

The impact of folding modes and deuteration on the atomic resolution structure of hen egg-white lysozyme

Joao Ramos,^{a,b,c} Valerie Laux,^{a,b} Michael Haertlein,^{a,b} V. Trevor Forsyth,^{a,b,d,e,f} Estelle Mossou,^{b,g} Sine Larsen^h and Annette E. Langkilde^{c*}

Received 13 August 2021

Accepted 20 October 2021

Edited by P. Langan, Oak Ridge National Laboratory, USA

Keywords: *in vitro* refolding; isotope effect; thermal stability; folding modes; deuteration; folding dynamics; hen egg-white lysozyme; enzymatic activity; X-ray crystallography.

PDB reference: hydrogenated refolded hen egg-white lysozyme, 7p6m

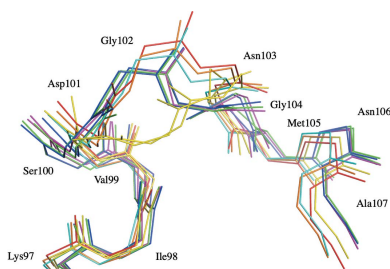
Supporting information: this article has supporting information at journals.iucr.org/d

^aLife Sciences Group, Institute Laue–Langevin, 71 Avenue des Martyrs, 38000 Grenoble, France, ^bPartnership for Structural Biology (PSB), 71 Avenue des Martyrs, 38000 Grenoble, France, ^cDepartment of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark, ^dFaculty of Natural Sciences, Keele University, Newcastle ST5 5BG, United Kingdom, ^eFaculty of Medicine, Lund University, 221 00 Lund, Sweden, ^fLINXS Institute for Advanced Neutron and X-ray Science, Scheelvagen 19, 223 70 Lund, Sweden, ^gEuropean Synchrotron Radiation Facility, 71 Avenue des Martyrs, 38000 Grenoble, France, and ^hDepartment of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. *Correspondence e-mail: annette.langkilde@sund.ku.dk

The biological function of a protein is intimately related to its structure and dynamics, which in turn are determined by the way in which it has been folded. *In vitro* refolding is commonly used for the recovery of recombinant proteins that are expressed in the form of inclusion bodies and is of central interest in terms of the folding pathways that occur *in vivo*. Here, biophysical data are reported for *in vitro*-refolded hydrogenated hen egg-white lysozyme, in combination with atomic resolution X-ray diffraction analyses, which allowed detailed comparisons with native hydrogenated and refolded perdeuterated lysozyme. Distinct folding modes are observed for the hydrogenated and perdeuterated refolded variants, which are determined by conformational changes to the backbone structure of the Lys97–Gly104 flexible loop. Surprisingly, the structure of the refolded perdeuterated protein is closer to that of native lysozyme than that of the refolded hydrogenated protein. These structural differences suggest that the observed decreases in thermal stability and enzymatic activity in the refolded perdeuterated and hydrogenated proteins are consequences of the macromolecular deuteration effect and of distinct folding dynamics, respectively. These results are discussed in the context of both *in vitro* and *in vivo* folding, as well as of lysozyme amyloidogenesis.

1. Introduction

Secreted eukaryotic proteins produced *in vivo* typically pass through a process that starts at the ribosome (for example, at the endoplasmic reticulum). The initial stages of translation involve the production of a pre-sequence or signal peptide that binds to the signal recognition particle (SRP), which in turn binds to the membrane-bound SRP receptor. The unfolded peptide then passes through the membrane, whereupon the pre-sequence is cleaved, with folding subsequently occurring through a pathway involving multiple chaperones (Shikano & Colley, 2013). In the case of hen egg-white lysozyme (HEWL), recombinant production in *Escherichia coli* is problematic because the reductive environment of the bacterial cytosol prevents the correct formation of the four disulfide bridges, resulting in the formation of inclusion bodies. For the production of perdeuterated protein, we have recently adopted an approach whereby large quantities of insoluble protein were produced as inclusion bodies in *E. coli* followed by an optimized *in vitro* refolding process (Ramos *et al.*, 2021).



HEWL has been shown to possess two distinct folding domains (Supplementary Fig. S1): the α -domain, constituted by four α -helices, the Val120–Arg125 3_{10} -helix and the loops Gly16–Gly22 and Ser100–Ala107, and the β -domain, which comprises a triple-stranded antiparallel β -sheet, the Pro79–Leu84 3_{10} -helix and the Arg61–Ile78 loop (Miranker *et al.*, 1991). While the α -domain, which contains most of the secondary-structure motifs, folds first, the β -domain is stabilized later in the folding process (Miranker *et al.*, 1991, 1993; Radford *et al.*, 1992). The *in vitro* folding mechanism of lysozyme was later shown to involve intermediate states (Radford *et al.*, 1992; Miranker *et al.*, 1993; Wildegger & Kiefhaber, 1997). Additionally, the unfolding process of an amyloidogenic variant of human lysozyme, highly homologous to HEWL, seems to involve local cooperativity (Canet *et al.*, 2002). An earlier study had also suggested identity between the unfolding and folding intermediates of lysozyme and its structural homologue α -lactalbumin (Ikeguchi *et al.*, 1986). Understanding protein folding and the impact of different chemical environments on folding pathways is essential for current efforts in predicting three-dimensional structure using *in silico* methods and for the study of amyloidogenic pathologies. In the case of human lysozyme, several amyloidogenic mutations have been identified, particularly in α -helix C (Ile88–Asp101) and the β -domain (Pepys *et al.*, 1993; Gillmore *et al.*, 1999; Valleix *et al.*, 2002; Yazaki *et al.*, 2003; Wooliver *et al.*, 2007; Girnius *et al.*, 2012; Jean *et al.*, 2014; Sperry *et al.*, 2016; Nasr *et al.*, 2017).

H atoms constitute approximately 50% of all atoms in a protein. They are essential to biological function, either through their role in protein folding, in protein interactions or by participating in the catalysis of enzymatic reactions. While the substitution of H by its heavier isotope deuterium (D) in solvent-exchangeable positions (*i.e.* those bound to N or O atoms) or at all positions (*i.e.* perdeuteration) usually results in closely isomorphous structures, the associated mass difference can be expected to have consequences for protein dynamics. This isomorphism is routinely used in neutron macromolecular crystallography (NMX; Cuyppers *et al.*, 2013; Manzoni *et al.*, 2018; Yee *et al.*, 2019; McGregor *et al.*, 2021; Gajdos *et al.*, 2021; Kelpšas *et al.*, 2021), small-angle neutron scattering (SANS; Breyton *et al.*, 2013; Maric *et al.*, 2014, 2015; Dunne *et al.*, 2017; Nitsche *et al.*, 2018; Kehlenbeck *et al.*, 2019), neutron reflectometry (NR; Wacklin *et al.*, 2016; Moulin *et al.*, 2018; Campbell *et al.*, 2018; Waldie *et al.*, 2018, 2019), neutron spectroscopy (Foglia *et al.*, 2019), neutron fiber diffraction (Gardner *et al.*, 2004; Wada *et al.*, 2011) and nuclear magnetic resonance (NMR; Varga *et al.*, 2007), where the different physical properties of the two isotopes are fully exploited. One of the common ways of testing this isomorphism is through a comparison of the X-ray analyses of analogous structures, and there is now a steadily growing database of these comparisons. For crystalline systems, the isomorphism usually holds good to high resolution (Gamble *et al.*, 1994; Cooper *et al.*, 1998; Meilleur *et al.*, 2004; Artero *et al.*, 2005; Liu *et al.*, 2007; Fisher & Helliwell, 2008; Cuyppers *et al.*, 2013; Yee *et al.*, 2016, 2019; Koruza *et al.*, 2019; Ramos *et al.*,

2021). For the HEWL work described here, this is, to our knowledge, the first time that a detailed comparative study has been made of hydrogenated and perdeuterated analogs of a refolded protein.

While hydrogenated and perdeuterated variants of a protein are usually close to identical in structure, several studies of the macromolecular H/D isotope effect have reported decreases in the thermal stability of perdeuterated proteins compared with their hydrogenated analogs (Berns, 1963; Hattori *et al.*, 1965; Brockwell *et al.*, 2001; Meilleur *et al.*, 2004; Koruza *et al.*, 2018; Nichols *et al.*, 2020; Ramos *et al.*, 2021). Observations have also been made regarding solvent H/D isotope effects which suggest that both perdeuterated and hydrogenated proteins are thermally more stable in D₂O than in H₂O (Hattori *et al.*, 1965; Harrington & von Hippel, 1961; Makhatazde *et al.*, 1995; Freyman *et al.*, 2001; Kuhlman & Raleigh, 1998; Sasisanker *et al.*, 2004; Efimova *et al.*, 2007). While the macromolecular isotope effect seems to be caused by changes in the strengths of hydrophobic interactions between residue side chains, as described by Hattori *et al.* (1965), the solvent isotope effect appears to be a consequence of variations in protein–protein and protein–water hydrogen bonds following H/D substitution in exchangeable positions and to different solvation potentials for D₂O and H₂O (Svergun *et al.*, 1998; Sasisanker *et al.*, 2004; Artero *et al.*, 2005; Efimova *et al.*, 2007; Jasnin *et al.*, 2008). Changes in protein solubility and crystallization conditions are also apparent (Budayova-Spano *et al.*, 2000; Hazemann *et al.*, 2005; Liu *et al.*, 2007; Petit-Haertlein *et al.*, 2009; Koruza *et al.*, 2018). Often the crystallization conditions used for the perdeuterated variant comprise reductions in the precipitant and/or protein concentration, reflecting decreases in protein solubility compared with that of the hydrogenated variant (Hazemann *et al.*, 2005; Petit-Haertlein *et al.*, 2009). Nevertheless, crystallographic studies have been performed in which the crystallization conditions for both proteins are substantially different (Liu *et al.*, 2007). These variations need to be considered when discussing any structural changes associated with deuteration and when attributing these to macromolecular isotope effects.

In our previous study (Ramos *et al.*, 2021), we compared the structure of a perdeuterated refolded hen egg-white lysozyme variant (D-HEWL_{EC}) with that of the native hydrogenated variant (H-HEWL). This analysis was mainly carried out in the context of establishing the viability and validity of a neutron crystallographic study, in which major technical benefits were gained through the use of perdeuterated protein. The available data did not permit the effects of macromolecular and solvent isotope substitution (H/D) to be clearly distinguished from differences associated with folding dynamics. However, protocols developed to further produce the hydrogenated refolded analog (H-HEWL_{EC}) allowed a detailed comparison of the hydrogenated and perdeuterated analogs. With the aid of new X-ray crystallographic data to 0.89 Å resolution for H-HEWL_{EC} (PDB entry 7p6m), along with thermal stability and activity assays, we characterize the different effects in detail. Both the H-HEWL_{EC} and D-HEWL_{EC} variants differ from native H-HEWL by an

additional glycine residue at the N-terminus. This residue disrupts hydrogen-bond interactions at the N-terminus of the protein, rendering this region more disordered (Ramos *et al.*, 2021), although it did not perturb protein folding or function. The observed changes in enzymatic activity have been correlated with structural differences in the three variants. Additionally, the thermal stabilities of the three variants were investigated by differential scanning fluorimetry (DSF) in hydrogenated as well as deuterated solutions. This systematic comparison enables us to address and separate the effects of *in vitro* refolding, as well as the macromolecular and solvent isotope effects.

2. Materials and methods

2.1. Protein materials

Commercially available H-HEWL (catalog No. L6876; Sigma–Aldrich) was used without additional processing. D-HEWL_{EC} was recombinantly expressed in *E. coli* BL21(DE3) cells in the form of inclusion bodies and was purified and refolded in-column as described previously (Ramos *et al.*, 2021). Using a similar approach, H-HEWL_{EC} was expressed in *E. coli* BL21(DE3) cells, purified and refolded. The buffer compositions and procedures employed in inclusion-body washing, protein purification by gel filtration and refolding in-column were identical for both proteins. The main differences between the protein-production methods used for the two HEWL_{EC} variants were the culture media used for protein overexpression and the final protein buffer exchange performed after refolding. D-HEWL_{EC} was expressed in *E. coli* cultures adapted to fully deuterated minimal medium, while H-HEWL_{EC} was produced in *E. coli* cultures grown in H₂O minimal medium. Additionally, D-HEWL_{EC} was buffer-exchanged into a D₂O solution of 50 mM sodium acetate pD 4.5 (pD = pH + 0.4; Glasoe & Long, 1960). Meanwhile, H-HEWL_{EC} was buffer-exchanged into a H₂O solution of 50 mM sodium acetate pH 4.5. The proteins were concentrated in their final buffer solutions to 20 mg ml⁻¹ for further experiments.

2.2. Differential scanning fluorimetry

Protein thermal stability was assessed by DSF using a Prometheus instrument from NanoTemper. Protein unfolding is measured through the intrinsic fluorescent signal of Trp residues (there are six Trp residues in HEWL), which is quenched when the protein is folded and is enhanced when these residues are exposed to the solvent in the unfolded state. H-HEWL, D-HEWL_{EC} and H-HEWL_{EC} were each tested in four different buffer solutions: (i) 50 mM sodium acetate pH 4.5 in H₂O, (ii) 50 mM sodium acetate pD 4.5 in D₂O, (iii) 0.1 M sodium phosphate pH 7.5, 0.1 M NaCl, 2 mM NaN₃ in H₂O and (iv) 0.1 M sodium phosphate pD 7.5, 0.1 M NaCl, 2 mM NaN₃ in D₂O. The samples were diluted from stock solutions at 20 mg ml⁻¹ in each buffer solution to a final protein concentration of approximately 0.2 mg ml⁻¹. The fluorescence was measured at 90% instrument excitation

power in temperature ramps from 20°C to 95°C with increments of 0.1°C min⁻¹. The results presented were obtained from three different experiments with triplicate measurements for each condition.

2.3. Activity assays

The enzymatic activity was measured using the method originally reported by Shugar (1952), which estimates the activity rates by following the decrease in absorbance at 450 nm when HEWL is added to a cell suspension of *Micrococcus lysodeikticus*. The conditions used were identical to those used in our previous study (Ramos *et al.*, 2021). The results presented in Supplementary Fig. S2 were obtained for all HEWL variants in three experiments, with triplicate measurements for each condition (technical replicates), and controls were included as buffer without protein. The technical replicates were averaged and plotted against time to obtain the activity curves. The initial velocities were retrieved from the linear phase ($R^2 > 0.95$) corresponding to the first 5 min of reaction. The final activity rates and corresponding standard deviations were obtained from averaging the results from the three experiments.

2.4. Protein crystallization

Triclinic H-HEWL_{EC} crystals were obtained by micro-seeding triclinic H-HEWL in microbatch under oil at 18°C. The crystallization drop consisted of 2.5 µl 20 mg ml⁻¹ refolded H-HEWL_{EC}, 2.5 µl 0.3 M NaNO₃ and 50 mM sodium acetate pH 4.5 in H₂O and 0.5 µl H-HEWL seeds. Crystals of approximately 0.1 mm³ appeared within one week.

2.5. X-ray diffraction data collection and processing

A triclinic crystal of H-HEWL_{EC} was flash-cooled in liquid nitrogen after soaking in a solution consisting of 30%(v/v) glycerol, 0.4 M NaNO₃, 50 mM sodium acetate pH 4.5 in H₂O. X-ray diffraction data at 100 K were measured on the BioMAX beamline (Ursby *et al.*, 2020) at MAX IV Laboratory. Data to 0.89 Å resolution were recorded at two different κ orientations with 180° scans to maximize reciprocal-space coverage and data completeness. Data reduction was performed using *XDS* (Kabsch, 2010) and the two 180° scans were scaled using *XSCALE* (Kabsch, 2010). The data were converted to *SHELX* format using *XDS CONV* (Kabsch, 2010) and 5% of the reflections were randomly flagged for R_{free} calculations.

2.6. Model refinement

PDB entry 7ave (Ramos *et al.*, 2021) was used as the initial model for the refinement of H-HEWL_{EC} in *SHELXL* (Sheldrick, 2015). Alternate protein residue conformations, water molecules and ions were removed from the initial model and the anisotropic ADPs were converted to isotropic ADPs. After the refinement of residue disorder, H atoms were added in idealized positions and fixed as riding atoms. Default parameters for the geometric restraints of the residues were employed and a few bond-angle outliers were allowed, rather

than imposing stricter restraints, during refinement using high-resolution data. Anisotropic ADPs were refined for all non-H atoms, including water molecules and ions. Occupancies were refined for water molecules and ions with B factors larger than 30 \AA^2 . SIMU (0.1) and XPND (0.001) restraints were removed in the last rounds of refinement.

2.7. Analysis of HEWL structures

The H-HEWL and D-HEWL_{EC} models (PDB entries 7avf and 7ave, respectively; Ramos *et al.*, 2021) were used in comparisons with the H-HEWL_{EC} structure. The structural alignments between the HEWL structures were performed using *GESAMT* from the *CCP4* suite (Winn *et al.*, 2011). The alignments of the Lys97–Gly104 and binding-cleft regions were performed in *PyMOL* (version 2.0; Schrödinger) using the *ALIGN* function with zero refinement cycles. All illustrations of protein structures were made using *PyMOL*.

3. Results

3.1. *In vitro* refolding has a stronger impact on the thermal stability of lysozyme than deuteration

The results from the DSF experiments show clear trends regarding the effects on protein thermal stability of *in vitro* refolding, protein perdeuteration, H/D solvent substitution and the pH of the buffer solution. The melting temperatures for H-HEWL and D-HEWL_{EC} in sodium acetate pD 4.5 in D₂O and in sodium phosphate pH 7.5 in H₂O were remeasured in this study and are in agreement with the values reported in our previous work (Ramos *et al.*, 2021).

As shown in Fig. 1 and Supplementary Table S1, H-HEWL_{EC} displays a decrease in melting temperature (T_m) in H₂O of 4.7°C and 3.5°C at pH 4.5 and 7.5, respectively, in comparison with H-HEWL. In D₂O, H-HEWL_{EC} is less thermally stable than H-HEWL by 4.5°C and 3.7°C at pD 4.5 and 7.5, respectively. The reductions in T_m seem to be constant regardless of the solvent isotope substitution, with minor variations likely to be due to the differences between pH and pD (pD = pH + 0.4; Glasoe & Long, 1960).

Perdeuteration of HEWL_{EC} appears to reduce the protein thermal stability by 1.6°C and 1.3°C at pH 4.5 and pH 7.5, respectively. Interestingly, these variations are smaller than those observed between H-HEWL and H-HEWL_{EC}, highlighting that the refolding process has a stronger impact on protein thermal stability than perdeuteration. Additionally, the differences in T_m are similar in H₂O and D₂O solutions, reflecting effects of protein perdeuteration rather than solvent isotope substitution.

Comparing the individual variants in H₂O and D₂O respectively, as shown in Fig. 1, the solvent isotope effect has similar magnitudes across the three HEWL proteins, varying only according to the buffer solution pH/pD. In the case of H-HEWL, the replacement of H₂O with D₂O increased its thermal stability by 1.9°C and 2.8°C at pH/pD 4.5 and 7.5, respectively. Both H-HEWL_{EC} and D-HEWL_{EC} were found

to be more thermally stable in D₂O than in H₂O, with identical variations in T_m of 2.1°C at pH/pD 4.5 and 2.6°C at pH/pD 7.5.

The results suggest that protein thermal stability is influenced by pH. However, since the buffer solutions at pH/pD 4.5 and 7.5 differ significantly in composition, it is not appropriate to perform a direct comparison of the pH/pD effects. Nevertheless, it is clear that the thermal stability is significantly reduced for the three HEWL variants by the increase in pH/pD from 4.5 to 7.5.

3.2. *In vitro* refolding leads to loss of enzymatic activity

Previous studies have reported reductions in the enzymatic activity of refolded HEWL compared with the native variant (Batas & Chaudhuri, 1996; Batas *et al.*, 1999). However, no structural explanations were provided for these observations. Here, we report activity assays performed in parallel on all three HEWL variants. The results show that H-HEWL_{EC} and D-HEWL_{EC} retain 66% and 69% of the activity of H-HEWL, respectively (Supplementary Fig. S2). In our previous work, we reported an activity of 51% for D-HEWL_{EC} (Ramos *et al.*, 2021) based on an analysis of the first 8 min of reaction for the extrapolation of initial velocities, resulting in $R^2 > 0.91$. In this study, the first 5 min were used in the subsequent analysis, resulting in $R^2 > 0.95$.

3.3. The protein fold is retained despite *in vitro* refolding and perdeuteration

The conditions used to obtain the triclinic H-HEWL_{EC} crystals were virtually identical to those reported for H-HEWL and D-HEWL_{EC} (PDB entries 7avf and 7ave, respectively; Ramos *et al.*, 2021) and comparable X-ray diffraction data extending to 0.89 Å resolution were collected at 100 K. The data-merging and model-refinement statistics are shown in Table 1. The HEWL crystal structure reported by Wang *et al.* (2007) obtained from X-ray diffraction data to

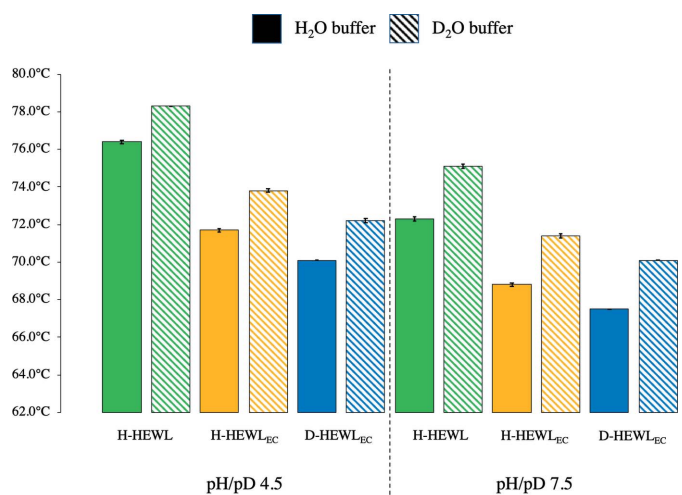


Figure 1 Melting temperatures measured in the DSF experiments for H-HEWL, H-HEWL_{EC} and D-HEWL_{EC} in buffer solutions at pH 4.5 and 7.5 in H₂O and at pD 4.5 and 7.5 in D₂O. The error bars correspond to the standard deviations of the measurements performed in three DSF experiments.

Table 1

X-ray diffraction data-collection and model-refinement statistics for H-HEWL_{EC} (PDB entry 7p6m).

Values in parentheses are for the outer resolution shell.

Temperature (K)	100
Source	BioMAX, MAX IV
Detector	EIGER 16M
Wavelength (Å)	0.700
Resolution range (Å)	22.46–0.89 (0.92–0.89)
Space group	<i>P</i> 1
<i>a</i> , <i>b</i> , <i>c</i> (Å)	26.12, 30.70, 33.57
α , β , γ (°)	88.968, 72.768, 69.505
Total reflections	234437 (23523)
Unique reflections	68446 (6717)
Multiplicity	3.4 (3.5)
Completeness (%)	96.2 (94.3)
Mean $I/\sigma(I)$	7.4 (2.2)
R_{merge}	0.083 (0.647)
R_{meas}	0.098 (0.762)
$R_{\text{p.i.m.}}$	0.052 (0.398)
$CC_{1/2}$	99.6 (79.9)
Reflections used in refinement with $F_o/\sigma(F_o) > 4$ /all reflections	50427/65023
Reflections used for R_{free} with $F_o/\sigma(F_o) > 4$ /all reflections	2635/3423
R_{work} for reflections with $F_o/\sigma(F_o) > 4$ /all reflections	11.16/12.51
R_{free} for reflections with $F_o/\sigma(F_o) > 4$ /all reflections	13.85/15.31
No. of non-H atoms (overall)	
Total	1496
Macromolecules	1304
Ligands	40
Solvent	152
Protein residues	130
R.m.s.d., bond lengths (Å)	0.023
R.m.s.d., angles (°)	3.32
Ramachandran favored (%)	95.7
Ramachandran allowed (%)	4.3
Ramachandran outliers (%)	0
Rotamer outliers (%)	3.6
Clashscore	8
Average <i>B</i> factor (Å ²)	
Overall	14.7
Macromolecules	14.0
Ligands	15.0
Solvent	20.9

0.65 Å resolution stands out as a reference among the plethora of HEWL structures deposited in the PDB (Berman *et al.*, 2000). However, a detailed comparison of the structures of several HEWL variants should only encompass those obtained using highly similar conditions for protein crystallization, X-ray data collection, processing and model refinement.

Structural alignments were thus performed for the following pairs of crystal structures: H-HEWL–H-HEWL_{EC} (highlighting refolding effects), D-HEWL_{EC}–H-HEWL_{EC} (highlighting deuteration effects) and H-HEWL–D-HEWL_{EC} (combining the effects previously observed). Overall, considering H-HEWL as a reference, the protein fold is retained in both recombinant variants, despite the *in vitro* refolding that was performed as well as the additional perdeuteration of D-HEWL_{EC} (Fig. 2). This evidence is supported by the fact that the two HEWL_{EC} variants could be successfully crystallized in highly comparable conditions, in an identical space group and with similar unit-cell dimensions. Interestingly,

H-HEWL and D-HEWL_{EC} are highly similar structurally and the main differences between the protein structures are observed in relation to H-HEWL_{EC}. While the r.m.s.d. of D-HEWL_{EC} from the H-HEWL structure is only 0.17 Å, H-HEWL_{EC} shows an r.m.s.d. of 0.56 Å. Additionally, H-HEWL_{EC} was found to have an r.m.s.d. of 0.55 Å when compared with D-HEWL_{EC}. As illustrated in Fig. 2, the main differences between H-HEWL_{EC} and the two other HEWL variants are observed in the Thr47–Gly49 and Lys97–Gly104 regions, both of which are at the protein surface and engage in crystal contacts.

3.4. Refolded H-HEWL_{EC} appears to be more compact and rigid than native H-HEWL

47 and 51 water molecules in the crystal structures of H-HEWL_{EC} and H-HEWL, respectively, are involved in important hydrogen-bond interactions that maintain the protein fold (*i.e.* are involved in two or more hydrogen bonds to protein atoms from nonconsecutive residues). While 34 of these water molecules are conserved in both structures, the absence of the remaining water molecules in either variant does not appear to affect their fold. The protein region where the water structures differ the most is the disordered Lys97–Gly104 loop. Differences in alternate conformations of side chains of residues such as Glu7, Arg68, Thr89 and Gln121 also promote changes to the location of structural waters without significantly perturbing the protein fold. For instance, the hydrogen-bond interactions facilitated by water molecules 314 and 387 in H-HEWL_{EC} are enabled in H-HEWL by different side-chain conformations of Glu7 and Asn44. Concomitantly, water 138 in H-HEWL is involved in identical hydrogen bonds to those established by Gly0 in H-HEWL_{EC}.

104 additional water molecules were observed in the hydration shells of H-HEWL, compared with 100 in H-HEWL_{EC}. This minor difference in the number of hydration water molecules can be explained by the $2F_o - F_c$ electron-density map cutoff at 1.5σ that was used to model structural waters. The fact that water molecules with weaker densities were not included in the models might have

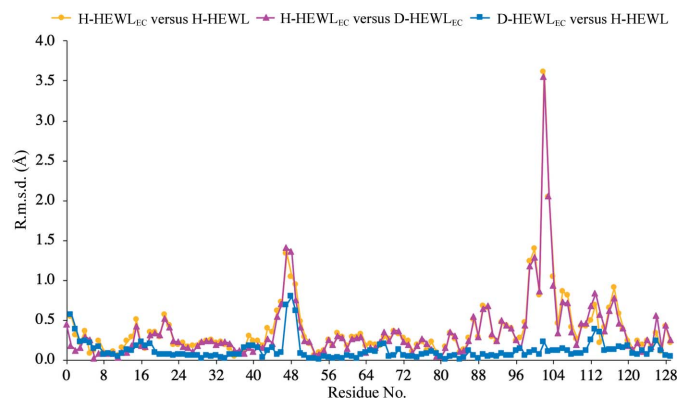


Figure 2
Pairwise structural comparisons. The structures were aligned using *GESAMT* from the *CCP4* suite (Winn *et al.*, 2011) and the r.m.s.d. between each pair of structures is plotted on the residue level.

precluded differences in hydration and surface area between H-HEWL and H-HEWL_{EC}. However, calculation of the molecular surface areas of both structures yielded values of 15 720 and 14 961 Å² for H-HEWL and H-HEWL_{EC}, respectively. As previously noted, H-HEWL_{EC} contains an additional Gly residue at the N-terminus compared with native H-HEWL. However, when Gly0 is removed from the structure the calculated molecular surface of H-HEWL_{EC} is only marginally affected (14 899 Å²).

The models of H-HEWL and H-HEWL_{EC} include 41 and 36 residues with alternate conformations, respectively (Supplementary Fig. S3). It is notable that although the resolution of the X-ray data is lower for H-HEWL (1.0 Å) than for H-HEWL_{EC} (0.89 Å), the latter displays less disorder of its protein residues and also a reduced number of structural waters, as shown previously. These discrepancies could be explained by differences in quality and completeness between the data sets; however, in this case both data sets are similar ($R_{p.i.m.}$ of ~5% and completeness of >95%). An example of the reduced flexibility of H-HEWL_{EC} is observed in the Arg112–Arg114 region, where both Arg residues display single conformations. In H-HEWL these Arg residues adopt alternate conformations, which apparently cause significant displacement of the protein backbone structure. Corroborating the decreased flexibility of H-HEWL_{EC} is the systematically lower mean residue *B* factor verified throughout the structure in comparison with that of H-HEWL (Supplementary Fig. S4). 74% of H-HEWL_{EC} residues display lower mean *B* factors for their main-chain atoms compared with H-HEWL, with exceptions being found in the following regions: Thr43–Asp48, Leu75–Ile78, Ser86–Cys94, Lys97–Trp108 and Gly126–Leu129.

3.5. The Lys97–Gly104 loop is crucial to HEWL folding

Our previous study revealed that the structure of the Lys97–Gly104 loop was the most affected upon *in vitro* refolding of D-HEWL_{EC}. The backbone disorder of the Lys97–Gly104 region was shown to be associated with the occurrence of an Asn103 peptide-plane flip (for example in PDB entry 2vb1; Wang *et al.*, 2007). While a partial Asn103 peptide-plane flip has been described in H-HEWL with a refined occupancy of 33%, it was found that this alternate conformation was present in 46% of the molecules of D-HEWL_{EC} (Ramos *et al.*, 2021).

The conformation of the Lys97–Gly104 loop in H-HEWL_{EC} differs significantly from its conformations in H-HEWL and D-HEWL_{EC} (Fig. 3). The r.m.s.d.s between this region of H-HEWL_{EC} and the same residues in H-HEWL and D-HEWL_{EC} are 2.38 and 2.33 Å, respectively, noting that the r.m.s.d. between the latter two is only 0.27 Å. Furthermore, according to the $2F_o - F_c$ electron-density map of H-HEWL_{EC}, both conformations of this loop are associated with a peptide-plane flip of Asn103 (Supplementary Fig. S5). Variations in side-chain conformations can also be observed, namely for Asp101 and Asn103. In the case of the Asn103 side chain, significant disorder is evident from the lack of density in the $2F_o - F_c$ electron-density map (contoured at 1σ). Mean-

while, for Asp101 the electron density suggests a different side-chain position compared with other HEWL variants (Ramos *et al.*, 2021), leading to the disruption of a hydrogen-bond crystal contact with Glu7.

Because crystal packing can perturb protein structure, particularly at its surface (as is the case for the Lys97–Gly104 loop), it is important to compare the triclinic structures with models of HEWL obtained in different space groups (Fig. 3). The Lys97–Gly104 loop appears to be disordered, regardless of the HEWL crystal system. This flexibility is promoted by the presence of Gly102 and Gly104, which increase the rotational freedom about the peptide bonds. Common to the monoclinic (PDB entry 3wl2), tetragonal (PDB entry 1iee; Sauter *et al.*, 2001) and orthorhombic (PDB entry 6f1o; Plaza-Garrido *et al.*, 2018) structures, Asn103 of native H-HEWL seems to be flipped, whereas in the triclinic system this flip is only partial, with the major conformation being the *trans* peptide bond. Superposition of these HEWL structures reveals that their main deviations are seen for residues Asp101–Gly104 (Fig. 3). While the triclinic HEWL models containing the partial peptide flip of Asn103 seem to be influenced by the crystal contact between the Asp101 and Glu7 side chains, the H-HEWL_{EC} structure lacks this interaction, similar to other HEWL crystal systems. This observation suggests that in solution the Lys97–Gly104 loop adopts a conformation closer to that seen in the monoclinic, tetragonal and orthorhombic structures, and one could therefore expect that H-HEWL_{EC} would adopt a similar arrangement. However, it is clear that its Lys97–Gly104 loop deviates significantly from the expected conformation, particularly between Asp101 and Gly104. Furthermore, in H-HEWL_{EC} a different main-chain hydrogen-bond pattern can be observed when compared with the other native H-HEWL models. While this loop is stabilized by an Asp101 N(H)–Ile98 O hydrogen bond in monoclinic, tetragonal and orthorhombic H-HEWL, in H-HEWL_{EC} this interaction is replaced by a Gly102 N(H)–Ile98 O hydrogen bond. Additionally, an

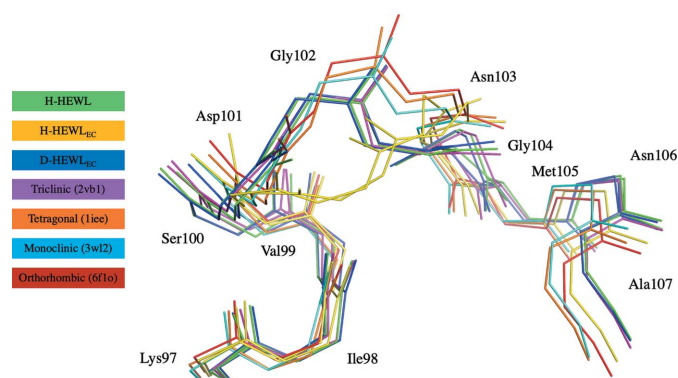


Figure 3
The backbone structure of the H-HEWL_{EC} Lys97–Gly104 loop (extended to Ala107) is considerably different from those of D-HEWL_{EC}, H-HEWL and of HEWL in different crystal systems. PDB entries 2vb1, 1iee, 3wl2 and 6f1o were chosen as representative of native H-HEWL crystallized in the triclinic, tetragonal, monoclinic and orthorhombic systems, respectively. The Lys97–Ala107 regions of the different variants were structurally aligned; for clarity, only the main-chain atoms are shown. This illustration was produced using PyMOL (version 2.0; Schrödinger).

unusual Gly104 N(H)–Gly102 O hydrogen bond further stabilizes the Lys97–Gly104 loop of H-HEWL_{EC}. It is noteworthy that H-HEWL in other crystal systems contains a water molecule (W312 in PDB entry 3wl2, W1010 in PDB entry liee and W316 in PDB entry 6f1o) that plays an important role in shaping this loop by mediating hydrogen-bond interactions between Val99 O, Gly102 N(H) and Gly104 N(H) (Supplementary Fig. S6). In the triclinic form of H-HEWL this water molecule is displaced and is only present when the Asn103 peptide plane is flipped; it no longer interacts with Gly102 N(H), which instead hydrogen bonds to Lys97 O.

Since X-ray structures can contain artifacts when studied at cryogenic temperatures (Halle, 2004), room-temperature models of native H-HEWL were also considered, namely the triclinic PDB entry 4lzt (Walsh *et al.*, 1998) and the tetragonal PDB entry 1bwj (Dong *et al.*, 1999). In the triclinic HEWL, a configuration of the Lys97–Gly104 loop similar to those of our native H-HEWL and D-HEWL_{EC} can be observed. Regarding the tetragonal model, the same water molecule (W148) as in the monoclinic, tetragonal and orthorhombic 100 K structures can be found hydrogen-bonding to Val99 O, Gly102 N(H) and Gly104 N(H) (Supplementary Fig. S6). These observations support the assumption that the structural changes present in our H-HEWL_{EC} are not artifacts from collecting X-ray data at cryogenic temperatures but rather reflect differences in the protein folding.

As noted above, in triclinic H-HEWL the Glu7 side chain is involved in crystal contacts with the Lys97–Gly104 loop through the Asp101 O^{δ2}–Glu7 O^{ε2} hydrogen bond. Meanwhile, in other H-HEWL crystal systems Glu7 adopts a different conformation, which is stabilized by hydrogen bonds to Lys1 N(H)^ε, Gly4 N(H) and Val2 O, with the latter being mediated by a water molecule. For clarity, the conformation of Glu7 found in the triclinic system will be referred to as conformation *A*, while the alternate conformation in other H-HEWL systems will be named conformation *B*.

Our triclinic H-HEWL structure displays Glu7 in conformation *A* with full occupancy. On the other hand, in D-HEWL_{EC} the Glu7 side chain is found in conformation *A* with a refined occupancy of 49% and in conformation *B* with

an occupancy of 51%. Interestingly, the refined occupancies of both Glu7 and Lys97–Gly104 in conformation *B* are similar (51% and 46%, respectively). These observations indicate that the displacement of the Lys97–Gly104 loop associated with the Asn103 peptide flip is linked to the disruption of the Asp101 O^{δ2}–Glu7 O^{ε2} hydrogen bond and the shift of the Glu7 side chain to conformation *B*. In the case of H-HEWL_{EC}, where there is a greater displacement of the Lys97–Gly104 loop and where the Asp101 O^{δ2}–Glu7 O^{ε2} hydrogen bond is completely disrupted, the Glu7 side chain is found to fully occupy conformation *B* (Fig. 4).

The conformation of Glu7 in H-HEWL_{EC} appears to be further linked to the displacement of the Thr47–Gly49 region via crystal contacts. Disorder is present in the Thr47–Gly49 region of the three HEWL variants; however, their hydrogen-bond patterns display noteworthy variations, in particular for H-HEWL_{EC} compared with both H-HEWL and D-HEWL_{EC}. In H-HEWL and D-HEWL_{EC} Thr47/*A* is involved in water-mediated hydrogen bonds to Leu75 O and Glu7/*A* O^{ε2} via crystal contacts, while Thr47/*B* appears to establish hydrogen bonds to Leu75 O and Asn74 O. On the other hand, in H-HEWL_{EC} the fact that Glu7 is in a single conformation, identical to that of Glu7/*B* in D-HEWL_{EC}, eliminates the water-mediated hydrogen bonds to Glu7/*A* O^{ε2} and Leu75 O via crystal contacts found in both H-HEWL and D-HEWL_{EC}. Moreover, this water molecule seems to be absent in the H-HEWL_{EC} model. Therefore, while Thr47/*A* O^{γ1} participates in weak hydrogen-bond interactions with Asp48 mediated by W367, Thr47/*B* O^{γ1} can form hydrogen bonds to W365, W314 and Glu7 O^{ε2} via crystal contacts (Supplementary Fig. S7).

3.6. *In vitro* refolding perturbs the configuration of the HEWL binding cleft

H-HEWL_{EC} displays significant differences in the position of a number of binding-cleft residues (Fig. 5). Both Asn46 and Asp48 appear to be displaced in comparison with H-HEWL and D-HEWL_{EC}, which can be explained by the crystal contacts previously described with Glu7, affecting the Thr47–Gly49 region. More importantly, the Lys97–Gly104 loop is considerably displaced, most noticeably affecting the residues

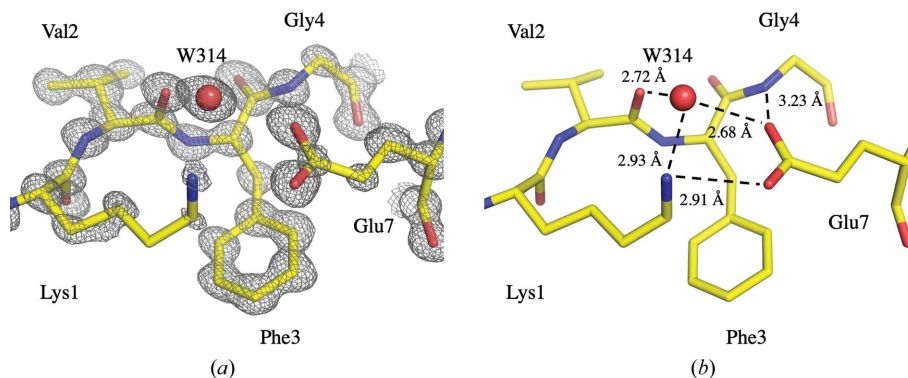


Figure 4

The single Glu7 conformation present in H-HEWL_{EC}. The $2F_o - F_c$ electron-density map shown was contoured at 2σ (a) and the respective hydrogen-bond interactions with Lys1 N^ε, Gly4 N and Val2 O are depicted in (b). W314 mediates the interaction between Glu7 and Val2 through hydrogen bonds. This illustration was produced using *PyMOL* (version 2.0; Schrödinger).

Asp101–Gly104. This variation in protein structure is directly linked to the distinct hydrogen-bond pattern found in the Lys97–Gly104 loop, which stabilizes an unusual conformation of the protein backbone when compared with other H-HEWL variants. This shift in the protein main-chain structure seems to be propagated until Val109, due to the Asn106 N(H)–Asn103 O and Ala107 N(H)–Gly104 O hydrogen bonds. Additionally, displacement of the water-molecule arrangements can be observed in the active site and binding cleft of the enzyme. It is also interesting to note the absence of a nitrate ion in the H-HEWL_{EC} binding cleft which is present in the other triclinic HEWL variants (Ramos *et al.*, 2021).

An alternate conformation of Asn44 is found in H-HEWL_{EC} (Supplementary Fig. S8), which has not been observed in either triclinic H-HEWL or D-HEWL_{EC} (Ramos *et al.*, 2021) or in the 0.65 Å resolution H-HEWL structure determined by Wang *et al.* (2007). The refined occupancy of Asn44 in this conformation is 55%, where it is engaged in a 2.68 Å hydrogen bond to Gln57 O^{δ1}. Additionally, the minor conformation, which can be found in the other triclinic HEWL variants, appears to be significantly disordered, with clear electron density only for the side-chain C^γ in the 2F_o – F_c electron-density map (contoured at 1σ, Supplementary Fig. S8). In addition, a fully occupied water molecule (W387) is present, forming hydrogen bonds to Asn46 O^{δ1} and the catalytic residue Asp52 O^{δ2}.

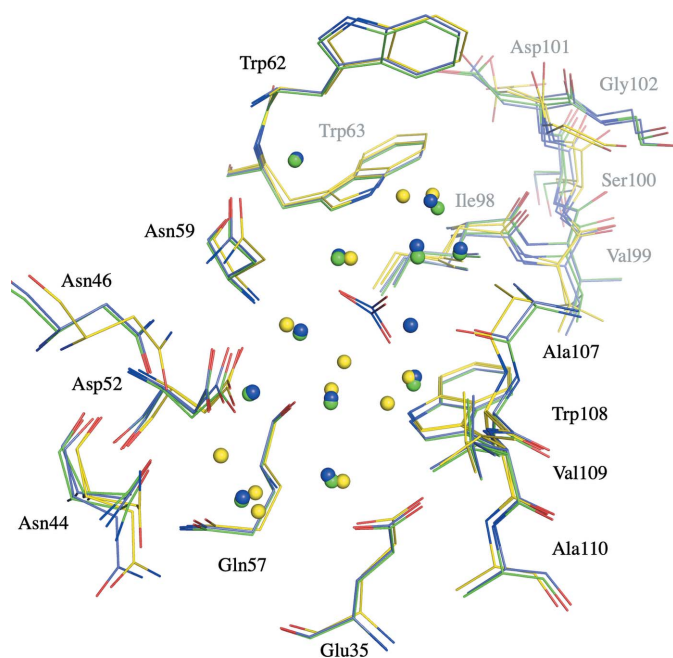


Figure 5
The configuration of the enzymatic binding cleft of H-HEWL_{EC} (C atoms in yellow) is altered compared with those of H-HEWL (C atoms in green) and D-HEWL_{EC} (C atoms in blue). The crystal structures of H-HEWL_{EC} and D-HEWL_{EC} were structurally aligned with that of H-HEWL and the active-site and binding-cleft residues of the enzyme are shown as sticks. The water molecules are represented as spheres colored according to the color of the respective model (*i.e.* the color of the model C atoms) and they are not labeled for clarity. The nitrate ions placed at the center of the figure belong to the D-HEWL_{EC} and H-HEWL models. This illustration was produced using PyMOL (version 2.0; Schrödinger).

4. Discussion

4.1. Deuteration and *in vitro* refolding differently affect the biophysical properties of HEWL

In this study, the contributions of *in vitro* refolding, protein perdeuteration and solvent isotope substitution to the thermal stability of HEWL have been quantified. Surprisingly, the observations suggest that *in vitro* refolding has a stronger impact on HEWL thermal stability than protein perdeuteration, with the respective decreases in T_m being greater than 3.5°C compared with variations of smaller than 1.6°C. These changes in HEWL thermal stability are consistent throughout different buffer compositions and also in H₂O and D₂O solvents, where the minor deviations can be explained by the aforementioned difference between pH and pD (Glasoe & Long, 1960). Moreover, the reduction in protein thermal stability associated with protein perdeuteration is in agreement with several previous studies (Berns, 1963; Hattori *et al.*, 1965; Brockwell *et al.*, 2001; Meilleur *et al.*, 2004; Koruza *et al.*, 2018; Nichols *et al.*, 2020; Ramos *et al.*, 2021).

The effects of solvent isotope substitution were measurable and of similar magnitude in all three HEWL variants. Protein thermal stability was increased by replacing the H₂O solvent by D₂O, with differences in T_m of between 1.9°C and 2.8°C. The small variations observed between H-HEWL and both HEWL_{EC} variants could be linked to distinct solvation structures relating to the protein-refolding process or to the addition of a Gly residue at the N-terminus of HEWL_{EC}. Nevertheless, the measurement of higher protein thermal stability in D₂O when compared with H₂O is consistent with previous reports (Hattori *et al.*, 1965; Harrington & von Hippel, 1961; Makhatadze *et al.*, 1995; Dong *et al.*, 1997; Kuhlman & Raleigh, 1998; Sasisanker *et al.*, 2004; Efimova *et al.*, 2007). It is noted that the solvent isotope effect appears to be significantly stronger than the macromolecular isotope effect regarding protein thermal stability. This observation suggests that weakening the hydrophobic interactions of the protein residues through perdeuteration perturbs protein structure and stability to a lower degree compared with the changes in protein solvation caused by solvent isotope substitution from H₂O to D₂O.

It is evident that all three HEWL variants are less thermally stable in buffer solutions at pH/pD 7.5 than at pH/pD 4.5. A direct comparison of the results obtained at different pH/pD values was not possible since the buffer solutions have very different compositions. Nevertheless, since the pI of HEWL is 11, it was expected that the different variants would be less thermally stable at pH/pD 7.5 compared with a more acidic buffer solution. This observation is also consistent with the optimal pH of 5 for HEWL activity.

4.2. Perdeuteration can enable different folding modes of HEWL

The atomic resolution H-HEWL_{EC} crystal structure provides valuable insight regarding the effects of both *in vitro* refolding and perdeuteration on protein structure, which can be linked to differences in thermal stability and enzymatic

activity. Although the protein fold is shown to be conserved in H-HEWL_{EC} (compared with H-HEWL), significant variations are observed that (as in the case of D-HEWL_{EC}) stem from the Lys97–Gly104 loop. In D-HEWL_{EC} the disorder of this region seems to be favored by protein perdeuteration and *in vitro* refolding, while similar observations were made for the perdeuterated HEWL variant expressed in *Pichia pastoris*, which was produced in fully deuterated medium (Ramos *et al.*, 2021). However, it was not possible to disentangle perdeuteration and refolding effects as the structure of H-HEWL_{EC} was unknown. The slower folding dynamics caused by a combination of more viscous solvents and protein perdeuteration is likely to contribute to the increased probability of the Asn103 peptide-plane flip occurring, as well as the subsequent displacement of the Lys97–Gly104 loop (Ramos *et al.*, 2021). Unexpectedly, the structure of H-HEWL_{EC} reveals that this loop is significantly perturbed by the conditions of *in vitro* refolding. The resulting unusual Lys97–Gly104 loop conformation has important consequences for the configuration of the H-HEWL_{EC} binding cleft and for the typical crystal contacts of triclinic HEWL. Interestingly, the elution profiles of H-HEWL_{EC} and D-HEWL_{EC} during refolding show variations (Supplementary Fig. S9), which may arise from the impact of the macromolecular isotope effect in terms of mass effects on dynamics. In the case of H-HEWL_{EC} the quantities of aggregated and unfolded or misfolded protein appear to be greater, and as a consequence the refolding yield is lower in comparison with D-HEWL_{EC}. This study thus revealed a greater perturbation by *in vitro* refolding than expected, and it appears that the isotope effects reduce or counteract most of these perturbations through changed dynamics.

Several crystallographic studies have provided insight into the unfolding and refolding processes of HEWL (Raskar *et al.*, 2016, 2019; Kita & Morimoto, 2016, 2020). The binding of guanidinium and urea molecules to HEWL is of particular interest since both H-HEWL_{EC} and D-HEWL_{EC} were refolded by gradually substituting a 6 M guanidine–HCl buffer solution with one containing 2 M urea. The structural changes observed in HEWL–guanidinium/urea complexes (Raskar *et al.*, 2016, 2019) are not found in any of our refolded HEWL_{EC} variants. However, an inspection of the $2F_o - F_c$ electron-density maps of the HEWL–urea complexes (PDB entries 5i4x, 5i54, 5i53 and 5i4y; Raskar *et al.*, 2016) revealed the potential binding of a urea molecule to the Lys97–Gly104 loop, specifically through hydrogen-bond interactions with Ile98 O, Asp101 O and Asn103 N(H) and O^{δ1}. This observation suggests that the folding pathways of both H-HEWL_{EC} and D-HEWL_{EC} could be influenced by interaction with urea molecules, particularly at the Lys97–Gly104 loop, which has been shown to differ significantly from the native H-HEWL structure.

The main difference in the production of H-HEWL_{EC} and D-HEWL_{EC} is that the latter contains 698 D atoms instead of H atoms, which leads to an increase in molecular weight of 702 Da. This discrepancy has consequences for protein dynamics and importantly for the effect of dynamics on refolding. One would expect that the lighter protein

(H-HEWL_{EC}) would experience a faster folding process. Additionally, the strength of hydrophobic interactions, essential to protein folding, is significantly affected by isotope substitution. Hydrophobic interactions are weakened by deuteration due to the reduced vibrational amplitude of the C–D bond compared with C–H (Hattori *et al.*, 1965). The combination of the aforementioned factors suggests that the hydrophobic core of H-HEWL_{EC} would collapse faster than that of D-HEWL_{EC} and result in more stable and restrictive folding intermediates. This is likely to be the reason why H-HEWL_{EC} displays a reduced molecular surface area and fewer solvation water molecules compared with H-HEWL and D-HEWL_{EC}. The formation of secondary structure through hydrogen-bond interactions is also facilitated in H-HEWL_{EC} due to the faster dynamics, permitting a larger conformational energy landscape than for the perdeuterated variant. It is at this stage that H-HEWL_{EC} would select the most favorable conformation of the Lys97–Gly104 loop, which, as shown in non-triclinic HEWL systems, involves a complete Asn103 peptide-plane flip. This conformation appears to be stabilized by unusual hydrogen bonds, including Gly102 N(H)–Ile98 O and Gly104 N(H)–Gly102 O. The fact that the Lys97–Gly104 loop is part of the HEWL α -domain reveals its importance for both protein folding and stability. In both D-HEWL_{EC} and native H-HEWL the Lys97–Gly104 loop is allowed greater flexibility, as shown by the observation of both *cis* and *trans* conformations of the Asn103 peptide bond in their crystal structures (Ramos *et al.*, 2021). Although it is assumed that peptide-plane flips occur in the early stages of protein folding and are later stabilized by hydrogen-bond formation (Hayward, 2001), HEWL crystal structures in various space groups suggest that these peptide flips are dynamical and play a role in protein flexibility. The more rigid configuration of the Lys97–Gly104 loop in H-HEWL_{EC} formed in the early stages of folding suggests a different pathway of folding than for H-HEWL and possibly even D-HEWL_{EC}. Takano *et al.* (2000) have shown that a deletion mutant of Arg101 in human lysozyme (corresponding to Ser100 in HEWL) presents different folding kinetics compared with the wild-type protein. Furthermore, the rigidity of this region in H-HEWL_{EC} demonstrates its greater local structural stability, which hinders the overall plasticity of the protein and subsequently its resistance to thermal denaturation, as shown by our biophysical data.

4.3. The disorder of the Lys97–Gly104 region may play a role in amyloidogenesis

Interestingly, the Lys97–Gly104 loop is suggested to be involved in the formation of amyloid fibrils of HEWL. Frare *et al.* (2006) have shown that the core structure in human lysozyme amyloids encompasses the entire α -helix C and part of the β -domain (residues 32–108). Specific fragments of HEWL (Gly49–Asp101, Tyr53–Ser100, Tyr53–Asp101 and Gln57–Ala107) were found to be highly amyloidogenic, while the remaining regions of the protein remain largely soluble (Frare *et al.*, 2004). Taken alongside our structural data, these findings

could indicate that HEWL α -helix C, particularly the region extending from Lys97 to Gly104, is of crucial importance in the protein-folding process and its destabilization may lead to different folding modes. This information is especially relevant in a clinical context due to the high homology between HEWL and human lysozyme. A possible link to the Lys97–Gly104 loop is seen in the amyloidogenic variants Trp64Arg and Trp112Arg of human lysozyme (corresponding to Trp63 and Trp111 in HEWL; Valleix *et al.*, 2002; Sperry *et al.*, 2016). Both Trp63 and Trp111 appear to play a crucial role in stabilizing the Lys97–Gly104 region. Trp63 enables CH– π interactions with Ile98 and the disorder of Ile98 is reflected in a similar Trp63 side-chain disorder. Meanwhile, Trp111 is part of a complex network of hydrophobic interactions that stabilize the conformation of Met105, also including Tyr23, Trp28 and Trp108.

4.4. The configuration of the Lys97–Gly104 loop impacts the activity of HEWL

The flexibility of the Lys97–Gly104 loop also plays a role in the activity of native HEWL, since this region is thought to be responsible for the interaction with the first three carbohydrate residues (A, B and C) of the substrate (Blake *et al.*, 1967; Phillips, 1967). Several X-ray crystal structures of HEWL bound to inhibitor carbohydrates have proven the importance of this region for enzymatic activity (Strynadka & James, 1991; Cheetham *et al.*, 1992; Maenaka *et al.*, 1995; Tanaka *et al.*, 2021). In particular, Asp101, which is conserved in several c-type lysozymes (Supplementary Fig. S10), is believed to make hydrogen bonds to substrate residues A and B (Blake *et al.*, 1967; Phillips, 1967). Although further clarification would require the structure of an H-HEWL_{EC}–inhibitor complex, one could interpret the observed changes of the protein backbone and water structures in the binding cleft of the apo form as an indication that its ability to interact with the substrate molecule is hindered.

The observations made for the structure of H-HEWL_{EC} can provide an explanation for the reduced enzymatic activity reported for both nonrecombinant and recombinant H-HEWL in previous studies (Batas & Chaudhuri, 1996; Batas *et al.*, 1999). In the present work, H-HEWL_{EC} and D-HEWL_{EC} were refolded in identical conditions and subsequent enzymatic assays were performed for all HEWL variants, including native H-HEWL. Despite the similar levels of function of H-HEWL_{EC} and D-HEWL_{EC}, the reasons for the reduction in activity by comparison with H-HEWL seem to differ according to their crystal structures. While the function of H-HEWL_{EC} appears to be hindered by differences in the configuration of the binding cleft of the enzyme, D-HEWL_{EC} activity might be primarily affected by the macromolecular isotope effect, which alters protein dynamics.

5. Conclusions

In summary, this study has shown that *in vitro* refolding may lead to significant changes in protein structure, affecting thermal stability and activity; these aspects are very important

in the context of the folding pathways that occur *in vivo* and *in vitro*. Nevertheless, this method of protein production from inclusion bodies is an attractive approach to facilitate recombinant protein production and improve protein yields (Ramos *et al.*, 2021). Intriguingly, the perdeuterated protein was found to be closer in structure to native H-HEWL than it was to its hydrogenated variant; this indicates that the folding dynamics are different for the hydrogenated and perdeuterated analogs. These observations emphasize the significance of structural data in validating *in vitro* refolding approaches and in understanding any differences in behavior and function. Finally, the effects of deuteration (macromolecular and solvent) in protein-folding dynamics have not been extensively studied to date; however, the underlying slower dynamics of perdeuterated proteins could be of interest in mimicking the crowded cellular environment and the action of chaperones in which proteins are natively folded.

Acknowledgements

JR thanks the Institut Laue–Langevin (ILL) for a PhD studentship. All authors thank the ILL for access to its Deuteration Laboratory within the Partnership for Structural Biology (PSB; <https://www.psb-grenoble.eu/>). We acknowledge MAX IV Laboratory for time on the BioMAX Beamline under Proposal 20200259. Research conducted at MAX IV, a Swedish national user facility, is supported by the Swedish Research council under contract 2018-07152, the Swedish Governmental Agency for Innovation Systems under contract 2018-04969 and Formas under contract 2019-02496.

Funding information

We thank DANSCATT for travel support to MAX IV, which was made possible through funding from the Danish Agency for Science, Technology and Innovation. AEL acknowledges funding from the Lundbeck Foundation through the BRAINSTRUC Structural Biology Initiative (R155-2015-2666). VTF acknowledges support from the UK Engineering and Physical Sciences Research Council (EPSRC), which funded the Deuteration Laboratory (D-Lab) in ILL's Life Sciences Group under grants GR/R993931 and EP/C015452/1.

References

- Artero, J.-B., Härtlein, M., McSweeney, S. & Timmins, P. (2005). *Acta Cryst.* **D61**, 1541–1549.
- Batas, B. & Chaudhuri, J. B. (1996). *Biotechnol. Bioeng.* **50**, 16–23.
- Batas, B., Schiraldi, C. & Chaudhuri, J. B. (1999). *J. Biotechnol.* **68**, 149–158.
- Berman, H., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I. & Bourne, P. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Berns, D. S. (1963). *Biochemistry*, **2**, 1377–1380.
- Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967). *Proc. R. Soc. London B.* **167**, 378–388.
- Breyton, C., Gabel, F., Lethier, M., Flayhan, A., Durand, G., Jault, J., Juillan-Binard, C., Imbert, L., Moulin, M., Ravaud, S., Härtlein, M. & Ebel, C. (2013). *Eur. Phys. J. E*, **36**, 71.

- Brockwell, D., Yu, L., Cooper, S., McClelland, S., Cooper, A., Attwood, D., Gaskell, S. J. & Barber, J. (2001). *Protein Sci.* **10**, 572–580.
- Budayova-Spano, M., Lafont, S., Astier, J.-P., Ebel, C. & Veessler, S. (2000). *J. Cryst. Growth*, **217**, 311–319.
- Campbell, R. A., Tummino, A., Varga, I., Milyaeva, O. Y., Krycki, M. M., Lin, S. Y., Laux, V., Haertlein, M., Forsyth, V. T. & Noskov, B. A. (2018). *Langmuir*, **34**, 5020–5029.
- Canet, D., Last, A. M., Tito, P., Sunde, M., Spencer, A., Archer, D. B., Redfield, C., Robinson, C. V. & Dobson, C. M. (2002). *Nat. Struct. Biol.* **9**, 308–315.
- Cheetham, J. C., Artymiuk, P. J. & Phillips, D. C. (1992). *J. Mol. Biol.* **224**, 613–628.
- Cooper, S. J., Raftery, J., Helliwell, J. R., Brockwell, D., Attwood, D. & Barber, J. (1998). *Chem. Commun.*, pp. 1063–1064.
- Cuypers, M. G., Mason, S. A., Blakeley, M. P., Mitchell, E. P., Haertlein, M. & Forsyth, V. T. (2013). *Angew. Chem. Int. Ed.* **52**, 1022–1025.
- Dong, A., Kendrick, B., Kreilgård, L., Matsuura, J., Manning, M. C. & Carpenter, J. F. (1997). *Arch. Biochem. Biophys.* **347**, 213–220.
- Dong, J., Boggon, T. J., Chayen, N. E., Raftery, J., Bi, R.-C. & Helliwell, J. R. (1999). *Acta Cryst. D* **55**, 745–752.
- Dunne, O., Weidenhaupt, M., Callow, P., Martel, A., Moulin, M., Perkins, S. J., Haertlein, M. & Forsyth, V. T. (2017). *Eur. Biophys. J.* **46**, 425–432.
- Efimova, Y. M., Haemers, S., Wierczynski, B., Norde, W. & Well, A. A. (2007). *Biopolymers*, **85**, 264–273.
- Fisher, S. J. & Helliwell, J. R. (2008). *Acta Cryst. A* **64**, 359–367.
- Foglia, F., Hazael, R., Meersman, F., Wilding, M. C., Sakai, V. G., Rogers, S., Bove, L. E., Koza, M. M., Moulin, M., Haertlein, M., Forsyth, V. T. & McMillan, P. F. (2019). *Sci. Rep.* **9**, 8716.
- Frare, E., Mossuto, M. F., Polverino de Laureto, P., Dumoulin, M., Dobson, C. M. & Fontana, A. (2006). *J. Mol. Biol.* **361**, 551–561.
- Frare, E., Polverino de Laureto, P., Zurdo, J., Dobson, C. M. & Fontana, A. (2004). *J. Mol. Biol.* **340**, 1153–1165.
- Freyman, T. M., Yannas, I. V. & Gibson, L. J. (2001). *Prog. Mater. Sci.* **46**, 273–282.
- Gajdos, L., Blakeley, M. P., Kumar, A., Wimmerová, M., Haertlein, M., Forsyth, V. T., Imberty, A. & Devos, J. M. (2021). *Structure*, **29**, 1003–1013.
- Gamble, T. R., Clauser, K. R. & Kossiakoff, A. A. (1994). *Biophys. Chem.* **53**, 15–25.
- Gardner, K. H., English, A. D. & Forsyth, V. T. (2004). *Macromolecules*, **37**, 9654–9656.
- Gillmore, J. D., Booth, D. R., Madhoo, S., Pepys, M. B. & Hawkins, P. N. (1999). *Nephrol. Dial. Transpl.* **14**, 2639–2644.
- Girnius, S., Skinner, M., Spencer, B., Prokaeva, T., Bartholomew, C., O'Hara, C., Seldin, D. C. & Connors, L. H. (2012). *Amyloid*, **19**, 182–185.
- Glasoe, P. K. & Long, F. A. (1960). *J. Phys. Chem.* **64**, 188–190.
- Halle, B. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 4793–4798.
- Harrington, W. F. & von Hippel, P. H. (1961). *Arch. Biochem. Biophys.* **92**, 100–113.
- Hattori, A., Crespi, H. L. & Katz, J. J. (1965). *Biochemistry*, **4**, 1213–1225.
- Hayward, S. (2001). *Protein Sci.* **10**, 2219–2227.
- Hazemann, I., Dauvergne, M. T., Blakeley, M. P., Meilleur, F., Haertlein, M., Van Dorsselaer, A., Mitschler, A., Myles, D. A. A. & Podjarny, A. (2005). *Acta Cryst. D* **61**, 1413–1417.
- Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986). *Biochemistry*, **25**, 6965–6972.
- Jasnin, M., Tehei, M., Moulin, M., Haertlein, M. & Zaccari, G. (2008). *Eur. Biophys. J.* **37**, 613–617.
- Jean, E., Ebbo, M., Valleix, S., Benarous, L., Heyries, L., Grados, A., Bernit, E., Grateau, G., Papo, T., Granel, B., Daniel, L., Harlé, J. & Schleinitz, N. (2014). *BMC Gastroenterol.* **14**, 159.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 133–144.
- Kehlenbeck, D.-M., Josts, I., Nitsche, J., Busch, S., Forsyth, V. T. & Tidow, H. (2019). *Biol. Chem.* **400**, 1509–1518.
- Kelpšas, V., Caldararu, O., Blakeley, M. P., Coquelle, N., Wierenga, R. K., Ryde, U., von Wachenfeldt, C. & Oksanen, E. (2021). *IUCrJ*, **8**, 633–643.
- Kita, A. & Morimoto, Y. (2016). *Mol. Biotechnol.* **58**, 130–136.
- Kita, A. & Morimoto, Y. (2020). *J. Appl. Cryst.* **53**, 837–840.
- Koruza, K., Lafumat, B., Nyblom, M., Mahon, B. P., Knecht, W., McKenna, R. & Fisher, S. Z. (2019). *Acta Cryst. D* **75**, 895–903.
- Koruza, K., Lafumat, B., Végvári, Á., Knecht, W. & Fisher, S. Z. (2018). *Arch. Biochem. Biophys.* **645**, 26–33.
- Kuhlman, B. & Raleigh, D. P. (1998). *Protein Sci.* **7**, 2405–2412.
- Liu, X., Hanson, B. L., Langan, P. & Viola, R. E. (2007). *Acta Cryst. D* **63**, 1000–1008.
- Maenaka, K., Matsushima, M., Song, H., Sunada, F., Watanabe, K. & Kumagai, I. (1995). *J. Mol. Biol.* **247**, 281–293.
- Makhatadze, G. I., Clore, G. M. & Gronenborn, A. M. (1995). *Nat. Struct. Mol. Biol.* **2**, 852–855.
- Manzoni, F., Wallerstein, J., Schrader, T. E., Ostermann, A., Coates, L., Akke, M., Blakeley, M. P., Oksanen, E. & Logan, D. T. (2018). *J. Med. Chem.* **61**, 4412–4420.
- Maric, S., Skar-Gislinge, N., Midtgaard, S., Thygesen, M. B., Schiller, J., Frielinghaus, H., Moulin, M., Haertlein, M., Forsyth, V. T., Pomorski, T. G. & Arleth, L. (2014). *Acta Cryst. D* **70**, 317–328.
- Maric, S., Thygesen, M. B., Schiller, J., Marek, M., Moulin, M., Haertlein, M., Forsyth, V. T., Bogdanov, M., Dowhan, W., Arleth, L. & Pomorski, T. G. (2015). *Appl. Microbiol. Biotechnol.* **99**, 241–254.
- McGregor, L., Földes, T., Bui, S., Moulin, M., Coquelle, N., Blakeley, M. P., Rosta, E. & Steiner, R. A. (2021). *IUCrJ*, **8**, 46–59.
- Meilleur, F., Contzen, J., Myles, D. A. A. & Jung, C. (2004). *Biochemistry*, **43**, 8744–8753.
- Miranker, A., Radford, S. E., Karplus, M. & Dobson, C. M. (1991). *Nature*, **349**, 633–636.
- Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T. & Dobson, C. M. (1993). *Science*, **262**, 896–900.
- Moulin, M., Strohmeier, G. A., Hirz, A., Mant, C., Thompson, K. C., Rennie, A. R., Campbell, R. A., Pichler, H., Maric, S., Forsyth, V. T. & Haertlein, M. (2018). *Chem. Phys. Lipids*, **212**, 80–87.
- Nasr, S. H., Dasari, S., Mills, J. R., Theis, J. D., Zimmermann, M. T., Fonseca, R., Vrana, J. A., Lester, S. J., McLaughlin, B. M., Gillespie, R., Highsmith, W. E., Lee, J. J., Dispenzieri, A. & Kurtin, P. J. (2017). *J. Am. Soc. Nephrol.* **28**, 431–438.
- Nichols, P. J., Falconer, I., Griffin, A., Mant, C., Hodges, R., McKnight, C. J., Vögeli, B. & Vugmeister, L. (2020). *Protein Sci.* **29**, 1641–1654.
- Nitsche, J., Josts, I., Heidemann, J., Mertens, H. D., Maric, S., Moulin, M., Haertlein, M., Busch, S., Forsyth, V. T., Svergun, D. I., Uetrecht, C. & Tidow, H. (2018). *Commun. Biol.* **1**, 206.
- Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M. & Hsuan, J. J. (1993). *Nature*, **362**, 553–557.
- Petit-Haertlein, I., Blakeley, M. P., Howard, E., Hazemann, I., Mitschler, A., Haertlein, M. & Podjarny, A. (2009). *Acta Cryst. F* **65**, 406–409.
- Phillips, D. C. (1967). *Proc. Natl Acad. Sci. USA*, **57**, 483–495.
- Plaza-Garrido, M., Salinas-Garcia, M. C. & Camara-Artigas, A. (2018). *Acta Cryst. D* **74**, 480–489.
- Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). *Nature*, **358**, 302–307.
- Ramos, J., Laux, V., Haertlein, M., Boeri Erba, E., McAuley, K. E., Forsyth, V. T., Mossou, E., Larsen, S. & Langkilde, A. E. (2021). *IUCrJ*, **8**, 372–386.
- Raskar, T., Khavnekar, S. & Hosur, M. (2016). *Sci. Rep.* **6**, 32277.
- Raskar, T., Koh, C. Y., Niebling, S., Kini, R. M. & Hosur, M. V. (2019). *Int. J. Biol. Macromol.* **122**, 903–913.
- Sasisanker, P., Oleinikova, A., Weingärtner, H., Ravindra, R. & Winter, R. (2004). *Phys. Chem. Chem. Phys.* **6**, 1899–1905.

- Sauter, C., Otálora, F., Gavira, J.-A., Vidal, O., Giegé, R. & García-Ruiz, J. M. (2001). *Acta Cryst. D* **57**, 1119–1126.
- Sheldrick, G. M. (2015). *Acta Cryst. C* **71**, 3–8.
- Shikano, S. & Colley, K. J. (2013). *Encyclopedia of Biological Chemistry*, 2nd ed., edited by W. J. Lennarz & M. D. Lane, pp. 203–209. New York: Academic Press.
- Shugar, D. (1952). *Biochim. Biophys. Acta*, **8**, 302–309.
- Sperry, B., Hanna, M., Ikram, A., Theis, J., Leung, N., Highsmith, W. E., Grogan, M. & Dispenzieri, A. (2016). *XVth International Symposium on Amyloidosis*, p. 432. Madison: International Society of Amyloidosis.
- Strynadka, N. C. J. & James, M. N. G. (1991). *J. Mol. Biol.* **220**, 401–424.
- Svergun, D. I., Richard, S., Koch, M. H. J., Sayers, Z., Kuprin, S. & Zaccari, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2267–2272.
- Takano, K., Yamagata, Y. & Yutani, K. (2000). *Biochemistry*, **39**, 8655–8665.
- Tanaka, I., Nishinomiya, R., Goto, R., Shimazaki, S. & Chatake, T. (2021). *Acta Cryst. D* **77**, 288–292.
- Ursby, T., Åhnberg, K., Appio, R., Aurelius, O., Barczyk, A., Bartalesi, A., Bjelčić, M., Bolmsten, F., Cerenius, Y., Doak, R. B., Eguiraun, M., Eriksson, T., Friel, R. J., Gorgisyan, I., Gross, A., Haghighat, V., Hennies, F., Jagudin, E., Norsk Jensen, B., Jeppsson, T., Kloos, M., Lidon-Simon, J., de Lima, G. M. A., Lizatovic, R., Lundin, M., Milan-Otero, A., Milas, M., Nan, J., Nardella, A., Rosborg, A., Shilova, A., Shoeman, R. L., Siewert, F., Sondhauss, P., Talibov, V. O., Tarawneh, H., Thånell, J., Thunnissen, M., Unge, J., Ward, C., Gonzalez, A. & Mueller, U. (2020). *J. Synchrotron Rad.* **27**, 1415–1429.
- Valleix, S., Drunat, S., Philit, J.-B., Adoue, D., Piette, J.-C., Droz, D., MacGregor, B., Canet, D., Delpech, M. & Grateau, G. (2002). *Kidney Int.* **61**, 907–912.
- Varga, K., Aslimovska, L., Parrot, I., Dauvergne, M. T., Haertlein, M., Forsyth, V. T. & Watts, A. (2007). *Biochim. Biophys. Acta*, **1768**, 3029–3035.
- Wacklin, H. P., Bremec, B. B., Moulin, M., Rojko, N., Haertlein, M., Forsyth, T., Anderluh, G. & Norton, R. S. (2016). *Biochim. Biophys. Acta*, **1858**, 640–652.
- Wada, M., Nishiyama, Y., Bellesia, G., Forsyth, T., Gnanakaran, S. & Langan, P. (2011). *Cellulose*, **18**, 191–206.
- Waldie, S., Lind, T. K., Browning, K., Moulin, M., Haertlein, M., Forsyth, V. T., Luchini, A., Strohmeier, G. A., Pichler, H., Maric, S. & Cárdenas, M. (2018). *Langmuir*, **34**, 472–479.
- Waldie, S., Moulin, M., Porcar, L., Pichler, H., Strohmeier, G. A., Skoda, M., Forsyth, V. T., Haertlein, M., Maric, S. & Cárdenas, M. (2019). *Sci. Rep.* **9**, 5118.
- Walsh, M. A., Schneider, T. R., Sieker, L. C., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1998). *Acta Cryst. D* **54**, 522–546.
- Wang, J., Dauter, M., Alkire, R., Joachimiak, A. & Dauter, Z. (2007). *Acta Cryst. D* **63**, 1254–1268.
- Wildegger, G. & Kiefhaber, T. (1997). *J. Mol. Biol.* **270**, 294–304.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst. D* **67**, 235–242.
- Wooliver, C., Coriu, D., Murphy, D., Kestler, D., Wang, S., Weiss, D. T. & Solomon, A. (2007). *XIth International Symposium on Amyloidosis*, edited by M. Skinner, J. L. Berk, L. H. Connors & D. C. Seldin, pp. 208–218. Boca Raton: CRC Press.
- Yazaki, M., Farrell, S. A. & Benson, M. D. (2003). *Kidney Int.* **63**, 1652–1657.
- Yee, A. W., Aldeghi, M., Blakeley, M. P., Ostermann, A., Mas, P. J., Moulin, M., de Sanctis, D., Bowler, M. W., Mueller-Dieckmann, C., Mitchell, E. P., Haertlein, M., de Groot, B. L., Boeri Erba, E. & Forsyth, V. T. (2019). *Nat. Commun.* **10**, 925.
- Yee, A. W., Moulin, M., Breteau, N., Haertlein, M., Mitchell, E. P., Cooper, J. B., Boeri Erba, E. & Forsyth, V. T. (2016). *Angew. Chem. Int. Ed.* **55**, 9292–9296.