Article

The Right-Handed Parallel β-Helix Topology of *Erwinia chrysanthemi* Pectin Methylesterase Is Intimately Associated with Both Sequential Folding and Resistance to High Pressure †

Jessica Guillerm [1], Jean-Marie Frère [1], Filip Meersman [2] and André Matagne [1,*]

[1] Centre for Protein Engineering, Laboratory of Enzymology and Protein Folding, InBioS, University of Liège, Building B6C, Quartier Agora, Allée du 6 Août, 13, 4000 Liège, Sart-Tilman, Belgium; jessica_guillerm@yahoo.fr (J.G.); jmfrere@uliege.be (J.-M.F.)

[2] Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium; filip.meersman@uantwerpen.be

[*] Correspondence: amatagne@ulg.ac.be; Tel.: +32-(0)-43663419; Fax: +32-(0)-43663364

[†] This manuscript is dedicated to the memory of Christopher Martin Dobson (1949–2019), a leading British scientist in the field of protein folding and its link with misfolding diseases. He was an inspiring mentor to FM and AM during their postdoctoral years with him and had a major impact on their respective careers. The whole CMD community was dismayed by his death, and the world is a smaller and much sadder place without Chris. He will be remembered as an outstanding scientist but also as a great human being. He had genuine interest in how former members of his group were doing both professionally and also personally. We miss him a lot.

Abstract: The complex topologies of large multi-domain globular proteins make the study of their folding and assembly particularly demanding. It is often characterized by complex kinetics and undesired side reactions, such as aggregation. The structural simplicity of tandem-repeat proteins, which are characterized by the repetition of a basic structural motif and are stabilized exclusively by sequentially localized contacts, has provided opportunities for dissecting their folding landscapes. In this study, we focus on the *Erwinia chrysanthemi* pectin methylesterase (342 residues), an all-β pectinolytic enzyme with a right-handed parallel β-helix structure. Chemicals and pressure were chosen as denaturants and a variety of optical techniques were used in conjunction with stopped-flow equipment to investigate the folding mechanism of the enzyme at 25 °C. Under equilibrium conditions, both chemical- and pressure-induced unfolding show two-state transitions, with average conformational stability (ΔG° = 35 ± 5 kJ·mol^{-1}) but exceptionally high resistance to pressure (P_m = 800 ± 7 MPa). Stopped-flow kinetic experiments revealed a very rapid (τ < 1 ms) hydrophobic collapse accompanied by the formation of an extended secondary structure but did not reveal stable tertiary contacts. This is followed by three distinct cooperative phases and the significant population of two intermediate species. The kinetics followed by intrinsic fluorescence shows a lag phase, strongly indicating that these intermediates are productive species on a sequential folding pathway, for which we propose a plausible model. These combined data demonstrate that even a large repeat protein can fold in a highly cooperative manner.

Keywords: protein folding; parallel β-helix; repeat proteins; circular dichroism; high pressure; kinetic intermediate; sequential pathway; contact order; dry molten globule

1. Introduction

Proteins are synthesized on ribosomes, in the form of long amino acid chains, with the sequence coding for their three-dimensional structure [1,2]. Unraveling the details of the mechanisms by which a disordered polypeptide chain folds so rapidly to a specific and functional structure, not only in the test tube but also, and most remarkably, in the crowded environment of the cell [3,4], remains a fundamental challenge in modern structural biology. The refolding of proteins in vitro from inclusion bodies [5] for biotechnological and

biomedical applications remains a major bottleneck in commercial and academic laboratories, and, moreover, the misfolding and aggregation of several proteins into toxic species is a hallmark of some of the most devastating human diseases [6]. In either situation, understanding what goes wrong in folding can be just as important as understanding how the right fold is achieved [7].

Most of the information available to date has been gathered using small soluble globular proteins (i.e., generally less than circa 100–150 residues), which often fold fast via two-state kinetic mechanisms [8–14], without a significant population of partially folded species, although complexities in some reactions have been observed [14–21]. Thus, despite the wide variety of native structures examined, common principles have emerged [22–30] and a unifying mechanism for protein folding has been proposed [31]. An important finding is that the folding rate of many proteins is inversely correlated to the average sequence distance between residues that form native contacts (known as "contact order") in the native state [32,33]. Thus, the folding of large, multi-domain proteins, which are characterized by intricate topologies and numerous interactions between residues distant in the sequence (i.e., long-range interactions), is relatively slow in comparison with that of small, single-domain proteins. Although the high cooperativity of protein globular structures arises in part from long-range interactions, large proteins often fold and unfold through partially folded intermediate species. These are prone to aggregation and render the study of these proteins very difficult. However, proteins made up of more than 150 residues constitute the major fraction of all proteomes [4] and thus, there is a need for a better description of the stability and folding of medium-to-large-size globular proteins [34], and also of their self-assembly into macromolecular complexes [35].

As an alternative to globular proteins, repeat proteins [36] offer an attractive model to investigate the folding and stability of large proteins [37–46]. They show a distinctive modular nature, characterized by the succession of homologous structural motifs (termed repeats or coils), which stack up to form generally elongated, non-globular structures. In contrast with typical globular proteins, repeat proteins display the architectural simplicity of their repeats and are dominated by short-range interactions. In particular, all-β repeat proteins have not been characterized in as much detail as all-α and mixed α/β repeat proteins, and it will be interesting to compare them with globular proteins and see how they differ in their folding properties [45].

Erwinia chrysanthemi is responsible for soft-rot diseases in a wide range of plant species [47]. Pectin methylesterase (PemA) (EC 3.1.1.11) from *E. chrysanthemi* 3937 is a large monomeric enzyme of 342 residues (M_r 36953), which catalyzes the essential first step in the bacterial invasion, namely the deesterification of the methylated α-(1-4)-linked D-galacturonosyl residue component in the pectin molecule of the plant cell wall [48,49]. The X-ray structure of PemA [49,50] reveals a right-handed parallel β-helix fold (Figure 1), with a deep cleft on the surface of the enzyme, where the two conserved catalytic aspartate residues, at positions 178 and 199, are found. The polypeptide backbone folds into three parallel β-sheets, with the strands of successive repeats stacking on top of each other, which form the β-helix core of the protein, and several loops of variable lengths, resulting in a large elongated right-handed coil.

The right-handed parallel β-helix architecture is common in both secreted and membrane-bound microbial proteins that mediate diverse interactions with the extracellular medium. In particular, it has been discovered in bacterial, fungal, and viral adhesins, and also in various enzymes involved in the degradation and modification of carbohydrates, some as a virulence factor [36,50–53]. Furthermore, the naturally occurring right-handed parallel β-helix fold was proposed [54–56] as a plausible model to describe primordial amyloid fibril structure (for a comprehensive review on amyloid formation and structure, see [6]) but this was ruled out on the basis of X-ray fiber diffraction data [57].
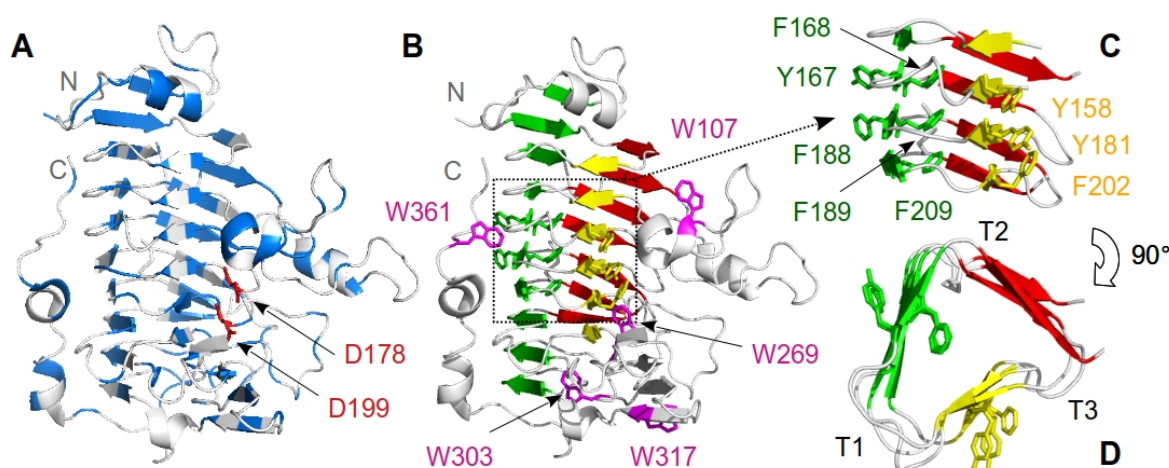
Figure 1. Schematic ribbon representation of the structure of *Erwinia chrysanthemi* 3937 pectin methylesterase (2NSP, [49]). The secondary structure elements are shown with (**A**) the hydrophobic residues in blue and the catalytic aspartate residues in red, (**B**) the three β-sheets in yellow (PB1), green (PB2), and red (PB3), and the tryptophan residues in purple. A closer view on the aromatic stacks in the central part of the β-helix is shown from (**C**) side and (**D**) top views. N and C indicate the N- and *C*-terminal ends of the polypeptide chain, respectively, and T1, T2, and T3 correspond to the turns following the three β-sheets. The figure was drawn using the open-source molecular graphics system PyMOL.

According to the common nomenclature [58], the three β-strands that make each repeat of a parallel β-helix are called PB1, PB2, and PB3 (highlighted in Figure 1B), and the connecting regions (turns or loops) following the β-strands are termed T1, T2, and T3 (Figure 1D), respectively. As with other β-helical proteins (see e.g., the pectate lyase PelC [59], the P22 tailspike protein (TSP) [60], and P69 pertactin [61]), the structure of PemA [49,50] reveals that the β-helix itself is compact and mostly hydrophobic inside, while long flexible peripheral T1 and T3 loops are found towards the *C*-terminal end of the protein. PemA displays an α-helix at the N-terminal end of the β-helix, and capping structures are found at both extremities, which protect the hydrophobic core from the solvent and prevent oligomerization between the monomers [62]. It has a deep active-site cavity along the parallel β-helix, formed by the T3-PB1-T1 region that contains the most conserved residues in pectin methylesterases [49,50]. Finally, PemA shows the distinctive extensive side-chain stacking observed with all parallel β-helix proteins [50,52,53,58,59]. Thus, aromatic stacks of phenylalanine and tyrosine (Figure 1C,D) are found between β-strands of different repeats in the central part of the parallel β-helix domain, which most likely contribute to the correct folding and the conformational stability of the protein [53,63–65]. In addition, an external asparagine stack of three residues is present [50] and a disulfide bridge between Cys192 and Cys212 might occur [50,53]. Interactions between repeats have been shown to play a critical role in the stability and cooperativity of repeat proteins [42].

TSP [60] is a large trimeric protein that arises from the assembly of identical 666-residue-long polypeptide chains. The major part (residues 143–540) of the protein shows a typical parallel β-helix fold [60], whereas both the *N*- and *C*-terminal domains display anti-parallel β-sheets. TSP was the first protein for which in vivo partially folded intermediates could be identified [66], and in vivo and in vitro folding could be directly compared [67–73]. The folding of PelC [74–76] and pertactin [77,78], two other proteins with a classical β-helix structure [59,61], has also been studied in some detail. Thus, with both proteins, equilibrium folding was reported to deviate from a two-state model and kinetic analysis indicated a multistep process with a significant population of on-pathway partially folded species.

In this work, the chemical- and pressure-induced unfolding transitions of PemA were studied using a combination of spectroscopic techniques. This analysis provides an estimate of the unfolding free energy of the protein and reveals a remarkable resistance to high pressure. Furthermore, kinetic data obtained with the help of a stopped-flow instrument indicate sequential folding, with significant population of productive partially

folded species. In particular, the fluorescence experiments provide support for the transient formation of at least two on-pathway kinetic species. This study brings further insights into the cooperativity of repeat-protein folding and the effect of native-state topology in protein folding as a whole. Moreover, it highlights structural aspects linked to the absence of long-range contacts and the simple topology of repeat proteins, which breaks the correlation between the rate of folding and the density of direct interactions between residues distant in sequence (i.e., contact order), as found for small globular proteins.

2. Materials and Methods

2.1. Chemicals

Ultrapure guanidinium chloride (GdmCl), pectin from citrus peel (76280), acrylamide, 8-anilino-1-naphtalene-sulfonic acid (ANS), and bromocresol purple were purchased from Sigma Chemical Co. All other chemicals were reagent grade.

2.2. Molecular Biology

Standard procedures for recombinant DNA technology were used as described by Sambrook and Russell [79]. The *pBCKSpemA* plasmid [49] was used as a template to amplify the *PemA* gene coding for *Erwinia chrysanthemi* 3937 pectin methylesterase, with its own 24 amino acid residue signal sequence. The forward (pemANdeI; 5′-ATTCATATGTTAAAAA-CGATCTCTGGAACC-3′) and reverse (pemAXhoI; 5′-ATTCTCGAGCGTCAGGGTAATG-TCGGCG-3′) primers contained sites for NdeI (CATATG) and XhoI (CTCGAG) restriction enzymes, respectively. A polymerase chain reaction (PCR) using Taq polymerase was performed for amplification, as follows: initial denaturation of the DNA at 94 °C for 4 min with subsequent amplification for 25 cycles of incubation at 94 °C for 1 min, 55 °C for 30 s, and 72 °C for 3 min; for the last cycle, incubation at 72 °C was extended to 7 min. The PCR product was purified and ligated to the pJET1.2 cloning vector (CloneJET PCR Cloning Kit, Fermentas). The presence of the insert within the pJET1.2 plasmid was checked by colony-PCR on randomly selected transformants, and plasmids from colonies carrying an insertion were amplified and extracted. Their sequences were determined by the Sanger method at the GIGA GenoTranscriptomics technology platform (Liège, Belgium).

To achieve a high level of expression, the *PemA* gene was cloned into the expression vector pET20b(+), containing the ampicillin-resistance gene, using the NdeI and XhoI restriction endonucleases. Finally, *Escherichia coli* DH5-α competent cells (Invitrogen, Paisley, UK) were transformed by the pET20bpemA vector and spread out on LB (lysogeny broth) agar plate with 100 µg·mL^{-1} of ampicillin (Sigma). Transformed *E. coli* DH5-α cells were grown on lysogeny broth (LB) medium containing 100 µg·mL^{-1} ampicillin. The recombinant plasmid was extracted and purified from DH5-α cells before use for protein expression.

2.3. Enzyme Expression and Purification

The recombinant PemA gene, cloned in a pET20b(+) plasmid, was transformed into *E. coli* BL21(DE3) cells (Novagen Inc., Madison, WI, USA). Cells were transformed with pET20bpemA plasmid and spread out on a LB agar plate with 100 µg·mL^{-1} of ampicillin and 0.2% of glucose (*w/v*) for 48 h at 28 °C. Transformed *E. coli* BL21(DE3) cells were grown overnight at 28 °C in 100 mL of yeast extract tryptone (YT) medium containing 100 µg·mL^{-1} of ampicillin and 0.2% of glucose (*w/v*). Following inoculation with 8 mL of an approximately 16 h preculture, the enzyme was expressed in 2 L bottles containing 500 mL of YT medium and 100 µg·mL^{-1} of ampicillin. The 6 L culture was grown at 28 °C, and cell development was followed by absorbance measurements at 600 nm. At an absorbance of circa 0.6, the temperature was fixed at 18 °C, and cell growth was allowed to proceed for an additional 12–14 h. The periplasmic proteins were then extracted by osmotic shock as described [80]. The supernatant containing the periplasmic proteins was filtered on a 0.22 µm membrane and dialyzed against 13 L of 20 mM MES, pH 6 (buffer A). The sample was then loaded onto a 25 mL SP Sepharose Fast Flow column (GE Healthcare),

equilibrated in buffer A. PemA was eluted with a linear NaCl gradient (0–190 mM) in buffer A, and fractions containing the protein (identified by SDS-PAGE) were pooled together to be dialyzed in 20 mM HEPES, pH 7 (buffer B). A second purification step was performed using a 1 mL MonoS HR 5/5 column (GE Healthcare), equilibrated in buffer B. PemA was eluted with a linear NaCl gradient (0–63 mM) in buffer B, and the purified protein was finally dialyzed against a 50 mM phosphate sodium buffer, pH 7. The quality of the recombinantly produced PemA enzyme was assessed in line with the best practice recommendations established by the ARBRE-MOBIEU and P4EU networks (see https://arbre-mobieu.eu/guidelines-on-protein-quality-control/) (accessed on 11 May 2021). Thus, the purity of the sample was checked by SDS-PAGE and was found to be above 98%. The size homogeneity of the protein was assayed by size exclusion chromatography (24 mL Superdex75 HR10/30; [PemA] = 50 μM, i.e., 1.85 mg·mL^{-1}), and no significant oligomerization ($\leq$ 1% dimer and no larger oligomeric species) was observed. Identity was confirmed by both intact protein mass determination (M_r 36954 $\pm$ 4; ESI-Q-TOF Ultima, Micromass, Manchester, UK) and N-terminal amino acid sequencing (ATTYN; Applied Biosystems 476A protein sequencer). The final PemA preparation (circa 30 mg) was stored at -20 °C. The PemA concentration was determined through absorbance measurements performed at 280 nm in a 1 cm pathlength cell, on the basis of the calculated [81] extinction coefficient value (49975 M^{-1}·cm^{-1}).

2.4. PemA Activity Assay

Enzymatic activity was measured at 25 °C by a colorimetric assay using bromocresol purple. This reagent is used as a pH indicator to follow the time-course of proton release during the deesterification of pectin from citrus fruit. The assay was performed in a 0.5 mL solution of 150 mM NaCl, pH 6, containing 0.75% pectin (*w/v*) and 37 μM bromocresol purple. A standard curve was established by the titration of the assay solution with 100 mM HCl, in order to correlate the change of absorbance measured at 590 nm with the amount of protons released. Enzymatic activity measurements were performed by monitoring the change in absorbance at 590 nm, following the addition of 4 μL of enzyme ($0.3 - 1.5$ μM, final concentration).

2.5. Circular Dichroism Measurements

Far-UV circular dichroism (CD) spectra were recorded with a Jasco J-810 spectropolarimeter at 20 °C, using a 1 mm pathlength quartz Suprasil cell (Hellma), with protein concentrations of circa 0.09 mg/mL. Five scans (10 nm/min, 1 nm bandwidth, 0.2 nm data pitch, and 4 s DIT) were averaged, base lines were subtracted, and no smoothing was applied. Data obtained with an applied high-tension voltage above 600 V were not considered. Data are presented as the molar residue ellipticity ($[\ominus]_{MRW}$) calculated using the molar concentration of protein and the number of residues. Secondary structure analysis was performed using the CDSSTR [82,83] algorithm provided in the DichroWeb analysis server [84,85].

2.6. Chemical-Induced Unfolding Transitions

Equilibrium folding was studied at 25 °C in a 50 mM sodium phosphate buffer, pH 7. With GdmCl, native or fully unfolded samples incubated at various final denaturant concentrations were allowed to unfold or refold, respectively, and to equilibrate for at least 9 days (under these conditions, equilibrium is reached throughout the transition; see the results). In contrast, with urea, only fully unfolded samples, incubated for 3 h in 8 M denaturant, were left to refold and equilibrate for 20 h. Under these conditions, equilibrium was also reached throughout the transition, and the risk of protein carbamylation was minimized. Unfolding and refolding transitions were obtained by monitoring the changes in intrinsic fluorescence emission (λ_{ex} = 280 nm; λ_{em} = 342 nm) and CD at 218 nm, using a Varian Cary Eclipse spectrofluorimeter and a Jasco J-810 spectropolarimeter, respectively, both equipped with a thermostatically controlled cell holder. With all samples, the data were cor-

rected for the contribution of the solution (buffer + denaturant). Denaturant concentrations were determined from refractive index measurements [86], using an Atago R5000 hand refractometer. A protein concentration of 0.12 mg·mL^{-1} (3.3 μM) was used throughout.

ANS-bound fluorescence measurements were performed using the same samples, with excitation at 350 nm, and emission spectra recorded from 420 to 600 nm. The fluorescence spectra were corrected for the background fluorescence of ANS. The ANS concentration (determined from the molar extinction coefficient of 4950 M^{-1}·cm^{-1} at 350 nm; Merck Index, Merck & Co., Whitehouse Station, NJ, USA) was 660 μM and hence, [ANS]/[PemA] = 200.

2.7. Quenching of Intrinsic Fluorescence by Acrylamide

All measurements were performed at 25 °C in a 50 mM sodium phosphate buffer, pH 7, using a protein concentration of 0.10 mg·mL^{-1} (2.7 μM). Native (in the presence of 0.4 M GdmCl) and unfolded (in the presence of 4 M GdmCl) PemA was diluted with increasing concentrations of acrylamide (ranging from 0 to 250 mM), and fluorescence emission spectra were recorded from 310 to 440 nm, following excitation at 295 nm. The solvent accessibility of tryptophan residues was estimated according to the Stern–Volmer equation [87]:

$$\frac{F_0}{F} = 1 + K_{SV}[Q] \tag{1}$$

where F_0 and F are the fluorescence intensity of the protein in the absence and presence of acrylamide, respectively; K_{sv} is the Stern–Volmer quenching constant, and $[Q]$ is the molar concentration of acrylamide.

2.8. Pressure-Induced Unfolding

PemA was dialyzed three times against 50 mM TrisDCl, pD 7.6, and concentrated to 18 mg·mL^{-1} (0.49 mM). The sample was stored overnight at 25 °C to ensure that all solvent-accessible protons were exchanged for deuterons. Just before measurement, the sample was briefly centrifuged to remove any possible insoluble aggregated form of the protein, although no evidence was found for aggregation of the enzyme. Pressure scans were performed in a diamond anvil cell (Diacell Products, Leicester, UK) at 25 °C, and IR spectra were recorded using a Bruker IFS66 FTIR spectrometer (Karlsruhe, Germany) as described [88]. A smoothing of 17 points and baseline correction were applied to the spectra before data analysis. Secondary structure content of the protein was determined as described [89]. Fourier self-deconvolution was used before curve-fitting analysis.

2.9. Kinetics of Unfolding and Refolding

All experiments were performed at 25 °C in a 50 mM sodium phosphate buffer, pH 7, using protein concentrations of 0.10 mg·mL^{-1} (2.7 μM) and 0.14 mg·mL^{-1} (3.7 μM) for fluorescence and CD measurements, respectively. Refolding reactions were initiated by a 10-fold dilution of PemA unfolded for circa 18 h in 4 M GdmCl (under these conditions PemA unfolds with τ ≈ 10 s), with the refolding buffer containing various GdmCl concentrations, yielding final concentrations in the 0.2 to 0.9 M range. Conversely, unfolding reactions were initiated by a 10-fold dilution of native PemA with the same buffer containing various amounts of GdmCl to yield final concentrations ranging from 2.5 to 4.5 M.

CD kinetics were monitored following the change in ellipticity at 218 nm in a 0.1 cm pathlength cell, whereas intrinsic fluorescence data were collected using excitation and emission wavelengths of 280 nm and 342 nm, respectively, in a 1 cm pathlength cell.

2.10. Stopped-Flow Experiments

All fast-mixing experiments were performed using a Bio-Logic (Claix, France) SFM-400 stopped-flow device, coupled with a MOS-450/AF-CD spectrophotometer, as described [90], with experimental dead times of ~3 ms and ~7 ms for fluorescence and CD measurements, respectively. For fluorescence quenching experiments, acrylamide was

added into the refolding buffer to yield a final concentration of 81 mM and total emission fluorescence above 320 nm was measured following excitation at 295 nm. Binding of the fluorescent dye ANS was performed as described [90]. ANS was included in the refolding buffer, leading to final ANS concentration of 270 µM (thus, [ANS]/[PemA] = 100). In all experiments, 6000 data points were sampled over the time course of one experiment.

2.11. Thermodynamic Analysis

The thermodynamic parameters for chemical- and pressure induced unfolding were obtained on the basis of a two-state (N $\rightleftharpoons$ U), according to [88,91], respectively.

2.12. Kinetic Analysis

Each kinetic trace resulted from the accumulation of approximately five and ten experiments for fluorescence and CD measurements, respectively. The resulting multiple data sets were fitted separately. These traces were analyzed according to a sum of exponential terms:

$$y_t = y_\infty + \sum A_i e^{-k_i t} \tag{2}$$

where y_t is an observable parameter (i.e., CD or fluorescence); A_i and k_i are the amplitude and the rate constant of the ith phase, respectively; t is the time, and y_∞ is the equilibrium value of the observed property. The rate constants were obtained by averaging the data sets, and errors were calculated as standard deviations throughout.

The dependence of the unfolding and folding rate constants on denaturant concentration was analyzed according to the following linear relationship [8,18,92–94]:

$$\ln(k_{obs}) = \ln\left(k_f^{H_2O} \, e^{\frac{-m_{kf}}{RT \cdot [denaturant]}} + k_u^{H_2O} e^{\frac{m_{ku}}{RT \cdot [denaturant]}} \right) \tag{3}$$

where k_{obs} is the rate of unfolding or refolding measured at various GdmCl concentrations; $k_f^{H_2O}$ and $k_u^{H_2O}$ are the values for folding and unfolding, respectively, in the absence of a denaturant; and m_{k_f}/RT and m_{k_u}/RT are proportionality constants which describe the denaturant dependence. R is the gas constant, and T is the absolute temperature.

The difference in free energy between the folded and unfolded conformations was also calculated using kinetic data:

$$\Delta G^0(H_2O)_{NU}^{kinetic} = -RT \cdot \ln\left(\frac{k_f^{H_2O}}{k_u^{H_2O}} \right) \tag{4}$$

The programs Sigmaplot 9.0 (SPSS Inc., Chicago, IL, USA), Bio-Kine 32 V4.45 (Bio-Logic), and QtiPlot 0.9.97.10 (ProIndep Serv S.R.L., Craiova, Romania) were used for nonlinear least-squares analysis of the data.

2.13. Analysis of the Folding Kinetics

For a sequential four-state reaction (scheme (5)), there are three macroscopic rate constants (λ_1, λ_2, and λ_3), which characterize the interconversion between the first and second, second and third, and third and fourth species, respectively:

$$U \underset{}{\overset{\lambda_1}{\rightleftharpoons}} I_1 \underset{}{\overset{\lambda_2}{\rightleftharpoons}} I_2 \underset{}{\overset{\lambda_3}{\rightleftharpoons}} N \tag{5}$$

An analytical solution [95,96] can be found for the time-course of the different species, which is given by:

$$\frac{d[U]}{dt} = A_0 e^{-\lambda_1 t} \tag{6}$$

$$\frac{d[\mathrm{I}_1]}{dt} = A_0\lambda_1 \frac{e^{-\lambda_2 t} - e^{-\lambda_1 t}}{\lambda_1 - \lambda_2} \tag{7}$$

$$\frac{d[\mathrm{I}_2]}{dt} = -A_0\lambda_1\lambda_2 \frac{(\lambda_2 - \lambda_3)e^{-\lambda_1 t} + (\lambda_3 - \lambda_1)e^{-\lambda_2 t} + (\lambda_1 - \lambda_2)e^{-\lambda_3 t}}{(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_1)} \tag{8}$$

$$\frac{d[\mathrm{N}]}{dt} = A_0\left(1 + \frac{\lambda_2\lambda_3(\lambda_2 - \lambda_3)e^{-\lambda_1 t} + \lambda_1\lambda_3(\lambda_3 - \lambda_1)e^{-\lambda_2 t} + \lambda_1\lambda_2(\lambda_1 - \lambda_2)e^{-\lambda_3 t}}{(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_1)}\right) \tag{9}$$

The length of the lag phase (τ_{lag}, termed lag time), of any given species (e.g., I_2), is given by the abscissa at the inflection point of the corresponding kinetic curve [95]. This parameter can be calculated by solving the following equation:

$$\frac{d^2[\mathrm{I}_2]}{dt^2} = 0 \tag{10}$$

The lag time also corresponds to the time necessary for the previous species in the reaction (e.g., I_1) to reaches its maximum concentration. This is given by:

$$\frac{d[\mathrm{I}_1]}{dt} = 0 \tag{11}$$

Similarly, the lag time for N (i.e., the time needed for I_2 to reach its maximum concentration) is given by:

$$\frac{d[\mathrm{I}_2]}{dt} = 0 \tag{12}$$

3. Results

3.1. Chemical-Induced Unfolding

The changes in secondary and tertiary structural content were followed as a function of the GdmCl concentration by CD at 218 nm and by tryptophan fluorescence emission at 342 nm. The five tryptophan residues in PemA, all located in peripheral loops of the β-helix (Figure 1B), give rise to a single broad fluorescence emission band with a maximum at 342 nm (data not shown), which is consistent with the partial burial of the indole groups into the native structure [49,50] (see also acrylamide quenching experiments below). Upon the addition of 2.5 M GdmCl, the intensity decreased by about ~75%, accompanied by a red-shift of the maximum to 355 nm, indicating that stable tertiary contacts are lost, and the tryptophan side chains are exposed to the solvent. The far-UV CD spectrum (Figure 2A) of the native enzyme is typical for a protein with a high fraction of β-sheet structure, as indicated by a broad minimum in molar ellipticity centered at circa 218 nm [97]. Deconvolution using CDSSTR [82,83] revealed a β-strand content of 37 ± 1 %, in good agreement with the 38 ± 2 % found in the X-ray structure of PemA [49,50]. In 2.5 M GdmCl, the band centered at 218 nm is lost and the CD spectrum indicates the formation of an unordered structure (Figure 2A).

The full recovery of the optical and, most significantly, catalytic properties of PemA after a complete unfolding/refolding cycle demonstrated unambiguously that unfolding by GdmCl is fully reversible. For enzymatic measurements, we have developed a method based on a colorimetric assay, using a pH indicator (i.e., bromocresol purple, pK_a ~6.4) with a pH transition zone of 6.8 to 5.2 (purple to yellow). The specific activity of the enzyme was found to be $(8 \pm 1) \times 10^3$ µmoles·min^{-1}·mg^{-1}. This assay can normally be used in this pH range for any reaction that leads to the release of protons.
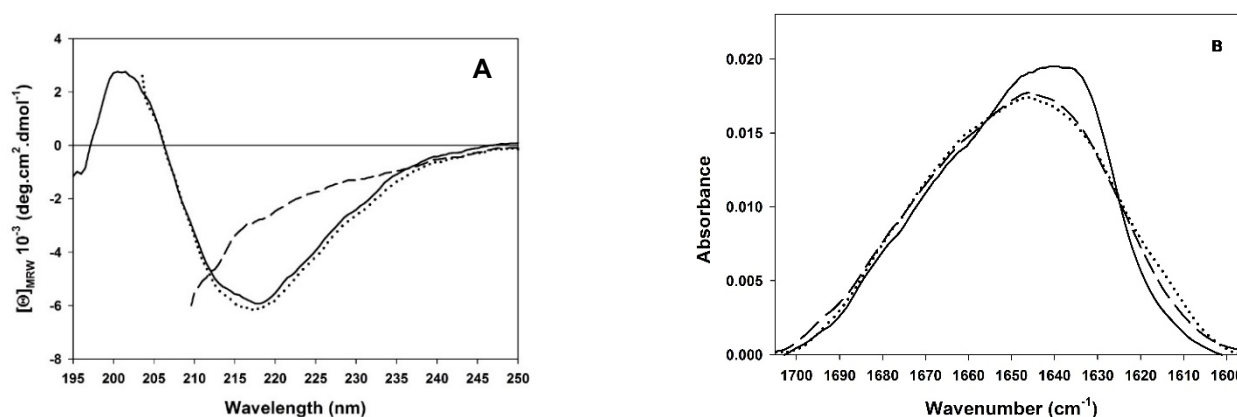
Figure 2. (**A**) Far-UV CD spectra of native (continuous line), unfolded in 2.5 M GdmCl (broken line) and refolded from 2.5 to 0.25 M GdmCl (dotted line) PemA. (**B**) FTIR spectra of PemA under native (33 MPa; continuous line), unfolding (1100 MPa; broken line), and refolding (i.e., 41 MPa after decompression; dotted line) conditions, with data normalized using the area under the amide I′ band equal to 1.

Both spectroscopic techniques (Figure 3) indicated that GdmCl induces a single co-operative transition between the native and unfolded states. Data revealed, however, that, following 20 h of GdmCl-induced unfolding and refolding, transition curves did not coincide (hysteresis) for GdmCl concentrations in the range from 0.75 to 2 M (Figure 3, insets), indicating that the samples had not yet reached equilibrium. Remarkably, unfolding proved to be extremely slow under these conditions, as it took about 9 days to reach equilibrium, while refolding was complete within 20 h.

Following circa 9 days of incubation in GdmCl (0 to 2.6 M), PemA was found to unfold with full thermodynamic reversibility, and the coincidence of the transition curves (Figure 3C) obtained by intrinsic fluorescence emission and far-UV CD measurements indicated that secondary and tertiary structures were destabilized concomitantly. The dye ANS binds preferentially to partially folded protein molecules with exposed hydrophobic patches [13,90,98–103]. We could not detect ANS binding in the transition region, confirming that PemA unfolds cooperatively in a single two-state transition ($N \rightleftharpoons U$), where only the native (N) and the unfolded (U) states are significantly populated, not any partially folded intermediates. Therefore, we used a simple two-state model to calculate the thermodynamic parameters for the PemA unfolding transition shown in Figure 3: $\Delta G°(H_2O)_{NU} = 31 \pm 2$ kJ·mol^{-1}, $m_{NU} = -30 \pm 2$ kJ·mol^{-1}·M^{-1} and $C_m = 1.0 \pm 0.1$ M. Note that a $\Delta G°(H_2O)_{NU}$ value identical within the error limits was obtained using urea as the denaturant (pH 7, 25°C; $\Delta G°(H_2O)_{NU} = 39 \pm 7$ kJ·mol^{-1}, $m_{NU} = -14 \pm 2$ kJ·mol^{-1}·M^{-1} and $C_m = 2.9 \pm 0.1$ M; data not shown).
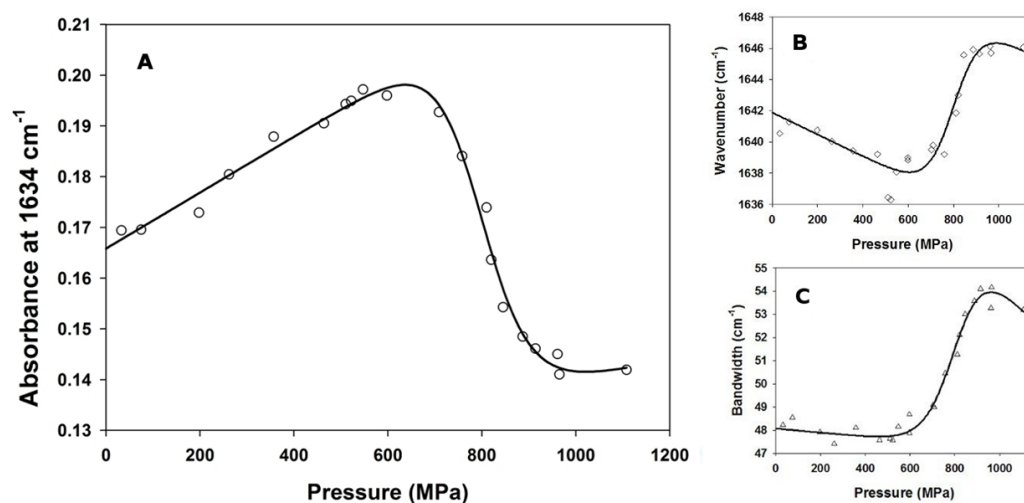
3.2. Pressure-Induced Unfolding

FTIR spectroscopy was used in combination with pressure-induced unfolding for measuring PemA stability. Very similar to other β-helix proteins [104], the FTIR spectrum of native PemA (Figure 2B) displays a maximum of the amide I′ band at 1634 cm^{-1}, typical of β-sheet structure [105]. At 1100 MPa, the broadening of the amide I′ band and the displacement of its maximum to around 1647 cm^{-1} revealed the formation of an unordered structure [105]. Three parameters were considered in the analysis (Figure 4), which provide details about the structural modifications of the protein upon pressure-induced unfolding [88], i.e., the absorbance at a fixed wavenumber, the wavenumber corresponding to the absorbance maximum of the amide band, and the width of the band. Thus, as the pressure was raised up to 600 MPa (Figure 4) no significant change in the IR spectrum was observed. In this pressure range, only a limited decrease in the wavenumber of the band maximum occurred (Figure 4B), which probably results from the effect of compression of the hydrogen bonds [106] and also from H/D exchange of the internal hydrogens [107]. Above 600 MPa, a cooperative displacement of the band maximum

towards higher wavenumber values (Figure 4B) took place, together with a broadening of the band (Figure 4C) and significant changes in intensity. In particular, the decrease in band intensity observed at 1634 cm^{-1} (Figure 4A) indicates the disappearance of the native sheet structure and the shift of the band maximum from 1634 to 1647 cm^{-1} is consistent with an increased number of unordered structures.
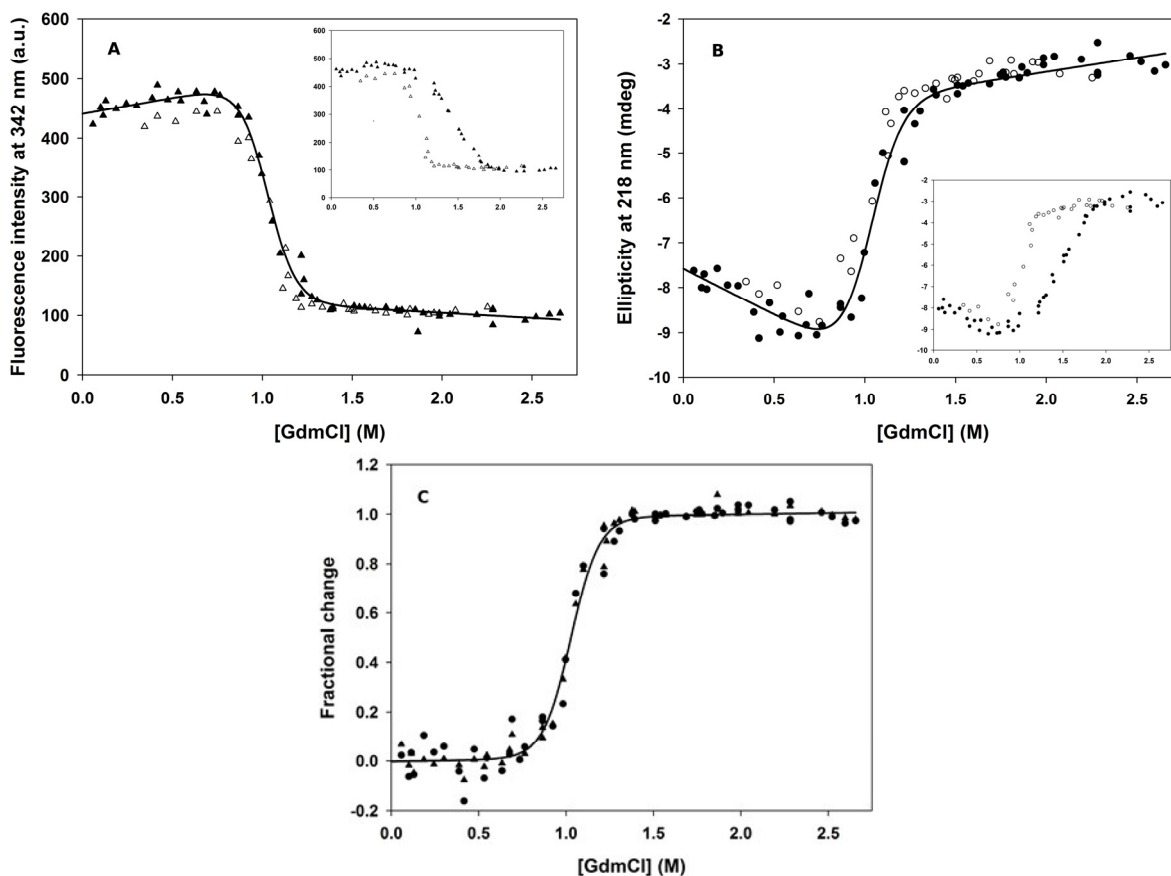
Figure 3. GdmCl-induced equilibrium unfolding transitions of PemA at pH 7 and 25 °C, monitored by (**A**) the change in fluorescence intensity at 342 nm and (**B**) the change in ellipticity at 218 nm. Closed symbols represent the addition of a denaturant to a solution of native protein, whereas open symbols are for the dilution of unfolded protein (in 3 M GdmCl) to the indicated concentrations of GdmCl. The samples were incubated at the indicated GdmCl concentrations for 9 days before the measurements. Data were analyzed on the basis of a two-state model [91], and the lines were drawn using (**A**) $\Delta G°(H_2O)_{NU} = 32 \pm 3$ kJ·mol^{-1}, $m_{NU} = -31 \pm 3$ kJ·mol^{-1}·M^{-1} and $C_m = 1.0 \pm 0.1$ M, and (**B**) $\Delta G°(H_2O)_{NU} = 29 \pm 4$ kJ·mol^{-1}, $m_{NU} = -28 \pm 4$ kJ·mol^{-1}·M^{-1} and $C_m = 1.0 \pm 0.1$ M. (**C**) Fractional change in signal [91] as a function of GdmCl concentration. Triangles and circles represent fluorescence and far-UV CD data, respectively. The line was drawn using the average value for the parameters obtained in fluorescence and CD experiments (see text). Insets in (**A**,**B**) represent data obtained after circa 20 h of incubation (see text).

Figure 4. Pressure-induced unfolding curves of PemA followed by FTIR spectroscopy at 25 °C, through monitoring (**A**) the absorbance at 1634 cm^{-1}, (**B**) the wavenumber corresponding to the band maximum and (**C**) the amide I′ bandwidth. The solid lines represent the best fits of Equation (3) in [88] to the data, yielding the following average apparent values for the parameters: $\Delta G°(H_2O)_{NU}$ = 33 ± 4 kJ·mol^{-1}, ΔV = −41 ± 5 mL·mol^{-1} and P_m = 800 ± 7 MPa.

Although the three parameters appeared to change cooperatively upon pressure increase, suggesting a simple two-state unfolding process as observed in chemical unfolding experiments, FTIR spectra (Figure 2B) indicated that pressure-induced unfolding of the enzyme is not reversible. Indeed, following the release of the pressure, the spectrum of PemA did not shift back to its original position and remained rather broad, indicating that the enzyme was left largely unordered. Nevertheless, the apparently cooperative transition between N and U was tentatively analyzed according to a two-state model (Figure 4), yielding the apparent values for the thermodynamic parameters given in the legend to Figure 4. Remarkably, the apparent unfolding free energy ($\Delta G°(H_2O)_{NU}$ = 33 ± 4 kJ·mol^{-1}) is identical within the error limits to that found using chemical denaturants (both GdmCl and urea), indicating that the pressure-induced unfolding transition between N and U can also be satisfactorily described according to a simple two-state equilibrium.

3.3. Fluorescence- and CD-Detected Folding Kinetics

The unfolding of the enzyme was monitored by fluorescence, in the presence of GdmCl concentrations in the range of 2.5 to 4.5 M. Following manual mixing, single exponential fluorescence decays were observed throughout and analyzed by using Equation (2) with i = 1. The refolding kinetics of PemA at 0.4 M GdmCl, monitored by intrinsic fluorescence and far-UV CD spectroscopy at pH 7.0, 25 °C, are shown in Figure 5. A sum of four and three exponential functions (Equation (2)), respectively, was fitted to the data, yielding the kinetic parameters in Table 1. Measurements of CD at 225 nm (Figure 5B) revealed that a substantial part (~60%) of the CD native signal is restored within the dead time of mixing (circa 7 ms). In the fluorescence measurements, however, a burst phase was not observed, indicating that the tryptophan emission did not change within the mixing dead time (circa 3 ms). Thus, these kinetic experiments suggest that a transient intermediate accumulates in the dead time of stopped-flow mixing, which shows a substantial amount of secondary structure, as indicated by the presence of ~60% of the CD signal at 225 nm but lacks stable tertiary contacts, as shown by the absence of changes in tryptophan fluorescence. After this burst phase, three kinetic phases (numbered 2 to 4) could be resolved with the two optical probes, which show identical time constants (see Table 1). The two slowest phases (i.e., 3 and 4) could also be measured reproducibly, by the two methods, following manual mixing (dead time ~10 s; time constants of 20 ± 5 s and 70 ± 20 s).
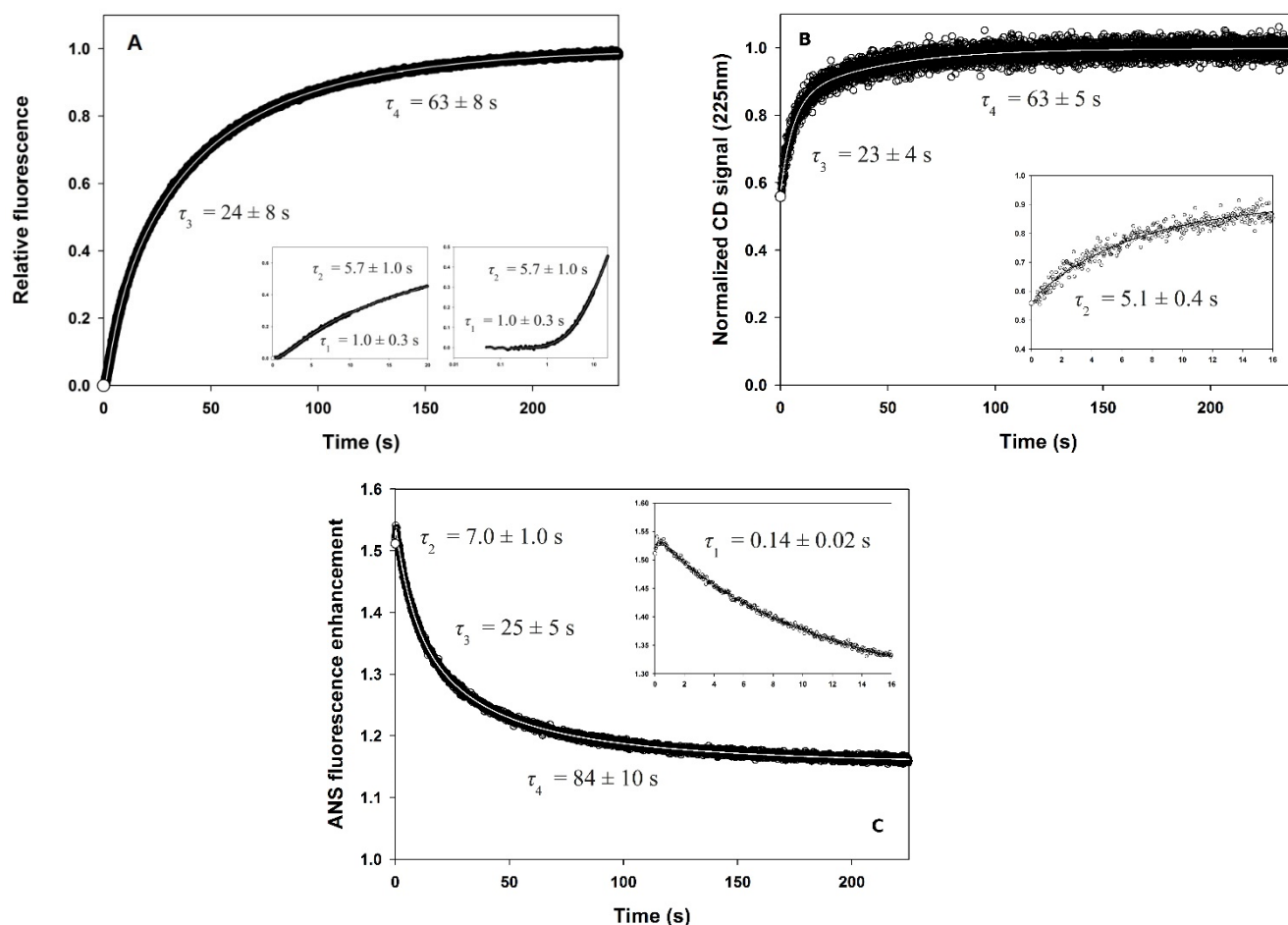
Figure 5. Refolding kinetics of PemA at pH 7, 25 °C, in 0.4 M GdmCl and 50 mM sodium phosphate, monitored by (**A**) total intrinsic fluorescence emission above 320 nm (after excitation at 295 nm), (**B**) CD at 225 nm, and (**C**) total ANS fluorescence emission above 450 nm. A sum of four and three exponential functions (Equation (2)) has been fitted to the fluorescence (both **A**,**C**) and CD data, respectively. Intrinsic fluorescence (**A**) and CD (**B**) data have been normalized to the total signal difference between the native (1) and unfolded (0) proteins under refolding conditions, whereas extrinsic fluorescence intensities in ANS binding kinetics (**C**) are expressed relative to the fluorescence of ANS in the presence of the unfolded protein in 4 M GdmCl. The resulting time constants are indicated. (○) Signal extrapolated to zero time from the kinetic data. Insets show the first 16 s of the reaction (note that intrinsic fluorescence data are displayed with both a linear and a logarithmic time scale).

In contrast to far-UV CD measurements, which detect three visible kinetic phases only, an additional exponential term was necessary to fit the stopped-flow fluorescence data satisfactorily. This term corresponds to the fastest detectable process (phase 1; $\tau_1 = 1.0 \pm 0.3$ s) and, remarkably, the sign of its amplitude (negative, see comment in the footnote to Table 1), which accounts for ~6% of the total signal change (Table 1), is opposite to that of the following phases (positive). This results in a lag phase in the refolding kinetics, which is clearly visible (Figure 5A, inset) due to the low and negative amplitude of the fluorescence change corresponding to this transition. This is consistent with the transient accumulation of an obligatory intermediate [108]. This species apparently lacks a specific fluorescence signature and hence cannot be distinguished from the unfolded state, but it results in a delay in the accumulation of the native or other highly fluorescent species [109,110].

Table 1. Kinetic parameters for the refolding of PemA after rapid mixing in 0.4 M GdmCl at pH 7, 25 °C.

| Experiment	Burst Phase	Phase 1	Phase 2	Phase 3	Phase 4
Int. fluo. [a]					
τ (s) [e]	<0.003 [g]	1.0 ± 0.3	5.7 ± 1.0	24 ± 8	63 ± 8
amplitude [f]	0	-0.06 ± 0.01	0.20 ± 0.01	0.31 ± 0.02	0.55 ± 0.04
CD225nm [b]					
τ (s) [e]	<0.007	nd [h]	5.1 ± 0.4	23 ± 4	63 ± 5
amplitude [f]	0.56 ± 0.09	nd	0.25 ± 0.01	0.12 ± 0.02	0.07 ± 0.02
ANS fluo. [c]					
τ (s) [e]	<0.003	0.14 ± 0.02	7.0 ± 1.0	25 ± 5	84 ± 10
amplitude [f]	0.52 ± 0.02	0.032 ± 0.003	-0.13 ± 0.02	-0.15 ± 0.01	-0.10 ± 0.02
Acryl. [d]					
τ (s) [e]	<0.003	nd	7.0 ± 1.5	20 ± 1	115 ± 20
amplitude [f]	-0.54 ± 0.05	nd	0.014 ± 0.006	0.10 ± 0.02	-0.018 ± 0.002

[a] Intrinsic fluorescence. [b] Circular dichroism at 225 nm. [c] ANS fluorescence. [d] Quenching of fluorescence by acrylamide. [e] Time constant τ = $1/k$, with k representing the macroscopic rate constants obtained using Equation (2). [f] Amplitude values are normalized as indicated in the text. The signs of the amplitudes derived from Equation (2) have been reversed, so that an increase of the measured parameter is denoted by a positive amplitude. [g] Lag phase. [h] nd, not detected. All values are the averages of results from three refolding experiments, and errors are calculated as standard deviations.

The folding reaction was monitored by tryptophan fluorescence at various GdmCl concentrations and a plot (known as a chevron plot [94]) of the natural logarithm of the apparent rate constants for both folding and unfolding against the denaturant concentration is shown in Figure 6. Refolding of the enzyme at [GdmCl] < 1 M showed the complex kinetics described above, with at least 4 phases (see Table 1), whereas unfolding in the presence of GdmCl concentrations in the range of 2.5 to 4.5 M showed single exponential kinetics. At concentrations of denaturant ranging from 0.6 to 4.5 M, the unfolding and slowest phase-related refolding branches of the plot were found to be close to linear. Thus, at concentrations above 0.6 M GdmCl, Equation (3) was used to fit the data, yielding values of the kinetic parameters given in the legend to Figure 6. Extrapolation of the unfolding rate constant (k_u) at 0 M GdmCl indicates that, in the absence of a denaturant, unfolding is remarkably slow ($k_u^{H_2O}$ = $1.2 \cdot 10^{-7}$ s^{-1} i.e., $t_{1/2} \approx$ 67 days). Analysis of the data highlights the strong influence of the denaturant concentration on the unfolding rate constant and explains why circa 9 days were required to reach apparent equilibrium in the transition region, starting from native PemA. Thus, it can be calculated that it takes about 15 and 2.5 days in the presence of 0.8 and 1.3 M GdmCl, respectively, for the unfolding reaction to be 95% complete. Note that the time constant value for unfolding at 2.5 M GdmCl was as high as circa 23 min (i.e., $t_{1/2} \approx$ 16 min), and hence, no data were collected at lower concentrations due to the experimental complications inherent to such long time scales.

3.4. ANS Binding

The time course of ANS fluorescence intensity during PemA folding (Figure 5C) was measured by the inclusion of the dye in the refolding buffer, and data reveal that ~95 % of the ANS fluorescence enhancement (i.e., binding) occurred in the dead time of the stopped-flow experiment and that the maximum was reached after circa 0.5 s. This was followed by a three-exponential decay of the intensity, reflecting structural changes associated with the exclusion of the dye from the surface of the refolding protein. These three phases show time constants (Table 1) very similar to those measured by the changes in tryptophan fluorescence and in far-UV CD, indicating that the presence of the dye did not significantly affect the kinetics. Furthermore, the increase in ANS fluorescence was found to be independent of the ANS-to-protein concentration ratio ([ANS]/[PemA]) in the range from 50 to 200 (data not shown).
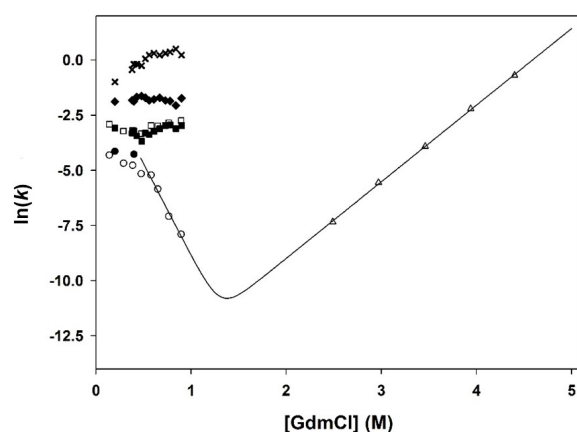
Figure 6. GdmCl concentration dependence of the rate constants for PemA folding at pH 7, 25 °C. Circles and squares represent the logarithm of the observed rate constants (in s^{-1}) for slow refolding (phases 3 and 4 in Table 1); diamonds and crosses are for fast refolding (phases 1 and 2), whereas triangles correspond to the unfolding rate constant. Data obtained after rapid mixing are represented by closed symbols, while those obtained following manual mixing are shown by open symbols. The solid line represents the fit of Equation (3) to the data obtained with GdmCl concentration in the range from 2.5 to 4.5 M (triangles), using $k_f^{H_2O} = 0.8 \pm 0.2$ s^{-1}, $m_{k_f} = 21 \pm 2$ kJ·mol^{-1}·M^{-1}, $k_u^{H_2O} = (1.2 \pm 0.1) \cdot 10^{-7}$ s^{-1}, and $m_{k_u} = 9 \pm 1$ kJ·mol^{-1}·M^{-1}.
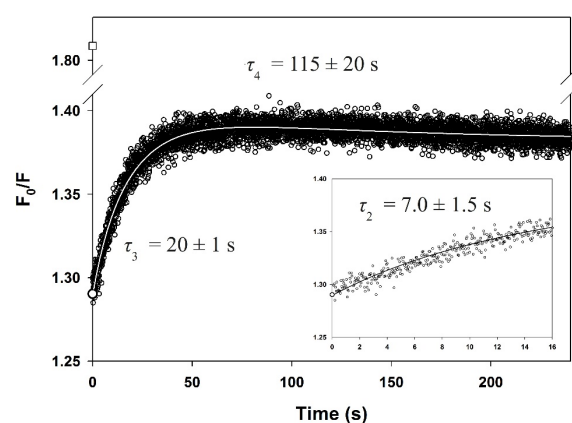
3.5. Quenching of Fluorescence by Acrylamide

The exclusion of tryptophan residues from the solvent as the protein folds can be followed by monitoring intrinsic fluorescence in the presence of acrylamide as a fluorescence quencher [103,111]. The X-ray crystal structure of PemA [49,50] shows that the five tryptophan residues of the enzyme are partially buried into the native structure, with about 37, 28, 2, 55, and 10% solvent accessible area for residues at positions 107, 269, 303, 317, and 361 (calculated using the Naccess program [112]), respectively. This is consistent with the significant, although moderate, red-shift observed as the protein unfolds (from 342 to 355 nm, data not shown). The accessibility of fluorophores to collisional quenching is analyzed in the form of Stern–Volmer plots, as shown in Figure 7. Linear regression analysis of the data using Equation (1) yielded an average value of the Stern–Volmer constant (K_{SV}) for the native state (5.0 ± 0.1 M^{-1}) that is half that of the unfolded state (10.0 ± 0.2 M^{-1}), in good agreement with significant exposure of tryptophan residues upon unfolding.

To investigate the change of solvent accessibility of the tryptophan residues during refolding, it was followed in the absence and in the presence of 81 mM acrylamide and the ratio (F_0/F) of the fluorescence intensity in the absence (F_0) and in the presence (F) of acrylamide was measured as a function of time, as shown in Figure 8. Data were analyzed according to a sum of three exponential functions (Equation (2)) and the resulting time constants were again consistent with those obtained using other spectroscopic probes (Table 1), indicating that all three methods monitor the same steps of folding. In the dead time of the rapid-mixing experiment, a large decrease in F_0/F was observed, which reflects the sequestering of tryptophan residues from the solvent during the first circa 3 ms. This was followed by a significant increase in the F_0/F ratio and thus, by an apparent re-exposure of the aromatic side chains, in two successive phases. Finally, the last phase shows a weak decrease in amplitude (see Table 1), probably due to a slight reduction in exposure of tryptophan indole groups as the protein adjusts to its native structure.
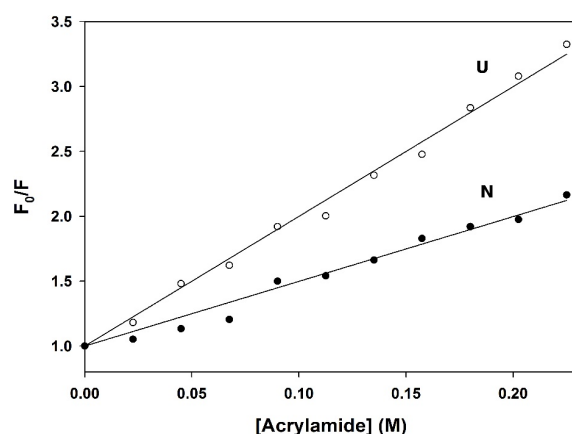
Figure 7. Stern–Volmer plots of F_0/F versus quencher concentration for the native state (N) of PemA in 0.4 M GdmCl (●) and the unfolded state (U) in 4 M GdmCl (○). F_0 and F are the fluorescence intensities in the absence and presence of acrylamide, respectively. Values of the average Stern–Volmer constant (K_{SV}), i.e., 5.0 ± 0.1 M^{-1} (N) and 10.0 ± 0.2 M^{-1} (U), are given by the slope of the straight line fitted to the data according to Equation (1).

Figure 8. The quenching of intrinsic fluorescence by 81 mM acrylamide during the refolding of PemA. Refolding was carried out in 0.4 M GdmCl and 50 mM sodium phosphate at pH 7 and 25 °C. The data are plotted as the ratio of the intrinsic fluorescence in absence of acrylamide (F_0) to that in the presence of the quencher (F). (□) is F_0/F for the unfolded protein measured as the ratio of the fluorescence intensities of the protein in 4 M GdmCl, with and without acrylamide; (○) represents the signal of F_0/F extrapolated to zero time, and the solid line represents the fit of a sum of three exponentials (Equation (2)) to the data. The resulting time constants are indicated, and the inset shows the first 16 s of the reaction.

3.6. Slow Phases in PemA Folding

A possible reason for the occurrence of slow phases during protein folding, with time constants in the 10–100 s range [91,113,114], is the *cis-trans* isomerization of Xaa-Pro peptide bonds [115]. PemA contains 11 prolines, out of which 10 are located in the peripheral loops of the β-helix [49,50]; they are all in the *trans* configuration, and it is thus possible that the isomerization of non-native prolyl peptide bonds in the unfolded state determines the folding kinetics of the enzyme. In particular, the two slowest phases (i.e., 3 and 4) with time constants around 25 and 80 s could be associated with this phenomenon. In order to test this hypothesis, a classical double mixing experiment [116] was performed. In this experiment, the protein was fully unfolded in 6 M GdmCl at 25 °C for 5 to 10 s only. Under these conditions, this was sufficient for complete conformational unfolding of the enzyme ($\tau \sim 7$ ms), but it was too short to allow significant isomerization of proline-containing peptide bonds, which thus remained in their native-like isomeric state. Under

such conditions, the intrinsic fluorescence of the native PemA was recovered in two well-resolved kinetic phases, with amplitudes and time constants (circa 25 and 80 s, data not shown) identical to those obtained in single mixing experiments. Thus, the dramatic reduction of the time left for the protein to unfold had no influence on folding, suggesting that the two slow phases are not associated with proline isomerization. In addition, no effect was observed when folding was followed in the presence (5:1 PemA to PPI ratio in the final refolding mixture; data not shown) of either trigger factor (modified W151F mutant; [117]) or SlyD [118], two peptidyl prolyl isomerases (PPI; [115]), which normally accelerate the isomerization of prolyl peptide bonds. In conclusion, although both the prolyl bonds might just not be accessible to the PPIs tested here, and *trans*-to-*cis* isomerization in the unfolded state might be too fast for the double mixing experiment, all kinetic phases observed in PemA folding (Figures 5 and 8) are likely to be associated with conformational refolding and not with the isomerization of incorrect *cis* forms of prolyl peptide bonds.

3.7. Analysis of the Folding Kinetics

The close similarity of the rate constants measured for the three phases (2 to 4, see Table 1) by the various probes suggests that they represent well-defined, cooperative folding transitions. The data are consistent with a sequential four-state model (scheme (5)) for folding, where λ_1 to λ_3 represent the macroscopic rate constants measured for phases 2 to 4. In this model, U represents the burst phase intermediate, considered as the ensemble of unfolded conformations under the refolding conditions and which already contains a significant fraction of secondary structure, whereas I_1 and I_2 corresponds to ensembles of transient, partially folded species, and N stands for the native state. Using the rate constants measured by tryptophan fluorescence, this model was used to simulate the time-courses of the various species (Figure 9). It shows a distinct lag phase in the formation of both I_2 and N, occurring in the first ~20 and ~100 s of the reaction, respectively. Lag times of 9 and 44 s are calculated using Equations (11) and (12), for I_2 and N, respectively. The experimental observation of a lag phase in the first few second of the fluorescence-detected reaction (Figure 5A) suggests that an intermediate is also significantly populated between the burst phase species (i.e., U in scheme (5)) and I_1. This species is not directly detected because its fluorescence intensity is not significantly different from that of both the unfolded and burst phase species, but it results in a delay in the formation of the three following kinetic species (i.e., I_1, I_2, and N), which are characterized by a gradual enhancement of the fluorescence intensity of the enzyme until the native signal is reached.
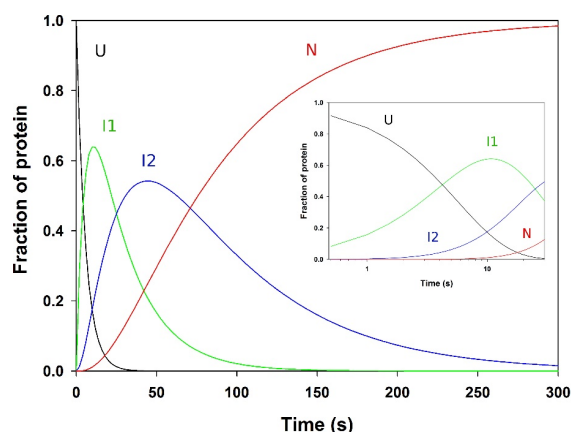
Figure 9. Simulation of the time-course of the fractional population of the various species for a four-state sequential model with two on-pathway intermediates (Scheme (5)). The simulation was performed using a program developed with the MATLAB computing platform (MathWorks Inc., Natick, MA, USA), with $\lambda_1 = 0.1754$ s^{-1} (i.e., $\tau_1 = 5.7$ s), $\lambda_2 = 0.0412$ s^{-1} (i.e., $\tau_2 = 24$ s) and $\lambda_3 = 0.0159$ s^{-1} (i.e., $\tau_3 = 63$ s). The inset shows the first 30 s of the simulated kinetics, on a logarithmic time scale.

4. Discussion

PemA displays a typical right-handed parallel β-helix fold (Figure 1) [49,50], with three parallel β-sheets (PB1, PB2 and PB3) forming the β-helix domain. This structural organization determines the spatial arrangement of the connecting peripheral loops (T1, T2, and T3), which can adopt their native conformation only after the β-sheets have formed. Furthermore, long flexible peripheral T1 and T3 loops, and an α-helix are located at the *C*- and *N*-terminal ends of the β-helix, respectively, which protect the hydrophobic interior from the solvent and prevent oligomerization. This topology causes special requirements for the folding of the polypeptide chain. Remarkably, the location of the five tryptophan residues of PemA (i.e., W107, W269, W303, W317, and W361) on peripheral loops (Figure 1B) might enable discriminating between the formation of the β-helix domain, monitored by far-UV CD, and the structural organization and docking of the lateral loops, monitored by intrinsic fluorescence measurements.

The conformational stability of PemA was probed by adding denaturants or by increasing pressure. GdmCl- and urea-induced unfolding was fully reversible as probed by intrinsic tryptophan fluorescence, far-UV CD and enzymatic activity measurements.

The denaturant-induced equilibrium unfolding of PemA (Figure 3) is well described by a two-state transition ($N \rightleftharpoons U$) between the native (N) and unfolded (U) states, without partially folded species. Remarkably, circa 9 days at 25 °C were needed to reach equilibrium all through the transition, because the protein unfolded and refolded extremely slowly in the transition region. Extremely slow unfolding kinetics were also observed for two other β-helical proteins [74,77] and, by analogy with proteins from hyperthermophilic organisms [119,120], this is likely associated with the compact, hydrophobic core formed by their β-helix domain.

Analysis of the data according to a two-state model yielded experimental cooperativity parameters (*m* values) for the unfolding transition (-30 ± 2 and -14 ± 2 kJ·mol^{-1}·M^{-1} in GdmCl and urea, respectively), which are consistent with those calculated (-33 and -15 kJ·mol^{-1}·M^{-1}, respectively) on the basis of the size of PemA [121]. The *m* value depends on the efficiency of the denaturant ($m_{GdmCl} \approx 2.3 \cdot m_{urea}$ [121]) and is proportional to the change in solvent exposed surface area upon unfolding. It is a good indicator for the two-state character of a transition, because hidden intermediates in a transition would decrease its apparent value substantially [121–123]. Thus, the coincidence between the unfolding curves obtained by fluorescence and CD measurements (Figure 3C), and the good agreement between calculated and experimentally determined *m* values indicate that the enzyme unfolds in a cooperative transition that involves the entire molecule. This allows the free energy change for unfolding ($\Delta G°_{NU} = 35 \pm 5$ kJ·mol^{-1}) to be calculated with confidence. The two-state unfolding of PemA by denaturants contrasts with the three-state transitions observed for pertactin [77] and for the plasmid-encoded toxin (Pet) from *E. coli* [124], two other β-helical proteins. For them, the population of a partially folded intermediate was observed at about 1.5 and 1.1 M GdmCl, respectively. Finally, for a large fragment (termed Bhx) of TSP (residues 109–544), which encompasses the entire β-helix domain, Seckler and collaborators [70,125] showed that its unfolding is a two-state process in urea, at low protein concentrations (10 μg·mL^{-1}, i.e., circa 0.2 μM), low ionic strength, and low temperature (10 °C), yielding (pH 7) $\Delta G°_{NU} = 32$ kJ·mol^{-1} and $m_{NU} = 12.7$ kJ·mol^{-1}·M^{-1}. In this case, however, the experimental *m* value was much lower than expected (19.8 kJ·mol^{-1}·M^{-1}) for a protein of this size, suggesting that the experimentally observed unfolding transitions involved only a part of the β-helix domain [70,125].

Thus, despite its large size and modular nature, PemA shows a remarkably high degree of cooperativity in equilibrium unfolding, as expressed in its close adherence to a simple two-state model. In comparison with PelC, which is very similar in size to PemA (353 and 342 residues, respectively), the data suggest a more homogeneous distribution of stabilities across the PemA β-helix length and also perhaps a higher stability of the repeat interfaces [46]. With pertactin, the larger size (i.e., 539 residues) and higher stability of

the *C*-terminal half [77] of the protein probably explains the non-cooperative equilibrium unfolding. Finally, for Pet [124], the occurrence of at least two structural domains besides the right-handed β-helical *C*-terminal domain [126] probably explains the complex multi-state transition.

The cooperative transition of PemA is also clearly evident from pressure-induced unfolding experiments. The FTIR measurements revealed that very high pressures (>600 MPa) are needed to unfold the enzyme. Pressure-induced unfolding was not reversible, but a tentative two-state analysis gave a $\Delta G°_{NU}$ value of 33 ± 4 kJ·mol^{-1}, which is identical within the error limits to that obtained from denaturant-induced unfolding. The apparent volume change for the unfolding of PemA ($\Delta V = -41 \pm 5$ mL·mol^{-1}) is relatively small for its size (it corresponds to 0.15% of its molecular volume, i.e., 27336 mL·mol^{-1}, as calculated with the help of the Brugel software package [127]), and it is in the lower range compared to other globular proteins [128,129]. In comparison, pressure induced-unfolding studies on staphylococcal nuclease, a small monomeric protein of circa 17 kDa, revealed a significantly larger volume change ($\Delta V = -75$ mL·mol^{-1}) [130]. This contrasts, however, with the volume change for another (all α) repeat protein, Nank1-7, containing seven ankyrin sequence repeats and consisting of 248 amino acids, which was found to be -44 mL·mol^{-1} [131]. Here, a systematic analysis suggested that the volume change is mainly due to the changes in cavity volume rather than differential hydration. This study suggested that the major contributing factors to pressure effects on proteins are their imperfect internal packing and the occurrence of dry (i.e., with very low solvent occupancy) cavities in the folded state. Furthermore, it supports the view that there is a strong correlation between internal cavity volume and the volume change for protein unfolding [130,132]. In PemA, the side chain alignment of aliphatic and aromatic residues inside the hydrophobic β-helix core does not allow the presence of extended cavities. Analysis of the enzyme using the Brugel software [127] reveals that it contains seven major cavities, with volumes ranging from 7.7 to 14.8 mL·mol^{-1}. This implies that the overall cavity volume is larger than the observed volume change upon folding. Clearly, relating this reaction volume solely to changes in cavities would be over simplistic. Nevertheless, PemA shows a higher stability towards pressure-induced unfolding ($P_m = 800 \pm 7$ MPa) than many other proteins [128,133,134].

A combination of complementary spectroscopic probes was used to monitor the folding kinetics of PemA (Figures 5 and 8). These experiments revealed that very fast changes occurred during the 3–7 ms dead time of stopped-flow mixing, followed by at least three resolved phases that were observed by all spectroscopic techniques. Within the dead time, the enzyme acquired one half (~56 %) of its native ellipticity at 225 nm (Figure 5B), accompanied by the formation of hydrophobic patches that were accessible for the dye ANS (Figure 5C). Tryptophan fluorescence did not change in the dead-time reaction (Figure 5A), but the access of indole side chains towards quenching by acrylamide was reduced. These observations are consistent with very fast ($\tau < 1$ ms) hydrophobic collapse, a partial burial of tryptophan side chains, and substantial secondary structure formation. Stable native tertiary contacts had presumably not yet formed, however, as indicated by the absence of fluorescence changes during the dead time of mixing. These features are reminiscent of the molten globule type of intermediates that have been observed in the folding of many proteins [135–138]. A species with partial formation of the β-helix domain, involving the stack of phenylalanine residues (F168, F189 and F209—Figure 1C) in the core of the enzyme, might be a plausible model for the burst phase intermediate. This hypothesis is based on the data [64] showing that the burial of the internal side chains of stacked phenylalanine residues in the central region of the parallel β-helix domain of TSP is critical for its folding. Furthermore, it is also supported by the finding that ANS binding during the burst phase is not very important (i.e., ~1.5-fold increase) when compared to the typical > 5-fold increase observed with other proteins (see e.g., [100–103]). Similar results were obtained with pertactin [78], and these observations are interpreted here in terms of the early formation of a dry molten globule intermediate [139], with a partially structured β-helical domain. This species provides only low surface binding for the dye and hence,

is apparently already quite dehydrated and probably offers little solvent access to the protein interior. Finally, the loops, although probably forming dynamic and fluctuating structures, show a significant degree of hydrophobic collapse, however, with burial of the five tryptophan side chains in a non-native environment, as indicated by the low degree of acrylamide quenching.

Following the kinetically unresolved changes in far-UV CD, ANS fluorescence, and quenching of intrinsic fluorescence by acrylamide, a further minor increase (~5%) in ANS fluorescence was observed (Figure 5C), which reaches a maximum at circa 1 s. The native optical signals were then regained in three visible, well-resolved kinetic phases (numbered 2 to 4 in Table 1), which were observed by all four optical probes. The various methods yielded very similar time constant values (i.e., 6 ± 1 s, 23 ± 2 s, and 80 ± 20 s) for these phases, which are all compatible with the *cis-trans* isomerization of Xaa-Pro peptide bonds [113–115]. PemA contains 11 prolines, but double jump experiments and refolding assays in the presence of two different PPIs did not provide evidence that prolyl isomerization is involved in the folding kinetics of the enzyme. These data, together with the observation of a lag phase (phase 1, Figure 5A) in the stopped-flow intrinsic fluorescence experiments, suggest that the kinetic changes observed in phases 2 to 4 represent well-defined cooperative conformational transitions. These can be described by a sequential folding pathway (scheme (5)), with the population of four species, i.e., U, I_1, I_2, and N, wherein I_1 and I_2 represent two obligatory, on-pathway intermediates. Using the macroscopic rate constants given in Table 1, the time-course of the four species was simulated (Figure 9) on the basis of scheme (5). This indicates that I_1 and I_2 reach a maximum amount (~60 and ~50%, respectively) after circa 9 and 44 s, respectively. These values are greater than the length of the lag phase (time constant of circa 1 s) observed in fluorescence experiments (Figure 5A), which therefore, suggests that an additional kinetic intermediate accumulates between the burst phase species and I_1. It exhibits a low fluorescence intensity, similar to the phase burst intermediate, and therefore makes it possible to observe a lag in the appearance of the following species, which exhibit a greater fluorescence intensity. Very few claims (see [109,110]) of obligatory intermediates in protein folding are supported by the observation of an appropriate lag period ([108]), because favorable circumstances are required, namely the formation of an early intermediate species with spectroscopic properties identical or nearly identical to the starting species. The lag phase (also sometimes referred to as an induction period) is simply due to the accumulation of this spectroscopically silent intermediate before the substantial formation of subsequent species, with native-like spectroscopic characteristics. Clearly, the observation of a lag phase (Figure 5A) in the stopped-flow fluorescence experiments provides convincing evidence for the significant population of at least two productive (i.e., kinetically competent) intermediate species on the folding route between the denatured and native states (scheme (5) and Figure 10). Observation of this phase is possible here because the fluorescence properties of the native state, and of the two intermediate species I_1 and I_2, are markedly different from those of the unfolded and early intermediate species. The analysis of PemA folding, wherein structurally defined intermediates accumulate on a preferred pathway, is consistent with the classical macroscopic view [22,23] that folding occurs through a statistically predominant route. The distinct structural species that pave the way to the native state correspond to large molecular ensembles of rapidly interconverting molecules, which display significant common structural features and a measurable degree of cooperativity [24,140].
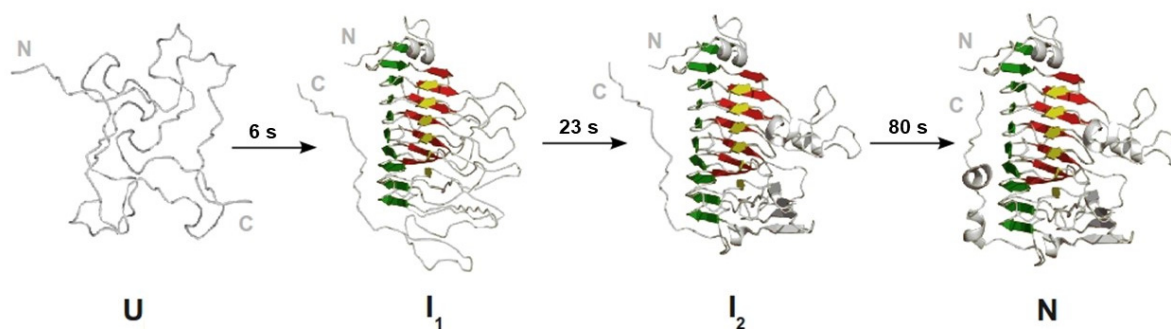
Figure 10. Representation of a possible folding pathway of PemA (pH 7, 25 °C). The sequential nature of the refolding process is indicated, and the two major intermediates are illustrated. The three β-sheets are in yellow (PB1), green (PB2), and red (PB3), and the *N*- and *C*-terminal ends of the polypeptide chain are indicated by N and C, respectively. Starting from a highly heterogeneous denatured state (U), the process of collapse (burst phase) and further minor structural rearrangements (phases 1 and 2) leads to a first intermediate species (I_1), with native-like structure in the entire β-helix domain but unstructured peripheral loops (see text for details). This is followed by the folding of most of the peripheral loops and the formation of a second intermediate species (I_2). Finally, the folding of the *C*-terminal loop and its docking against the parallel β-helix (PB2) (phase 4) leads to the final native structure. The values of the time constants for each step are indicated. Note that the various states in the model need to be described in terms of ensembles of conformers, and the experimental parameters are averaged over these ensembles.

The restoration of the native optical signals (i.e., ellipticity, fluorescence emission intensity, and access to acrylamide), together with the release of the dye (ANS) in three exponential phases, can be interpreted as follows (illustrated in Figure 10). The first event (phase 2) leads to the regaining of a large percentage (~80%) of native ellipticity at 225 nm and a mere ~15% of native fluorescence intensity. This could reflect the formation of virtually the entire β-helix domain with the peripheral loops remaining in a non-native collapsed state (represented by I_1 in Figure 10). A species (termed "native-like") with similar structural features has been observed [75] in the early stage of PelC folding, suggesting a common intermediate on the folding pathway of these two right-handed parallel β-helix proteins. Following the population of I_1, a second species (I_2, Figure 10) accumulates (phase 3) before the formation of the native structure. Although it has recovered ~93% of the native CD signal at 225 nm, it shows only ~45% of the native fluorescence signal, indicating that it is missing a significant fraction of the native tertiary contacts. The unexpected increase in acrylamide quenching demonstrates, however, that a substantial change in the environment of one or more of the five tryptophan residues takes place during this phase. An attractive model would involve the docking of the tryptophan indole side chains close to their native orientations, particularly those at positions 107, 269, and 317 (see Figure 1). Thus, W107 and W317 display the two most solvent-exposed side chains, with ~37% and ~55% accessibility, respectively [112]. The increase in acrylamide quenching in phase 3 might be caused by the reexposure of these tryptophan residues during the final folding of the peripheral loops. In contrast, the large (~30%) increase in tryptophan fluorescence in this phase might originate from the burial of W269 in its final hydrophobic environment, close to the central PB1 β-sheet. W269 shows enhanced fluorescence probably because it resides in a hydrophobic environment and because the neighboring tyrosines Y158 and Y181 (Figure 1) transfer energy to W269. Finally, we speculate that the strong (~55%) increase in intrinsic fluorescence and ~7% change in ellipticity (phase 4), which characterize the final conformational adjustments to achieve the native structure, might be associated with the folding of the *C*-terminal loop (residues 345 to 359). The latter, which forms two short helices, packs anti-parallelly against the PB2 β-sheet of the parallel β-helix [49,50]. This results in the transfer of the indole side chain of W361 to a more hydrophobic environment, in close proximity to Y167 (Figure 1). As described for W269, the final positioning of W361 could result in a substantial increase in fluorescence intensity.

The chevron plot in Figure 6 highlights the complexity of the refolding kinetics of PemA, which contrasts with the apparent simplicity of unfolding. Thus, the absence of curvature in the unfolding limb of the chevron plot suggests a single cooperative unfolding, without intermediates, i.e., no sequential unraveling of the repeat array, repeat by repeat [45]. In refolding at low denaturant concentrations, only the rate of the slowest of the four exponential phases varies significantly with GdmCl concentration. This supports our analysis (schematically summarized in Figure 10) that, from the burst phase intermediate to I_2, only minor local conformational rearrangements occur, without major changes in the surface area accessible to the denaturant. Strong changes in solvent accessibility occur during the last phase of folding and therefore, its rate shows strong denaturant dependence [141]. Between 0.6 and 4.5 M, the unfolding and refolding branches of the plot for the slowest phase (i.e., phase 4 in Table 1) were found to be linear. Thus, in this concentration range, Equation (3) was used to fit the data, yielding the kinetic rate constants given in the legend to Figure 6. Using the values of $k_f{}^{H_2O}$ and $k_u{}^{H_2O}$, Equation (4) allows calculation of a $\Delta G°_{kin}$ value of 39 ± 8 kJ·mol^{-1}, which is identical within the error limits to the value (32 ± 3 kJ·mol^{-1}) obtained from the equilibrium unfolding transitions.

Interestingly, the value of the folding rate constant (i.e., 0.8 s^{-1}) extrapolated at 0 M GdmCl is about five orders of magnitude lower than that (circa $2 \cdot 10^5$ s^{-1}) estimated on the basis of relative protein contact order [32,33] (calculated with the help of Baker Laboratory Tool http://depts.washington.edu/bakerpg/contact_order/, accessed on 11 May 2021). Other repeat proteins (see e.g., [78,142]) also fold with rates much lower than predicted based on their native state topology. As discussed in [42], this is due at least in part to the unusual local topology of all repeat proteins, which lack sequence-distant contacts. This lack and the simple topology of repeat proteins result in uniform and low contact order, and hence, the calculation based on contact order leads to folding rates that are orders of magnitude too high. Clearly, contact-order models, as developed for small globular proteins, are inadequate for repeat proteins [42].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230. [CrossRef]
2. Ellis, R.J.; Dobson, C.M.; Hartl, U. Sequence does specify protein conformation. *Trends Biochem. Sci.* **1998**, *23*, 468. [CrossRef]
3. Gierasch, L.M.; Gershenson, A. Post-reductionist protein science, or putting Humpty Dumpty back together again. *Nat. Chem. Biol.* **2009**, *5*, 774–777. [CrossRef]

4. Balchin, D.; Hayer-Hartl, M.; Hartl, F.U. In vivo aspects of protein folding and quality control. *Science* **2016**, *353*, aac4354. [CrossRef] [PubMed]

5. Ventura, A.; Villaverde, A. Protein quality in bacterial inclusion bodies. *Trends. Biotechnol.* **2006**, *24*, 179–185. [CrossRef]

6. Chiti, F.; Dobson, C.M. Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade. *Annu. Rev. Biochem.* **2017**, *86*, 27–68. [CrossRef] [PubMed]

7. Dobson, C.M. Finding the right fold. *Nat. Struct. Biol.* **1995**, *2*, 513–517. [CrossRef]

8. Jackson, S.E.; Fersht, A.R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **1991**, *31*, 10428–10435. [CrossRef]

9. Viguera, A.R.; Martínez, J.C.; Filimonov, V.V.; Mateo, P.L.; Serrano, L. Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* **1994**, *33*, 2142–2150. [CrossRef] [PubMed]

10. Huang, G.S.; Oas, T.G. Submillisecond folding of monomeric lambda repressor. *Proc. Natl. Acad. Sci. USA* **1995**, *95*, 6878–6882. [CrossRef]

11. Schindler, T.; Herrler, M.; Marahiel, M.A.; Schmid, F.X. Extremely rapid protein folding in the absence of intermediates. *Nat. Struct. Biol.* **1995**, *2*, 663–673. [CrossRef] [PubMed]

12. Kragelund, B.B.; Robinson, C.V.; Knudsen, J.; Dobson, C.M.; Poulsen, F.M. Folding of a four-helix bundle: Studies of acyl-coenzyme a binding protein. *Biochemistry* **1995**, *34*, 7217–7224. [CrossRef]

13. Van Nuland, N.A.; Meijberg, W.; Warner, J.; Forge, V.; Scheek, R.M.; Robillard, G.T.; Dobson, C.M. Slow cooperative folding of a small globular protein HPr. *Biochemistry* **1998**, *37*, 622–637. [CrossRef]

14. Friel, C.T.; Beddard, G.S.; Radford, S.E. Switching two-state to three-state kinetics in the helical protein Im9 via the optimisation of stabilising non-native interactions by design. *J. Mol. Biol.* **2004**, *342*, 261–273. [CrossRef]

15. Radford, S.E.; Dobson, C.M.; Evans, P.A. The folding pathway of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **1992**, *358*, 302–307. [CrossRef] [PubMed]

16. Forge, V.; Wijesinha, R.T.; Brew, K.; Robinson, C.V.; Redfield, C.; Dobson, C.M. Rapid collapse and slow structural reorganisation during the refolding of bovine α-lactalbumin. *J. Mol. Biol.* **1999**, *288*, 673–688. [CrossRef] [PubMed]

17. Fersht, A.R. The sixth Datta Lecture. Protein folding and stability: The pathway of folding of barnase. *FEBS Lett.* **1993**, *325*, 5–16. [CrossRef]

18. Khorasanizadeh, S.; Peters, I.D.; Roder, H. Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nat. Struct. Biol.* **1996**, *3*, 193–205. [CrossRef] [PubMed]

19. Friel, C.T.; Smith, D.A.; Vendruscolo, M.; Gsponer, J.; Radford, S.E. The mechanism of folding of Im7 reveals competition between functional and kinetic evolutionary constraints. *Nat. Struct. Mol. Biol.* **2009**, *16*, 318–324. [CrossRef] [PubMed]

20. Arai, M.; Iwakura, M.; Matthews, C.R.; Bilsel, O. Microsecond subdomain folding in dihydrofolate reductase. *J. Mol. Biol.* **2011**, *410*, 329–342. [CrossRef]

21. Houwman, J.A.; Westphal, A.H.; Visser, J.W.G.; Borst, J.W.; van Mierlo, C.P. Concurrent presence of on- and off-pathway folding intermediates of apoflavodoxin at physiological ionic strength. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7059–7072. [CrossRef]

22. Kim, P.S.; Baldwin, R.L. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **1990**, *59*, 631–660. [CrossRef]

23. Matthews, C.R. Pathways of protein folding. *Annu. Rev. Biochem.* **1993**, *62*, 653–683. [CrossRef]

24. Dobson, C.M.; Šali, A.; Karplus, M. Protein folding: A perspective from theory and experiment. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 868–893. [CrossRef]

25. Chan, H.S.; Dill, K.A. Protein folding in the landscape perspective: Chevron plots and non-Arrhenius kinetics. *Proteins* **1998**, *30*, 2–33. [CrossRef]

26. Radford, S.E. Protein folding: Progress made and promises ahead. *Trends Biochem. Sci.* **2000**, *25*, 611–618. [CrossRef]

27. Szilágyi, A.; Kardos, J.; Osváth, S.; Barna, L.; Závodszky, P. Protein Folding. In *Handbook of Neurochemistry and Molecular Neurobiology*, 3rd ed.; Lajtha, A., Banik, N., Eds.; Springer: Boston, MA, USA, 2007; pp. 303–343.

28. Udgaonkar, J.B. Multiple routes and structural heterogeneity in protein folding. *Annu. Rev. Biophys.* **2008**, 489–510. [CrossRef] [PubMed]

29. Dill, K.A.; Ozkan, S.B.; Shell, M.S.; Weikl, T.R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289–316. [CrossRef] [PubMed]

30. Gruebele, M.; Dave, K.; Sukenik, S. Globular protein folding in vitro and in vivo. *Annu. Rev. Biophys.* **2016**, *45*, 233–251. [CrossRef]

31. Daggett, V.; Fersht, A.R. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* **2003**, *28*, 18–25. [CrossRef]

32. Plaxco, K.W.; Simons, K.T.; Ruczinski, I.; Baker, D. Topology, Stability, Sequence, and Length: Defining the Determinants of Two-State Protein Folding Kinetics. *Biochemistry* **2000**, *39*, 11177–11183. [CrossRef]

33. Ivankov, D.N.; Garbuzynskiy, S.O.; Alm, E.; Plaxco, K.W.; Baker, D.; Finkelstein, A.V. Contact order revisited: Influence of protein size on the folding rate. *Prot. Sci.* **2003**, *12*, 2057–2062. [CrossRef]

34. Walters, B.T.; Mayne, L.; Hinshaw, T.R.; Englander, S.W. Folding of a large protein at high structural resolution. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18898–18903. [CrossRef]

35. Clarke, J.; Pappu, R.V. Protein Folding and Binding, Complexity Comes of Age. *Curr. Opin. Struct. Biol.* **2017**, *42*, 5–7. [CrossRef] [PubMed]

36. Kajava, A.V. Tandem repeats in proteins: From sequence to structure. *J. Struct. Biol* **2012**, *179*, 279–288. [CrossRef]

37. Mello, C.C.; Barrick, D. An experimentally determined protein folding energy landscape. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 14102–14107. [CrossRef] [PubMed]
38. Ferreiro, D.U.; Cho, S.S.; Komives, E.A.; Wolynes, P.G. The energy landscape of modular repeat proteins: Topology determines folding mechanism in the ankyrin family. *J. Mol. Biol.* **2005**, *354*, 679–692. [CrossRef]
39. Main, E.R.G.; Lowe, A.R.; Mochrie, S.G.J.; Jackson, S.E.; Regan, L. A recurring theme in protein engineering: The design, stability and folding of repeat proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 464–471. [CrossRef] [PubMed]
40. Löw, C.; Weininger, U.; Neumann, P.; Klepsch, M.; Lilie, H.; Stubbs, M.L.; Balbach, J. Structural insights into an equilibrium folding intermediate of an archaeal ankyrin repeat protein. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3779–3784. [CrossRef] [PubMed]
41. Barrick, D.; Ferreiro, D.U.; Komives, E.A. Folding landscapes of ankyrin repeat proteins: Experiments meet theory. *Curr. Opin. Struct. Biol.* **2008**, *18*, 27–34. [CrossRef]
42. Kloss, E.; Courtemanche, N.; Barrick, D. Repeat-protein folding: New insights into origins of cooperativity, stability, and topology. *Arch. Biochem. Biophys.* **2008**, *469*, 83–99. [CrossRef] [PubMed]
43. Wetzel, S.K.; Settanni, G.; Kenig, M.; Binz, H.K.; Plückthun, A. Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J. Mol. Biol.* **2008**, *376*, 241–257. [CrossRef]
44. Javadi, Y.; Itzhaki, L. Tandem-repeat proteins: Regularity plus modularity equals design-ability. *Curr. Opin. Struct. Biol.* **2013**, *23*, 622–631. [CrossRef] [PubMed]
45. Rowling, P.J.E.; Sivertsson, E.M.; Perez-Riba, A.; Main, E.R.G.; Itzhaki, L.S. Dissecting and reprogramming the folding and assembly of tandem-repeat proteins. *Biochem. Soc. Trans.* **2015**, *43*, 881–888. [CrossRef]
46. Perez-Riba, A.; Synakewicz, M.; Itzhaki, L.S. Folding cooperativity and allosteric function in the tandem-repeat protein class. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2018**, *373*. [CrossRef]
47. Toth, I.K.; Bell, K.S.; Holeva, M.C.; Birch, P.R.J. Soft rot erwiniae: From genes to genomes. *Mol. Plant Pathol.* **2003**, *4*, 17–30. [CrossRef]
48. Kazemi-Pour, N.; Condemine, G.; Hugouvieux-Cotte-Pattat, N. The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*. *Proteomics* **2004**, *4*, 3177–3186. [CrossRef] [PubMed]
49. Fries, M.; Ihrig, J.; Brocklehurst, K.; Shevchik, V.E.; Pickersgill, R.W. Molecular basis of the activity of the phytopathogen pectin methylesterase. *EMBO J.* **2007**, *26*, 3879–3887. [CrossRef]
50. Jenkins, J.; Mayans, O.; Smith, D.; Worboys, K.; Pickersgill, R.W. Three-dimensional structure of *Erwinia chrysanthemi* pectin methylesterase reveals a novel esterase active site. *J. Mol. Biol.* **2001**, *305*, 951–960. [CrossRef]
51. McDonnel, A.V.; Menke, M.; Palmer, N.; King, J.; Cowen, L.; Berger, B. Fold recognition and accurate sequence-structure alignment of sequences directing β-sheet proteins. *Proteins* **2006**, *63*, 976–985. [CrossRef] [PubMed]
52. Kajava, A.V.; Steven, A.C. β-Rolls, β-helices, and other β-solenoid proteins. *Adv. Protein Chem.* **2006**, *73*, 55–96.
53. Jenkins, J.; Pickersgill, R. The architecture of parallel β-helices and related folds. *Prog. Biophys. Mol. Biol.* **2001**, *77*, 111–175. [CrossRef]
54. Wetzel, R. Ideas of order for amyloid fibrils. *Structure* **2002**, *10*, 1031–1036. [CrossRef]
55. Wille, H.; Michelitsch, M.; Guenebaut, V.; Supattapone, S.; Serban, A.; Cohen, F.; Agard, D.; Prusiner, S. Structural studies of the scrapie prion protein by electron crystallography. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3563–3568. [CrossRef] [PubMed]
56. Pickersgill, R.W. A primordial structure underlying amyloid. *Structure* **2003**, *11*, 137–138. [CrossRef]
57. Jahn, T.R.; Makin, O.S.; Morris, K.L.; Marshall, K.E.; Tian, P.; Sikorski, P.; Serpell, L.C. The Common Architecture of Cross-β Amyloid. *J. Mol. Biol.* **2010**, *395*, 717–727. [CrossRef] [PubMed]
58. Yoder, M.D.; Lietzke, S.E.; Jurnak, F. Unusual structural features in the parallel β-helix in pectate lyases. *Structure* **1993**, *1*, 241–251. [CrossRef]
59. Yoder, M.D.; Keen, N.T.; Jurnak, F. New domain motif: The structure of pectate lyase C, a secreted plant virulence factor. *Science* **1993**, *260*, 1503–1507. [CrossRef]
60. Steinbacher, S.; Miller, S.; Baxa, U.; Budisa, N.; Weintraub, A.; Seckler, R.; Huber, R. Phage P22 tailspike protein: Crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. *J. Mol. Biol.* **1997**, *267*, 865–880. [CrossRef] [PubMed]
61. Emsley, P.; Charles, I.G.; Fairweather, N.F.; Isaacs, N.W. Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* **1996**, *381*, 90–92. [CrossRef] [PubMed]
62. Bryan, A.W., Jr.; Starner-Kreinbrink, J.L.; Hosur, R.; Clark, P.L.; Berger, B. Structure-based prediction reveals capping motifs that inhibit β-helix aggregation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11099–11104. [CrossRef]
63. McGaughey, G.B.; Gagné, M.; Rappé, A.K. Pi-Stacking interactions. Alive and well in proteins. *J Biol Chem.* **1998**, *273*, 15458–15463. [CrossRef]
64. Betts, S.; Haase-Pettingell, C.; Cook, K.; King, J. Buried hydrophobic side-chains essential for the folding of the parallel β-helix domains of the P22 tailspike. *Protein Sci.* **2004**, *13*, 2291–2303. [CrossRef] [PubMed]
65. Simkovsky, R.; King, J. An elongated spine of buried core residues necessary for in vivo folding of the parallel β-helix of P22 tailspike adhesin. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3575–3580. [CrossRef] [PubMed]
66. Betts, S.; King, J. There's a right way and a wrong way: In vivo and in vitro folding, misfolding and subunit assembly of the P22 tailspike. *Structure* **1999**, *7*, R131–R139. [CrossRef]

67. King, J.; Haase-Pettingell, C.; Robinson, A.S.; Speed, M.; Mitraki, A. Thermolabile folding intermediates: Inclusion body precursors and chaperonin substrates. *FASEB J.* **1996**, *10*, 57–66. [CrossRef] [PubMed]

68. Haase-Pettingell, C.; King, J. Prevalence of temperature sensitive folding mutations in the parallel β coil domain of the phage P22 tailspike endorhamnosidase. *J. Mol. Biol.* **1997**, *267*, 88–102. [CrossRef]

69. Seckler, R. Folding and function of repetitive structure in the homotrimeric phage P22 tailspike protein. *J. Struct. Biol.* **1998**, *122*, 216–222. [CrossRef] [PubMed]

70. Schuler, B.; Seckler, R. P22 tailspike folding mutants revisited: Effects on the thermodynamic stability of the isolated β-helix domain. *J. Mol. Biol.* **1998**, *281*, 227–234. [CrossRef] [PubMed]

71. Benton, C.B.; King, J.; Clark, P.L. Characterization of the protrimer intermediate in the folding pathway of the interdigitated β-helix tailspike protein. *Biochemistry* **2002**, *41*, 5093–5103. [CrossRef]

72. Jain, M.; Evans, M.S.; King, J.; Clark, P.L. Monoclonal antibody epitope mapping describes tailspike β-helix folding and aggregation intermediates. *J. Biol. Chem.* **2005**, *280*, 23032–23040. [CrossRef] [PubMed]

73. Evans, M.S.; Sander, I.M.; Clark, P.L. Cotranslational folding promotes β -helix formation and avoids aggregation in vivo. *J. Mol. Biol.* **2008**, *383*, 683–692. [CrossRef]

74. Kamen, D.E.; Griko, Y.; Woody, R.W. The Stability, Structural Organization, and Denaturation of Pectate Lyase C, a Parallel β-Helix Protein. *Biochemistry* **2000**, *39*, 15932–15943. [CrossRef] [PubMed]

75. Kamen, D.E.; Woody, R.W. Folding kinetics of the protein pectate lyase C reveal fast-forming intermediates and slow proline isomerization. *Biochemistry* **2002**, *41*, 4713–4723. [CrossRef] [PubMed]

76. Kamen, D.E.; Woody, R.W. Identification of proline residues responsible for the slow folding kinetics in pectate lyase C by mutagenesis. *Biochemistry* **2002**, *41*, 4724–4732. [CrossRef]

77. Junker, M.; Schuster, C.C.; McDonnell, A.V.; Sorg, K.A.; Finn, M.C.; Berger, B.; Clark, P.L. Pertactin β-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4918–4923. [CrossRef] [PubMed]

78. Junker, M.; Clark, P.L. Slow formation of aggregation-resistant β-sheet folding intermediates. *Proteins* **2010**, *78*, 812–824. [CrossRef]

79. Sambrook, J.; Russel, D.W. *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2001.

80. Vandevenne, M.; Filée, P.; Scarafone, N.; Cloes, B.; Gaspard, G.; Yilmaz, N.; Dumoulin, M.; François, J.-M.; Frère, J.-M.; Galleni, M. The *Bacillus licheniformis* BlaP β-lactamase as a model protein scaffold to study the insertion of protein fragments. *Protein Sci.* **2007**, *16*, 2260–2271. [CrossRef] [PubMed]

81. Pace, C.N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **1995**, *4*, 2411–2423. [CrossRef]

82. Manavalan, P.; Johnson, W.C., Jr. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **1987**, *167*, 76–85. [CrossRef]

83. Sreerama, N.; Woody, R.W. Estimation of protein secondary structure from CD spectra: Comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **2000**, *287*, 252–260. [CrossRef] [PubMed]

84. Whitmore, L.; Wallace, B.A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **2004**, *32*, W668–W673. [CrossRef]

85. Whitmore, L.; Wallace, B.A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **2008**, *89*, 392–400. [CrossRef]

86. Nozaki, Y. The preparation of guanidine hydrochloride. *Methods Enzymol.* **1972**, *26*, 43–50.

87. Lakowicz, J.R. Quenching of Fluorescence. In *Principles of Fluorescence Spectroscopy*, 2nd ed.; Kluwer Academic/Plenum Publishers: New York, NY, USA, 1999; pp. 237–265.

88. Dumoulin, M.; Conrath, K.; Van Meirhaeghe, A.; Meersman, F.; Heremans, K.; Frenken, L.G.; Muyldermans, S.; Wyns, L.; Matagne, A. Single-domain antibody fragments with high conformational stability. *Protein Sci.* **2002**, *11*, 500–515. [CrossRef]

89. Dirix, C.; Duvetter, T.; Loey, A.V.; Hendrickx, M.; Heremans, K. The in situ observation of the temperature and pressure stability of recombinant *Aspergillus aculeatus* pectin methylesterase with Fourier transform IR spectroscopy reveals an unusual pressure stability of β-helices. *Biochem. J.* **2005**, *392*, 565–571. [CrossRef] [PubMed]

90. Di Paolo, A.; Balbeur, D.; De Pauw, E.; Redfield, C.; Matagne, A. Rapid Collapse into a Molten Globule Is Followed by Simple Two-State Kinetics in the Folding of Lysozyme from Bacteriophage λ. *Biochemistry* **2010**, *49*, 8646–8657. [CrossRef]

91. Vandenameele, J.; Lejeune, A.; Di Paolo, A.; Brans, A.; Frère, J.-M.; Schmid, F.X.; Matagne, A. Folding of class A β-lactamases is rate-limited by peptide bond isomerization and occurs via parallel pathways. *Biochemistry* **2010**, *49*, 4264–4275. [CrossRef] [PubMed]

92. Tanford, C. Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* **1970**, *24*, 1–95.

93. Baldwin, R.L. On-pathway versus off-pathway folding intermediates. *Fold. Des.* **1996**, *1*, R1–R8. [CrossRef]

94. Fersht, A.R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; WH. Freeman and Co.: New York, NY, USA, 1998.

95. Moore, J.W.; Pearson, R.G. Complex reactions. In *Kinetics and Mechanism*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 1981; pp. 284–333.

96. Doetsch, G. *Introduction to the Theory and Application of the Laplace Transformation*; Springer: Berlin/Heidelberg, Germany, 1974.

97. Johnson, W.C., Jr. Protein secondary structure and circular dichroism: A practical guide. *Proteins* **1990**, *7*, 205–214. [CrossRef]

98. Ptitsyn, O.B.; Pain, R.H.; Semisotnov, G.V.; Zerovnik, E.; Razgulyaev, O.I. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **1990**, *262*, 20–24. [CrossRef]

99. Semisotnov, G.V.; Rodionova, N.A.; Razgulyaev, O.I.; Uversky, V.N.; Gripas, A.F.; Gilmanshin, R.I. Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **1991**, *31*, 119–128. [CrossRef] [PubMed]

100. Jones, B.E.; Jennings, P.A.; Pierre, R.A.; Matthews, C.R. Development of nonpolar surfaces in the folding of Escherichia coli dihydrofolate reductase detected by 1-anilinonaphthalene-8-sulfonate binding. *Biochemistry* **1994**, *33*, 15250–15258. [CrossRef] [PubMed]

101. Itzhaki, L.S.; Evans, P.A.; Dobson, C.M.; Radford, S.E. Tertiary interactions in the folding pathway of hen lysozyme: Kinetic studies using fluorescent probes. *Biochemistry* **1994**, *33*, 5212–5220. [CrossRef]

102. Engelhard, M.; Evans, P.A. Kinetics of interaction of partially folded proteins with a hydrophobic dye: Evidence that molten globule character is maximal in early folding intermediates. *Protein Sci.* **1995**, *4*, 1553–1562. [CrossRef] [PubMed]

103. Vanhove, M.; Lejeune, A.; Guillaume, G.; Virden, R.; Pain, R.H.; Schmid, F.X.; Frère, J.-M. A collapsed intermediate with nonnative packing of hydrophobic residues in the folding of TEM-1 β-lactamase. *Biochemistry* **1998**, *37*, 1941–1950. [CrossRef]

104. Khurana, R.; Fink, A.L. Do parallel βhelix proteins have a unique Fourier transform infrared spectrum? *Biophys. J.* **2000**, *78*, 994–1000. [CrossRef]

105. Jackson, M.; Mantsch, H.H. The use and misuse of FTIR spectroscopy in the determination of protein structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95–120. [CrossRef]

106. Sandroff, C.J.; King, H.E.; Herschbach, D.R. High pressure study of the liquid/solid interface: Surface enhanced raman scattering from adsorbed molecules. *J. Phys. Chem.* **1984**, *88*, 5647–5653. [CrossRef]

107. Haris, P.I.; Chapman, D. The conformational analysis of peptides using Fourier transform IR spectroscopy. *Biopolymers* **1995**, *37*, 251–263. [CrossRef] [PubMed]

108. Creighton, T.E. Protein folding. *Biochem. J.* **1990**, *270*, 1–16. [CrossRef] [PubMed]

109. Walkenhorst, W.F.; Green, S.M.; Roder, H. Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease. *Biochemistry* **1997**, *36*, 5795–5805. [CrossRef]

110. Heidary, D.K.; O'Neill, J.C., Jr.; Roy, M.; Jennings, P.A. An essential intermediate in the folding of dihydrofolate reductase. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5866–5870. [CrossRef]

111. Garvey, E.P.; Swank, J.; Matthews, C.R. A hydrophobic cluster forms early in the folding of dihydrofolate reductase. *Proteins* **1989**, *6*, 259–266. [CrossRef]

112. Hubbard, S.J.; Thornton, J.M. *'NACCESS', Computer Program*; Department Biochemistry and Molecular Biology, University College: London, UK, 1993.

113. Schmid, F.X.; Baldwin, R.L. Acid catalysis of the formation of the slow-folding species of RNase A: Evidence that the reaction is proline isomerization. *Proc. Nat. Acad. Sci. USA* **1978**, *75*, 4764–4768. [CrossRef]

114. Kiefhaber, T.; Kohler, H.H.; Schmid, F.X. Kinetic coupling between protein folding and prolyl isomerization. I. Theoretical models. *J. Mol. Biol.* **1992**, *224*, 217–229. [CrossRef]

115. Schmid, F.X. Prolyl isomerase: Enzymatic catalysis of slow protein-folding reactions. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 123–142. [CrossRef] [PubMed]

116. Brandts, J.F.; Halvorson, H.R.; Brennan, M. Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **1975**, *14*, 4953–4963. [CrossRef]

117. Maier, R.; Scholz, C.; Schmid, F.X. Dynamic association of trigger factor with protein substrates. *J. Mol. Biol.* **2001**, *314*, 1181–1190. [CrossRef]

118. Weininger, U.; Haupt, C.; Schweimer, K.; Graubner, W.; Kovermann, M.; Brüser, T.; Scholz, C.; Schaarschmidt, P.; Žoldák, G.; Schmid, F.X.; et al. NMR solution structure of SlyD from Escherichia coli: Spatial separation of prolyl isomerase and chaperone function. *J. Mol. Biol.* **2009**, *387*, 295–305. [CrossRef]

119. Sterner, R.; Liebl, W. Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* **2001**, *36*, 39–106. [CrossRef]

120. Feller, G. Protein folding at extreme temperatures: Current issues. *Semin. Cell Dev. Biol.* **2018**, *84*, 129–137. [CrossRef]

121. Myers, J.K.; Pace, C.N.; Scholtz, J.M. Denaturant *m* values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **1995**, *4*, 2138–2148. [CrossRef] [PubMed]

122. Pace, C.N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **1986**, *131*, 266–280.

123. Spudich, G.; Marqusee, S. A change in the apparent *m* value reveals a populated intermediate under equilibrium conditions in *Escherichia coli* ribonuclease HI. *Biochemistry* **2000**, *39*, 11677–11683. [CrossRef] [PubMed]

124. Renn, J.P.; Clark, P.L. A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. *Biopolymers* **2008**, *89*, 420–427. [CrossRef]

125. Miller, S.; Schuler, B.; Seckler, R. A reversibly unfolding fragment of P22 tailspike protein with native structure: The isolated beta-helix domain. *Biochemistry* **1998**, *37*, 9160–9168. [CrossRef]

126. Meza-Aguilar, J.D.; Fromme, P.; Torres-Larios, A.; Mendoza-Hernández, G.; Hernandez-Chiñas, U.; Monteros, R.A.A.-E.D.L.; Campos, C.A.E.; Fromme, R. X-ray crystal structure of the passenger domain of plasmid encoded toxin(Pet), an autotransporter enterotoxin from enteroaggregative Escherichia coli (EAEC). *Biochem. Biophys. Res. Commun.* **2014**, *445*, 439–444. [CrossRef]

127. Delhaise, P.; Bardiaux, M.; De Maeyer, M.; Prévost, M.; Vanbelle, D.; Donneux, J.; Lasters, I.; Vancustem, E.; Alard, P.; Wodak, S.J. The brugel package—Toward computer-aided-design of macromolecules. *J. Mol. Graph.* **1988**, *6*, 219. [CrossRef]

128. Gross, M.; Jaenicke, R. Proteins under pressure. The influence of high hydrostatic pressure on structure, function and assembly of proteins and protein complexes. *Eur. J. Biochem.* **1994**, *221*, 617–630. [CrossRef] [PubMed]

129. Royer, C.A. Revisiting volume changes in pressure-induced protein unfolding. *Biochim. Biophys. Acta.* **2002**, *1595*, 201–209. [CrossRef]

130. Roche, J.; Caro, J.A.; Norberto, D.R.; Barhe, P.; Roumestand, C.; Schlessman, J.L.; Garcia, A.E.; Garcia-Moreno, B.E.; Royer, C. Cavities determine the pressure unfolding of proteins. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6945–6950. [CrossRef]

131. Rouget, J.B.; Aksel, T.; Roche, J.; Saldana, J.L.; Garcia, A.E.; Barrick, D.; Royer, C.A. Size and sequence and the volume change of protein folding. *J. Am. Chem. Soc.* **2011**, *133*, 6020–6027. [CrossRef]

132. Ando, N.; Barstow, B.U.; Baase, W.A.; Fields, A.; Matthews, B.W.; Gruner, S.M. Structural and Thermodynamic Characterization of T4 Lysozyme Mutants and the Contribution of Internal Cavities to Pressure Denaturation. *Biochemistry* **2008**, *47*, 11097–11109. [CrossRef] [PubMed]

133. Silva, J.L.; Foguel, D.; Royer, C.A. Pressure provides new insights into protein folding, dynamics and structure. *Trends Biochem. Sci.* **2001**, *26*, 612–618. [CrossRef]

134. Meersman, F.; Dobson, C.M.; Heremans, K. Protein unfolding, amyloid fibril formation and configurational energy landscapes under high pressure conditions. *Chem. Soc. Rev.* **2006**, *35*, 908–917. [CrossRef] [PubMed]

135. Kuwajima, K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* **1989**, *6*, 87–103. [CrossRef] [PubMed]

136. Christensen, H.; Pain, R.H. Molten globule intermediates and protein folding. *Eur. Biophys. J.* **1991**, *19*, 221–229. [CrossRef]

137. Ptitsyn, O.B. Molten globule and protein folding. *Adv. Protein. Chem.* **1995**, *47*, 83–229.

138. Redfield, C. Molten Globules. *Curr. Biol.* **1999**, *9*, R313. [CrossRef]

139. Baldwin, R.L.; Frieden, C.; Rose, G.D. Dry molten globule intermediates and the mechanism of protein unfolding. *Proteins* **2010**, *78*, 2725–2737. [CrossRef] [PubMed]

140. Matagne, A.; Jamin, M.; Chung, E.W.; Radford, S.E.; Robinson, C.V.; Dobson, C.M. Thermal unfolding of an intermediate is associated with non-Arrhenius kinetics in the folding of hen lysozyme. *J. Mol. Biol.* **2000**, *297*, 193–210. [CrossRef] [PubMed]

141. Wallace, L.A.; Matthews, C.R. Sequential vs. parallel protein-folding mechanisms: Experimental tests for complex folding reactions. *Biophys. Chem.* **2002**, *101–102*, 113–131. [CrossRef]

142. Mello, C.C.; Bradley, C.M.; Tripp, K.W.; Barrick, D. Experimental characterization of the folding kinetics of the notch ankyrin domain. *J. Mol. Biol.* **2005**, *352*, 266–281. [CrossRef]