prolidase (PEPD, Xaa-Pro dipeptidase, ECnonbreakingspace3.4.13.9) is the only known enzyme in humans capable of hydrolyzing dipeptides containing an imino acid on their C terminus [1]. As a result, free proline is produced, which can be salvaged for reuse by the cell. This activity is of particular importance in collagen metabolism [2–4]. Decreased or absent prolidase activity frequently leads to the development of a syndrome referred to as prolidase deficiency (PD), which is characterized by a broad spectrum of clinical symptoms with the most common ones being chronic skin ulcerations, recurring respiratory system infections, and mental retardation [4–8]. PD is a rare recessive genetic disorder with only a little over a hundred patients diagnosed to date (www.orpha.net; ORPHA: 742). In PD patients, 35 mutations could be mapped to the PEPD gene including 16 missense mutations and nine insertions/deletions [9]. A series of eight mutations that result in single amino acid deletion or substitution has recently been analyzed by X-ray crystallography. In three of the resulting crystal structures, namely in the structures of the Ser202Phe, Gly287Asn, and Gly448Arg variants of the enzyme, significant level of protein disorder in the region winding around the active site was observed [10].

Abbreviations
ADP, atomic displacement parameters; LOF, loss of function; PD, prolidase deficiency; wt, wild-type.

The prevention or reversion of structural disorder as a cure for a disease has been little studied, especially at the molecular level. In spite of this, one could envision that a drug molecule, such as for instance a shape-specific antibody, an aptamer or a small molecule, has a stabilizing effect on the native structure. In fact, such small molecular agents referred to as pharmaceutical chaperones [11] have been investigated for a number of targets related to protein misfolding [12,13] resulting in several drug candidates and approved drugs [14]. Yet, the development of such a stabilizing agent is usually a tedious project [11–13,15] and the utmost importance of structural biology was underlined in such investigations [16].

Another possibility is to help the protein to fold correctly from the very beginning, that is, during translation, to assume a more stable structure. In 2013, Besio *et al*. [17] showed that the induction of the expression of Hsp70 and Hsp90 in human fibroblasts led to protein stabilization and partial recovery of prolidase activity (up to 40% depending on the mutant in question). Molecular chaperones are specialized proteins that assist cotranslational protein folding. They also play a role in refolding of misfolded proteins and can help in dissolving protein aggregates and target misfolded proteins for degradation [18]. Several approaches to modulate chaperone activity were developed and investigated primarily in neurodegenerative disorders [19,20], and the pharmacological induction of chaperone proteins was recently shown to be a potent way of ameliorating amyloid-like aggregation involving protein kinase Cγ [21]. Chaperone co-expression is often also advantageous for heterologous protein expression [22–24]. All these effects of chaperones have their basis in the functioning of chaperones as folding catalysts, usually with a rather broad substrate spectrum. However, to the best of our knowledge, there is no study, which had analyzed the impact of the presence or absence of chaperones during protein expression on the 3D structure of proteins.

In this study, following the report of Besio *et al*. as well as our previous crystallographic studies, we analyzed the crystal structures and the activity profiles of three selected, structurally disordered prolidase variants co-expressed in the presence of chaperones, and compared our findings to previous results based on protein preparations from an expression background which is devoid of increased chaperone activity. We demonstrate a significant effect of the chaperones on both the final protein structure and its activity.

Materials and methods

Protein expression and purification

The three selected prolidase variants were expressed in *Escherichia coli* Arctic Express (DE3) (Agilent, Waldbronn, Germany) cells. The cells were transformed with a pET28a plasmid bearing the prolidase gene with the desired mutation. The cells were grown in TB at 37 °C. When the OD_{600} value was about 1.0, the temperature was decreased to 10 °C and protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM. The cells were left overnight and then harvested by centrifugation and suspended in lysis buffer (50 mM Tris/HCl pH 7.8, 300 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, 10 mM β-ME). They were then disrupted by sonication (15 min, 5-s pulse/3-s pause cycles), and the soluble fraction was separated from cell debris by centrifugation (45 min, 4 °C, 53 000 ***g***) and loaded onto an equilibrated 5 mL HisTrap column (GE Healthcare Europe GmbH, Freiburg, Germany). Unspecifically bound contaminants were washed from the affinity column with buffer A [50 mM Tris/HCl pH 7.8, 200 mM NaCl, 40 mM imidazole, 5% (v/v) glycerol, 5 mM β-ME] until no further signal decrease was observable. Specifically bound protein was eluted with buffer B [50 mM Tris/HCl pH 7.8, 200 mM NaCl, 400 mM imidazole, 5% (v/v) glycerol, 5 mM β-ME] and dialyzed against storage buffer (50 mM Tris/HCl pH 7.8, 200 mM NaCl, 5 mM β-ME). During dialysis, 2 mg of TEV protease per 100 mg of prolidase was added. The dialyzed protein was concentrated and subjected to size exclusion chromatography using a Superdex 200 (16/600) pg column (GE Healthcare Europe GmbH) equilibrated with storage buffer. A single peak corresponding to prolidase was collected, concentrated to ~ 50 $mg{\cdot}mL^{-1}$, and used for crystallization trials or aliquoted and flash-frozen in liquid nitrogen (LN_2) for later use. Prolidase expression in Rosetta (DE3) cells followed the same procedure and was previously described [10,25].

Protein crystallization

Purified prolidase was crystallized as previously reported [10,25]. Protein solution in storage buffer (13–18 $mg{\cdot}mL^{-1}$) was set up in sitting drop 96-well Intelli low-profile crystallization plates as a 1 : 1 mix with reservoir solution (10 mM sodium tetraborate and 720–1050 mM sodium citrate, pH: 7.5–8.5). Well-shaped 3D protein crystals of ~ 100 μm in the longest direction appeared after 3–6 days of incubation at room temperature in most wells on the plate. The best crystals were mounted using litholoops, cryopreserved in 1200 mM sodium citrate, 20% (v/v) glycerol, supplemented with 20 mM $MnCl_2$ and 20 mM GlyPro, and flash-cooled in LN_2.

Diffraction data collection and structure solution and refinement

X-ray diffraction data were collected at beamline BL14.1 of the BESSY electron storage ring operated by the Helmholtz-Zentrum Berlin. The data were recorded at 13.5 keV, which is the critical energy of the beamline's insertion device [26,27]. The phase problem was solved by molecular replacement using the program PHASER [28] and the wild-type (wt) prolidase structure (PDB-ID 5M4J [25]) as a search model. Optimal model placement in the asymmetric unit was verified with the program ACHESYM [29]. In order to ensure reproducibility, for each variant expressed from both expression backgrounds several (≥ 5) crystals originating from different crystallization drops were analyzed by X-ray diffraction. Since no significant differences could be observed between crystals from the same variant and the same expression background, only one model per variant and per expression background was subjected to full refinement and reported here. Structures were submitted to several model rebuilding and refinement cycles using COOT [30] and phenix.refine [31,32]. Refined models and reduced experimental data were deposited in the Protein Data Bank (PDB) [33,34] under the accession codes 6SRE, 6SRF, and 6SRG. Relevant data collection and refinement statistics are summarized in Table S1 for the structures derived from Arctic Express cells. For the structures derived from Rosetta cells, the relevant parameters have been reported previously [10,25].

Structure analysis and comparison

Structural analysis and comparisons were executed in COOT [30] and PYMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC, New York, NY, USA), and atomic displacement parameter (ADP) variabilities were analyzed with the program Baverage from the CCP4 suite [35]. The CαRMSD for the regions of interest was plotted in Figs 2–4, and the full list of numerical values is reported as Table S2. Normalized B factor values were obtained by dividing the local ADP by the average ADP for the entire model using the values derived from Baverage. The phenix.ensemble_refinement [36] was run for each analyzed model using the implementation in Phenix GUI with default parameters and the deposited model as an input. The residue average root mean square fluctuations (RMSFs) for atom coordinates and ADPs in the final ensemble model were calculated using ens_rmsf command from PYMOL ens_tools plugin [37] and were plotted for the region of interest.

Protein stability test

In order to investigate protein stability, the melting temperatures of all protein variants were determined by a thermal shift assay (TSA). The experiment was performed as described previously [38]. The protein solution (4 $mg \cdot mL^{-1}$) was incubated with 1 : 500 diluted Sypro orange dye and in an assay buffer (50 mM Tris, pH 8.0, 250 mM NaCl, 1 mM $MnCl_2$). The fluorescence signal (λex = 492 nm, λem = 610 nm) from Sypro orange was determined as a function of temperature between 5 and 95 °C in increments of 0.5 $°C \cdot min^{-1}$. The melting temperature was calculated as the negative inflection point of the fluorescence curve. Each experiment was performed in triplicate.

Prolidase activity assay

A prolidase activity assay was performed according to the previously described method [39], but was adjusted to measurements using a plate reader. In short, 50 µL of 10 µM of the appropriate prolidase variant in 50 mM Tris, pH 8.0, 250 mM NaCl, and 1 mM $MnCl_2$ buffer was incubated at 37 °C in the presence of 20 µL of 250 mM GlyPro. After 30 min, the reaction was stopped by the addition of 135 µL of 100% solution of TCA and 150 µL of Chinard's reagent. The mix was then incubated for 5 min at 95 °C. Denatured protein samples were centrifuged for 5 min at 17 000 ***g***, and 100 µL of each sample was transferred onto a transparent 96-well plate and the absorbance at 515 nm was measured. To make sure that no more than 5% of substrate was consumed, the results were compared against a calibration curve prepared by measuring the absorption from increasing proline concentrations in 5 mM HCl-treated alike protein samples. During data analysis, the values were normalized to the activity of wt prolidase. Each experiment was performed in triplicate.

Results

Wild-type hsProl—the reference

Wild-type human prolidase is a homodimer composed of two subunits. Each monomer harbors two domains with active sites located at the bottom of so-called pita-bread fold, which is characterized by two highly bent β-sheets flanked by four α-helices [40]. It forms a deep cleft at the bottom of which two manganese ions are bound. Upon substrate binding, the active site is sealed from the top by a flexible helix from the opposing N-terminal domain [25]. Figure 1 gives an overview of the 3D structure of prolidase. From the electron density shown in the inset of Fig. 1, it is clearly evident that all three mutation sites discussed in this work are completely ordered in the wt structure.

Ser202Phe—mild stabilization

The analysis of the Ser202Phe prolidase variant expressed in Rosetta cells (Prol_S202F_Ros) shows

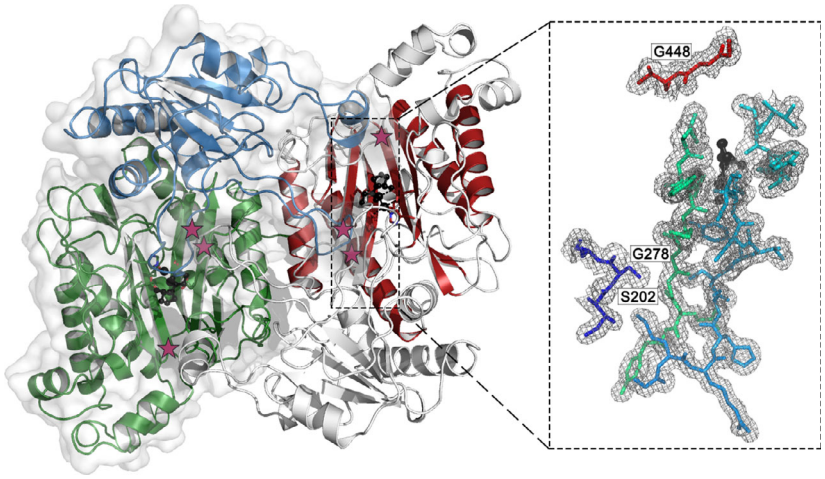

Fig. 1. Structural overview of human prolidase. A C_2-symmetric homodimer of prolidase is drawn with cartoon representation. In addition, a semi-transparent surface representation is shown for monomer A. Monomer A is colored according to its domain organization with the N-terminal domain in blue and the catalytic C-terminal domain in green. Monomer B is colored in light gray, while the pita-bread motif is highlighted in red. Substrate molecules and metal ions are shown in black ball-and-stick representation. The stars indicate the relative positions of the three discussed mutations. The enlargement shows a fragment of human prolidase shown in stick representation, which forms part of the active site cavity and is directly affected by the discussed mutations. The mutation sites are indicated by residue names and numbers, and the electron density map shown is a composite omit electron density map contoured at 1.0 σ. The enlarged fragment is colored blue (N-) to red (C-) following the chain trace to ease the localization of the described mutants. The substrate and metal ions are drawn as black ball-and-stick representation to indicate the location of the active sites. The image was prepared based on the PDB entry 5M4J [24].

that the introduction of the bulky, hydrophobic phenylalanine side chain replacing a small serine pushes the neighboring β-strand (encompassing residues 271–281) which harbors two glycine residues (277–278). The effect of this shift is amplified in the next strand (residues 237–248) where the electron density map becomes discontinuous and main chain could not be traced completely (Fig. S2). It can also be observed that the ADP values are increased significantly for the entire strand, including the two maxima around residues 240 and 260 (Figs 2C and S2). This region is also characterized by highly elevated flexibility as indicated by coordinate and B factor RMSF obtained from ensemble refinement analysis. This part of the chain harbors the residues Tyr241 and His256, which were previously shown to play a role in the stabilization of the active site and the transition state during catalysis. In the structure obtained from the protein produced in Arctic Express cells (Prol_S202F_ArEx), the ADP values are also increased for one of the strands, but the increase is not as strong as in the Prol_S202F_Ros case and electron density remains continuous and well defined even for the Arg side chains protruding over to the opposite subunits (Fig. S2). Also, the flexibility is lowered and concentrated around two residues (240 and 260) rather than for the entire region (Fig. 2D,E). The mutation caused a significant decrease in protein stability expressed as $T_{M,}$ but no significant differences in this regard were observed due to expression system (Fig. 5A).

Gly278Asp—significant stabilization

Gly278 is located in one of the β-strands (271–281) next to the active site of prolidase. The introduction of an Asp side chain in this position disrupts the trace of the neighboring strand (237–248), which leads to significant disorder in the proximity of the active site of the enzyme expressed in Rosetta cells (Prol_G278D_Ros). The loss of electron density details hampers the chain tracing for part of this strand (Fig. S3). Concomitantly, the ADP values are increased for one of the strands and part of the chain around 252–263 has been displaced (Figs 3A,C, and S3). The large increase in protein flexibility for these strands and loops connecting them is also evident from the RMSF plots (Fig. 3D,E). For the structure of the protein produced in Arctic Express (Prol_G278D_ArEx), much clearer electron density can be observed, as well as lower ADP values for some parts, such as for instance for residue Leu274 or Arg237. Above all, the chain fragments 252–263 become completely ordered and follow the trace expected from the wt HsProl (Figs 3 and S3). These differences are also reflected in the B normalized

plot (Fig. 3C). The coordinate and B factor fluctuations remain elevated yet to much lesser extent than in the Prol_G278D_Ros model. Interestingly, the point of substitution differs significantly between two models. In the Prol_G278D_Ros model, the main chain seems unaltered around Asp278 and the aspartate itself displays elevated ADPs for side chain only. In contrast, in the Prol_G278D_ArEx model the entire Asp278 and its neighboring residues display significantly higher than average ADPs which are also reflected as a sharp peak on the B normalized plot (Figs 3C and S3). A striking difference is the conformation of the introduced side chain. In the Rosetta model, it is directed toward the neighboring strand, while in Arctic Express one it is swung outwards at the cost of an increase in the local ADP. Apparently, this allows for a better overall preservation of the β-sheet structure (Fig. 3A). Regardless of the expression system, the T_M of G278D variant is decreased with respect to the wt protein (Fig. 5A).

Gly448Arg—no structural differences

The substitution of Gly448 by an Arg residue leads to the displacement of a large portion of polypeptide chain connecting two antiparallel β-strands forming one side of the pita-bread fold. This is the largest disorder described for prolidase variants covering a dozen residues and also affecting flexibility of active site residues (e.g., Tyr241 and crucial His255). Surprisingly, it has relatively little effect on the overall protein stability (Fig. 5A). In this case, no significant structure stabilization was achieved by co-expression with chaperonins as the crystal structure of protein obtained from Arctic Express is virtually identical with the one obtained from Rosetta cells (Fig. 4). Above others, it maintains structural disorder manifested as diffuse electron density, gaps in the model (Fig. S4), and large increase in protein flexibility (Fig. 4D,E).

In vitro enzymatic activity

In order to estimate the effect of the chaperone on the final structure and function of prolidase, the relative enzymatic activities of all prolidase preparations *in vitro* were analyzed. In the enzymatic assay, the amount of proline released in a time unit was measured by a colorimetric reaction using Chinard's reagent. The result obtained for wt prolidase was taken as the reference and defined as 100% (see the Materials and methods section for more details). No significant proline release was measured in neither of the Ser202Phe preparations, showing that the chaperone activity was not able to rescue the enzymatic function of the enzyme, despite its effect on the 3D structure. For the other two variants, a higher proline release was measured for the prolidase preparations from Arctic Express compared with the preparations from Rosetta cells (Fig. 5B). No significant impact on the overall protein stability was observed, and the reaction was conducted at least 8 °C below T_M of the least stable variant. This demonstrates that the chaperone activity was able to partially rescue the enzymatic activity of prolidase.

Discussion

Human prolidase plays an important role in proline turnover. Despite PD being a rare disease with estimated occurrence of 1–2 cases per million births, more than 30 mutations were associated with PD, OMIM 170100 [17], and new ones are still being identified (Natália Duarte Linhares, personal communication). Several studies showed the potential increase in total prolidase activity in cultured fibroblasts [41–44] and various therapeutic therapies were investigated [9,45–51], but to date no efficient treatment was introduced. In recent years, we have structurally characterized wt prolidase [25] and a series of eight single amino acid substitution or deletions [10] and we could show that one of the major causes of loss of function (LOF) was structural destabilization induced by bulky side chains introduced in places of small ones. Previously, Besio *et al.* showed that the induction of the chaperonins Hsp90 and Hsp70 in cultured fibroblasts derived from patients leads to partial rescue of prolidase activity. The effect seems to be case-dependent and varies from negligible to about 40% [17]. This could potentially improve the patients' well-being by for example changing disease manifestations from acute to mild. LOF variants of proteins are generally little studied, and their activation poses bigger challenge that inhibition of a given activity. Nevertheless, several approaches have been investigated with some success [16]. One of the approaches was the development of so-called pharmaceutical chaperones [12], but to the best of our knowledge activation of natural protein chaperones as a way of stabilizing LOF enzymes has not been widely tested and such effect was never proved by means of structural biology.

In this study, the aim was to investigate the effect of co-expression of chaperones with several LOF prolidase variants. Our primary focus was on structure stabilization. Since chaperones are known to facilitate proper protein folding, we have selected three mutants for which significant structural alterations could be

identified. For the remaining ones, the stabilizing effect observed by Besio *et al*. could not be linked to protein structure. Also, the wt protein which was shown to be active and fully folded was limited to the standard expression protocol, since we did not expect any further influence of the chaperone-rich expression background. The simplest system allowing us to obtain protein expressed in the presence of chaperones in an amount sufficient for crystallization was *E. coli* Arctic Express strain. This strain constitutively expresses a variant of cold-adapted chaperonins (Cpn10 and Cpn60) from *Oleispira antarctica* [24,23], which are homologues of *E. coli* GroEL/GroES. Similar to human Hsp70/Hsp90, Cpn10 and Cpn60 can utilize a broad spectrum of substrates and require ATP for the activity. By using the *E. coli* expression system and identical purification and crystallization protocols as previously, we minimized the source of differences between the analyzed crystal structures and we have good reasons to believe that the presence of chaperones during protein expression is the main source of observed effects.

In all three analyzed structures, a small amino acid side chain was replaced by a bigger, bulkier side chain which led to the repulsion of a neighboring main chain segment, ultimately leading to its disordering. This effect was limited to a fragment encompassing two of the active site-forming antiparallel β-strands (Arg237-Glu248 and Met271-Tyr281) and the loop connecting them, also containing elements forming the active site. Interestingly, the Phe side chain in the Ser202Phe variant collides with the glycines 278 and 279, that is, with

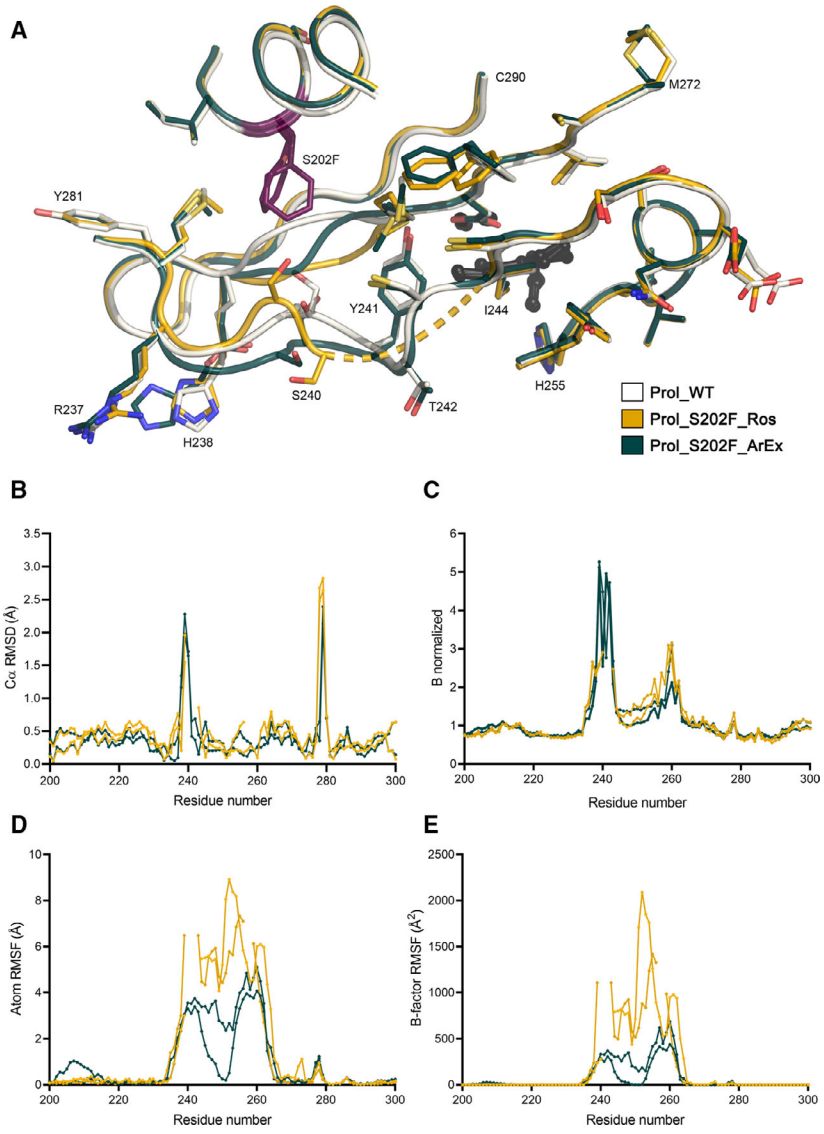

Fig. 2. Comparison of the Ser202Phe variant of prolidase derived from two different expression systems with the wt prolidase structure. (A) Superposition of mutant models (Prol_S202F_Ros—yellow, Prol_S202F_ArEx—green) with wt human prolidase in light gray. The site of substitution is highlighted in purple, and the substrate is shown in black. The discontinuity in the Prol_S202F_Ros model comprising residues 241–243 is approximated by a dashed line. (B) Differences between Ca atoms (CαRMSD) based on the superposition of the mutant model with wt human prolidase. (C) Variability of normalized B factors drawn for discussed region in Prol_S202F models. (D, E) RMSF of average residue coordinates and ADPs, respectively, based on the ensemble refinement. The coloring scheme is the same as for the structural models. Each line on the plot represents a single protein chain.

a fragment, where the second substitution (Gly278Asp) occurs. Analyzing the differences in the main chain trace with respect to wt prolidase structure, in both cases one notes a two distinctive maxima of Cα deviation from wt structure around residue 278 (primary interaction or substitution) and around residue 240, where polypeptide is pushed by displaced 277–279 fragments (Figs 2B and 3B). The destabilizing effect of the substitutions is further seconded by an observed decrease in the thermal stability of the variants, even though the expression strain used had negligible effect on the T_M (Fig. 5A). Crystal structures are time and space averages of all the molecules building a given crystal and as such may not fully represent the local heterogeneity of the sample. Ensemble refinement is an approach utilizing molecular dynamic simulation restrained by the experimental component derived from the X-ray diffraction experiment and can be used to estimate local structural heterogeneity. This method has been employed to all of the discussed model, and it was observed that the fragment encompassing two strands, referenced to throughout the text, and a loop connecting them (roughly residue range 235–281) exhibit a slightly elevated flexibility in the wt prolidase (Fig. S1C). In the case of all the variants, this flexibility is highly enhanced as indicated by RMSF values calculated based on the phenix.ensemble_refinement. It is of note that this enhancement is significantly smaller in case of variants expressed in Arctic Express cells (see panels D and E on Figs 2–4).

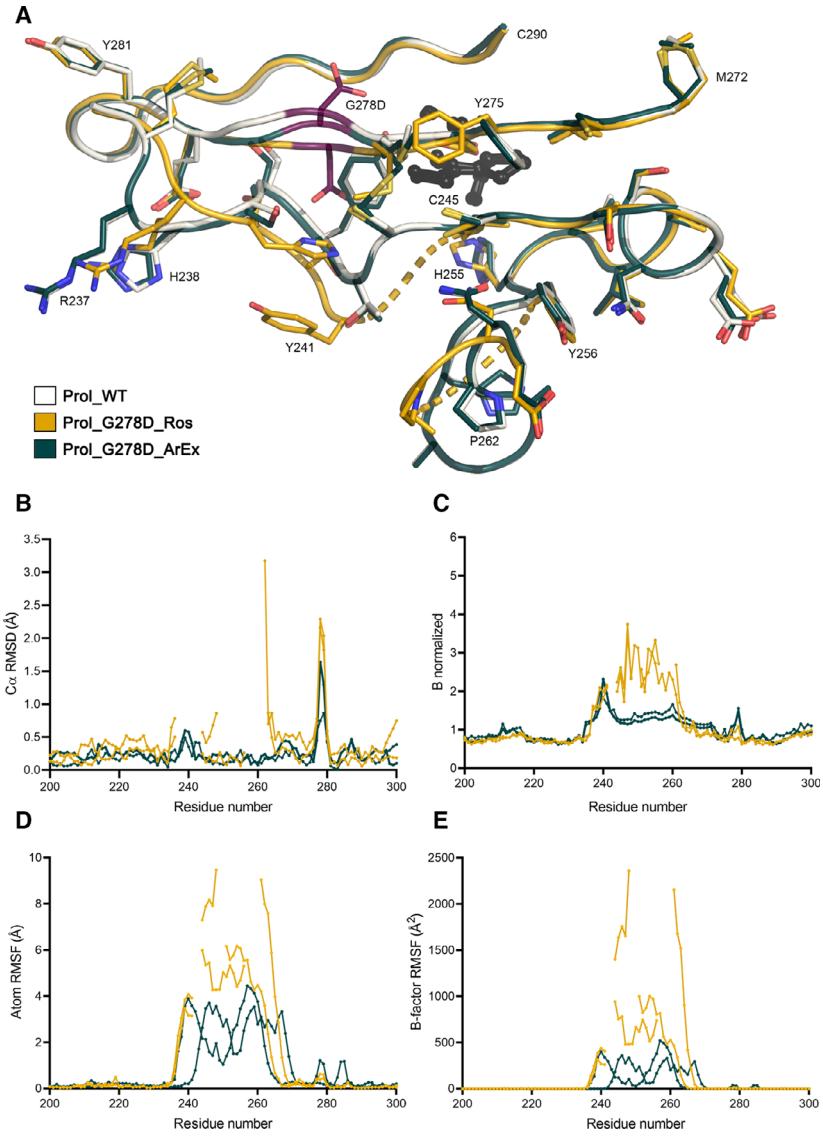

Fig. 3. Comparison of the Gly278Asp variant of prolidase derived from two different expression systems with the wt prolidase structure. (A) Superposition of mutant models (Prol_G278D_Ros—yellow, Prol_G278D_ArEx—marine) with wt human prolidase in light gray. Point of substitution is highlighted in purple, and substrate is shown in black. Discontinuity in the model is approximated by a dashed line (residues 242–243 and residues 257–260 for Prol_G278D_Ros). (B) Differences between Ca atoms (CαRMSD) based on the superposition of the mutant model with wt human prolidase. (C) Variability of normalized B factors drawn for discussed region in Prol_S202F models. (D, E) RMSF of average residue coordinates and ADPs, respectively, based on the ensemble refinement. The coloring scheme is the same as for the structural models. Each line on the plot represents single protein chain.

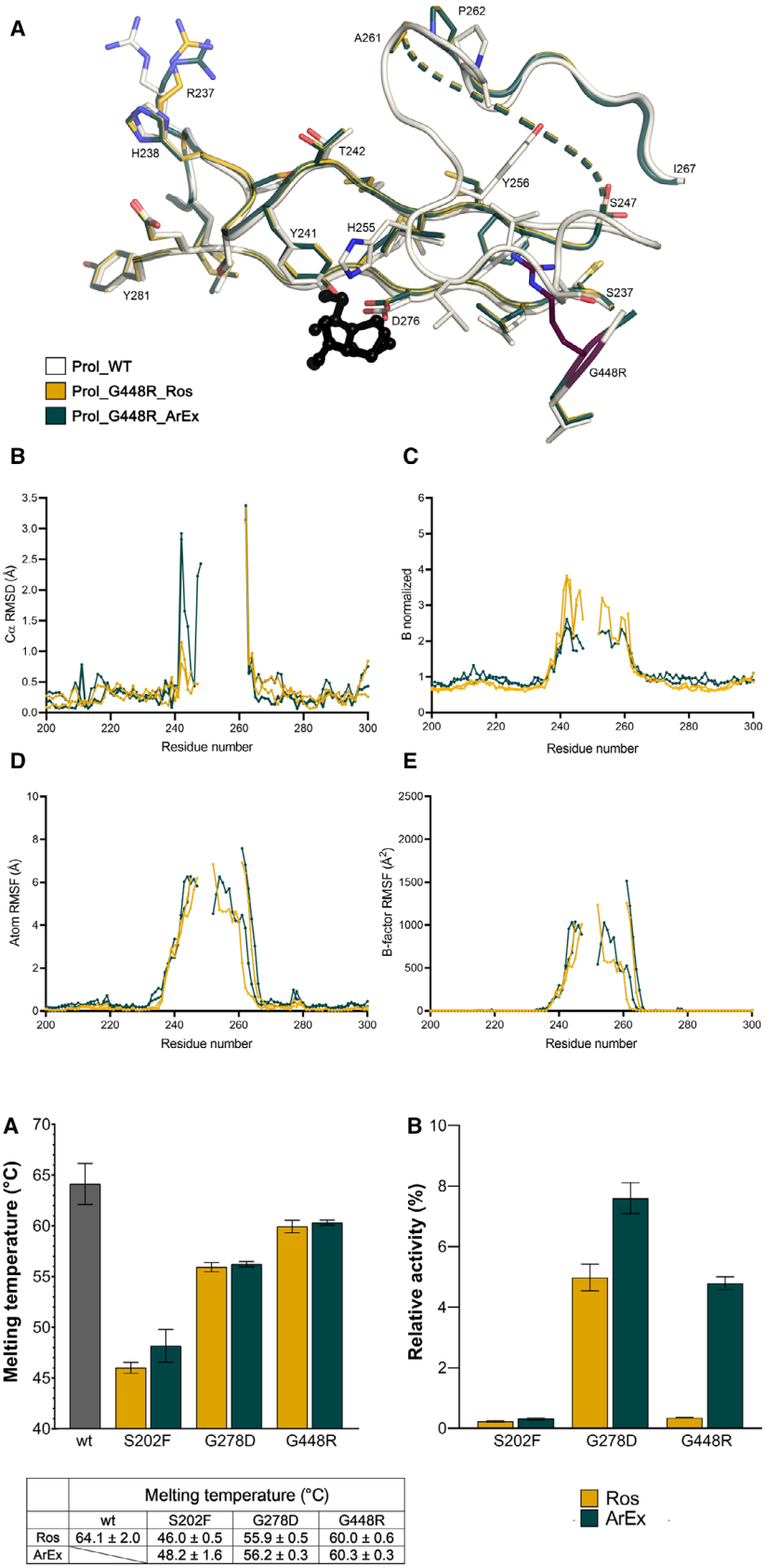

Fig. 4. Comparison of the Gly448Arg variant of prolidase derived from two different expression systems with the wt prolidase structure. (A) Superposition of mutant models (Rosetta—yellow, Arctic Express—marine) with wt (white) human prolidase. Point of substitution is highlighted in purple, and substrate is shown in black. Discontinuity in the model is approximated by a dashed line (residues 248–260 for both Prol_G448R_Ros and Prol_G448R_ArEx). (B) Differences between Ca atoms (CαRMSD) based on the superposition of the mutant model with wt human prolidase. (C) Variability of normalized B factors drawn for the discussed region in the Prol_G448R models. (D, E) RMSF of average residue coordinates and ADPs, respectively, based on the ensemble refinement. The coloring scheme is the same as for the structural models. Each line on the plot represents single protein chain.

	Melting temperature (°C)			
	wt	S202F	G278D	G448R
Ros	64.1 ± 2.0	46.0 ± 0.5	55.9 ± 0.5	60.0 ± 0.6
ArEx		48.2 ± 1.6	56.2 ± 0.3	60.3 ± 0.3

Fig. 5. Relative prolidase stability and activity. (A) Plot of T_M of prolidase variants measured by TSA. Numerical values are reported in the table below. (B) The relative enzymatic activity measured as the amount of proline released from GlyPro substrate upon reaction with prolidase. Plot scaled to activity of wt enzyme as 100%. Scale bars represent ±½SD errors from at least three independent measurements.

In the case of the Ser202Phe variant, one can notice a small rotational difference in the side chain position between the Prol_S202F_Ros and Prol_S202F_ArEx structures. This seems sufficient to allow a more wt-like arrangement of the Prol_S202F_ArEx variant (Fig. 2A). Additionally, decreased chain flexibility (Fig. 2C) resulted in clearer maps and allowed tracing of the entire model (Fig. S2). This, however, was not sufficient to restore the enzymatic activity. Of note is that for this variant not even residual activity was reported in previous studies [52]. Since the enzymatic assay used is based on end-point measurements, it cannot be excluded with certainty that deterioration of the enzymatic activity over the incubation time may also have some influence on the measured activity data. By performing the assay at a temperature, which was significantly below the measured T_M of the least stable variant, however, we have tried to minimize this potential effect.

An even bigger stabilizing effect was observed for the Gly278Asp substitution, where in the ArEx model no disordered fragments, characteristic for the Ros model, were observed. Here, the biggest difference is direction of the side chain of the introduced Asp278. In the Prol_G278D_Ros model, it is directed toward a neighboring strand where it hinders its proper folding. In contrast, in the Prol_G278D_ArEx it is swung away in the opposite direction where it introduces no disorder (Fig. 3A). In both models, the Arg237-Glu248 β-strand exhibits a higher ADP, but the pattern is not uniform (Figs 3C and S3). In the Prol_G278D_Ros structure, a small increase in B normalized is observed for the site of the substitution at the cost of large disorder in other chain fragments. In contrast, in Prol_G278D_ArEx the β-strand exhibits less elevated ADP at a cost of forcing the side chain of Asp into a conformationally unfavorable position with a high B factor (Fig. 3). Interestingly, despite retaining the increased flexibility the Prol_G278D_ArEx preparation has significantly higher enzymatic activity, reaching 8% of wt activity (Fig. 5). Of note is that in comparison with Prol_S202F_ArEx, the Cα deviation around Tyr241 is very small (Fig. 3B) and could highlight the importance of this residue, which was previously identified as one of the residues stabilizing the architecture of the active site [10].

Among the investigated prolidase variants, the Gly448Arg variant bears the biggest difference in the size of the side chain and causes the largest disorder of the enzyme structure. In the case of the Gly448Arg substitution, a structural disorder affects mainly a loop region connecting two β-strands rather than β-sheet itself and we could not identify any expression-related differences in obtained models. We can speculate that it is more likely to restore order in fragments with high secondary structure propensity than in naturally more flexible regions, such as loops. Interestingly, our activity assay indicates that Prol_G448R_ArEx exhibits higher *in vitro* activity than Prol_G448_Ros. This is a discrepancy that we cannot explain based on structural investigation, and it shows that the effect of chaperons may be even more elusive. This is in line with the results of Besio *et al.* [17], who reported a stabilizing effect of Hsp70/90 also on prolidase variants for which no structural stabilization can be expected (Prol_231-delY and Prol-E412K).

Conclusions

Prolidase deficiency is a rare recessive disorder caused by LOF mutations in PEPD gene. Unfortunately, such mutations are relatively little studied structurally. Several therapies were tested, but to date no efficient treatment for PD is available. Previously, it was reported that the induction of Hsp70/90 in cultured fibroblasts can partially restore the activity of a subset of prolidase LOF variants. Here, we investigated the effect of chaperone co-expression on a series of previously identified prolidase LOF related to structural disorder. Our crystallographic studies prove that in two of three analyzed prolidase variants prone to structural disorder, the expression in the presence of elevated concentration of chaperones significantly stabilizes protein and reverts its native-like conformation. We also show that enzymatic activity was partially restored. Our results suggest that the induction of chaperone activity may lead to stabilization and partial recovery of enzymatic activity of LOF mutants and therefore be considered as potential treatment. Both Cpn10/Cpn60 used in this study and Hsp70/Hsp90 analyzed previously in human fibroblasts are broad-spectrum chaperones, and therefore, we believe that observed effects are generic rather than chaperone-specific. Our studies also show that an expression system should certainly be considered as one of the variabilities in structural analyses of proteins.

Acknowledgements

X-ray data were collected at the BESSY II 14.1 beamline at Helmholtz-Zentrum Berlin für Materialien und Energie (HZB). We thank the HZB for the allocation of synchrotron radiation beamtime. We gratefully acknowledge the support of the Joint Berlin MX Laboratory, which made this work possible. We would particularly like to acknowledge the support of Dr.

Christian Feiler, Dr. Martin Gerlach, and Ronald Förster. MR's scholarship was financed by the National Science Centre, Poland (2018/28/T/ST5/00233). We acknowledge the MCB Structural Biology Core Facility (supported by the TEAM TECH CORE FACILITY/2017-4/6 grant from the Foundation for Polish Science) for valuable support.

Author contributions

PW designed the experiments. EW, MR, and PW prepared the protein samples, crystallized the protein, and performed the diffraction experiments. EW and PW solved and refined the structures. MSW contributed to structure refinements. EW conducted activity assay. PW analyzed the results. MSW and PW coordinated the study. EW and PW prepared figures and PW and MSW wrote the manuscript with contributions from all authors.

References

1 Lupi A, Della Torre S, Campari E, Tenni R, Cetta G, Rossi A and Forlino A (2006) Human recombinant prolidase from eukaryotic and prokaryotic sources: expression, purification, characterization and long-term stability studies. *FEBS J* **273**, 5466–5478.

2 Besio R, Baratto MC, Gioia R, Monzani E, Nicolis S, Cucca L, Profumo A, Casella L, Basosi R, Tenni R *et al.* (2013) A Mn(II)-Mn(II) center in human prolidase. *Biochim Biophys Acta Proteins Proteomics* **1834**, 197–204.

3 Kitchener RL and Grunden AM (2012) Prolidase function in proline metabolism and its medical and biotechnological applications. *J Appl Microbiol* **113**, 233–247.

4 Lupi A, Tenni R, Rossi A, Cetta G and Forlino A (2008) Human prolidase and prolidase deficiency: an overview on the characterization of the enzyme involved in proline recycling and on the effects of its mutations. *Amino Acids* **35**, 739–752.

5 Forlino A, Lupi A, Vaghi P, Cornaglia AI, Calligaro A, Campari E and Cetta G (2002) Mutation analysis of five new patients affected by prolidase deficiency: the lack of enzyme activity causes necrosis-like cell death in cultured fibroblasts. *Hum Genet* **111**, 314–322.

6 Royce PM and Steinmann B (2002) Prolidase deficiency. *Connect Tissue and its Heritable Disorders*, 727–743.

7 Bissonnette R, Friedmann D, Giroux JM, Dolenga M, Hechtman P, Der Kaloustian VM and Dubuc R (1993) Prolidase deficiency: a multisystemic hereditary disorder. *J Am Acad Dermatol* **29**, 818–821.

8 Di Rocco M, Fantasia AR, Taro M, Loy A, Forlino A and Martini A (2007) Systemic lupus erythematosus-like disease in a 6-year-old boy with prolidase deficiency. *J Inherit Metab Dis* **30**, 814.

9 Spodenkiewicz M, Spodenkiewicz M, Cleary M, Massier M, Fitsialos G, Cottin V, Jouret G, Poirsier C, Doco-Fenzy M and Lèbre AS (2020) Clinical genetics of prolidase deficiency: an updated review. *Biology* **9**, 1–17.

10 Wilk P, Uehlein M, Piwowarczyk R, Dobbek H, Mueller U and Weiss MS (2018) Structural basis for prolidase deficiency disease mechanisms. *FEBS J* **285**, 3422–3441.

11 Leidenheimer NJ and Ryder KG (2014) Pharmacological chaperoning: a primer on mechanism and pharmacology. *Pharmacol Res* **83**, 10–19.

12 Liguori L, Monticelli M, Allocca M, Mele BH, Lukas J, Cubellis MV and Andreotti G (2020) Pharmacological chaperones: a therapeutic approach for diseases caused by destabilizing missense mutations. *Int J Mol Sci* **21**, 489.

13 Convertino M, Das J and Dokholyan NV (2016) Pharmacological chaperones: design and development of new therapeutic strategies for the treatment of conformational diseases. *ACS Chem Biol* **11**, 1471–1489.

14 Moran N (2018) FDA approves Galafold, a triumph for Amicus. *Nat Biotechnol* **36**, 913.

15 Sean Froese D, Michaeli A, McCorvie TJ, Krojer T, Sasi M, Melaev E, Goldblum A, Zatsepin M, Lossos A, Álvarez R *et al.* (2015) Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. *Hum Mol Genet* **24**, 5667–5676.

16 Yue WW (2016) From structural biology to designing therapy for inborn errors of metabolism. *J Inherit Metab Dis* **39**, 489–498.

17 Besio R, Gioia R, Cossu F, Monzani E, Nicolis S, Cucca L, Profumo A, Casella L, Tenni R, Bolognesi M *et al.* (2013) Kinetic and structural evidences on human prolidase pathological mutants suggest strategies for enzyme functional rescue. *PLoS One* **8**, e58792.

18 Chaudhuri TK and Paul S (2006) Protein-misfolding diseases and chaperone-based therapeutic approaches. *FEBS J* **273**, 1331–1349.

19 Nagai Y, Fujikake N, Popiel H and Wada K (2010) Induction of molecular chaperones as a therapeutic strategy for the polyglutamine diseases. *Curr Pharm Biotechnol* **11**, 188–197.

20 Friesen EL, De Snoo ML, Rajendran L, Kalia LV and Kalia SK (2017) Chaperone-based therapies for disease modification in Parkinson's disease. *Parkinsons Dis* **2017**, 5015307.

21 Nakazono A, Adachi N, Takahashi H, Seki T, Hamada D, Ueyama T, Sakai N and Saito N (2018) Pharmacological induction of heat shock proteins ameliorates toxicity of mutant PKCγ in spinocerebellar ataxia type 14. *J Biol Chem* **293**, 14758–14774.

22 Rosano GL, Morales ES and Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: a 5-year update. *Protein Sci* **28**, 1412–1422.
23 Schumann W and Ferreira LCS (2004) Production of recombinant proteins in *Escherichia coli*. *Genet Mol Biol* **27**, 442–453.
24 Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN and Timmis KN (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures [2]. *Nat Biotechnol* **21**, 1266–1267.
25 Wilk P, Uehlein M, Kalms J, Dobbek H, Mueller U and Weiss MS (2017) Substrate specificity and reaction mechanism of human prolidase. *FEBS J* **284**, 2870–2885.
26 Mueller U, Förster R, Hellmig M, Huschmann FU, Kastner A, Malecki P, Pühringer S, Röwer M, Sparta K, Steffien M *et al*. (2015) The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: current status and perspectives. *Eur Phys J Plus* **130**, 141.
27 Gerlach M, Mueller U and Weiss MS (2016) The MX Beamlines BL14.1-3 at BESSY II. *JLSRF* **2**, 1–6.
28 Bunkóczi G, Echols N, McCoy AJ, Oeffner RD, Adams PD and Read RJ (2013) Phaser.MRage: automated molecular replacement. *Acta Crystallogr D Biol Crystallogr* **69**, 2276–2286.
29 Kowiel M, Jaskolski M and Dauter Z (2014) ACHESYM: an algorithm and server for standardized placement of macromolecular models in the unit cell. *Acta Crystallogr D Biol Crystallogr* **70**, 3290–3298.
30 Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
31 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH and Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352–367.
32 Afonine PV, Grosse-Kunstleve RW, Adams PD and Urzhumtsev A (2013) Bulk-solvent and overall scaling revisited: faster calculations, improved results. *Acta Crystallogr D Biol Crystallogr* **69**, 625–634.
33 Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S *et al*. (2002) The protein data bank. *Acta Crystallogr D Biol Crystallogr* **58**, 899–907.
34 Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, Christie C, Dalenberg K, Duarte JM, Dutta S *et al*. (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* **47**, D464–D474.
35 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al*. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.
36 Tom Burnley B, Afonine PV, Adams PD and Gros P (2012) Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* **2012**, 1–29.
37 Burnley BT and Gros P (2013) phenix.ensemble_refinement: a test study of apo and holo BACE1. *Comput Crystallogr Newsl* **4**, 51–58.
38 Reinhard L, Mayerhofer H, Geerlof A, Mueller-Dieckmann J and Weiss MS (2013) Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallogr F Struct Biol Cryst Commun* **69**, 209–214.
39 Besio R, Monzani E, Gioia R, Nicolis S, Rossi A, Casella L and Forlino A (2011) Improved prolidase activity assay allowed enzyme kinetic characterization and faster prolidase deficiency diagnosis. *Clin Chim Acta* **412**, 1814–1820.
40 Lowther WT and Matthews BW (2002) Metalloaminopeptidases: common functional themes in disparate structural surroundings. *Chem Rev* **102**, 4581–4607.
41 Karna E, Miltyk W and Pałka JA (2006) Butyrate-induced collagen biosynthesis in cultured fibroblasts is independent on α2β1 integrin signalling and undergoes through IGF-I receptor cascade. *Mol Cell Biochem* **286**, 147–152.
42 Karna E, Trojan S and Pałka JA (2009) The mechanism of butyrate-induced collagen biosynthesis in cultured fibroblasts. *Acta Pol Pharm Drug Res* **66**, 229–233.
43 Surazyński A, Sienkiewicz P, Wołczyński S and Pałka J (2005) Differential effects of echistatin and thrombin on collagen production and prolidase activity in human dermal fibroblasts and their possible implication in β1-integrin-mediated signaling. *Pharmacol Res* **51**, 217–221.
44 Karna E, Szoka L and Palka JA (2013) The mechanism of hydralazine-induced collagen biosynthesis in cultured fibroblasts. *Naunyn Schmiedebergs Arch Pharmacol* **386**, 303–309.
45 Sheffield LJ, Schlesinger P, Faull K, Halpern BJ, Schier GM, Cotton RGH, Hammond J and Danks DM (1977) Iminopeptiduria, skin ulcerations, and edema in a boy with prolidase deficiency. *J Pediatr* **91**, 578–583.
46 Berardesca E, Fideli D, Bellosta M, Dyne KM, Zanaboni G and Cetta G (1992) Blood transfusions in the therapy of a case of prolidase deficiency. *Br J Dermatol* **126**, 193–195.
47 Lupi A, Casado B, Soli M, Bertazzoni M, Annovazzi L, Viglio S, Cetta G and Iadarola P (2002) Therapeutic apheresis exchange in two patients with prolidase deficiency. *Br J Dermatol* **147**, 1237–1240.

48 Yasuda K, Ogata K, Kariya K, Kodama H, Zhang J, Sugahara K, Sagara Y and Kodama H (1999) Corticosteroid treatment of prolidase deficiency skin lesions by inhibiting iminodipeptide-primed neutrophil superoxide generation. *Br J Dermatol* **141**, 846–851.
49 Monafo V, Marseglia GL, Maghnie M, Dyne KM and Cetta G (2000) Transient beneficial effect of GH replacement therapy and topical GH application on skin ulcers in a boy with prolidase deficiency. *Pediatr Dermatol* **17**, 227–230.
50 Ogata A (1981) Autosomal recessive prolidase deficiency. *Arch Dermatol* **117**, 689.
51 Perugini P, Hassan K, Genta I, Modena T, Pavanetto F, Cetta G, Zanone C, Iadarola P, Asti A and Conti B (2005) Intracellular delivery of liposome-encapsulated prolidase in cultured fibroblasts from prolidase-deficient patients. *J Control Release* **102**, 181–190.
52 Hershkovitz T, Hassoun G, Indelman M, Shlush LI, Bergman R, Pollack S and Sprecher E (2006) A homozygous missense mutation in PEPD encoding peptidase D causes prolidase deficiency associated with hyper-IgE syndrome. *Clin Exp Dermatol* **31**, 435–440.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Data collection and refinement statistics.

Table S2. Per residue numerical values of CαRMSD [Å] of each variant vs wild type prolidase.

Fig. **S1**. (A) Structural aspects of human prolidase. Fragment of human prolidase close to the dimer interface is shown as sticks with composite omit map countered at 1.0 σ. Monomer A is colored according to its local B factor (blue: low B, red: high B) and monomer B is colored following the chain trace to ease localization of described mutants. The substrate and metal ions are drawn with black ball-and-stick representation to indicate location of the active sites. (B) variation of B factors for both subunits for the discussed residue range. The scale of the plot is kept constant to ease comparison with later figures. (C) B-factor and coordinate RMSF derived from ensemble refinement of the wild-type prolidase are plotted for comparison with panels D & E on Figures 2–4. Here the scale was adjusted as using the same as for the mutants lines plotted here would appear flat.

Fig. **S2**. Comparison of the Ser202Phe variant of prolidase derived from two different expression systems.

Fig. **S3**. Comparison of the Gly278Asp variant of prolidase derived from two different expression systems.

Fig. **S4**. Comparison of the Gly448Arg variant of prolidase derived from two different expression systems.

Fig. **S5**. Overlay of the full prolidase variant dimers on the wild type prolidase drawn in the cartoon representation.