Pea Seed Lectin Folds and Oligomerizes via an Intermediate Not Represented in the Structural Hierarchy†

Frank Küster‡ and Robert Seckler*

Physical Biochemistry, Potsdam University, Potsdam, Germany

Received September 18, 2007; Revised Manuscript Received December 11, 2007

ABSTRACT: Large oligomeric proteins are usually thought to fold and assemble hierarchically: Domains fold and coalesce to form the subunits, and folded subunits can then associate to form the multimeric structure. We have investigated the refolding pathway of the β-sheet protein pea seed lectin using spectroscopic and hydrodynamic techniques. In vivo, it is proteolytically processed post-translationally, so that the single-domain subunits of the initial homodimer themselves become heterodimers of intertwined fragment polypeptide chains. Despite this complex topology, mature pea seed lectin reassembles with considerable efficiency at low total protein concentration (10 μg/mL) and low temperature (10 °C), albeit very slowly ($t_{1/2} \approx 2$ days). Contrary to expectations, the primary assembly product is not the intact β-sheet domain, but the larger fragment chains first dimerize to form the native-like subunit interface. The smaller fragment chains then associate with this preformed dimer.

Structure formation of multimeric or multidomain proteins is usually seen as a hierarchic process: First, domains fold, either independently or depending on other domains being structured. They then assemble to form multidomain proteins. Although in many cases no structured monomers of oligomeric proteins are populated at equilibrium, the folding process usually involves association of partially or completely structured monomers (*1*, *2*).

Legume lectins have attracted interest because of their high variability in quaternary association of the same monomer scaffold, a β-sandwich (*3*). No monomeric legume lectin structure has been reported, but there are different kinds of dimers and tetramers (*4*, *5*). The most frequently observed mode of association is that of the "canonical dimer" (*4*), where the back sheets of both subunits form a contiguous 12-stranded sheet. The legume lectin monomer scaffold is a common fold in only distantly related proteins, some of them monomeric (*6*), and unfolding via a monomeric intermediate has been described (*7*).

In some legume lectins, among them pea seed lectin (Psl[1]), each subunit is post-translationally cleaved into a short α-chain ($\approx$ 52 residues) and a longer β-chain (181 residues). As a consequence, the structure of the conserved "monomer" is in fact formed by two polypeptide chains and is itself a dimer. Thus, a hierarchy of subunits is created, because the processed chains forming the conserved subunit are more closely related in terms of structure than the two pairs that associate at the "canonical" dimer interface. The structure of the processed protein (see Figure 1) shows that the α- and β-chains are closely interdigitated in the structure of one subunit. The interface to the other subunit in the so-called "canonical dimer" is composed almost exclusively of the β-chain.

The physiological role of processing is not known (*4*). Because of the interdigitation, the fragment chains cannot attain their native fold in the absence of their partner. Therefore, the folding of the domains would be expected to be coupled to the association of the fragments and to precede domain association, as has been described for example for arc repressor (*8*). On the other hand, due to the complex topology, it would not be very suprising if the processed form was not able to fold at all. This would pose a built-in limit on the lifetime of Psl in the seeds during maturation and storage determined by the rate of unfolding, which is different from zero even for very stable proteins. We here show that processed Psl is in fact able to refold in vitro, albeit with moderate yields limited by off-pathway aggregation. Furthermore, our data indicate that the expected hierarchy of structure formation is not followed in folding of Psl. The formation of the "canonical" dimer interface upon association of two β-chains precedes the formation of the intact β-sandwich domain. We propose that the association of the α-chain with the β-chain dimer involves insertion of parts of the α-chain into preformed β-sheets, as has been described for some protein−protein complexes (*9*).

MATERIALS AND METHODS

Material. Dry green pea seeds were purchased from a local food store. GdmCl and urea were from ICN (Ohio), and all other chemicals were of highest purity available. Ultrapure water (Millipore ion exchanger) was used in all experiments.

† Supported by the German Research Foundation (DFG, Grant Se517-17 to R.S.).

* To whom correspondence should be addressed: Universität Potsdam, Physikalische Biochemie, Karl-Liebknecht-Str. 24-25, Haus 25, D-14476 Potsdam-Golm, Germany. Tel: +49-331-977-5240. Fax: +49-331-977-5062. E-mail: seckler@uni-potsdam.de.

‡ Present address: Eyesense GmbH, Stockstädter Str. 17, 63762 Grossostheim, Germany.

[1] Abbreviations: Psl, pea seed lectin; PslP, mature processed Psl; PslS, recombinant single-chain Psl; GdmCl, guanidinium chloride; MOPS, 3-(*N*-morpholino)-propanesulfonic acid.

10.1021/bi7019047 CCC: $40.75
Published on Web 01/24/2008

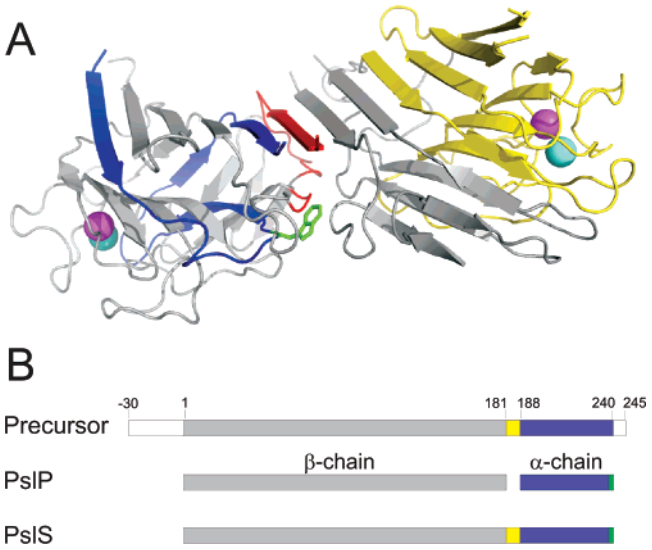

FIGURE 1: Structure of pea seed lectin. A: Three-dimensional structure of the canonical Psl-dimer (PDB entry 1bqp). In the left subunit, the α-chain is colored blue, the two segments of the *β*-chain involved in contacts to the other subunit are colored red, and the side-chain of Trp206 is colored green. In the other subunit, the parts corresponding to the minimal region according to (*6*) are colored yellow (residues 62−181 of the *β*-chain and 188−196 of the α-chain, see Discussion). The flat *back sheet* is located on top of the structure and forms a contiguous 12-stranded sheet over the dimer interface. The curved *front sheet* (7 strands per subunit) is at the bottom, and the small 4-stranded *top sheet* is in the back in the left subunit and pointing to the viewer in the right one. The metal ions are shown as spheres (Ca, cyan; Mn, magenta). The figure was prepared using Pymolpt (*44*). B: Primary structures of PslP and PslS. The polypeptide chain segment (residues 182−187) connecting the *β*- and α-chains in the Psl precursor and in PslS is indicated in yellow, and the carboxy-terminal heterogeneity distinguishing Psl isoforms is indicated in green.

Sephadex G100 superfine was from Pharmacia (Uppsala, Sweden).

Purification of PslP. PslP was purified following published procedures (*10*, *11*) with minor modifications. A fine meal was prepared from 500 g of dry pea seeds and extracted with water overnight at 4 °C. Acid precipitation and affinity chromatography were carried out as described (*11*). Mature PslP consists of two isoforms (see Results). On ion exchange materials used previously (*10*, *11*), we did not observe complete separation of the isoforms. Therefore, after dialysis against 20 mM Tris/HCl, pH 8.8, residual impurities were removed and the isoforms were separated on a Source 15Q anion exchange column (1.7 mL bed volume, flow 1 mL/min), using a linear gradient from 0−200 mM NaCl.

Preparation of PslS. The fragment coding for Thr 1 to Gln 241 of Psl (with the termini corresponding to the isoform B of PslP (*12*)) was amplified from pMP2809 (*13*), ligated into the expression plasmid pET21a (*14*), and *Escherichia coli* BL21 (DE3) cells were transformed with the resulting vector.

For purification, a protocol adapted from (*15*), (*16*) and (*13*) was used. Bacteria were grown in LB full medium (5 L), induced for protein production at an optical density of 1 at 550 nm, and cells were harvested by centrifugation after overnight incubation at 30 °C. The frozen pellets were resuspended in 20 mL of 10 mM Tris/HCl, 150 mM NaCl, 0.5 M Pefabloc SC, pH 6.8, and cells were disrupted by high-pressure lysis. The insoluble fraction was suspended in 30 mL of solubilization buffer (8 M GdmCl, 10 mM Tris, pH 6.8) and gently shaken overnight at 4 °C. The supernatant was diluted 1:200 with ice-cold refolding buffer (10 mM Tris/HCl, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 6.8) containing 1.5 M urea and stirred overnight at 4 °C. The renatured protein was concentrated by ammonium sulfate precipitation and further purified by affinity chromatography as described (*11*). Purity of PslP and PslS was tested by SDS gel electrophoresis with silver staining. The purified proteins were stored at −70 °C.

Separation of α- and β-Chains. The processed chains were separated by preparative size-exclusion chromatography on a Superdex 75 column (diameter 1 cm, length 30 cm, Pharmacia) under denaturing conditions. Up to 0.5 mg of isoform B, denatured as described below, was applied to the column using 10 mM MOPS buffer at pH 6.8 with 3 M GdmCl as eluent, and the fragments eluted as baseline-separated peaks.

Spectroscopy. The extinction coefficient at 280 nm was determined by the method of Pace et al. (*17*) to be 36000 M^{-1} cm^{-1} for PslS and both isoforms of PslP. The extinction coefficient of isolated chains at 3 M GdmCl was taken to be equal to the value calculated at 6 M according to (*18*). All molar concentrations given refer to a single subunit in PslS and the respective structural entity, the processed subunit, in PslP; concentrations of α- or *β*-chains refer to monomeric chains. Fluorescence measurements were carried out in a Spex Fluoromax II with a thermostated cell holder and stirred cells, with the excitation at 280 ± 3 nm and emission at 315 ± 5 nm. Except for some fast kinetics, single measurements were at least 7 s apart and the slits were closed between measurements to reduce photobleaching. The concentration of GdmCl was calculated from the refractive index increment (*19*).

Hemagglutination Assay. The assay was carried out according to (*13*) and (*16*). In a microtiter plate with 96 round-bottomed wells, a series of 2-fold dilutions of the tested protein solution in PBS (10 mM sodium phosphate, 150 mM NaCl, pH 6.8) in a final volume of 250 *μ*L was prepared. Then 25 *μ*L of a 10% (v/v) suspension of human AB+ erythrocytes in PBS was added, and the plates were incubated for 1.5−2 h at room temperature and analyzed. In wells with sufficient concentration of Psl, the protein cross-links the cells, and the cells form a thin layer over the whole cross-section of a well, whereas they sediment to the middle of the round-bottom if the dilution is too high. Often there is one well that shows partial agglutination. In this case the series was counted as showing $N + 0.3$ or $N + 0.7$ agglutinated wells, meaning there was "some agglutination" detectable in a sedimented well or "some sedimentation" in an agglutinated well, respectively.

From each sample analyzed, at least 3 series of 2-fold dilution were prepared, and a mean N of the number of agglutinated wells was calculated. The concentration was calculated as $\ln c = (N + a) \ln 2$, where a is determined from controls of known concentration, analyzed in the same experiment.

Analytical Size-Exclusion Chromatography. Using a 50 *μ*L sample loop, renaturing samples where injected onto a TSK-Gel G2000SW XL300 column (5 *μ*m particle size, dimensions 7.8 × 300 mm, Toso Haas, Stuttgart, Germany) with a guard column. Chromatography was carried out at about

15 °C with refolding buffer at a flow rate of 0.9 mL/min. A Merck Hitachi L-7580 fluorescence detector was used at an excitation wavelength of 280 nm and emission wavelength of 320 nm. Peak areas were calculated using PeakFit v4.06 (SPSS Inc.)

Denaturation and Renaturation. To prepare unfolded stock solutions, a concentrated solution of native protein (≈10 mg/mL) was denatured by 1:4 dilution into denaturing buffer (8 M GdmCl, 10 mM MOPS, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 6.8) and incubated at least 5 h at room temperature. Samples of denatured protein at lower GdmCl concentrations were prepared by dilution from this fully unfolded protein stock solution, if needed. For unfolding and refolding transitions, aliquots of native and denatured protein solutions were added to 2 mL of buffer containing the desired GdmCl concentration thermostated at 10 °C in Eppendorf cups.

Refolding Kinetics. For fluorescence measurements of renaturation kinetics, a denatured stock solution was pipetted into the stirred thermostated fluorescence cell containing refolding buffer and the desired concentration of GdmCl.

Otherwise, renaturation was initiated in 50 mL plastic tubes with the thermostated refolding buffer. For size-exclusion chromatography and double jumps, samples were taken and directly used in the test. For the hemagglutination assay, 45 mL was concentrated to 1−2 mL using Centriplus YM30 centrifuge tubes (Millipore, Bedford, MA). This took about 12 h. Refolding buffer was added to get an equal volume for each sample, and aliquots were used in the hemagglutination assay (see above).

For double jumps, samples (625 μL) were taken after the refolding times indicated and diluted into 1875 μL of acid urea (≈ 9 M urea, adjusted to pH 1.8 with HCl) thermostated at 10 °C in the fluorescence cell. Unfolding kinetics were monitored for 7 min, and the amplitudes of monoexponential decays fitted to the data were taken as a measure of renatured protein, relative to a native control. Refolded protein unfolded with rates identical to the native control, and no other kinetical phases were observed. This indicates that any folding intermediates unfolded in the dead time and did not contribute to the measured amplitudes. Chemical cross-linking with glutaraldehyde was performed as described in (*20*), using 1.5 mL aliquots from the refolding sample. The precipitate was resuspended in alkaline gel loading buffer containing 0.6 M Tris/HCl pH 8.8 and analyzed on SDS gels.

RESULTS

Mature Psl exists in two isoforms, discriminated by the presence or absence of lysine 240 at the C-terminus (*12*, *21*). To be able to work with a defined preparation, we have separated the two isoforms. We followed published procedures for the purification of Psl from pea seeds (*10*, *11*), but used a high-resolution anion exchange column in the final purification step. Psl eluted in three well-separated peaks, the first corresponding to isoform B, the third to isoform A and the middle one containing mixed dimers of both isoforms (F.K. and R.S., unpublished observations). For all experiments in this paper, pure isoform B was used.

In order to investigate the effect of the processing on the folding process, we studied the properties of the isolated fragments separated by gel filtration under denaturing conditions. In addition, we used a recombinant, single-chain form that resembles isoform B of complete Psl (with Lys 240 present) except that there is no internal processing. We abbreviate this recombinant, single-chain variant as PslS and the mature, processed form as PslP.

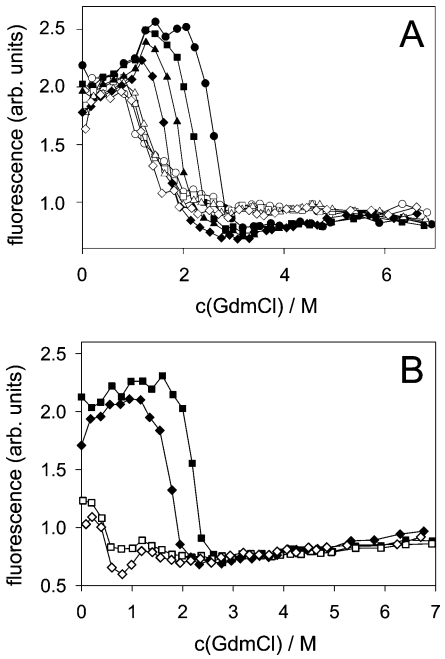

FIGURE 2: A: Comparison of GdmCl-induced unfolding and refolding transitions of recombinant PslS at a protein concentration of 10 μg/mL (0.39 μM) after 1 day (circles), 1 week (squares), 1 month (triangles) and 3 months (diamonds). Filled symbols denote denaturation, open symbols renaturation. B: Transitions for processed PslP under the same conditions, after 1 week and 3 months, symbols as in A.

Unfolding and Refolding Transitions. To investigate the stability of PslS and PslP, we measured the intrinsic protein fluorescence of samples incubated at varied concentrations of the denaturant GdmCl. Figure 2A shows unfolding and refolding transitions of PslS measured after different incubation times, at a protein concentration of 0.39 μM (10 μg/mL), at 10 °C. This temperature was used for all equilibrium and kinetic measurements to reduce aggregation. Denaturation (filled symbols) was exceedingly slow in the transition region around 2−3 M GdmCl, and samples did not reach equilibrium even after 3 months of incubation. At lower protein concentrations, denaturation was marginally accelerated (data not shown). This suggests that dissociation to monomers precedes unfolding, and the small fraction of monomeric molecules in the transition region slows down unfolding.

For PslS complete renaturation was possible at residual GdmCl concentrations below 1 M (cf. Figure 2A). At intermediate denaturant concentrations, however, the renaturation and denaturation curves did not coincide, probably because of the slow unfolding. Unfolding was somewhat faster at higher temperatures (up to 4-fold at 30 °C), but refolding yields decreased strongly with increasing temperature. Thus, evaluation of the thermodynamic stability of PslS was not possible on an experimentally accessible time scale (*22*).

Denaturation transition curves of the processed form, PslP, coincided with PslS data at the same times (Figure 2B).

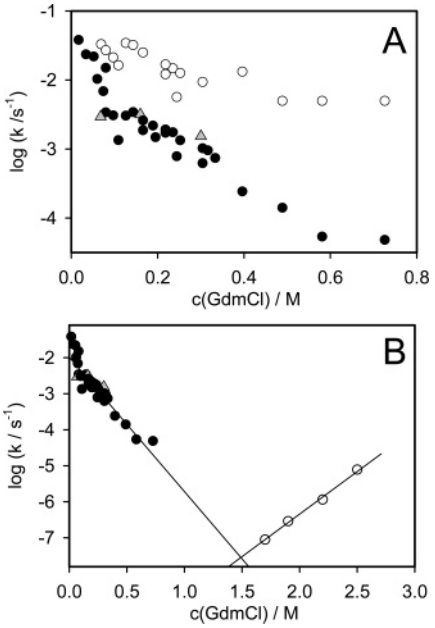

FIGURE 3: A: Rate constants of PslS refolding as measured by fluorescence at 10 °C and 0.39 μM (10 μg/mL) protein in refolding buffer. Filled circles: slow phase or (if monophasic) single phase. Open circles: fast phase. Gray triangles: formation of native protein measured with double-jumps. B: Rate constants of the slow phase of refolding redrawn from panel A together with rate constants of unfolding (open circles) as estimated from the midpoints of the unfolding transitions depicted in Figure 2A. Note that extrapolation lines meet at the midpoint of the refolding transition (see Figure 2A).

Renaturation, however, was incomplete even under conditions strongly favoring the folded form. The yield was around 25−30% at the lowest GdmCl concentrations. The drop in signal around 0.75 M is indicative for aggregation competing with folding, and in fact formation of aggregates was observed in samples renatured under these conditions.

Renaturation of PslS. To get insight into the folding mechanism of the recombinant form, PslS, we have monitored refolding kinetics by tryptophan fluorescence intensity. Observed rates of fluorescence change were independent of protein concentration between 2.5 and 30 μg/mL and thus were fit to exponential decays. Refolding was at least biphasic. The fast phase had a small positive amplitude at low and a small negative amplitude at higher concentrations of GdmCl. Rate constants obtained by fitting the kinetic traces are plotted in Figure 3. In some cases the fast phase could not be detected at concentrations where the sign of its amplitude changes (see overlap of filled and open circles around 0.1 M in Figure 3A). Kinetics measured with double jumps which directly monitor formation of native protein (see below) are monophasic and have the same rate constants as the slow phase (gray triangles in Figure 3). Thus, native protein is only formed in the slow phase, and the fast phase corresponds to formation of a folding intermediate.

Under certain conditions, aggregation was observed to compete with PslS refolding, thus interfering with the folding kinetics observed by fluorescence. At GdmCl concentrations above 0.25 M, no endpoint was reached, because at late time points the fluorescence decreased again through aggregation. Even at very low denaturant concentrations, the signal did not reach native values. This is in contrast to the results of the fluorescence transitions and may be due to the effects of stirring in the fluorescence cells or interactions with the walls of the quartz vessels. However, a native value for the fluorescence emission maximum was reached and the protein renatured under these conditions exhibited full specific hemagglutination activity (data not shown). This and the matching rates from fluorescence and double jump measurements show that native protein was formed.

The rates for the formation of native PslS correspond to half-times ranging from 1 min at 0.06 M GdmCl to 4 h at 0.6 M. This is slow compared to other proteins of comparable size (*1*). The fact that renaturation rates were independent of protein concentration indicates a conformational change rather than association of the canonical dimer is rate limiting for PslS refolding. This is also supported by comparing rates of folding and unfolding (Figure 3B): at the midpoint of the refolding transition (1.5 M GdmCl), the extrapolated rate of the slow folding phase and the extrapolated rate of unfolding estimated from the midpoints of the unfolding transitions are identical within experimental error.

Renaturation of PslP. Monitoring the refolding of PslP by fluorescence was not feasible even after extensive optimization of refolding conditions and at very low residual concentrations of GdmCl. Instead, under refolding conditions, and with the mixing and stirring necessary for technical reasons, the fluorescence signal was dominated by aggregation. Therefore, to elucidate the folding mechanism of processed PslP, double-jump measurements were employed to monitor the extent of renaturation: A denatured protein solution was allowed to refold for a certain amount of time and then transferred back to unfolding conditions. The observable amplitude of denaturation then corresponds to the amount of native protein accumulated during the refolding time (*23*), because any folding intermediates unfold in the dead time of mixing. To validate that indeed only native protein is quantified with this technique, we compared the results with a test for functional, sugar-binding Psl. The physiological binding partner of Psl is not known, and exploring its specificity toward monosaccharides, small oligosaccharides and modified sugars has not led to the discovery of a high affinity binding partner (*11, 24*−*28*). The fact, however, that Psl binds to surface oligosaccharides of red blood cells and agglutinates them (i.e., cross-links and precipitates them) has been used as a test of its activity (*13, 16*). Unfortunately, this test had a low time resolution of about 1 day when applied to the refolding experiments, because it was necessary to concentrate the sample from the dilute refolding solutions. Therefore only late time points in the refolding kinetics of PslP could be resolved, and refolding kinetics of PslS could not be followed by this method.

Figure 4A shows data obtained with both methods. Refolding of PslP was exceedingly slow, being completed only after almost one week. Interestingly, some native protein was formed very fast and was present even at the shortest renaturation time tested, 20 s. It can also be seen that the results from double-jump experiments and the hemagglutination assay agree well, with the agglutination assay showing more variation in the data.

To deduce a kinetic model for folding and association of PslP, we tried to describe the data quantitatively by appropriate functions. The data could be fitted quite well with a sum of two exponentials with rates of 11 day^{-1} and 0.25 day^{-1}, respectively (solid line in Figure 4A). A simple

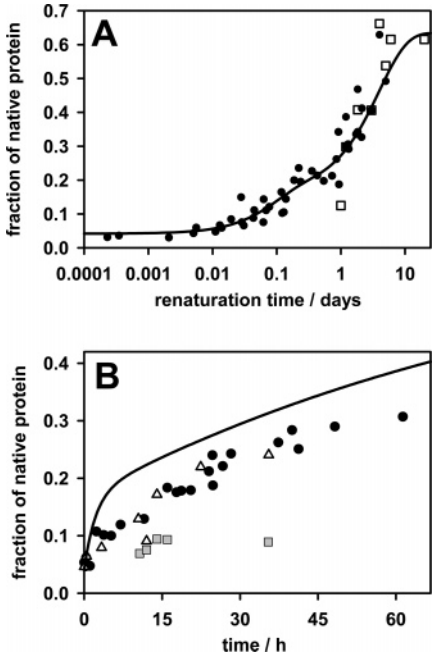

FIGURE 4: A: Renaturation of 10 μg/mL (0.39 μM) PslP in renaturation buffer at 10 °C and 0.2 M GdmCl. Filled circles are data from double-jump experiments, open squares from the hemagglutination assay. The line is a fit to a sum of two exponentials. B: Refolding of PslP at different polypeptide chain concentration ratios. A mixture of 0.78 μM α-chains and 0.39 μM β-chains (filled circles) or 0.39 μM α-chains and 0.78 μM β-chains (open triangles) were renatured at 10 °C at 0.4 M residual GdmCl concentration. A renaturation experiment at an α:β ratio of 1:1 (0.39 μM each) at 0.4 M GdmCl is shown for comparison (gray squares). The solid line is the two-exponential fit from A.

hyperbolic curve characteristic for a single, rate-limiting second-order association reaction did not fit the data well. Because, however, processed Psl is in fact a tetramer, more than one folding intermediate may be involved and more complex kinetics might be expected. Fitting to any more complex model would increase the number of parameters and thus improve the goodness of fit just for mathematical reasons, without providing evidence for the correctness of the chosen model. Thus, it was necessary to vary the concentration of one or both reactants.

Measuring refolding kinetics of PslP at different protein concentrations proved to be difficult. Any significant increase in protein concentration (e.g., from 10 to 20 μg/mL) resulted in a strong decrease in refolding yield due to aggregation. Below 10 μg/mL, on the other hand, the signal-to-noise ratio in the final unfolding jump was too low. Better renaturation yields were obtained by increasing only the concentration of one of the two types of chains.

Kinetics obtained that way are shown in Figure 4B. Preparation of the isolated chains under denaturing conditions in 3 M GdmCl yielded them in concentrations of at most 14 μM for the α-chain and 6 μM for the β-chain which elutes in a broader peak. Therefore the renaturation experiments had to be performed at a higher residual concentration of GdmCl of 0.4 M. Under these conditions, 1:1 mixtures of the chains renatured with a bad yield so that kinetics could hardly be measured (gray squares in Figure 4B, compare the beginning drop in yield in Figure 2). In the 2:1 or 1:2 mixtures, however, renaturation could still be achieved with considerable yield even at these concentrations of GdmCl

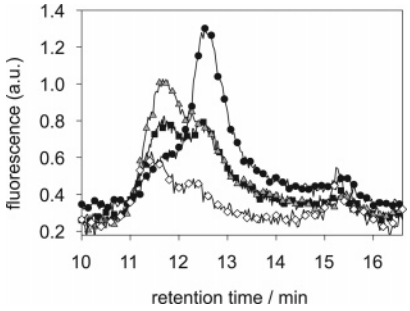

FIGURE 5: Size-exclusion chromatograms of the isolated β-chain (0.39 μM) in refolding buffer, at 0.2 M GdmCl and 10 °C. The samples were injected onto the column immediately after the initiation of refolding (line with filled circles), after 60 min (filled squares), 90 min (gray triangles) and 3 h (open diamonds).

and even though the total protein concentration was increased compared to the 1:1 control (Figure 4B). This is consistent with an association reaction competing with off-pathway aggregation.

Properties of Isolated Chains under Refolding Conditions. To further investigate the mechanism of refolding of PslP and to identify folding intermediates, we used size-exclusion chromatography of refolding samples and of the isolated chains in conjunction with fluorescence spectroscopy. Size exclusion is well suited to monitor slow association events as observed with PslP.

Native PslP eluted from the column in one single peak at 11.5 min. Samples of the isolated α-chain showed no sign of structure formation under refolding conditions. Its fluorescence emission maximum was at 355 nm, characteristic for fully solvent-exposed tryptophan, and it eluted from the size-exclusion column in one peak at about 14 min. The peak slowly decreased with time after dilution from the denaturant, indicating aggregation, but the majority of the α-chains remained soluble for nearly 1 h at 0.39 μM.

The isolated β-chain behaved differently. Immediately after dilution, its fluorescence emission maximum shifted from 355 to 346 nm, indicating structure formation. Its elution pattern from the size-exclusion column slowly changed with time after dilution (Figure 5). Upon injection immediately after transfer to folding conditions, one major peak at 12.5 min was observed. It had a shoulder at 11.5 min which grew and was the main peak after 90 min of renaturation. It was only after hours that the amount of soluble polypeptide decreased by aggregation. Whereas the molecules in peak at 12.5 min were clearly smaller than native Psl, the growing peak eluted at the same retention time as native protein.

The simplest explanation for the observed changes in fluorescence and in the Stokes radius is the association of two structured β-chains to a dimeric intermediate with a native-like interface but a less compact packing.

Cross-Linking Experiments. Because of the limited solubility of the β-chain in the absence of denaturant, we could not determine its secondary structure by CD-spectroscopy. However, the formation of a homodimer of the isolated β-chains could be confirmed by chemical cross-linking. As a control, native PslS was cross-linked with glutaraldehyde and analyzed by SDS gel electrophoresis. Only dimers and no higher oligomers were observed, indicating that the reagent cross-linked only polypeptides specifically associated to dimers and not proteins meeting by diffusion.

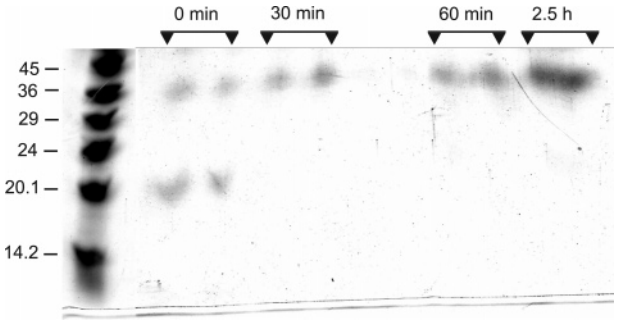

FIGURE 6: Cross-linking of β-chains under native conditions. The peptide was transferred to refolding buffer (at 0.39 μM, 10 °C, 0.2 M GdmCl) and cross-linked by glutaraldehyde after the times indicated. The first lane is a molecular weight marker.

To test for homodimerization of the β-chain of PslP, we transferred it to renaturing conditions and cross-linked samples after different "renaturation" times. As shown in Figure 6, we indeed observed dimers of the β-chain. After 30 min and at later time points, there were only dimers. Since no higher oligomers were observed, dimerization must be a specific process. We propose that it involves the native interface of the canonical dimer.

Size-Exclusion Chromatography of Refolding PslP. In order to investigate whether this β2-intermediate is also present during refolding of PslP, we also subjected samples of complete PslP to size-exclusion chromatography under refolding conditions. The first chromatogram was similar to a superposition of the respective chromatograms of the isolated chains (Figure 7A, 0 min). Later, the shoulder at 11.5 min, the retention time of both the β-chain dimer and the native PslP, became the prominent peak. The increase was much more pronounced than that observed with the isolated β-chain, and its rate was only slightly faster than the formation of native PslP (see Figure 7B). Concomitant with the increase of the peak at 11.5 min, the peak of the α-chain at 14 min decreased.

Taken together, size-exclusion chromatography and cross-linking data suggest that the native structure of PslP is formed by association of α-chains with a dimeric folding intermediate of the β-chain.

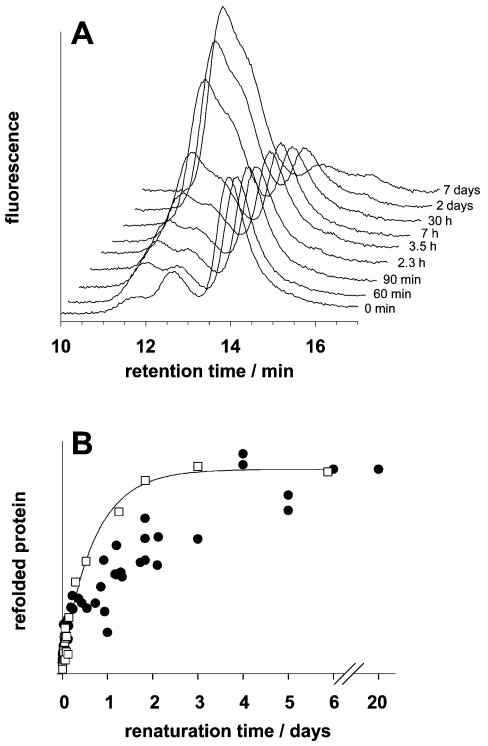

FIGURE 7: A: 3-D-plot of size-exclusion chromatograms of refolding PslP. The *x*-axis is the retention time on the column, and the *z*-axis is the time between onset of renaturation and the injection on the column. Refolding times are indicated on the right side of each curve. The peak at 14 min corresponds to the α-chain. B: Comparison of refolding kinetics monitored by size-exclusion chromatography and formation of native protein. Filled circles are all the data from Figure 4A, and open squares are absolute peak areas of the peak at 11.5 min, all scaled arbitrarily. The solid line is a monoexponential fit to the peak areas.

DISCUSSION

Stability of Psl. There has been ongoing debate on whether it is possible to determine the thermodynamic stability of legume lectins. Thermodynamic analysis of differential scanning calorimetry data of legume lectins (*29*−*32*) has been criticized because the thermal denaturation of these proteins is irreversible (*33*), cf. also (*34*). Denaturant-induced unfolding transitions of legume lectins have been published (*7*, *35*), but in no case has it been shown unambiguously that an unfolding−refolding equilibrium was reached. Ahmad et al. (*36*) have investigated urea-induced unfolding transitions of PslP at different protein concentrations and temperatures and have evaluated them according to a two-state equilibrium model, although only unfolding transitions, and no refolding transitions, were measured and only partial reversibility of unfolding was observed at a single, low denaturant concentration. In the present study using GdmCl as the denaturant, we have observed slow and incomplete refolding of the native, proteolytically processed PslP and extremely slow unfolding in the transition region (see Figure 2). It is clear that such data cannot be interpreted thermodynamically, because they represent time points in slow unfolding kinetics rather than unfolding−refolding equilibrium transitions. When using urea as the denaturant, we found even slower unfolding, and attempts to speed up denaturation by addition of EDTA to remove the cofactor ions Ca^{2+} and Mn^{2+} were not successful (data not shown). In addition, cyanate accumulates in urea solutions and chemically modifies proteins at near neutral pH values on the time scale required for PslP unfolding (*37*). Taken together, the results suggest that the thermodynamic analysis published by Ahmad et al. (*36*) should be reconsidered.

Refolding and Association of PslP. The results above prove that the isolated β-chain of mature, proteolytically processed pea seed lectin forms dimers. During reassembly of the $\alpha_2\beta_2$-heterotetramer, dimerization of the β-chains appears to precede the addition of the α-chains. The fact that only β-dimers are observed, but no higher oligomers, implies that this association is specific. As subunit association in the native protein is also mediated by the β-chains (see below), the specificity of dimerization is likely achieved by a native-like interaction between two β-chains. What could such a dimer look like?

In Figure 1A the structure of PslP is shown with the α-chain in blue. It contributes two strands to the contiguous *back sheet* (on top in the figure), one near the subunit

interface and one at the outer end of the sheet. To the *front* and *top sheets*, the α-chain contributes one strand each.

Chandra et al. (*6*) have analyzed structures of legume lectins as well as of members of many other protein families with the legume lectin fold. They all have the typical *jelly roll* fold with two major and one small sheet, but different numbers of strands in each sheet and different insertions in the loops. The authors identified the strands that are common to all members of the superfamily, and they termed this part the "minimal determining region" of the fold. The functions of the proteins in this superfamily are quite diverse, although most of them are known to bind saccharides. Sugar binding is mainly performed by the loops decorating the structure. Therefore it is likely that the common minimal region represents the structural scaffold onto which the different functions are build (*6*), and that it has evolved mainly for stability and foldability.

The strands of Psl belonging to this minimal region are shown in yellow in Figure 1. Except for the N-terminal strand of the α-chain at the outer end of the *back sheet*, this minimal region is composed solely of the *β*-chain. Our data seem to indicate that this major part of the *β*-chain can fold into a native-like conformation even without the missing strand.

Brinda et al. (*38*) have analyzed structure graphs of legume lectins. They identified clusters of amino acids that interact at a certain level of interaction strength. In this picture, clusters are formed by residues that interact with many other side chains in their neighborhood. They found that in the canonical dimer, the largest cluster of residues that stabilize the structure, is located in the interface and formed by residues from either monomer. This points to an important role of this interface in stabilizing the monomer structure, and makes it likely that this interface can be formed even if other parts of the monomer are not yet correctly structured.

The parts of the protein that form contacts to the other subunit are colored red in Figure 1. This interface is composed mainly of side chains and backbone from the carboxy-terminal section of the two *β*-chains, whereas the only contribution of the α-chains is from the side chain of Trp90 (Trp206 according to the numbering in 2bqp, Trp88 in the ConA sequence). Although Trp90 is a central "hub" residue within the interface (*38*), it is the carboxy-terminal section of the *β*-chain which forms most contacts within the subunit interface. The contacting residues are located in two sequence stretches, separated by 30 residues folded into four strands with the topology of an up-and-down beta-barrel that might form an independently folding subdomain.

We propose that this C-terminal interface region is native-like in the *β*-chain dimer and stabilized by homodimerization. The two subdomains—the minimal region and the subunit interface—have coalesced to form contiguous *β*-sheets lacking the strands that are contributed by the α-chain in the native structure. For the α-chain to associate with the β_2 dimer, it has to insert into the preformed sheets requiring dissociation of the subdomains of one *β*-chain or the stepwise insertion of the α-chain into each of the sheets.

Strand insertion into preformed *β*-sheets is central to the function of the serpin class of protease inhibitors (*39*, *40*) and to the evolution and assembly of bacterial pilus proteins (*41*, *42*) and is suggested by several crystal structures of heteromeric protein complexes (*9*). Upon cleavage by the protease, the serpin active-site loop inserts into a mixed parallel−antiparallel *β*-sheet which thus achieves a regular antiparallel structure (*39*). The same holds for each of the three sheets in PslP when the α-chain is inserted (see Figure 1), even though the insertion reaction occurs intramolecularly in the case of the serpins, whereas is requires chain association in the assembly of PslP. The insertion process in serpins occurs on the seconds time scale (*43*), much slower than other protein conformational changes. A requirement for three strands to insert into preformed sheets would readily explain the exceptionally slow folding of PslP.

ACKNOWLEDGMENT

We thank J. W. Kijne for providing the plasmid pMP2809, U. Baxa for helpful discussion, and J. Kramer, D. Hobson and C. Pankoke for technical assistance.

REFERENCES

1. Jaenicke, R. (1987) Folding and association of proteins, *Prog. Biophys. Mol. Biol. 49*, 117−237.
2. Lilie, H., and Seckler, R. (2005) Folding and association of multidomain and oligomeric proteins, in *Protein Folding Handbook* (Buchner, J., and Kiefhaber, T., Eds.) pp 32−72, Wiley VCH, Weinheim, Germany.
3. Srinivas, V. R., Reddy, G. B., Ahmad, N., Swaminathan, C. P., Mitra, N., and Surolia, A. (2001) Legume lectin family, the 'natural mutants of the quaternary state', provide insights into the relationship between protein stability and oligomerization, *Biochim. Biophys. Acta 1527*, 102−111.
4. Loris, R., Hamelryck, T., Bouckaert, J., and Wyns, L. (1998) Legume lectin structure, *Biochim. Biophys. Acta 1383*, 9−36.
5. Bettler, E., Imberty, A., and Loris, R. (1998) The 3D lectin database, URL: http://www.cermav.cnrs.fr/glyco3d.
6. Chandra, N. R., Prabu, M. M., Suguna, K., and Vijayan, M. (2001) Structural similarity and functional diversity in proteins containing the legume lectin fold, *Protein Eng. 14*, 857−866.
7. Reddy, G. B., Srinivas, V. R., Ahmad, N., and Surolia, A. (1999) Molten globule-like state of peanut lectin monomer retains its carbohydrate specificity. Implications in protein folding and legume lectin oligomerization, *J. Biol. Chem. 274*, 4500−4503.
8. Srivastava, A. K., and Sauer, R. T. (2000) Evidence for partial secondary structure formation in the transition state for arc repressor refolding and dimerization, *Biochemistry 39*, 8308−8314.
9. Remaut, H., and Waksman, G. (2006) Protein-protein interaction through beta-strand addition, *Trends Biochem. Sci. 31*, 436−444.
10. Entlicher, G., Kostir, J. V., and Kocourek, J. (1970) Studies on phytohemagglutinins. 3. Isolation and characterization of hemagglutinins from the pea (Pisum sativum L.), *Biochim. Biophys. Acta 221*, 272−281.
11. Trowbridge, I. S. (1974) Isolation and chemical characterization of a mitogenic lectin from Pisum sativum, *J. Biol. Chem. 249*, 6004−6012.
12. Hoedemaeker, F. J., Richardson, M., Diaz, C. L., de Pater, B. S., and Kijne, J. W. (1994) Pea (Pisum sativum L.) seed isolectins 1 and 2 and pea root lectin result from carboxypeptidase-like processing of a single gene product, *Plant Mol. Biol. 24*, 75−81.
13. van Eijsden, R. R., Hoedemaeker, F. J., Diaz, C. L., Lugtenberg, B. J., de Pater, B. S., and Kijne, J. W. (1992) Mutational analysis of pea lectin. Substitution of Asn125 for Asp in the monosaccharide-binding site eliminates mannose/glucose-binding activity, *Plant Mol. Biol. 20*, 1049−1058.
14. Dubendorff, J. W., and Studier, F. W. (1991) Cloning and expression of the gene for bacteriophage T7 RNA polymerase under control of its cognate promotor, *J. Mol. Biol. 219*, 61−68.
15. Stubbs, M. E., Carver, J. P., and Dunn, R. J. (1986) Production of pea lectin in Escherichia coli, *J. Biol. Chem. 261*, 6141−6144.
16. Prasthofer, T., Phillips, S. R., Suddath, F. L., and Engler, J. A. (1989) Design, expression, and crystallization of recombinant lectin from the garden pea (Pisum sativum), *J. Biol. Chem. 264*, 6793−6796.
17. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci. 4*, 2411−2423.

18. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
19. Nozaki, Y. (1972) The preparation of guanidine hydrochloride, *Methods Enzymol. 26*, 43−50.
20. Rudolph, R., Böhm, G., Lilie, H., and Jaenicke, R. (1997) Folding Proteins, in *Protein function* (Creighton, C. E., Ed.) pp 57−99, Oxford University Press, Oxford, U.K.
21. Higgins, T. J., Chrispeels, M. J., Chandler, P. M., and Spencer, D. (1983) Intracellular sites of synthesis and processing of lectin in developing pea cotyledons, *J. Biol. Chem. 258*, 9550−9552.
22. Pace, C. N., and Scholtz, J. M. (1997) Measuring the conformational stability of a protein, in *Protein structure* (Creighton, C. E., Ed.) pp 299−321, Oxford University Press, Oxford, U.K.
23. Kiefhaber, T. (1995) Protein folding kinetics, *Methods Mol. Biol. 40*, 313−341.
24. Van Wauwe, J. P., Loontiens, F. G., and De, Bruyne, C. K. (1975) Carbohydrate binding specificity of the lectin from the pea (Pisum sativum), *Biochim. Biophys. Acta 379*, 456−461.
25. Schwarz, F. P., Puri, K. D., Bhat, R. G., and Surolia, A. (1993) Thermodynamics of monosaccharide binding to concanavalin A, pea (Pisum sativum) lectin, and lentil (Lens culinaris) lectin, *J. Biol. Chem. 268*, 7668−7677.
26. Küster, F. (2002) Das Lektin aus der Erbse *Pisum sativum*: Bindungsstudien, Monomer-Dimer-Gleichgewicht und Rückfaltung aus Fragmenten. Doctoral thesis, 130 pp, Universität Potsdam, Potsdam, Germany.
27. Schwarz, F. P., Misquith, S., and Surolia, A. (1996) Effect of substituent on the thermodynamics of D-glucopyranoside binding to concanavalin A, pea (Pisum sativum) lectin and lentil (Lens culinaris) lectin, *Biochem. J. 316*, 123−129.
28. Schlick, K. H., Udelhoven, R. A., Strohmeyer, G. C., and Cloninger, M. J. (2005) Binding of mannose-functionalized dendrimers with pea (Pisum sativum) lectin, *Mol. Pharmaceutics 2*, 295−301.
29. Reddy, G. B., Bharadwaj, S., and Surolia, A. (1999) Thermal stability and mode of oligomerization of the tetrameric peanut agglutinin: a differential scanning calorimetry study, *Biochemistry 38*, 4464−4470.
30. Srinivas, V. R., Singha, N. C., Schwarz, F. P., and Surolia, A. (1998) Differential scanning calorimetric studies of the glycoprotein, winged bean acidic lectin, isolated from the seeds of Psophocarpus tetrogonolobus, *FEBS Lett. 425*, 57−60.
31. Surolia, A., Sharon, N., and Schwarz, F. P. (1996) Thermodynamics of monosaccharide and disaccharide binding to Erythrina corallodendron lectin, *J. Biol. Chem. 271*, 17697−17703.
32. Schwarz, F. P., Puri, K., and Surolia, A. (1991) Thermodynamics of the binding of galactopyranoside derivatives to the basic lectin from winged bean (Psophocarpus tetrogonolobus), *J. Biol. Chem. 266*, 24344−24350.
33. Marcos, M. J., Chehin, R., Arrondo, J. L., Zhadan, G. G., Villar, E., and Shnyrov, V. L. (1999) pH-dependent thermal transitions of lentil lectin, *FEBS Lett. 443*, 192−196.
34. Kurganov, B. I., Lyubarev, A. E., Sanchez-Ruiz, J. M., and Shnyrov, V. L. (1997) Analysis of differential scanning calorimetry data for proteins - Criteria of validity of one-step mechanism of irreversible protein denaturation, *Biophys. Chem. 69*, 125−135.
35. Mitra, N., Srinivas, V. R., Ramya, T. N., Ahmad, N., Reddy, G. B., and Surolia, A. (2002) Conformational stability of legume lectins reflect their different modes of quaternary association: solvent denaturation studies on concanavalin A and winged bean acidic agglutinin, *Biochemistry 41*, 9256−9263.
36. Ahmad, N., Srinivas, V. R., Reddy, G. B., and Surolia, A. (1998) Thermodynamic characterization of the conformational stability of the homodimeric protein, pea lectin, *Biochemistry 37*, 16765−16772.
37. Stark, G. R. (1965) Reactions of cyanate with functional groups of proteins. 3. Reactions with amino and carboxyl groups, *Biochemistry 4*, 1030−1036.
38. Brinda, K. V., Surolia, A., and Vishveshwara, S. (2005) Insights into the quaternary association of proteins through structure graphs: a case study of lectins, *Biochem. J. 391*, 1−15.
39. Huntington, J. A., Read, R. J., and Carrell, R. W. (2000) Structure of a serpin-protease complex shows inhibition by deformation, *Nature 407*, 923−926.
40. Huntington, J. A. (2006) Shape-shifting serpins−advantages of a mobile mechanism, *Trends Biochem. Sci. 31*, 427−435.
41. Sauer, F. G., Remaut, H., Hultgren, S. J., and Waksman, G. (2004) Fiber assembly by the chaperone-usher pathway, *Biochim. Biophys. Acta 1694*, 259−267.
42. Nishiyama, M., Horst, R., Eidam, O., Herrmann, T., Ignatov, O., Vetsch, M., Bettendorff, P., Jelesarov, I., Grutter, M. G., Wuthrich, K., Glockshuber, R., and Capitani, G. (2005) Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD, *EMBO J. 24*, 2075−2086.
43. Lee, C., Park, S. H., Lee, M. Y., and Yu, M. H. (2000) Regulation of protein function by native metastability, *Proc. Natl. Acad. Sci. U.S.A. 97*, 7727−7731.
44. DeLano, W. L. (2002) The Pymol molecular graphics system. http://www.pymol.org.

BI7019047