METHODS AND APPLICATIONS

Engineered solubility tag for solution NMR of proteins

Amy M. Ruschak,[1]* Justine D. Rose,[1] Michael P. Coughlin,[2] and Tomasz L. Religa[2]*

[1]Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio, 44106
[2]Department of Physiology and Biophysics, Case Western Reserve University, Cleveland, Ohio 44106

Received 28 June 2013; Revised 6 August 2013; Accepted 8 August 2013
DOI: 10.1002/pro.2337
Published online 21 August 2013 proteinscience.org

Abstract: The low solubility of many proteins hinders large scale expression and purification as well as biophysical measurements. Here, we devised a general strategy to solubilize a protein by conjugating it at a solvent-exposed position to a 6 kDa protein that was re-engineered to be highly soluble. We applied this method to the CARD domain of Apoptosis-associated speck-like protein containing a CARD (ASC), which represents one member of a class of proteins that are notoriously prone to aggregation. Attachment of the tag to a cysteine residue, introduced by site-directed mutagenesis at its self-association interface, improved the solubility of the ASC CARD over 50-fold under physiological conditions. Although it is not possible to use nuclear magnetic resonance (NMR) to obtain a high quality 2D correlation spectrum of the wild type domain under physiological conditions, we demonstrate that NMR relaxation parameters of the solubilized variant are sufficiently improved to facilitate virtually any demanding measurement. The method shown here represents a straightforward approach for dramatically increasing protein solubility, enabled by ease of labeling as well as flexibility in tag placement with minimal perturbation to the target.© 2013 The Protein Society

Keywords: protein solubility; protein engineering; protein aggregation; nuclear magnetic resonance; protein A; CARD domain

Introduction

Solution nuclear magnetic resonance (NMR) has proven to be a powerful technique for studies of structure, stability, and dynamics of proteins.[1] Over the last decade we have observed significant advances in hardware and methodologies that improve the sensitivity of solution NMR.[2,3] While nowadays, 1D measurements and 2D ^{1}H-^{15}N and ^{1}H-^{13}C correlation maps can often be achieved at concentrations of tens of micromolar, a number of other experiments, such as backbone and side-chain assignments, studies of dynamics and acquisition of NOESY datasets for structure determination, require protein concentrations in excess of hundreds of micromolar.[4,5] While this has not been a problem for a

Abbreviations: ASC CARD, Apoptosis-associated speck-like protein containing a CARD; BdpA, B domain of protein A; CARD, caspase activation and recruitment domain; CD, circular dichroism; DTT, dithiothreitol; NHS, *N*-hydroxysuccinimide; NMR, nuclear magnetic resonance; PA, protein A; pI, isoelectric point.

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Amy M. Ruschak, Department of Biochemistry, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106. E-mail: amy.ruschak@case.edu or Tomasz L. Religa, Department of Physiology and Biophysics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106. E-mail: tomasz.religa@case.edu

number of model systems, for many proteins achieving solubility in this range has often proved challenging. Even if a protein can be concentrated without visible aggregation, it is rare that these costly samples are stable for the weeks that it would take to enable data collection and analysis for many triple resonance measurements. Moreover, self-association can lead to slower tumbling and shorter relaxation times than would be obtained from a monomeric protein. Sample degradation can change the spectral appearance over time, adding complexity to data analysis. The above factors collectively increase the demand on the time required to conduct NMR measurements.[6]

Caspase recruitment domains, or Caspase activation and recruitment domains (CARDs), are interaction motifs found in a wide array of proteins, typically those involved in processes relating to inflammation and apoptosis.[7,8] These domains mediate the formation of large protein complexes via direct interactions between individual CARDs. Likely because its self-association is inherent to its function to build scaffolds, CARD domains (and death fold domains, in general) are highly insoluble, and only a handful of NMR studies can be performed for a few of the more soluble variants.[9–12] Apoptosis-associated speck-like protein containing a CARD (ASC CARD) is a 10 kDa protein with a typical six-helix bundle death domain-fold. A protein construct that consists of the CARD, corresponding to residues 107–195 of ASC, is maximally soluble at 10 μM in a pyroptosis buffer (20 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 20 mM KCl pH 7.4) that mimics the environment in which it functions.[13]

Lack of solubility could be due to: (a) low protein stability, resulting in transient exposure of aggregation-prone regions; (b) presence of an intrinsic dimerization (oligomerization) interface; and (c) performing studies in solutions at a pH that is at or close to the protein's isoelectric point (pI).[14,15] Protein stability is usually improved by mutagenesis[16] or by optimizing buffer composition using thermal shift assays.[17] If identified, the oligomerization interface can be removed using mutagenesis,[18] or oligomerization propensity can be reduced by optimizing the ionic strength or pH of the buffer.[19,20] Oligomerization properties can also be reduced by changing the pI: by performing the measurements without cleaving the solubility tag used during recombinant protein expression,[21] by adding several charged residues at the N- or C- termini,[18] by adding short peptides that would bind the solubility-enhancing protein[22] or by linking the protein using intein chemistry.[23,24]

All of the aforementioned methods have had great success in decreasing the oligomerization propensities of proteins, but each of them is specialized and unlikely to provide a general approach. The major limitation of the chimeric fusions that have been proposed is that the solubility tag was attached only at the protein termini,[18,21–24] resulting in non-optimal tag placement that is limited to only two positions. Moreover, for some of them, the attachment results in isotopic labeling identical for the tag and the target protein, complicating the NMR spectra with undesired resonances from the tag.[21]

Here, we introduced mutations to *Staphylococcus aureus* B domain of protein A (BdpA) to create a highly negatively charged, 6 kDa variant of protein A (PA) that can be attached to, and improve the solubility of, a target partner protein. Specifically, the N-terminus of PA was linked to a cysteine located at a solvent-exposed position in the ASC CARD that was introduced by site-directed mutagenesis using a commercially available heterobifunctional crosslinking molecule, SM-$(PEG)_2$. The ASC CARD was produced with isotopic labeling for NMR, while the PA solubility tag was not, enabling NMR measurements to be conducted selectively for ASC CARD residues without obstruction by the tag. The PA-SM-$(PEG)_2$-ASC CARD fusion which was soluble up to 500 μM, yielded high quality 2D ^{1}H-^{15}N NMR spectra that does not change over time and had relaxation properties consistent with 17 kDa molecular weight species. This method can also be applied to increase the protein solubility of many other targets.

Results

Design of the fusion tag

Our goal was to develop a generalized method for straightforward attachment of a solubility tag that can be made NMR-invisible for protein NMR applications. The design of the tag is based on a 60 residue, three helix bundle fold, exemplified by domains of *Staphylococcus aureus* PA, which were used as a fusion tag for increasing protein solubility *in vivo*.[25] Protein A has five independent domains, which have over 80% sequence identity and the same structure.[26,27] The B domain of PA consists of three helices that span K9-L20 (H1), E25-D38 (H2), Q41-A57 (H3) [Fig. 1(A)]. This domain represents a model system for protein folding, where the thermodynamic stability and kinetics of folding have been characterized for more than 100 mutants.[28–31]

We decided to create a highly charged, solubility-enhanced version of the fold that would have greater resistance to aggregation. Charged residues were introduced using available mutagenesis data[28,29,31] so that the resulting PA variant: (a) has a low pI; (b) remains folded under solution conditions where most solution NMR studies are performed ((0–50°C), as well as low ionic strength, where cryo-probe benefits are the greatest[32]); (c) does not contain any lysines (due to the linker chemistry). The details of how the solubility tags increase

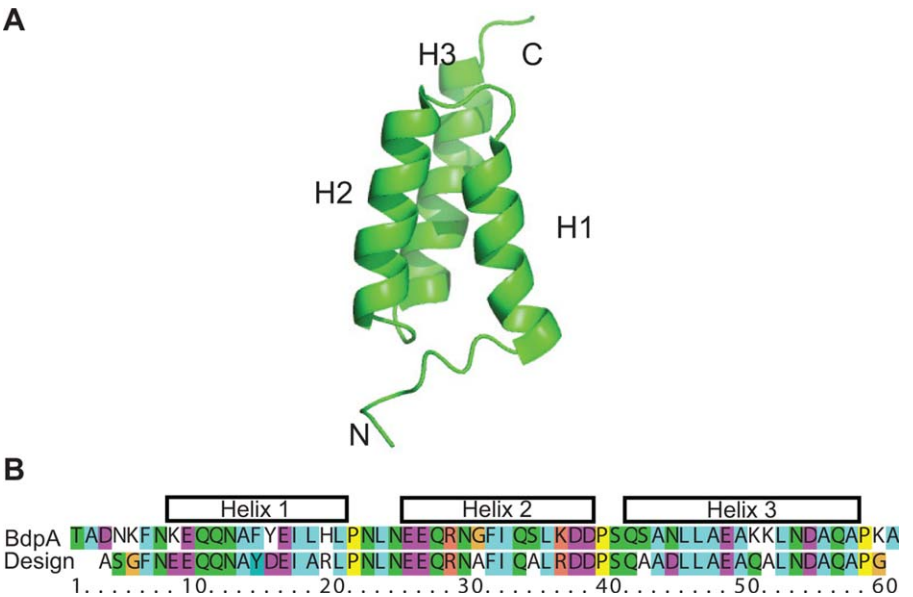

Figure 1. Design of PA solubility tag. A. Structure of BdpA (PDB: 1SS1),[28] used as a template for protein re-design. B. Sequence alignment of the designed PA and BdpA. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the soluble protein concentration are not fully understood, but it is well-known that increased negative surface charge correlates with increased solubility.[6,14,33] Therefore, many negatively charged residues were introduced into PA to give it a pI that is lower than the vast majority of proteins. The solubility of the protein being targeted is increased with the formation of the PA-target construct because repulsive electrostatic forces between PAs will oppose the driving force for aggregation between targets. Since most of the proteins have pIs in the 5–9 range,[34,35] attachment of PA will effectively decrease the pI of the target, which may be beneficial for improving solubility under the solution conditions of interest.

To re-engineer PA, we mutated 12 amino-acids (all at solvent-exposed positions) out of 53 residues available in our construct (23%) using the above principles as a guide [Fig. 1(B)]. Generally, a mutation was made if it: (a) has been found to stabilize the PA fold (G30A, S34A, S42A); (b) likely has negligible effect on stability, but removes a hydrophobic residue from the surface (Y15D, L18A, H19R); (c) likely has negligible effect on stability, but removes a lysine or lowers the pI (K8E, K36R, N44D, K50Q, K51A). In addition, we introduced a F14Y mutation to facilitate concentration determination by ultraviolet absorbance at 280 nm. Since no long-range NOEs have been found between the first five residues of the N-terminus and the rest of the protein as well as the last three on the C-terminus,[28] we shortened the construct by three residues on the N-terminus and one residue on the C-terminus. Mutations such as H19R, K36R, and K50Q, which are located at positions within helices, increase helical propensity through which it increases global protein stability. Finally, the H19R and K36R mutations introduce two new types of interactions to stabilize the protein. First, it creates a D-R (i, i + 4) salt bridge; and second, the charged side chain has favorable electrostatic interactions with the helix dipole.[36] Similarly, K50Q creates a Q-D (i, i + 4) hydrogen bond.[37] The final construct had 57 residues, a predicted pI of 3.9, and an overall net charge of −8 under physiological conditions, in comparison with the WT protein of 60 residues, a pI of 5.2, and a net charge of −2.

A codon-optimized gene corresponding to the above PA construct was synthesized, cloned into the pET29a vector, and expressed in *E. coli*. The peaks in ^{1}H-^{15}N HSQC spectra of the purified protein had high chemical shift dispersion similar to that of the WT protein,[27,28] indicative of structural similarity [Fig. 2(A)]. Circular dichroism (CD) spectrum was typical for a folded, helical protein and similar to that of BdpA [Fig. 2(B)]. To ensure that PA remains folded under the conditions of most NMR experiments, we measured its thermal denaturation midpoint (T_m) using CD spectroscopy [Fig. 2(C)]. Since stability of proteins that have charged residues in close proximity is often affected by ionic strength,[38–40] we focused on low and high ionic strength buffers. The T_m of PA in a low ionic strength buffer (10 m*M* potassium phosphate, pH 7.4) was 70°C and in a high-ionic strength buffer (10 m*M* potassium phosphate, 500 m*M* NaCl, pH 7.4) was 78.5°C, therefore the protein will remain >98% folded in the temperature range of most NMR experiments. In comparison, the T_m of BdpA is 77°C.[28] The ionic strength

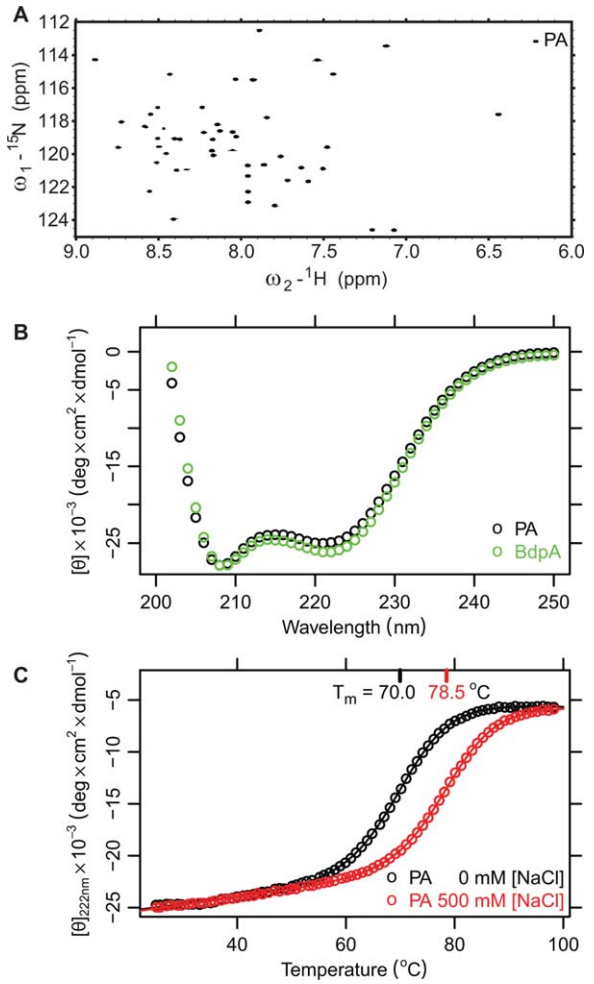

Figure 2. Biophysical properties of the designed PA tag. A. ^{1}H-^{15}N HSQC spectrum of designed PA in pyroptosis buffer. B. Circular dichroism spectrum of the B domain of PA (green) and re-designed PA (black) at low ionic strength. C. Thermal denaturation of designed PA monitored by CD at two ionic strengths.

dependence of protein stability might result from unfavorable electrostatic interactions between negatively charged residues, which are more effectively screened at the higher ionic strength.[39,40]

Applying PA to solubilize an aggregation-prone CARD domain under physiological conditions

Attachment of modified PA to the target protein was accomplished using a heterobifunctional crosslinker, succinimidyl-[(N-maleimidopropionamido)-diethyleneglycol] ester (SM-(PEG)$_2$) (Thermo Scientific) [Fig. 3(A)]. This crosslinker contains N-hydroxysuccinimide (NHS) ester and maleimide groups that allow covalent conjugation of amine- and sulfhydryl-containing molecules, which in this case correspond to PA and the target, respectively. This linker has arm length of 18 Å, equivalent to a N- to C- distance of ~5–6 disordered amino-acids, which is long enough so that PA does not directly interact with the CARD domain, but short enough, that its negative charge will decrease the aggregation properties of its partner. Numerous other possibilities are available which range in the arm length of 4.4–95 Å if, depending upon the specific need, shorter or longer linkers might be desired.[41]

The solubility of the ASC CARD is likely limited by self-association, as it has been observed to assemble into filamentous structures in COS cells.[42] The theoretical pI of the domain is 6.9, and therefore the solubility is expected to be low under physiological conditions. Moreover, experiments in living cells have identified residues on the surface on the ASC CARD that are essential for mediating its assembly into signal transduction complexes called inflammasomes. Based on these results, we selected E130, D134 and Y137 as positions for introducing a cysteine,[43] where conjugation of PA might sterically block and electrostatically repel self-association as well as lower the pI of the overall complex. Conjugation of PA to all mutants had the desired effect (see below),

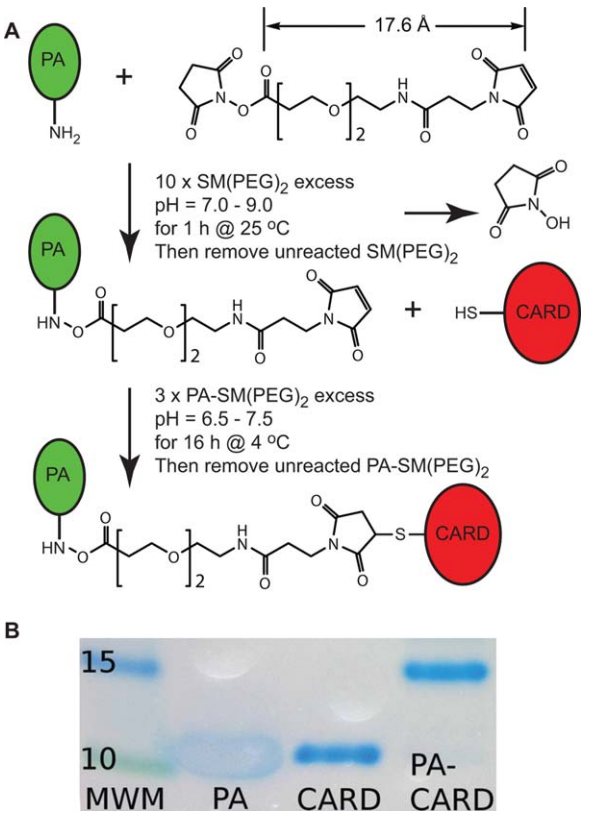

Figure 3. Conjugation of ASC CARD to the PA solubility tag. A. Overview of the labeling reaction. B. 4–20% SDS-PAGE gel showing successful labeling. PA was resistant to staining with Coomassie G-250 stain, and runs at higher apparent molecular weight than expected. Unusual amino acid distribution, with scarcity of positively charged and aromatic residues[53] and low pI[54] are likely responsible for the observed behavior.

and ultimately we chose to focus our analysis on the D134C C173S construct.

As mentioned above, PA was linked to ASC CARD D134C C173S using the commercially available crosslinker, SM-$(PEG)_2$. NHS esters react with primary amines at pH 7–9 to form amide bonds, while the maleimides react with sulfhydryl groups at pH 6.5–7.5 to form stable thioether bonds.[41] Since the target protein is likely to have a number of primary amine sites, corresponding to lysine side chains and the N-terminus, the SM-$(PEG)_2$ linker is first conjugated to the N-terminus of PA. Specifically, PA was incubated with 10-fold molar excess SM-$(PEG)_2$ at room temperature for 1 h in a buffer at pH 7.3 [Fig. 3(A), details in Materials and Methods]. Unreacted linker was then removed using a centrifugal concentrator (Millipore), as well as exchanged into a buffer conducive to refolding of the ASC CARD domain. Specifically, ASC CARD D134C C173S was denatured in guanidine-HCl and refolded by rapid dilution into a solution containing PA-SM-$(PEG)_2$. The conjugation reaction between the proteins was allowed to proceed overnight at 5°C. The PA-SM-$(PEG)_2$-ASC CARD was purified from unreacted components using size exclusion chromatography (details in Materials and Methods). The labeling was confirmed using SDS-PAGE gel [Fig. 3(B)], where we observed a band corresponding to the combined molecular weights of the ASC CARD and PA (~17 kDa).

Characterization of the ASC CARD-PA fusion by NMR and fluorescence

The suitability of PA-SM-$(PEG)_2$-ASC CARD was assessed for NMR applications. Enhanced solubility of ASC CARD was immediately obvious for the PA conjugate, since the protein did not aggregate during buffer exchange for NMR applications. A ^{1}H-^{15}N HSQC spectrum acquired at a protein concentration of 500 μ*M* at 25°C in pyroptosis buffer[13] has 85 resolved peaks [Fig. 4(A)] of 88 that would be expected on the basis of the primary structure. Moreover, the chemical shifts of the peaks are well dispersed, indicative of a folded protein. Narrow linewidths further suggest that the relaxation properties are satisfactory for most NMR applications. Finally, the sample was stable for weeks at room temperature and the appearance of the ^{1}H-^{15}N HSQC spectrum remained unchanged (data not shown).

We measured ^{15}N T_1, ^{15}N T_2 and {^{1}H}-^{15}N NOE for ASC CARD in PA-SM-$(PEG)_2$-CARD in pyroptosis buffer at 25°C, allowing for protein tumbling time determination.[44] The average ^{15}N R_2 and R_1 values were 14.4 ± 3.2 s^{-1}, and 1.48 ± 0.14 s^{-1}, respectively, on a spectrometer operating at 600 MHz [Fig. 4(B)]. The correlation time estimated from R_2/R_1 ratios was 9.20 ± 0.05 ns. This is consistent with a monomeric 17 kDa complex. Importantly, it should be possible to acquire all standard triple resonance experiments on this sample.[45]

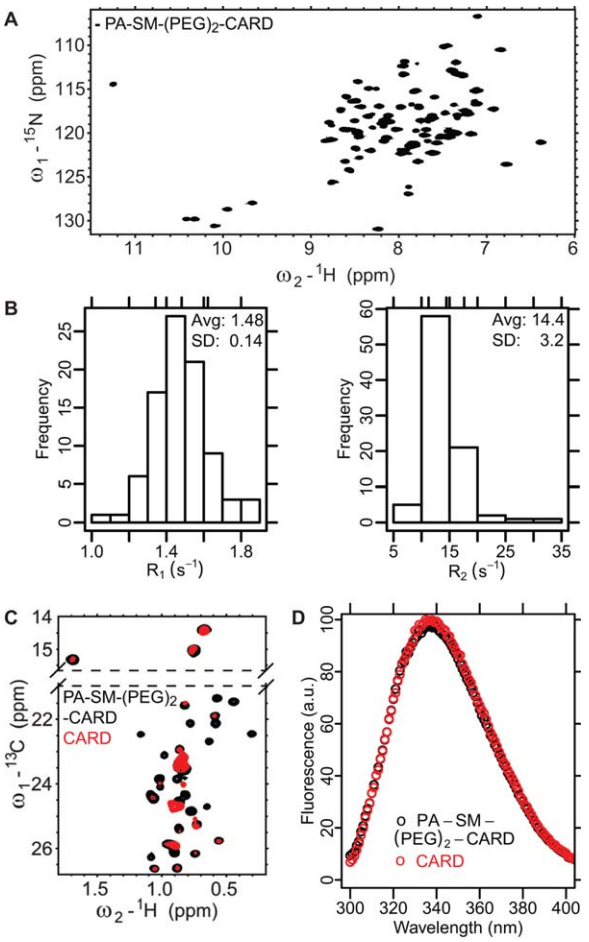

Figure 4. Biophysical analysis of PA-SM-$(PEG)_2$-CARD. A.^{1}H-^{15}N HSQC spectrum of ASC CARD D134C C173S with PA solubility tag. 85 well dispersed resonances are visible in the spectrum, suggesting a well folded protein. B. Distribution of R_1 (left) and R_2 (right) values for ASC CARD in PA-SM-$(PEG)_2$-CARD. The relaxation properties are consistent with monomeric protein. C. ^{1}H-^{13}C HSQC spectrum of ASC CARD with and without the PA solubility tag. The spectrum of ASC CARD only was acquired for 16 h on a 10 μ*M* sample; the spectrum of CARD with the solubility tag took only 20 min on a 500 μ*M* sample. Signal:noise on the 20 min spectrum is still >10-fold higher. D. Fluorescence emission spectrum of ASC CARD with and without the PA solubility tag. The excitation wavelength was 290 nm to ensure that tyrosines do not contribute significantly to this emission spectrum. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

In contrast to the PA-SM-$(PEG)_2$-CARD sample, we were unable to concentrate the wild type ASC CARD by itself in pyroptosis buffer above 10 μ*M*, since the protein formed visible aggregates above this concentration. We attempted to obtain ^{1}H-^{13}C HSQC spectrum on a sample where Ile, Leu, Met, and Val methyl groups are ^{13}C, ^{1}H labeled. Methyl groups are some of the most sensitive probes

available for NMR analysis of proteins and can be visible in species with unfavorable relaxation properties (such as high molecular weight oligomers) when other types of probes are not.[3] While peaks in the spectrum of the free protein are broadened, suggesting self-association even at 10 μM, the chemical shifts of visible resonances for both PA-SM-$(PEG)_2$-CARD and ASC CARD are very similar; thus labeling does not appear to fundamentally change the structure of CARD domain [Fig. 4(C)]. Differences in relative peak intensities between the two samples are most likely due to aggregation of the ASC CARD sample, albeit a dynamic process on a millisecond time scale cannot be excluded. If an NMR spectrum of the native protein cannot be obtained under the same solution conditions, tryptophan fluorescence is an alternative technique for determining the effects of labeling on the target protein structure. PA has been designed to have no tryptophans and a single tyrosine, and therefore tryptophan fluorescence measurements reflect properties of the target protein. Tryptophan fluorescence spectra of the ASC CARD are not affected by conjugation to PA, suggesting that labeling does not alter the conformation of ASC CARD [Fig. 4(D)].

Discussion

The approach presented here for increasing protein solubility has a number of obvious advantages: (1) there are a number of commercially available linkers that can be used that vary in chemical composition and length,[41] (2) the solubility tag can be introduced at any position on the target protein, (3) nearly all the target protein was labeled in the conjugation reaction, (4) ability to selectively introduce isotopic labels into the target protein, as well as deuterate the solubility tag, (5) the labeling reactions can be conducted under a wide range of solution conditions that are compatible with downstream NMR applications, and (6) highly sensitive fluorescence-based analysis determining whether or not the native fold is preserved upon labeling. The drawback of the method is the requirement for a solvent-exposed cysteine. If a protein structure is unknown, Ellman's reagent can be used to establish the number of solvent-exposed cysteines.[46] PA could potentially be conjugated to any solvent-exposed cysteine that is available. In the case that multiple cysteine side chains are solvent exposed, PA could be conjugated at all of these positions, with a concomitant increase in molecular weight of just over 6 kDa for every cysteine that is labeled. In contrast, if none are naturally available, cysteines could be introduced at positions close to the N- or C-termini immediately adjacent to the protein, where the mutation would not likely affect the protein fold. Positions not conserved in sequence alignments and predicted to be on the surface of the protein would also be excellent choices for cysteine substitution.

Other protein folds aside from the BdpA may provide even better platforms for designing solubility tags that can be attached using heterobifunctional crosslinkers. Small, single domain proteins that have proven to enhance the solubility of its partners *in vivo* and *in vitro* are the best targets, such as GB1 domain, SUMO or ubiquitin.[25,47] Similar to PA, all of them will require sequence re-design to remove lysines, while lowering or raising the pI to extreme values and maintaining stability of the fold. Extensive mutagenesis data are available for most of them.[48,49]

We showed that attaching a small, 6 kDa protein to the CARD domain of ASC with a heterobifunctional crosslinker resulted in dramatic improvements in protein solubility, so that effectively any NMR measurement is possible on this protein under physiological conditions. The method used here has the advantage that the solubility tag can be conjugated to the protein at any position, and not just at the N- or C-terminus as a recombinant fusion. The tag would most effectively be conjugated at positions on the target protein that would break an oligomerization interface. Also, the tag effectively lowers the pI of the PA-target protein conjugate, increasing the solubility of proteins that have pIs close to 7 at physiological pH. Finally, selective deuteration of PA in the PA-ASC CARD conjugate will eliminate any background signals arising from the tag that might otherwise complicate interpretation of NMR data. The method is sufficiently simple that it can be applied to any protein solubility problem. It will be of broad and general interest to anybody struggling with proteins that form dimers, oligomers, or aggregates.

Materials and Methods

Protein expression

Re-designed PA was expressed in *E. coli* Rosetta 2 (DE3) cells (Novagen) from the pET29a vector in either terrific broth (TB) when unlabeled protein was required, or M9 minimal media when ^{15}N labeling was needed. The cells were grown at 37°C in TB or M9 H_2O medium supplemented with 100 mg/L of kanamycin until $OD_{600} \sim 2$, when expression was induced by addition of 1 mM IPTG and allowed to proceed overnight at 25°C. The protein was purified using a two-step protocol that included Ni affinity and Q ion exchange columns. Specifically, the bacteria were harvested by centrifugation and resuspended in Ni-A buffer (10 mM imidazole, 50 mM potassium phosphate, 500 mM sodium chloride, pH 8.0) supplemented with PMSF, lysozyme and DNAse I. The cells were lysed by sonication and spun down at 45,000 *g* for 30 min. The filtered supernatant was

applied to a column with Ni Sepharose beads (GE Healthcare), equilibrated and washed with Ni-A buffer. PA was eluted using Ni-B buffer (500 m*M* imidazole, 50 m*M* potassium phosphate, 500 m*M* sodium chloride, pH 8.0). After overnight dialysis into 20 m*M* Tris pH 8.0 in 3k MWCO dialysis membrane (Spectrum Laboratories) with 1 mg of TEV protease,[50] the protein was further purified on the HiTrap Q HP 5 mL column (GE Healthcare) using 1%/mL NaCl gradient from 0 to 1 *M* in 20 m*M* Tris, pH 8.0. The purity and identity of the protein was confirmed using SDS-PAGE, NMR and electrospray ionization mass spectrometry.

BdpA was expressed from pRSET A vector, where the domain was fused to lipoyl fusion protein, separated by a thrombin protease cleavage site, and produced as described previously.[28]

The ASC CARD was expressed in *E. coli* C43(DE3) cells grown in M9 minimal media containing ^{15}N ammonium chloride to yield protein that is isotopically enriched with ^{15}N. In addition, methyl groups were selectively labeled with ^{13}C at δ_1 of isoleucine, δ_1 and δ_2 of leucine, γ_1 and γ_2 of valine, and ϵ_1 of methionine by adding 60 mg/L of 2-keto-3,3-D_2-4-^{13}C-butyrate (Ile), 80 mg/L 2-keto-3-methyl-D_3-3-D_1-4-^{13}C-butyrate (Leu,Val), or 80 mg/L ^{13}C-methyl methionine (Met) to the media 1 h before induction, as previously described.[18] Protein expression was induced at an OD_{600} of 0.8 and proceeded at 37°C for 6 h. After centrifugation, cells were resuspended in 7 *M* guanidine HCl, 50 m*M* potassium phosphate, 300 m*M* sodium chloride, 10 m*M* imidazole, 1 m*M* tris(2-carboxyethyl)phosphine (TCEP) pH 8.0 (Ni-C) and lysed by freezing overnight at −80°C. Cells were briefly sonicated and spun at 45,000 *g* for 45 min. The supernatant was applied to a HisTrap FF column (GE Healthcare) equilibrated with Ni-C and washed with ~200 mL Ni-C. To remove TCEP and shift the pH to 7.0, the column was washed with three column volumes of 7 *M* guanidine HCl, 50 m*M* potassium phosphate, 300 m*M* sodium chloride, 10 m*M* imidazole, pH 7.0, and then eluted with three column volumes of 7 *M* guanidine HCl, 50 m*M* potassium phosphate, 300 m*M* sodium chloride, 250 m*M* imidazole, pH 7.0.

Labeling

A stock of SM-$(PEG)_2$ (Thermo Scientific) was prepared in dimethyl sulfoxide (DMSO) as described by the manufacturer. It was diluted into a ~1 m*M* solution of PA in 25 m*M* HEPES pH 7.3, 100 m*M* NaCl at a 10-fold molar excess of of SM-$(PEG)_2$ to PA and incubated for 1 h at 25°C. Unreacted linker was removed by buffer exchanging into 50 m*M* potassium phosphate pH 7.0, 300 m*M* NaCl using a centrifugal concentrator (Millipore, 3 kDa cutoff). The ASC CARD was refolded at 5 μ*M* into a solution containing PA-SM-$(PEG)_2$ at 15 μ*M*, 5°C. The reaction between the linker and ASC CARD was allowed to proceed overnight. The next morning, the reaction was quenched by addition of 1 m*M* dithiothreitol (DTT), and the N-terminal His tag on the ASC CARD was cleaved by addition of TEV protease for 1 day incubation at 5°C. TEV was removed using a HisTrap FF column equilibrated in the refolding buffer, where the flow through contained the PA-SM-$(PEG)_2$–ASC CARD. The solution was concentrated using a centrifugal concentrator (Millipore, 3 kDa cutoff) and then purified by size exclusion on a HiLoad 16/60 Superdex 200 prep grade gel filtration column (GE Healthcare) equilibrated in 50 m*M* potassium phosphate pH 7.0, 300 m*M* NaCl.

This labeling reaction was performed by refolding the protein into the labeling buffer, but the labeling reaction for most of the target proteins should not require refolding. The solubility tags used for protein expression generally lack solvent-exposed cysteines,[47] so the reaction could easily be performed when the fusion tag used in protein expression is still attached to the target protein. After successful labeling, the reaction could be quenched with addition of DTT, and the fusion tag removed by adding appropriate protease. The target protein should remain in solution due to the attachment of PA.

SDS-PAGE gel electrophoresis

The labeling was confirmed on Any kD Mini-PROTEAN TGX precast polyacrylamide gel (Bio-rad) run for 90 min at 100 V. The proteins were visualized using Commassie Brilliant Blue G-250 staining protocol.[51]

NMR

Samples for NMR measurements were prepared in 20 m*M* HEPES, 20 m*M* potassium chloride, 0.02% sodium azide, 2% D_2O buffer that mimics physiologically-relevant conditions under which ASC self assembles into a regulatory complex called the pyroptosome to initate an inflammatory form of cell death.[13] All NMR measurements were performed at 25°C on Bruker Avance 600 MHz spectrometer equipped with 5 mm TXI triple resonance cryo-probe and single-axis gradients. The 2D ^{15}N T_1, T_2 and {^{1}H}–^{15}N nuclear Overhauser effect (NOE) experiments were acquired using Bruker built-in pulse sequences. The relaxation data was analyzed using Sparky 3.115 and fit using the non-linear least-squares Levenberg-Marquardt algorithm. The correlation time was determined using Tensor 2 software from R_2/R_1 ratio, considering only residues for which {^{1}H}–^{15}N NOE>0.65 and R_1, R_2 values are within one standard deviation from the mean.[44] This minimized the effect of motions that are faster than the overall rotational diffusion but slower than ~100 ps (which alter R_1 values), and motions occurring on a

micro-second to milli-second time scale (which affect R_2 values).[52]

Circular dichroism

Far-ultraviolet CD spectra were acquired on a Aviv Model 215 at 25°C using 50 μ*M* protein in a cuvette with 1-mm path length in two buffers: 10 m*M* potassium phosphate, pH 7.4 and 10 m*M* potassium phosphate, 500 m*M* sodium chloride, pH 7.4. Temperature scans were acquired under above conditions at 222 nm using bandwith of 1.5 nm at 1°C intervals between 25 and 98°C with 30 s equilibration at each temperature. Thermal denaturation curves were fitted to a two state model assuming ΔC_p of 650 cal mol^{-1} K^{-1} (Refs. 16,28) using R 2.14.1 statistical analysis software.

Fluorescence

Measurements of tryptophan fluorescence were performed on a Tecan Infinite M1000 PRO plate reader in black 96 full-well plates at 25°C. Fluorescence spectra of samples containing 3 μ*M* protein in pyroptosis buffer were recorded with an excitation wavelength of 290 nm and a bandwidth of 2.5 nm.

References

1. Barrett PJ, Chen J, Cho M-K, Kim J-H, Lu Z, Mathew S, Peng D, Song Y, Van Horn WD, Zhuang T, Sönnichsen FD, Sanders CR (2013) The quiet renaissance of protein nuclear magnetic resonance. Biochemistry 52:1303–1320.
2. Kovacs H, Moskau D, Spraul M (2005) Cryogenically cooled probes—a leap in NMR technology. Prog Nucl Mag Res Sp 46:131–155.
3. Atreya HS, editor. (2012). Isotope labeling in Biomolecular NMR. Adv. Exp. Dordrecht: Springer Netherlands.
4. Bieri M, Kwan AH, Mobli M, King GF, Mackay JP, Gooley PR (2011) Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination. FEBS J 278:704–715.
5. Kwan AH, Mobli M, Gooley PR, King GF, Mackay JP (2011) Macromolecular NMR spectroscopy for the non-spectroscopist. FEBS J 278:687–703.
6. Zhou P, Wagner G (2010) Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies. J Biomol NMR 46:23–31.
7. Kersse K, Verspurten J, Vanden Berghe T, Vandenabeele P (2011) The death-fold superfamily of homotypic interaction motifs. Trends Biochem Sci 36: 541–552.
8. Park HH, Lo Y-C, Lin S-C, Wang L, Yang JK, Wu H (2007) The death domain superfamily in intracellular signaling of apoptosis and inflammation. Annu Rev Immunol 25:561–586.
9. Zhou P, Chou J, Olea RS, Yuan J, Wagner G (1999) Solution structure of Apaf-1 CARD and its interaction with caspase-9 CARD: a structural basis for specific adaptor/caspase interaction. Proc Natl Acad Sci USA 96:11265–11270.
10. Chou JJ, Matsuo H, Duan H, Wagner G (1998) Solution structure of the RAIDD CARD and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment. Cell 94:171–180.
11. Ferrage F, Dutta K, Nistal-Villán E, Patel JR, Sánchez-Aparicio MT, De Ioannes P, Buku A, Aseguinolaza GG, García-Sastre A, Aggarwal AK (2012) Structure and dynamics of the second CARD of human RIG-I provide mechanistic insights into regulation of RIG-I activation. Structure 20:2048–2061.
12. De Alba E (2009) Structure and interdomain dynamics of apoptosis-associated speck-like protein containing a CARD (ASC). J Biol Chem 284:32932–32941.
13. Fernandes-Alnemri T, Alnemri ES (2008) Assembly, purification, and assay of the activity of the ASC pyroptosome. Methods Enzymol 442:251–270.
14. Kramer RM, Shende VR, Motl N, Pace CN, Scholtz JM (2012) Toward a molecular understanding of protein solubility: increased negative surface charge correlates with increased solubility. Biophys J 102:1907–1915.
15. Schein CH (1990) Solubility as a function of protein structure and solvent components. Biotechnology 8: 308–315.
16. Fersht AR (1999) Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding. W. H. Freeman and Company, New York, USA.
17. Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc 2: 2212–2221.
18. Ruschak AM, Kay LE (2012) Proteasome allostery as a population shift between interchanging conformers. Proc Natl Acad Sci USA 109:E3454–E3462.
19. Bagby S, Tong K, Ikura M (2001) Optimization of protein solubility and stability for protein nuclear magnetic resonance. Methods Enzymol 339:20–41.
20. Lepre CA, Moore JM (1998) Microdrop screening: a rapid method to optimize solvent conditions for NMR spectroscopy of proteins. J Biomol NMR 12:493–499.
21. Zhou P, Lugovskoy AA, Wagner G (2001) A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. J Biomol NMR 20:11–14.
22. Durst FG, Ou H Der, Löhr F, Dötsch V, Straub WE (2008) The better tag remains unseen. J Am Chem Soc 130:14932–14933.
23. Kobashigawa Y, Kumeta H, Ogura K, Inagaki F (2009) Attachment of an NMR-invisible solubility enhancement tag using a sortase-mediated protein ligation method. J Biomol NMR 43:145–150.
24. Kobayashi H, Swapna GVT, Wu K-P, Afinogenova Y, Conover K, Mao B, Montelione GT, Inouye M (2012) Segmental isotope labeling of proteins for NMR structural study using a protein S tag for higher expression and solubility. J Biomol NMR 52:303–313.
25. Bogomolovas J, Simon B, Sattler M, Stier G (2009) Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. Protein Exp Purif 64:16–23.
26. Jansson B, Uhlén M, Nygren P (1998) All individual domains of staphylococcal protein A show Fab binding. FEMS Immunol Med Microbiol 20:69–78.
27. Zheng D, Aramini J, Montelione G (2004) Validation of helical tilt angles in the solution NMR structure of the Z domain of Staphylococcal protein A by combined analysis of residual dipolar coupling. Prot Sci 13:549–554.
28. Sato S, Religa TL, Daggett V, Fersht AR (2004) Testing protein-folding simulations by experiment: B domain of protein A. Proc Natl Acad Sci USA 101:6952–6956.

29. Sato S, Religa TL, Fersht AR (2006) Phi-analysis of the folding of the B domain of protein A using multiple optical probes. J Mol Biol 360:850–864.
30. Fersht AR, Sato S (2004) Phi-value analysis and the nature of protein-folding transition states. Proc Natl Acad Sci USA 101:7976–7981.
31. Sato S, Fersht AR (2007) Searching for multiple folding pathways of a nearly symmetrical protein: temperature dependent phi-value analysis of the B domain of protein A. J Mol Biol 372:254–267.
32. Kelly AE, Ou HD, Withers R, Dotsch V (2002) Low-conductivity buffers for high-sensitivity NMR measurements. J Am Chem Soc 124:12013–12019.
33. Trevino S, Scholtz J, Pace C (2008) Measuring and increasing protein solubility. J Pharm Sci 97:4155–4166.
34. Kiraga J, Mackiewicz P, Mackiewicz D, Kowalczuk M, Biecek P, Polak N, Smolarczyk K, Dudek MR, Cebrat S (2007) The relationships between the isoelectric point and: length of proteins, taxonomy and ecology of organisms. BMC Genomics 8:163.
35. Schwartz R, Ting CS, King J (2001) Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. Genome Res 11:703–709.
36. Huyghues-Despointes B, Scholtz JM, Baldwin RL (1993) Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. Protein Sci 2:80–85.
37. Huyghues-Despointes BMP, Klingler TM, Baldwin RL (1995) Measuring the strength of side-chain hydrogen bonds in peptide helixes: the Gln. cntdot. Asp (i, i+ 4) interaction. Biochemistry 34:13267–13271.
38. Horovitz A, Serrano L, Avron B, Bycroft M, Fersht AR (1990) Strength and co-operativity of contributions of surface salt bridges to protein stability. J Mol Biol 216: 1031–1044.
39. Religa TL, Markson JS, Mayor U, Freund SMV, Fersht AR (2005) Solution structure of a protein denatured state and folding intermediate. Nature 437:1053–1056.
40. Arbely E, Neuweiler H, Sharpe TD, Johnson CM, Fersht AR (2010) The human peripheral subunit-binding domain folds rapidly while overcoming repulsive Coulomb forces. Prot Sci 19:1704–1713.
41. Hermanson GT (2008). Bioconjugate techniques. Elsevier London, United Kingdom.
42. Masumoto J, Taniguchi S, Sagara J (2001) Pyrin N-terminal homology domain- and caspase recruitment domain-dependent oligomerization of ASC. Biochem Biophys Res Commun 280:652–655.
43. Proell M, Gerlic M, Mace PD, Reed JC, Riedl SJ (2013) The CARD plays a critical role in ASC foci formation and inflammasome signalling. Biochem J 449:613–621.
44. Dosset P, Hus JC, Blackledge M, Marion D (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. J Biomol NMR 16: 23–28.
45. Rossi P, Swapna GVT, Huang YJ, Aramini JM, Anklin C, Conover K, Hamilton K, Xiao R, Acton TB, Ertekin A, Everett JK, Montelione GT (2010) A microscale protein NMR sample screening pipeline. J Biomol NMR 46:11–22.
46. Riener CK, Kada G, Gruber HJ (2002) Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4′-dithiodipyridine. Anal Bioanal Chem 373: 266–276.
47. Waugh DS (2005) Making the most of affinity tags. Trends Biotechnol 23:316–320.
48. Went HM, Jackson SE (2005) Ubiquitin folds through a highly polarized transition state. Protein Eng Des Sel 18:229–237.
49. Morrone A, Giri R, Toofanny RD, Travaglini-Allocatelli C, Brunori M, Daggett V, Gianni S (2011) GB1 is not a two-state folder: identification and characterization of an on-pathway intermediate. Biophys J 101:2053–2060.
50. Blommel PG, Fox BG (2007) A combined approach to improving large-scale production of tobacco etch virus protease. Protein Exp Purif 55:53–68.
51. Pink M, Verma N, Rettenmeier AW, Schmitz-Spanke S (2010) CBB staining protocol with higher sensitivity and mass spectrometric compatibility. Electrophoresis 31:593–598.
52. Tjandra N, Wingfield P, Stahl S, Bax A (1996) Anisotropic rotational diffusion of perdeuterated HIV protease from 15N NMR relaxation measurements at two magnetic fields. J Biomol NMR 8:273–284.
53. Compton SJ, Jones CG (1985) Mechanism of dye response and interference in the Bradford protein assay. Anal Biochem 151:369–374.
54. Shi Y, Mowery R a, Ashley J, Hentz M, Ramirez AJ, Bilgicer B, Slunt-Brown H, Borchelt DR, Shaw BF (2012) Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. Prot Sci 21:1197–1209.