RESEARCH PAPER

Intrinsically Disordered Proteins 1, e25464; 2013; © 2013 Landes Bioscience

A method to trap transient and weak interacting protein complexes for structural studies

Vishnu Priyanka Reddy Chichili, Veerendra Kumar and J Sivaraman*

Department of Biological Sciences; National University of Singapore; Singapore

Keywords: protein-protein interactions, transient binding, linked complex, co-crystallization

Several key biological events adopt a "hit-and-run" strategy in their transient interactions between binding partners. In some instances, the disordered nature of one of the binding partners severely hampers the success of co-crystallization, often leading to the crystallization of just one of the partners. Here, we discuss a method to trap weak and transient protein interactions for crystallization. This approach requires the structural details of at least one of the interacting partners and binding knowledge to dock the known minimum binding region (peptide) of the protein onto the other using an optimal-sized linker. Prior to crystallization, the purified linked construct should be verified for its intact folding and stability. Following structure determination, structure-guided functional studies are performed with independent, full-length unlinked proteins to validate the findings of the linked complex. We designed this approach and then validated its efficacy using a 24 amino acid minimum binding region of the intrinsically disordered, neuron-specific substrates, Neurogranin and Neuromodulin, joined via a Gly-linker to their interacting partner, Calmodulin. Moreover, the reported functional studies with independent full-length proteins confirmed the findings of the linked peptide complexes. Based on our studies, and in combination with the supporting literature, we suggest that optimized linkers can provide an environment to mimic the natural interactions between binding partners, and offer a useful strategy for structural studies to trap weak and transient interactions involved in several biological processes.

Introduction

Protein-protein interactions comprise the underlying mechanism of a number of biological processes, such as DNA replication, transcription, translation, signal transduction and other complex cellular processes.[1] More often, intracellular protein-protein interactions are transient, wherein two or more proteins are involved and exist as either strong transient or weak transient interactions.[2] Elucidating these interactions is critical for understanding the various biological processes and disease conditions for the development of therapeutic interventions. There are many techniques available to elucidate protein-protein interactions, among which X-ray crystallography and NMR spectroscopy are the common methods employed to study the interactions at atomic level. However, these techniques, particularly co-crystallization experiments, require the protein-protein complex to exist as a stable and homogeneous species. In some cases, it has been observed that the presence of certain crystallization reagents can destabilize the protein-protein interaction, leading to crystallization of only one of the proteins. In other instances, one of the binding partners might be disordered, and may gain a stable secondary structure upon transiently interacting with its binding partner to fulfill its biological role.[3] This "hit-and-run" strategy thus poses a challenge for trapping the event in co-crystallization.

Transient protein-protein interactions are commonly studied with the use of chemical cross linking.[4] This involves the formation of covalent bonds between 2 proteins using bifunctional reagents that react with functional groups, such as primary amines and sulfhydryls, of amino acid residues.[5] For instance, glutaraldehyde, one of the most common chemical cross linkers, is able to maintain the structural rigidity of proteins. However, utilizing chemical cross linkers depends on the proximity of particular amino acids (such as Asp, Cys, Glu and Lys) to the site of the interaction.[6] Moreover, optimization of cross linking reactions is required to reduce higher order oligomer formation.[7]

Besides chemical cross linking, the interacting peptide from one binding partner can be covalently linked to other binding partner using a Gly-rich linker to obtain a structurally well-ordered complex. A protein data bank (PDB) search (http://www.pdb.org/pdb/home/home.do) shows that several linked protein-peptide complexes have been studied. For instance, a kinetic study showed that the association between the TCR (T cell receptor) and a peptide-MHC (major histocompatibility complex) was slow, but dissociation was fast, making it difficult to trap this transient interaction for structural studies.[8]

*Correspondence to: J Sivaraman; Email: dbsjayar@nus.edu.sg
Submitted: 06/19/13; Accepted: 06/19/13
http://dx.doi.org/10.4161/idp.25464
Citation: Chichili V, Kumar V, Sivaraman J. A method to trap transient and weak interacting protein complexes for structural studies. Intrinsically Disordered Proteins; 2013; 1:e25464

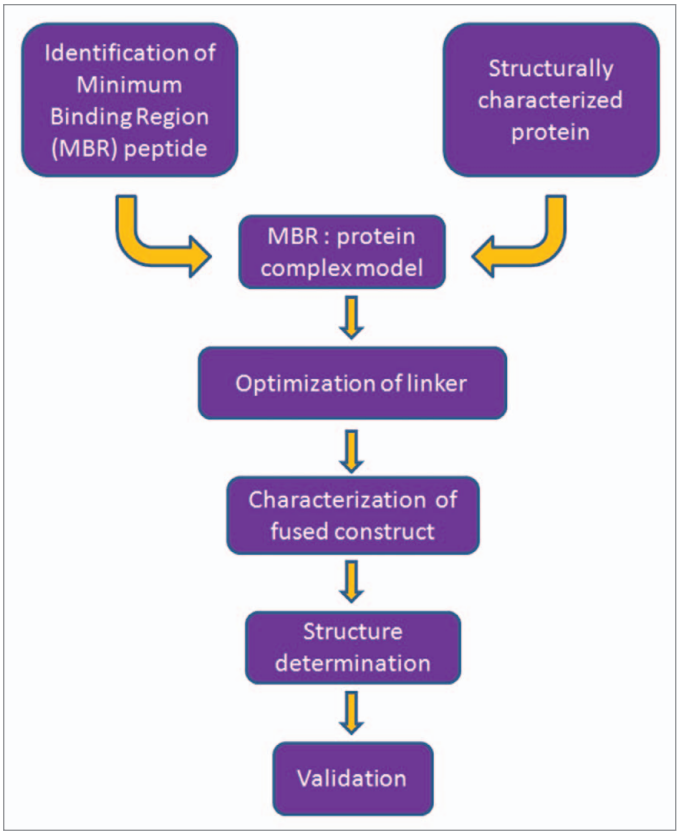

Figure 1. Flowchart representation of the various key steps involved in the linked peptide complex method.

Later studies, however, were able to exploit the use of linked peptides to study this interaction.[9-12] In order to mimic the natural interactions between binding partners and to trap the complex formed, numerous glycine-rich linkers have been used, with lengths that vary from 5 to 31 amino acids, as reviewed previously.[13] Although several linked complexes have been studied, there is no report available in the literature on the methodology by which this can be achieved.

Here, we describe a method to trap transient protein-protein interactions using an example for which one of the partners is intrinsically disordered but gains a stable secondary structure upon binding with its partner. For several years, our efforts to obtain a stable complex for this intrinsically disordered binding partner were unsuccessful. Now, using this approach, we are able to show that stable protein complexes can be obtained by linking a peptide of the minimum binding region (MBR) of the disordered partner to a structurally well-ordered protein. Using a computational model of the complex (generated with the available structural and binding details), we optimized the length and type of a linker to retain natural binding. To demonstrate this approach, here we report the interactions between the intrinsically disordered neuron-specific substrate proteins, Neurogranin (Ng) and Neuromodulin (Nm) with Calmodulin (CaM), using an optimized 5aa-Gly linker to fuse the MBR peptides (24 aa) of these two proteins with the C-terminal of CaM. Subsequently, the structures of these fused constructs were determined, and the importance of the key interacting residues was validated with the unlinked full-length proteins. Our results, combined with the supporting literature, suggest that optimized flexible polypeptide linkers can provide an environment that mimics the natural interactions between 2 binding partners. The results also show that the linker itself plays no role in dictating the interactions between the partners. This method can be employed to study other transient protein-protein interactions in several key biological events that have previously been unattainable.

Results

We adopted the strategy that has been described in the Materials and Methods section to study the transient and weak protein-protein interactions between the intrinsically disordered, neuron-specific substrate proteins, Ng and Nm, with Calmodulin (CaM). **Figure 1** summarizes the key steps employed in this approach. Our previous attempts to co-crystallize either of these proteins with CaM, using the full-length proteins or the commercially synthesized Ng/Nm IQ peptides (24 aa) did not yield complex crystals. When crystals were obtained under certain conditions, the crystals contained only CaM. This was likely due to the combined effect of the disordered nature of these proteins/peptides and the weak or transient interactions between these proteins and CaM. As such, the Ng/Nm and CaM interaction was deemed an appropriate model to demonstrate the efficacy of this methodology.

Identification of minimum binding region. Prior knowledge on the CaM binding regions of Nm/Ng were used to design the minimum binding region (MBR) peptides of the IQ motifs. Initially, a sequence analysis was performed for CaM binding IQ peptides, whose structures were solved in complex with CaM. It was identified that the average length of the binding peptides used for complex formation was 25 amino acids where the IQ motif was located in the middle of the sequence. Previously, different lengths of IQ motif peptides of Nm and Ng have been used to study their interaction with CaM by NMR and fluorescence techniques.[14,15] Thus, based on the literature and the structural analysis of the CaM complexes, we designed 2 MBRs (19 aa and 24 aa long) each for Nm and Ng to mimic the natural interactions of its full-length counterparts with CaM. The ITC analysis with CaM and these MBRs showed that the 24 aa MBR of both Nm and Ng had a higher affinity for CaM than the 19 aa MBRs.[16] Thus, a 24 aa MBR IQ motif peptide was selected to fuse with CaM.

Computational analysis and docking studies. Linking of Calmodulin Binding Domain (CBD) of calcineurin to the C-terminus of CaM using $(Gly)_5$ has been previously reported.[17] This CaM binding partner does not possess an IQ motif and hence could not be used as an example for modeling in this study. As such, we employed computational modeling for the CaM-IQ motif complexes using Deep View.[18] The template model was derived from a 2.5 Å crystal structure of the *apo*-CaM and IQ motif of myosin V complex (PDB code 2IX7).[19] A modeling analysis showed that the distance between N-terminus of the MBR peptide and the C-terminus of CaM was ~17 Å in the case of Nm and ~19 Å in the case of Ng (**Fig. 2A**).

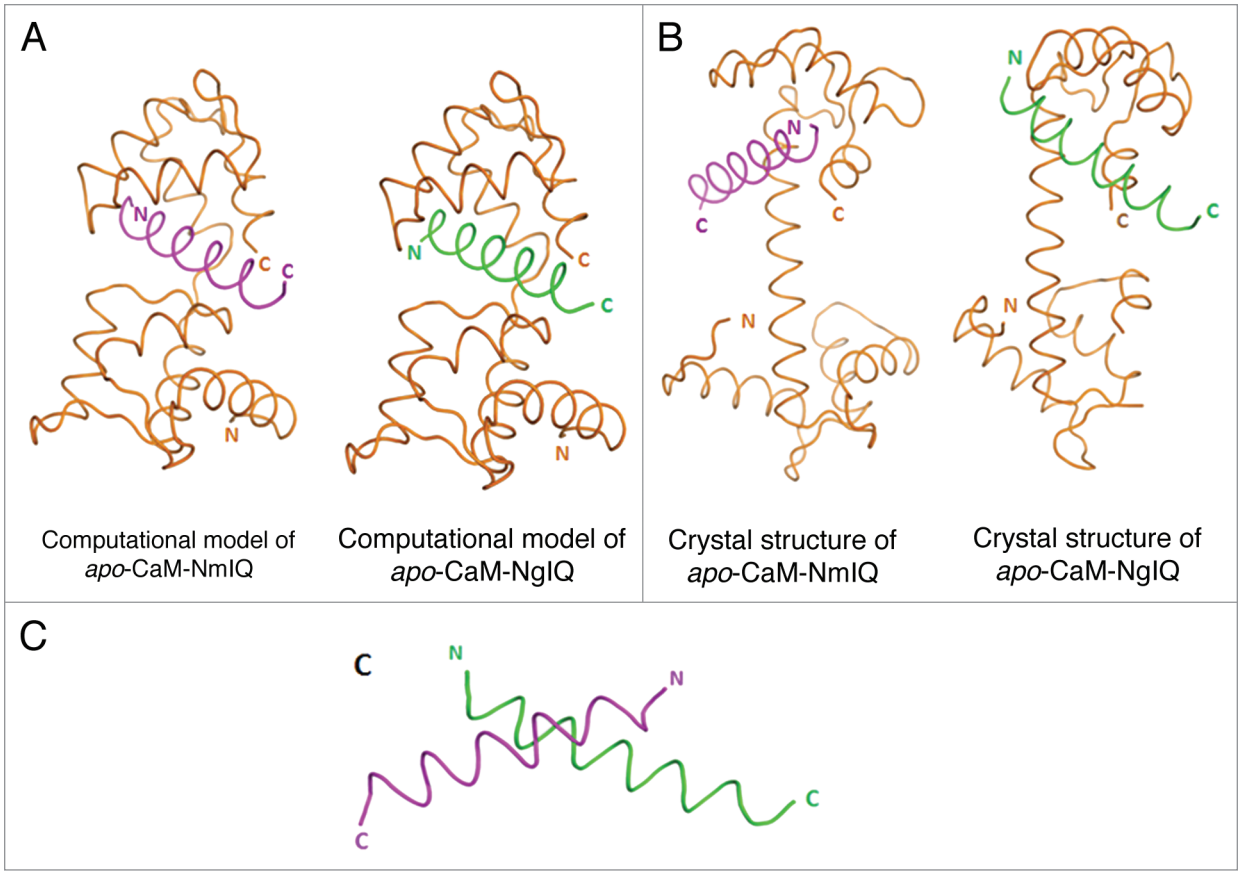

Figure 2. (**A**) Computational models of *apo*-CaM-Nm IQ and *apo*-CaM-Ng IQ complexes obtained with Deep View analysis using *apo*-CaM-myosin V IQ peptide complex as a template (PDB code: 2IX7). Computational models show that the IQ motif peptide interacts mainly with the C-lobe of CaM. (**B**) Crystal structures of *apo*-CaM-Nm (PDB code: 4E53) and *apo*-CaM-Ng (PDB code: 4E50) complexes obtained by linking the binding partners using a $(Gly)_5$ linker. CaM adopts an extended conformation and the IQ motif peptides interact with the C-lobe of CaM. (**C**) Ribbon representation of different orientations adopted by Nm (magenta) and Ng (green) IQ peptides when CaM-Nm IQ (PDB code: 4E53) and CaM-Ng IQ (PDB code: 4E50) structures were superimposed.

Linking the MBR peptide to protein. Previously, it was shown that the IQ motifs of Ng and Nm interact with the C-lobe of *apo*-CaM;[20-22] thus, the IQ peptides were fused to the C-terminus of CaM for this study. Moreover, the structural analysis of the known CaM-peptide complexes also suggested linking the peptide to the C-terminus of CaM using a $(Gly)_5$ linker.[17] While we believed that the $(Gly)_5$ linker would be sufficient, we also developed complexes using a longer linker ($[Gly]_8$) to avoid any possible hindrances that could arise due to the shorter $(Gly)_5$ linker. In brief, the IQ motifs (24 aa) of Nm and Ng were linked to the C-terminus of CaM via a flexible 5/8-Gly linker (CaM-$[Gly]_{5/8}$-NmIQ or NgIQ) using a three-step fusion PCR procedure, as described by Ye et al.[17] The first round of PCR involved amplification of the CaM gene with primers that incorporated the Gly linker at the C-terminus of CaM. The second round of PCR involved amplification of the Nm/Ng IQ motif gene with primers incorporating the Gly linker at the N-terminus of the IQ motif. The final round of PCR was done using CaM and IQ motif genes that were amplified in previous PCR rounds as templates, with a forward primer corresponding to the CaM gene and a reverse primer corresponding to the IQ motif genes. The recombinant proteins were expressed in *Escherichia coli BL21 (DE3)* cells grown in LB broth medium and were purified using Ni-NTA affinity column (Qiagen). Eluted proteins from the Ni-NTA column were further purified by size exclusion chromatography (SEC) (Superdex™75; GE Healthcare).

Characterization of the fused protein and crystallization. SEC showed that the CaM-$(Gly)_5$-Nm IQ motif had a similar elution profile to that of CaM, indicating the existence of a well-folded intact complex (**Fig. 3**). However, the CaM-$(Gly)_5$-Ng IQ motif eluted a little earlier than the CaM, suggesting the possibility of a non-interacting linked peptide. Consistently, the DLS results suggested that CaM and CaM-$(Gly)_5$-Nm-IQ existed as possible monomers in solution, whereas CaM-$(Gly)_5$-Ng-IQ existed in a size between monomer and dimer. Nonetheless, both linked complexes were determined to be homogeneous in solution and thus suitable for crystallization. Interestingly, we observed that the complexes with the longer $(Gly)_8$ linker eluted in the void volume in the SEC, revealing the possible disordered nature of the chimeric proteins. Thus, optimizing the length of the linker is important for obtaining stable protein-protein complexes. However, if the linker length is not sufficient to maintain

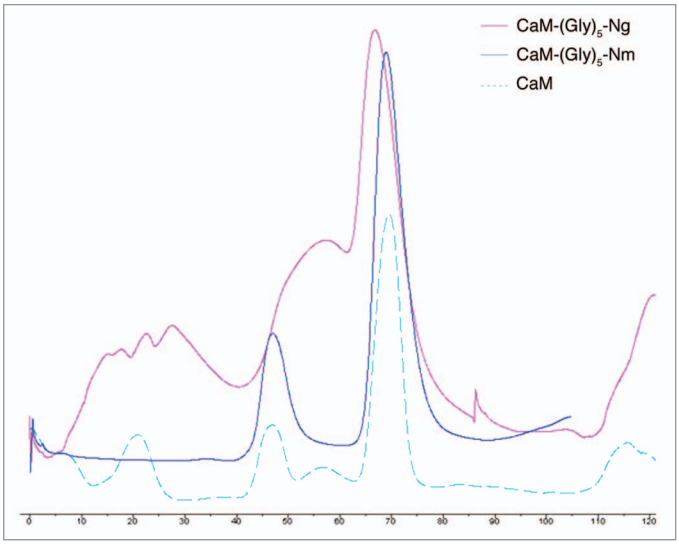

Figure 3. Comparison among Size Exclusion Chromatography elution profiles for CaM, CaM-$(Gly)_5$-Nm and CaM-$(Gly)_5$-Ng using 16/60 Superdex 75 prep grade column. A similar elution profile for CaM and CaM-$(Gly)_5$-Nm shows that the CaM and CaM linked Nm complex have similar hydrodynamic radii and indicates that the linked complex is well-folded. The CaM-$(Gly)_5$-Ng shows a small difference in the elution profile, indicating the possibility of a non-interacting linked complex. This was further verified using dynamic light scattering (DLS) (data not shown).

the natural interactions between the binding partners, the binding partners will adopt intermolecular interactions instead of intramolecular interactions to retain the natural interaction. Similar intermolecular binding was observed between the phosphoprotein (P) and nucleocapsid protein (N) in another study (**Fig. 4B**).[23] Collectively, linkers do not interfere with the natural interactions between the binding partners.[13]

We attempted to crystallize the linked complex in the presence and absence of Ca^{2+}. However, crystals were obtained only for the *apo* form (Ca^{2+}-free form).[16] The co-crystallization of the unlinked CaM and MBR peptides of Ng or Nm did not yield crystals, even in the same conditions in which the linked complex proteins were crystallized. The Seleno-L-methionine-labeled crystals of the linked complexes were produced and the structure was solved by ShelxC/D/E program.[24] The model was manually built using COOT[25] when necessary, and refinement was performed in Refmac5.[26] Both complex structures were refined to R-factors less than 30%, with good stereochemical parameters.[16]

Structures of the linked complexes. The Nm and Ng IQ peptides bound to the C-lobe of CaM and gained a secondary structure. Interestingly, the direction of the bound peptides in these complexes was different (**Fig. 2C**).[16] Moreover, the computational model and the determined crystal structure were different in the case of CaM-Nm complex and are similar in the case of CaM-Ng complex. Nm IQ peptide bound to the C-lobe of CaM is in a different orientation compared with the computational model and CaM-Ng complex (**Fig. 2**). Although the modeling analysis identified the binding region of the peptide based on known CaM-IQ peptide complex structures, the exact orientation and direction of the bound peptide could not be established and this was true in the case of CaM-Nm complex.

Validation. To further confirm that the interactions between CaM and the linked MBR peptides of Ng or Nm were the same as those in the unlinked proteins, the key interacting residues of Ng/Nm and CaM were mutated and validated using in vitro and in vivo experiments.[16] ITC experiments were performed with full-length, unlinked proteins to determine the binding affinity between various mutants of CaM and Ng/Nm. Further, the physiological role of the identified key residues of Ng, which are involved in its interactions with CaM, were determined using electrophysiological experiments. The key residues were identified as Ser36, Ile33, Arg38 of Ng, and Ser41, Phe42 and Arg43 of Nm. These residues were previously shown to be involved in interactions with the full-length, unlinked proteins by us and others.[16,20,27-29] This confirmed that the linker had no impact on dictating the orientation and interactions of the binding partners.

Discussion

Here, we have demonstrated a methodology to understand the transient protein-protein interactions, for which one partner is an intrinsically unstructured protein; i.e., Nm/Ng. Often, the intrinsically unstructured proteins gain secondary structure upon binding with their partners. For instance, this is true for the Myelin Basic Protein (MBP), another CaM binding partner, that is intrinsically disordered on its own and attains secondary structure upon its interaction with CaM.[30]

It is known that CaM interacts with the intrinsically disordered proteins Ng and Nm.[31,32] Based on the literature combined with the structure analysis of CaM-IQ motif complexes and binding studies, MBR peptides from Nm and Ng were derived. Moreover, computational modeling and literature analysis suggested that an appropriate linker length to allow for a natural interaction between CaM and Nm or Ng. These linked constructs were purified, crystallized and their structures were determined and validated.[16]

It is worth mentioning here that the crystal structure of CaM linked to the CBD peptide of calcineurin was previously determined.[17] In this crystal structure, the chimeric protein formed a tight dimer, where each CBD peptide was surrounded by the C-lobe of CaM to which the peptide is fused and the N-lobe of an adjoining molecule.[17] The crystal structure of the CaM-CBD of calcineurin in the absence of a linker was also determined. Both crystal structures (in the presence and absence of linker) illustrated this tight dimer formation.[33,34] However, the solution structure determined using Small Angle X-ray Scattering (SAXS) showed that the complex was a monomer with the peptide wrapped by the N- and C-lobes of a single CaM. Further, the complex required bending of the central helix between the 2 lobes of CaM.[34] A comparison of these 2 structures revealed that domain swapping was required to obtain the crystal structure of the same complex, and it is possible that the dimer crystal structure was stabilized by the crystal contacts.[34] Nonetheless, the interactions identified between CaM and CBD from both solution structure and crystal structure were the same. This further

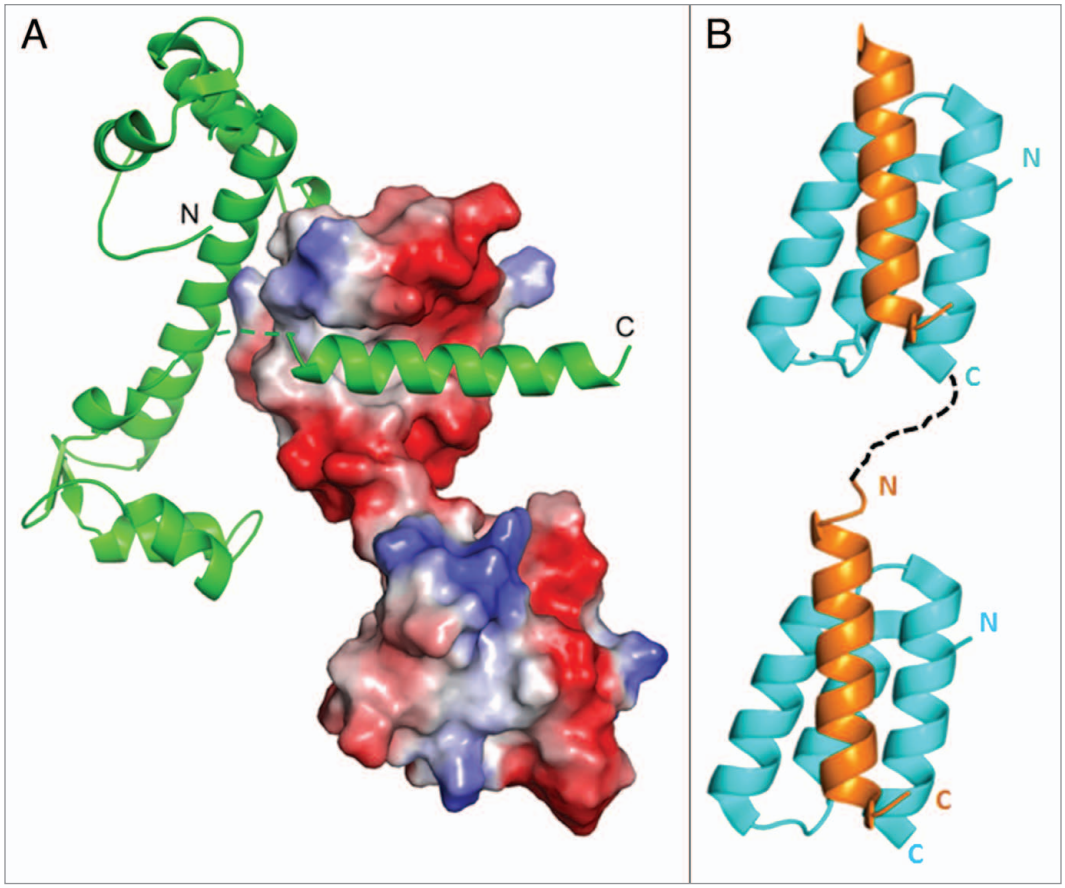

Figure 4. (**A**) Molecular surface and ribbon representations for the binding of Ng MBR peptide on CaM in the CaM-$(Gly)_5$-Ng linked peptide complex (PDB code: 4E50). CaM interacts with the linked MBR peptide of the nearby molecules (symmetry-related molecule) to mimic their natural interactions. Predicted position of the linker is shown as a green dotted line. (**B**) Ribbon representation of phosphoprotein ($P_{457–507}$) of paramyxoviral polymerase (cyan) - nucleocapsid protein ($N_{486–505}$) (orange) (PDB code: 1T6O). C-terminus of $P_{457–507}$ is linked using an 8 amino acid-long Gly-rich linker to the N-terminus of $N_{486–505}$. But, $N_{486–505}$ linked to $P_{457–507}$ of an adjacent symmetry-related molecule is involved in the interaction. Predicted position of the linker is shown as a black dotted line. This clearly indicates that the linker does not restrict the orientation of the binding.

confirmed that the linker did not affect the interactions between the binding partners.

Advantages. There are several advantages of the linked peptide complex method. The first advantage is the most obvious: the ability to trap transient interactions. Most cellular processes involve protein-protein interactions, which are transient in nature[3] and thus are difficult to study at the atomic level. The linked peptide complex method alleviates these concerns by covalently linking the 2 binding partners.

Second, the ease and simplicity of sample preparation in this method means that it is not difficult to replicate. The process requires a three-step PCR procedure to create the linked peptide fusion protein, and purification is simplified by purifying a single construct instead of two different constructs. Furthermore, since the peptide is fused to the protein, it will result in a homogeneous protein-peptide complex with a 1:1 ratio, which will help facilitate the crystallization process. Tying into this, the third advantage is that crystals of the fused peptide are obtained, thus guaranteeing the presence of the linked complex of the 2 interacting species, not just crystallization of one protein. This can then be verified using mass spectrometry and SDS-PAGE on the crystals. In some cases, if the length of the linker is not sufficient to mimic the natural binding, intermolecular protein-peptide interactions can form to stabilize the complex. This was observed in the case of *apo* CaM-$(Gly)_5$-NgIQ, where the NgIQ peptide interacted with CaM from the nearby symmetry-related molecule in the crystal (**Fig. 4A**). Previously, it was reported that where 2 proteins are linked there is a possibility for an intermolecular interaction. Phosphoprotein (P) is a component of viral RNA polymerase involved in facilitating the binding between RNA polymerase and the nucleocapsid protein (N). In order to understand the interaction between these 2 proteins, a glycine-rich linker (GSGSGSGS) was used to fuse the C-terminal of phosphoprotein ($P_{459–507}$) to the N-terminal of nucleocapsid protein ($N_{486–505}$). However, the structure of the complex revealed that, to maintain its natural interactions, the $N_{486–505}$ helix from the adjacent symmetry-related molecule formed intermolecular interactions with $P_{459–507}$ (**Fig. 4B**).[23]

Despite the highly flexible nature of the poly-Gly linker, the length of the linker may prevent the peptide from flipping to interact with its partner. However, as shown in the previous studies and in CaM-Ng IQ complex, if the linker is not long enough

to interact with the specific region of the binding partner it may engage with the nearest molecule in order to retain its natural interactions.

Collectively, this indicates that the linker is not dictating the mode of binding and orientation of the linked peptide; rather, the linked peptide interaction will mimic the binding between the respective full-length independent proteins. If the linker is not properly optimized, it may introduce conformational heterogeneity and will inhibit the crystallization.

Finally, the linked peptide complex method offers a more cost-effective method for crystallization than the independent protein-peptide complex method, which requires purified or commercially synthesized peptides. Besides applying this method for crystallographic studies, this approach may also be employed to study protein interactions using NMR and other spectroscopic methods where trapping the transient/weak complex is also challenging.

Limitations. Although this method has many advantages, there are some limitations too. First, it requires some structural knowledge of at least one of the binding partners in order to design and optimize the linker. In addition, it is essential that there is some knowledge of the site of the interaction on the protein surface, such that the peptide is fused to the correct terminus with an appropriate number of glycine residues in the glycine-rich linker to provide adequate flexibility. The third limitation of this method is the inability to provide any post-translational modifications to the fused protein if it is expressed using a bacterial expression system. While in vitro post-translation methods can be attempted (e.g., phosphorylation), it is recommended that the linked peptide complex be expressed either in yeast or baculovirus expression systems if post-translational modifications are required for the interaction.

The fourth limitation concerns crystallization. Sometimes the crystallization of the linked peptide complex will be challenging due to proteolytic cleavage. This may happen during crystallization, and the chance of only getting one partner in the crystal cannot be ruled out. If this happens, adequate measures can be taken to prevent this proteolytic cleavage. In addition, as the 2 interacting partners are linked, the stoichiometric ratio is limited to 1:1. Thus, this method is not recommended if the ratio is more than 1:1; however, we propose that a 1:2 stoichiometric ratio can be attempted by fusing the peptide to 2 appropriate linkers from both the N and C-terminals of the full-length protein. However, this technique should be approached cautiously and needs further experimental validation.

Materials and Methods

The following sections describe the various steps involved in studying transient and weak protein-protein interactions using our proposed methodology of employing an optimized flexible polypeptide linker (**Fig. 1**). Although the interacting partners vary in size, the interacting region, herein called the minimum binding region (MBR), will be relatively small in most cases. The identification of the MBR from one of the binding partners is a key step for elucidating the binding.

Identification of the minimum binding region. The MBR peptide should be preferably derived from the structurally unknown or relatively unstructured binding partner. To determine the MBRs between the binding partners, deletion studies and point mutational analyses can be employed, based on their existing functional knowledge. Moreover, chemical crosslinking and limited proteolysis, combined with mass spectrometry analysis, can provide information about the various protein domains or regions that are involved in the interaction.[35] Further, the role of these regions in protein-protein interactions can be verified using co-immunoprecipitation, gel-shift assays and biophysical binding studies, such as Surface Plasmon Resonance (SPR) and Isothermal Titration Calorimetry (ITC), by comparing the binding of the wild-type and respective mutant proteins. This would lead to the identification of the MBR from structurally unknown binding partners. It is also recommended to examine the MBR peptides of various lengths to identify the peptide that will most closely mimic the interaction attained by the full-length proteins. Notably, biophysical binding experiments will provide the stoichiometric ratio and binding affinity between the protein and MBR peptide.

Computational analysis and docking studies. Once the boundaries of the MBR peptides are identified through experimental methods, various investigations should be performed to predict the interaction site of the MBR peptide on the surface of the stable protein (other binding partner): this includes, but is not limited to, a literature search, sequence analysis, mutagenesis experiments, docking experiments, limited proteolysis and Hydrogen/Deuterium (H/D) exchange experiments.[36] Moreover, a DALI search[37] against the PDB database using the coordinates of the structurally known binding partner will lead to the identification of other similar protein-peptide complexes and perhaps more information as to the potential site of interaction. Besides, the positioning of the MBR peptide in the appropriate orientation and direction at the site of the interaction requires some clue about the residues from both partners, which are likely to mediate the interaction. The docked complex model of the MBR peptide with the full-length binding partner can be obtained using software such as 3D-Garden,[38] 3D-Dock Suite,[39] and Cluspro 2.0.[40] Previously, studies have used the program Auto Dock[41] to identify the binding sites of MBR peptides on a protein, with or without prior knowledge of their site of interaction.[42] Together, these investigations can be used to predict the approximate distance between the docked peptide at the site of interaction and the N- or C-terminus of the protein, which will ultimately lead to the design of an optimum linker length.

Linking the MBR peptide to the protein. Using a glycine-rich linker, the MBR peptide can be linked to its binding partner either on the N- or C-terminus, whichever is closest to the site of the interaction. Each glycine residue in the linker corresponds to an approximate length of ~3.8 Å.[43] If the site of the interaction is known, an appropriate linker length can be designed to facilitate the positioning of the MBR peptide in the binding region of the stable protein. If the site of the interaction is unknown, various linker lengths can be tested to identify an

appropriate length to facilitate the formation of a stable, intact complex.

Characterization of fused proteins and structure determination. Next, additional experiments should be performed to verify the integrity of the intact complex. Size Exclusion Chromatography (SEC) and Dynamic Light Scattering (DLS) experiments reveal the compactness of the fused protein.[44,45] If required, analytical ultra-centrifugation (AUC) can also be performed to analyze the intact fold of the fused MBR peptide-protein construct. Further, Circular Dichroism (CD) can be used to assess the fold of the fused construct. Any significant changes in the SEC/DLS/AUC/CD profiles compared with that of the unlinked full-length proteins may indicate improper folding of the fused construct. If this is the case, the length of the glycine-rich linker needs to be optimized before any further steps are taken. Characterization of the fused proteins is followed by crystallization and structure determination.

Validation. Once the structure of the linked complex is obtained, it is necessary to perform structure-guided validation experiments to confirm that the observed interactions are the same as the full-length unlinked protein-protein complex using in vitro and in vivo experiments. The key interacting residues should be mutated to study their implication in the interaction. This could be verified using various in vitro methods, such as ITC, SPR and/or pull-down assays. Besides, the appropriate in vivo studies should be performed to validate the role of the key residues in the protein-protein interaction.

Conclusions

We believe that the advantages of this method offer an excellent solution for trapping and characterizing transient and weak complexes that are otherwise unattainable. Here, we have demonstrated for the first time the study of the interactions of 2 intrinsically disordered proteins, both of which participate in transient binding with their partner and only gain secondary structure upon binding. This method has wide application for studying similar, transient protein-protein interactions involved in various catalytic and biological processes that have remained, until now, a challenge for structural studies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed. The authors declare no competing financial interests.

Acknowledgments

We are grateful to the Academic Research Funding (AcRF grant No. R154000563112), National University of Singapore (NUS), Singapore, for the support of this study. VPR is a graduate scholar in receipt of a research scholarship from the NUS.

References

1. Phizicky EM, Fields S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev 1995; 59:94-123; PMID:7708014.
2. Ozbabacan SE, Engin HB, Gursoy A, Keskin O. Transient protein-protein interactions. Protein Eng Des Sel 2011; 24:635-48; PMID:21676899; http://dx.doi.org/10.1093/protein/gzr025.
3. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. Curr Opin Struct Biol 2002; 12:54-60; PMID:11839490; http://dx.doi.org/10.1016/S0959-440X(02)00289-0.
4. Back JW, de Jong L, Muijsers AO, de Koster CG. Chemical cross-linking and mass spectrometry for protein structural modeling. J Mol Biol 2003; 331:303-13; PMID:12888339; http://dx.doi.org/10.1016/S0022-2836(03)00721-6.
5. Sinz A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrom Rev 2006; 25:663-82; PMID:16477643; http://dx.doi.org/10.1002/mas.20082.
6. Pierce Biotechnology I. Crosslinking Reagents Technical Handbook. 2006.
7. Trakselis MA, Alley SC, Ishmael FT. Identification and mapping of protein-protein interactions by a combination of cross-linking, cleavage, and proteomics. Bioconjug Chem 2005; 16:741-50; PMID:16029014; http://dx.doi.org/10.1021/bc050043a.
8. Alam SM, Travers PJ, Wung JL, Nasholds W, Redpath S, Jameson SC, et al. T-cell-receptor affinity and thymocyte positive selection. Nature 1996; 381:616-20; PMID:8637599; http://dx.doi.org/10.1038/381616a0.
9. Hennecke J, Carfi A, Wiley DC. Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. EMBO J 2000; 19:5611-24; PMID:11060013; http://dx.doi.org/10.1093/emboj/19.21.5611.
10. Hennecke J, Wiley DC. Structure of a complex of the human alpha/beta T cell receptor (TCR) HA1.7, influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): insight into TCR cross-restriction and alloreactivity. J Exp Med 2002; 195:571-81; PMID:11877480; http://dx.doi.org/10.1084/jem.20011194.
11. Li Y, Huang Y, Lue J, Quandt JA, Martin R, Mariuzza RA. Structure of a human autoimmune TCR bound to a myelin basic protein self-peptide and a multiple sclerosis-associated MHC class II molecule. EMBO J 2005; 24:2968-79; PMID:16079912; http://dx.doi.org/10.1038/sj.emboj.7600771.
12. Sethi DK, Schubert DA, Anders AK, Heroux A, Bonsor DA, Thomas CP, et al. A highly tilted binding mode by a self-reactive T cell receptor results in altered engagement of peptide and MHC. J Exp Med 2011; 208:91-102; PMID:21199956; http://dx.doi.org/10.1084/jem.20100725.
13. Reddy Chichili VP, Kumar V, Sivaraman J. Linkers in the structural biology of protein-protein interactions. Protein Sci 2013; 22:153-67; PMID:23225024; http://dx.doi.org/10.1002/pro.2206.
14. Cui Y, Wen J, Hung Sze K, Man D, Lin D, Liu M, et al. Interaction between calcium-free calmodulin and IQ motif of neurogranin studied by nuclear magnetic resonance spectroscopy. Anal Biochem 2003; 315:175-82; PMID:12689827; http://dx.doi.org/10.1016/S0003-2697(03)00007-1.
15. Chapman ER, Au D, Alexander KA, Nicolson TA, Storm DR. Characterization of the calmodulin binding domain of neuromodulin. Functional significance of serine 41 and phenylalanine 42. J Biol Chem 1991; 266:207-13; PMID:1824693.
16. Kumar V, Chichili VP, Zhong L, Tang X, Velazquez-Campoy A, Sheu FS, et al. Structural basis for the interaction of unstructured neuron specific substrates neuromodulin and neurogranin with Calmodulin. Sci Rep 2013; 3:1392; PMID:23462742; http://dx.doi.org/10.1038/srep01392.
17. Ye Q, Li X, Wong A, Wei Q, Jia Z. Structure of calmodulin bound to a calcineurin peptide: a new way of making an old binding mode. Biochemistry 2006; 45:738-45; PMID:16411749; http://dx.doi.org/10.1021/bi0521801.
18. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997; 18:2714-23; PMID:9504803; http://dx.doi.org/10.1002/elps.1150181505.
19. Houdusse A, Gaucher JF, Krementsova E, Mui S, Trybus KM, Cohen C. Crystal structure of apo-calmodulin bound to the first two IQ motifs of myosin V reveals essential recognition features. Proc Natl Acad Sci U S A 2006; 103:19326-31; PMID:17151196; http://dx.doi.org/10.1073/pnas.0609436103.
20. Ran X, Miao HH, Sheu FS, Yang D. Structural and dynamic characterization of a neuron-specific protein kinase C substrate, neurogranin. Biochemistry 2003; 42:5143-50; PMID:12718558; http://dx.doi.org/10.1021/bi0271751.
21. Cui Y, Wen J, Hung Sze K, Man D, Lin D, Liu M, et al. Interaction between calcium-free calmodulin and IQ motif of neurogranin studied by nuclear magnetic resonance spectroscopy. Anal Biochem 2003; 315:175-82; PMID:12689827; http://dx.doi.org/10.1016/S0003-2697(03)00007-1.
22. Zhang M, Vogel HJ, Zwiers H. Nuclear magnetic resonance studies of the structure of B50/neuromodulin and its interaction with calmodulin. Biochem Cell Biol 1994; 72:109-16; PMID:7818844; http://dx.doi.org/10.1139/o94-017.
23. Kingston RL, Hamel DJ, Gay LS, Dahlquist FW, Matthews BW. Structural basis for the attachment of a paramyxoviral polymerase to its template. Proc Natl Acad Sci U S A 2004; 101:8301-6; PMID:15159535; http://dx.doi.org/10.1073/pnas.0402690101.
24. Sheldrick GM. A short history of SHELX. Acta Crystallogr A 2008; 64:112-22; PMID:18156677; http://dx.doi.org/10.1107/S0108767307043930.

25. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004; 60:2126-32; PMID:15572765; http://dx.doi.org/10.1107/S0907444904019158.
26. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr D Biol Crystallogr 2004; 60:2184-95; PMID:15572771; http://dx.doi.org/10.1107/S0907444904023510.
27. Prichard L, Deloulme JC, Storm DR. Interactions between neurogranin and calmodulin in vivo. J Biol Chem 1999; 274:7689-94; PMID:10075657; http://dx.doi.org/10.1074/jbc.274.12.7689.
28. Chapman ER, Au D, Alexander KA, Nicolson TA, Storm DR. Characterization of the calmodulin binding domain of neuromodulin. Functional significance of serine 41 and phenylalanine 42. J Biol Chem 1991; 266:207-13; PMID:1824693.
29. Cui Y, Wen J, Hung Sze K, Man D, Lin D, Liu M, et al. Interaction between calcium-free calmodulin and IQ motif of neurogranin studied by nuclear magnetic resonance spectroscopy. Anal Biochem 2003; 315:175-82; PMID:12689827; http://dx.doi.org/10.1016/S0003-2697(03)00007-1.
30. Majava V, Petoukhov MV, Hayashi N, Pirilä P, Svergun DI, Kursula P. Interaction between the C-terminal region of human myelin basic protein and calmodulin: analysis of complex formation and solution structure. BMC Struct Biol 2008; 8:10; PMID:18284662; http://dx.doi.org/10.1186/1472-6807-8-10.
31. Andreasen TJ, Luetje CW, Heideman W, Storm DR. Purification of a novel calmodulin binding protein from bovine cerebral cortex membranes. Biochemistry 1983; 22:4615-8; PMID:6313045; http://dx.doi.org/10.1021/bi00289a001.
32. Neuner-Jehle M, Denizot JP, Mallet J. Neurogranin is locally concentrated in rat cortical and hippocampal neurons. Brain Res 1996; 733:149-54; PMID:8891262.
33. Ye Q, Wang H, Zheng J, Wei Q, Jia Z. The complex structure of calmodulin bound to a calcineurin peptide. Proteins 2008; 73:19-27; PMID:18384083; http://dx.doi.org/10.1002/prot.22032.
34. Majava V, Kursula P. Domain swapping and different oligomeric States for the complex between calmodulin and the calmodulin-binding domain of calcineurin a. PLoS One 2009; 4:e5402; PMID:19404396; http://dx.doi.org/10.1371/journal.pone.0005402.
35. Renzone G, Salzano AM, Arena S, D'Ambrosio C, Scaloni A. Mass Spectrometry-Based Approaches for Structural Studies on Protein Complexes at Low-Resolution. Current Proteomics 2007; 4:1-16; http://dx.doi.org/10.2174/157016407781387384.
36. Englander SW, Sosnick TR, Englander JJ, Mayne L. Mechanisms and uses of hydrogen exchange. Curr Opin Struct Biol 1996; 6:18-23; PMID:8696968; http://dx.doi.org/10.1016/S0959-440X(96)80090-X.
37. Holm L. Sander C. Protein structure comparison by alignment of distance matrices. J Mol Bio 1993; 233:123-38; PMID:8377180; http://dx.doi.org/10.1006/jmbi.1993.1489.
38. Lesk VI, Sternberg MJ. 3D-Garden: a system for modelling protein-protein complexes based on conformational refinement of ensembles generated with the marching cubes algorithm. Bioinformatics 2008; 24:1137-44; PMID:18326508; http://dx.doi.org/10.1093/bioinformatics/btn093.
39. Smith GR, Sternberg MJ; CAPRI blind trial. Evaluation of the 3D-Dock protein docking suite in rounds 1 and 2 of the CAPRI blind trial. Proteins 2003; 52:74-9; PMID:12784370; http://dx.doi.org/10.1002/prot.10396.
40. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 2004; 20:45-50; PMID:14693807; http://dx.doi.org/10.1093/bioinformatics/btg371.
41. Morris GMG, Halliday DS, Huey RS, Hart R, Belew WE, Olson AJ. Automated docking using a Lamarckian genetic algorithm andan empirical binding free energy function. J Comput Chem 1998; 19:1639-62; http://dx.doi.org/10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.
42. Hetényi C, van der Spoel D. Efficient docking of peptides to proteins without prior knowledge of the binding site. Protein Sci 2002; 11:1729-37; PMID:12070326; http://dx.doi.org/10.1110/ps.0202302.
43. Iwakura M, Nakamura T. Effects of the length of a glycine linker connecting the N-and C-termini of a circularly permuted dihydrofolate reductase. Protein Eng 1998; 11:707-13; PMID:9749924; http://dx.doi.org/10.1093/protein/11.8.707.
44. Dev S, Surolia A. Dynamic light scattering study of peanut agglutinin: size, shape and urea denaturation. J Biosci 2006; 31:551-6; PMID:17301492; http://dx.doi.org/10.1007/BF02708406.
45. Uversky VN. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. Biochemistry 1993; 32:13288-98; PMID:8241185; http://dx.doi.org/10.1021/bi00211a042.