

**Conformational flexibility of GRASP proteins and their constituent
PDZ subdomains reveals structural basis of their promiscuous
interactome**

Luis Felipe S. Mendes^{1,2}; Mariana R. B. Batista¹; Peter J. Judge²; Anthony Watts²; Christina Redfield²; Antonio J. Costa-Filho¹

¹ Molecular Biophysics Laboratory, Ribeirão Preto School of Philosophy, Sciences and Literature, Physics Department, University of São Paulo, Ribeirão Preto – Brazil

² Department of Biochemistry, University of Oxford, South Parks Road, Oxford – UK. OX1 3QU

To whom correspondence should be addressed: Antonio J. Costa-Filho, DF/FFCLRP/USP, Av. Bandeirantes 3900, Ribeirão Preto, SP, Brazil, 14040-901, E-mail: ajcosta@usp.br, Phone: +551633153665.

Running title: GRASP flexibility and its promiscuous interactome

1	Abbreviations
2	αPS1- Integrin alpha-PS1
3	ABF – Adaptive Biasing Force
4	APC - Adenomatous Polyposis Coli
5	CD83 - Ig-like domain-containing protein
6	CFTR – Cystic fibrosis transmembrane conductance regulator
7	DGRASP – GRASP domain
8	ER – Endoplasmic Reticulum
9	IL1-β - Interleukin 1β
10	GRASP – Golgi Reassembly and Stacking Protein
11	GRIP1 - Glutamate receptor-interacting protein 1
12	GXM - Glucuronoxylomannan
13	HSQC – Heteronuclear Single Quantum Coherence
14	JAM – Junctional adhesion molecule
15	LAMP2 - Lysosome-associated membrane glycoprotein 2
16	LC3-II - microtubule-associated protein light chain 3
17	NMR – Nuclear Magnetic Resonance
18	PDZ – <u>P</u>SD-95, <u>D</u>LG and <u>Z</u>O-1
19	PTP-BL - Protein Tyrosine Phosphatase
20	PTPN13 - Tyrosine-protein phosphatase non-receptor type 13
21	RAB2 – Ras-related protein 2
22	SEC16A – Protein transport protein Sec16A
23	SRCD – Synchrotron Radiation Circular Dichroism
24	TOCSY – Total Correlation Spectroscopy
25	TGFα - Transforming Growth Factor-α

Abstract

The Golgi complex is a central component of the secretory pathway, responsible for several critical cellular functions in eukaryotes. The complex is organized by the Golgi matrix that includes the Golgi Reassembly and Stacking Protein (GRASP), which was shown to be involved in cisternae stacking and lateral linkage in metazoan. GRASPs also have critical roles in other processes, with an unusual ability to interact with several different binding partners. The conserved N-terminus of the GRASP family includes two PDZ domains. Previous crystallographic studies of orthologues suggest that PDZ1 and PDZ2 have similar conformations and secondary structure content. However PDZ1 alone mediates nearly all interactions between GRASPs and their partners. In this work, NMR, Synchrotron-Radiation Circular Dichroism and Molecular Dynamics were used to examine the structure, flexibility and stability of the two constituent PDZ domains. GRASP PDZs are structured in an unusual $\beta_3\alpha_1\beta_4\beta_5\alpha_2\beta_6\beta_1\beta_2$ secondary structural arrangement and NMR data indicates that the PDZ1 binding pocket is formed by a stable β_2 -strand and a more flexible and unstable α_2 -helix, suggesting an explanation for the higher PDZ1 promiscuity. The conformational free energy profiles of the two PDZ domains were calculated using Molecular Dynamics simulations. The data suggest that, after binding, the protein partner significantly reduces the conformational space that GRASPs can access by stabilizing one particular conformation, in a partner-dependent fashion. The structural flexibility of PDZ1, modulated by PDZ2, and the coupled, coordinated movement between the two PDZs enable GRASPs to interact with multiple partners, allowing them to function as promiscuous, multitasking proteins.

Keywords: GRASP, NMR, ABF simulation, PDZ asymmetry, Molten Globule

1 Introduction

2
3 The Golgi complex is a membranous organelle located in the heart of the secretory pathway,
4 responsible for several critical cellular functions [1,2]. In vertebrates, this organelle is assembled
5 as a stack of flattened cisternae that can be also laterally connected in a ribbon-like shape [1].
6 Visually, the Golgi complex is one of the most immediately recognizable organelles inside higher
7 eukaryotic cells. Proteins that constitute the so-called Golgi matrix hold the complex together;
8 although the mechanism by which they accomplish this task is still not fully understood [3]. It is
9 likely that some proteins belonging to the Golgi matrix were present in the last eukaryotic
10 common ancestor (LECA), suggesting that eukaryotic cells might share some of the mechanisms
11 controlling the Golgi structure [4]. One of the most important proteins of the Golgi matrix
12 (probably reminiscent from LECA [4]) is the family of peripheral membrane proteins called Golgi
13 Reassembly and Stacking Proteins (GRASPs) [5, 6].

14 GRASPs were first observed to be essential factors for Golgi reassembly *in vitro* [7,8]
15 and *in vivo* [9]. They participate in the lateral linkage of Golgi stacks in vertebrates (building the
16 Golgi ribbon), and in the Golgi remodelling needed in migrating cells [10]. Current models
17 suggest that GRASP-mediated Golgi stacking occurs via GRASP dimerization in a *trans*
18 orientation [11], whereby a monomeric GRASP on one cistern face interacts with a second
19 GRASP monomer on an opposite cistern [11]. The mechanism of its oligomerization and whether
20 GRASPs have a decisive role to play in Golgi stacking are still unclear and topics of intense
21 debate [11,12,13].

22 Although first identified as potential stacking factors, other functions of GRASPs have
23 also been observed. GRASP can interact directly with the C-terminus of the Transforming Growth
24 Factor- α (TGF- α) and with several members of the p24/TMED family, facilitating the
25 conventional protein secretion pathway [14,15]. GRASPs also have a direct and essential role in
26 the unconventional secretion of a large number of proteins, such as: (1) the soluble acyl-coenzyme
27 A binding protein (ACBP) in *Dictyostelium* [16], *Saccharomyces cerevisiae* and *Pichia pastoris*
28 [17,18], (2) the α PS1 integrin during the stage 10B of *Drosophila* embryogenesis [19], (3) the
29 mutant Δ F508 of the cystic fibrosis transmembrane conductance regulator (CFTR) in an ER stress
30 situation [20], (4) the IL1- β [21], among others [22,23]. GRASP is apparently also required for
31 the export of non-protein molecules. In *Cryptococcus neoformans*, deletion of GRASP
32 (CnGRASP) results in inefficient secretion of its main virulence factor GXM, followed by
33 reduced capsule formation and attenuated virulence [24].

34
35 Although many functions associated with GRASPs have been identified, their detailed
36 molecular mechanisms are still unclear. The crystal structures of the N-terminal domain of *Homo*
37 *sapiens* GRASP55 (25), *Rattus norvegicus* GRASP55 (26) and *Rattus norvegicus* GRASP65 (26)

1 have only recently been reported. GRASPs are composed of two structurally similar PDZ domains
2 (named after the three first PDZ-proteins found: PSD-95, DLG and ZO-1) with an atypical
3 protozoa-like arrangement, which together comprise the so-called GRASP domain (DGRASP).
4 This finding also suggests that a new sequence analysis criterion might reveal a host of
5 unidentified eukaryotic PDZ domains [25]. A second region in the GRASP structure consists of
6 a poorly conserved domain that is rich in serine and proline (SPR domain) and that has a
7 regulatory function [27,28] and intrinsically disordered characteristics [29].

8
9 The PDZ domains have been classically divided in different subfamilies according to
10 their preferential binding partner [30]. The affinity for the C-terminal end is primarily dictated by
11 the carboxy residue and by the P (-2) residue of the protein partner sequence, leading to the PDZ
12 subfamily classification based on the three last aminoacids of the binding partner [31,32]. Class
13 I PDZ domains are known to recognize C-terminal peptides with Thr/Ser-X-Ψ (with X and Ψ
14 being any residue and a hydrophobic residue, respectively) motifs, whereas class II and III PDZ
15 domains preferentially bind Ψ -X- Ψ and Asp/Glu-X-Ψ motifs, respectively [30]. One of the most
16 intriguing properties of GRASPs is their ability to interact with a large variety of different protein
17 partners, which could then be divided according to that subfamily classification (Figure 1). In
18 addition to the Golgins, mammalian GRASPs also interact with cargo receptors from the
19 conventional secretory pathway [14] and form part of a complex with RAB2, essential for normal
20 protein transport and Golgi structure [33]. GRASPs have been also shown to interact with: TGF-
21 α [15], the potassium channel tetramerization domain-containing protein 5 (KCTD5) [34],
22 SEC16A and ΔF508 CFTR in response to ER stress [35,20], JAM-B and JAM-C during
23 spermatogenesis [36], LC3-II on the autophagosomes and LAMP2 on late endosomes/lysosomes
24 [37]. Besides, mammalian GRASPs can interact with CD83 in dendritic cells [38], with the
25 Membrane-type 1 matrix metalloproteinase [39], furin [39] and also with itself during
26 dimerization (involving PDZ1 binding pocket and a PDZ2 internal loop) [27]. The numerous
27 binding partners of GRASP55 span all three classes of the PDZ binding partner classification,
28 while GRASP65 would fit as classes I and II (Figure 1). These interactions preferentially involve
29 the PDZ1 domain rather than PDZ2, despite the structural similarity seen in the reported crystal
30 structures [12,13,20,26,36]. PDZ1 has been reported to bind even to a non-native truncated SPR
31 C-terminus [26].

32
33 It has been also observed that PDZ-tandem proteins might have different allosteric effects
34 after binding of a partner to one of the PDZs [40,41]. For instance, the modified binding
35 preference of PTPN13 PDZ2 domain for APC could be explained by an allosteric modulation
36 induced by the interaction of PDZ3 with PDZ2 [41]. Moreover, the potential of PTP-BL PDZ2
37 domain to interact with class III ligands was found to be modulated by the presence of PDZ1,

1 suggesting that the PDZ microenvironment is determinant for their ligand binding specificity
2 [42,43]. This allosteric control seems to take place even with recent evidence suggesting that
3 PDZ-tandem protein have structurally independent PDZs within the full-length scaffold protein
4 [44]. Therefore, how PDZs connected in tandem are organized and what is their impact on each
5 other functions are still open issues.

6
7 In a previous report, we showed that GRASPs from the fungal pathogen *C. neoformans*
8 [29] and *S. cerevisiae* [45] present structural features usually observed in molten globule proteins,
9 which exist even in the absence of any mild denaturing conditions. As expected for a collapsed
10 intrinsically disordered protein, many physicochemical perturbations induced by the crystal
11 packing (including local dehydration and changes in the dielectric constant, along with increased
12 PEG concentration used in the crystallographic reservoir) promote disorder-to-order transitions
13 in GRASPs [45,46]. Even though GRASPs appear to have a special affinity for di-valine-
14 containing C-terminal proteins [28], the overall sequences of their binding partners do not have a
15 conserved sequence or structural feature in common (Figure 1), suggesting GRASPs might be
16 highly promiscuous in terms of protein-protein interactions. Therefore, this apparent functional
17 asymmetry of the PDZs and their contrasting promiscuity of interactions with binding partners
18 might benefit from detailed studies of GRASPs in solution.

19
20 Since previous published data have not given a clear rationale as to why the two PDZ
21 subdomains behave so differently *in vivo*, we used high-field solution-state nuclear magnetic
22 resonance (NMR), synchrotron radiation circular dichroism (SRCD), conventional circular
23 dichroism (CD), steady-state fluorescence and molecular dynamics (MD) simulations to probe
24 the flexibility and stability of the different subdomains of the GRASP structure. Our data show
25 that the two PDZs behave differently in solution, with most of the flexible and promiscuous
26 regions located right in the binding pocket of PDZ1, whereas PDZ2 is significantly more ordered.
27 The coordinated movement between the two subdomains covers a wide conformational space,
28 accessible from the free energy present at physiological temperatures, which collapses to a small
29 area after binding with the protein partner. We propose that the orchestrated movement of the
30 PDZs is used by the protein to control access to the PDZ1 binding pocket, and an array of
31 secondary interactions with PDZ2 dictates the overall stability and the strength of the protein-
32 ligand complex. Our data give a convincing explanation as to why PDZ1 is capable of mediating
33 almost all the interactions with other protein partners, and suggest a new role for GRASPs as
34 highly promiscuous, multitasking proteins.

Results and Discussion

DGRASP as a molten globule-like protein

¹⁵N- or ¹⁵N/¹³C-labeled and unlabeled samples were prepared using the protocol described in the Materials and Methods section and protein purity was monitored using SDS-PAGE (Figure 2A). CnGRASP, like the orthologue GRASP from *S. cerevisiae* [45], has previously been identified as a molten globule-like protein in solution, and belongs to the family of collapsed intrinsically disordered proteins [29]. We have previously suggested that the isolated DGRASP domain of CnGRASP shares some of the molten globule features observed in the full-length structure [29]. Here we extended that analysis by showing that molten globule-like behavior is indeed observed for the isolated GRASP domain (DGRASP) (Figure 2). In the absence of urea, native DGRASP bound the fluorophore ANS with high affinity, compared to the fully unfolded protein (Figure 2B), consistent with the partitioning of ANS molecules into the hydrophobic interior in the native state of the protein. The gradual changes in the ANS fluorescence spectra at different urea concentrations indicate that the hydrophobic interior unfolds in a linear manner, exhibiting low cooperativity (Figure 2C). This low cooperativity during unfolding was also observed in the loss of regular secondary structure, as monitored by CD (Figure 2D) and in the change of the tryptophan fluorescence anisotropy (Figure 2E). Our data indicate that DGRASP is a molten globule-like protein, and that its three-dimensional structure unfolds in a low-cooperative way.

Although molten globules may have significant secondary structure content, the side chains in the hydrophobic cores are not well packed and their three-dimensional conformation structure fluctuates over a characteristic timescale of μ s-ms [47,48]. Intrinsically disordered proteins are frequently observed as important components of the cellular signaling machinery [49] and may function as central hubs [50], especially due to their unique structural plasticity [47]. It is becoming clear that the understanding of the structural behavior of molten globules in solution is of pharmacological importance. For example, prion-forming sequences, which are especially enriched in asparagine, have been shown to promote molten globule-like structures, in which amyloid-nucleating contacts can be made [51,52]. Mutant forms of the tumor suppressor protein, p53, have recently been shown to behave as molten globules while exhibiting prion-like properties and, although its standard functions are strongly connected to tumor suppression, p53 mutants and aggregates are involved in cancer progression [53,54].

PDZ asymmetry

1 It has been previously observed that PDZs in tandem might assist each other during
2 folding [55,56]. For example, for the PDZ12 tandem of the GRIP1 PDZ1–3 cassette, the folding
3 of PDZ1 strictly depends on the covalent attachment to PDZ2 [55]. In addition, PDZ5 of GRIP is
4 completely unstructured in solution but, when covalently bound to PDZ4, both PDZ domains
5 become well folded and stable [56]. We explored the behavior of the PDZs of DGRASP in
6 solution using solution NMR methods. The DGRASP ^1H - ^{15}N HSQC spectrum showed around
7 10% of the total expected resonance lines, as previously observed for other induced molten
8 globule-proteins (Figure 3 – 0 M urea) [48,57]. A broad, poorly dispersed NMR spectrum is
9 characteristic of a molten globule due to the structural dynamics in the μs -ms timescale, which
10 makes most of the resonance lines broadened beyond the limit to be detected by the technique
11 [58]. Furthermore, the observable peaks in DGRASP ^1H - ^{15}N HSQC spectrum showed low proton
12 dispersion (separation of less than 0.7 ppm), a strong indication that these peaks might be located
13 in disordered, highly dynamic regions.

14
15 The absence of resonance lines in the ^1H - ^{15}N HSQC spectrum of DGRASP is consistent with
16 the presence of a molten globule-like fluctuating structure in the characteristic timescale window
17 of μs -ms [48]. To recover the missing peaks, the structure is disrupted by increasing urea
18 concentration, to either increase the local dynamics or decrease the number of “conformational
19 states” the protein visits. The strategy of using urea unfolding to gain structural information by
20 NMR has been previously used to study the molten globule states of α -lactalbumin and p53
21 [59,60].

22
23 Urea was titrated from 0 to 9 M in 1 M steps and selected ^1H - ^{15}N HSQC spectra are shown
24 in Figure 3. The resonance peaks start to appear during the titration indicating that the
25 corresponding region is denaturated in that particular condition. Two groups of peaks appeared at
26 different urea concentrations: one around 3 M and another one at ~ 6 M, suggesting there are at
27 least two different regions of the protein, whose unfolding is independent and exhibits moderate
28 cooperativity (Figure 3). The maximum number of resonance lines appeared at 9 M urea (Figure
29 3 – 9 M urea), so this condition was chosen for the assignment of the NMR spectrum.

30
31 The DGRASP resonance assignments in 9 M urea were based on the two pairs of ^1H - ^{15}N -
32 ^{13}C 3D experiments for sequential linking CBCA(CO)NH + CBCANH and HN(CA)CO + HNCO,
33 in addition to HSQC-TOCSY, HSQC-NOESY-HSQC, (H)CC(CO)NH and
34 HBHA(CBCACO)NH. It was possible to unambiguously assign 96% of the total non-proline HN
35 resonances, together with 95%, 94.5% and 94.4% of the CO, $\text{C}\alpha$ and $\text{C}\beta$, respectively (Figure 4).
36 The assignment transfer to the remaining nine conditions (from 9 to 0 M urea) was done by
37 spectral superposition. Based on this logical approach, all the assignments that could be visually

1 tracked were transferred through the titration spectra. This approach has been successfully used
2 with other non-native molten globule structures [59,60].

3
4 Not all peaks were present at the very beginning of the urea titration and only started to
5 appear as the urea concentration was raised, a common feature of molten globule states (Figure
6 3). The gradual appearance of the peaks is attributed to the unfolding of the corresponding region
7 and all the analyses from here on rely on this assumption: if the resonance peak of a non-proline
8 residue X was only observed in the ^1H - ^{15}N HSQC at Y M urea, this is the urea concentration
9 where this residue is now assumed to be in a fully unfolded structure. This assumption is valid,
10 since these peaks are all located in a very narrow chemical shift region close to 8 ppm in the
11 proton dimension, typical of unfolded structures and intrinsically disordered proteins [61].

12
13 To further explore our data, we monitored every peak that could be unambiguously
14 tracked during the titration, and the urea concentration where each peak appeared during the
15 titration was plotted as a function of the amino acid sequence, giving the local unfolding plot
16 (Figure 5A). Figure 5A clearly shows the differences in the unfolding patterns of the two PDZs
17 in solution, in contrast to the available crystal structures, which suggest that they are structurally
18 identical to each other [13,25]. In general, ~60% of the PDZ1 resonances appear at urea
19 concentrations below 2 M, and the average urea concentration for the unfolding of each amino
20 acid is 2 M (Figure 5A). The most stable part of this subdomain seems to be the region between
21 the amino acids 70 and 95 with urea stability of ca. 4 M. According to Figure 5B, the second
22 region of higher stability in PDZ1 is the one in close contact with the PDZ2, which has similar
23 overall stability. This data suggest that the contact with PDZ2 might enhance the structural
24 stability of some PDZ1 regions, especially those at the interface of the PDZs. PDZ1 unfolding is
25 therefore weakly cooperative and at least two main unfolding transitions are expected.

26
27 Eukaryote PDZs are formed by a $\beta_1\beta_2\beta_3\alpha_1\beta_4\beta_5\alpha_2\beta_6$ secondary structure arrangement [62],
28 but GRASP PDZs have a circular permutation where the first two β -strands are located at the end
29 of the sequence ($\beta_3\alpha_1\beta_4\beta_5\alpha_2\beta_6\beta_1\beta_2$) [13,25]. The PDZ binding groove is therefore formed by the
30 final, rather than the second, β -strand within the fold [25]. PDZ binding grooves are structured
31 into a short α -helix (α_2) coupled with the abovementioned β_2 and our NMR data therefore suggest
32 that the PDZ1 binding pocket is formed by a stable β -strand and a more flexible and unstable α -
33 helix (Figure 5B). Since the PDZ interaction mechanism involves the formation of a β -sheet
34 between the β_2 and the disordered C-terminus of the protein partner [63], a more flexible α -helix
35 might be used to control the overall accessibility to this domain.

1 The opposite behavior was observed for PDZ2, where most of the resonance peaks
2 appeared at 5-7 M urea, in a much sharper transition (Figure 5A-B). The average urea
3 concentration for PDZ2 unfolding is 5.7 M. It is clear that the GRASP domain is formed by the
4 much more stable PDZ2 and a PDZ1 with at least two regions of different stability, where the
5 more unstable one is predominant (>60%) (Figure 5B).

6
7 The existing crystal structure of the human GRASP65 GRASP domain (DGRASP65)
8 indicates that PDZ1 has more conformational freedom than PDZ2 based on the B-factor values
9 (Figure 5C). Instability during chemical unfolding may be closely related to structural flexibility,
10 especially in the region close to and surrounding the binding pocket of PDZ1.

11 12 13 *Dissecting the GRASP domain*

14 Using the domain predictor Pfam (<http://pfam.xfam.org>, accessed in 2017), the disorder
15 prediction published previously [29] and the information from our own solution NMR
16 experiments, we designed constructs for each individual PDZ domain. Residues 1-115 were
17 assigned to “PDZ1” and 116-220 to “PDZ2” and the two constructs were expressed and purified
18 by following the same protocol used for the full-length DGRASP.

19
20 SRCD data indicate that the PDZs have different secondary structure content (Figure 6A,
21 Table 1). In solution, PDZ1 possesses lower helical content (approximately three times less than
22 PDZ2) and more strand structures, in contrast to predictions based on the crystal structures, in
23 which similar secondary structure content was observed [25,26]. It has been previously shown
24 that GRASP structure is sensitive to changes in the physicochemical parameters of the medium,
25 in particular the variation in the dielectric constant. [46]. High PEG concentrations have also been
26 shown to induce disorder-to-order transitions in GRASP, so we could expect that more disordered
27 regions might become more ordered in conditions used for crystallization [46]. Interestingly,
28 when we compare the secondary structure contents previously obtained for the DGRASP domain
29 with those found for its PDZ subdomains, the amount of helices and strands for the DGRASP are
30 closer to those found for PDZ2 (Table 1). The most remarkable differences are a decrease in turns
31 and an increase of disordered structures. The data suggest that the DGRASP is not simply
32 constituted by putting together two independent PDZ subdomains. The assembly of the PDZs in
33 the DGRASP somehow leads to more disordered regions. Since the amount of helices and strands
34 in DGRASP are similar to PDZ2, we could speculate that PDZ1 is the subdomain that would get
35 more disordered upon assembly of the full GRASP domain.

Changes in protein tertiary structure may be monitored using near-UV (250-350 nm) CD spectra [64]. If a protein lacks a well-defined structure (e.g. IDPs or "molten-globule" structures), the signals in the near-UV region will be close to zero [65]. On the other hand, the presence of significant near-UV signals is a good indication that the protein is folded into a well-defined tertiary structure [64]. PDZ1 (PDZ2) has 4 (6) tyrosines, 4 (3) phenylalanines, and 1 (3) tryptophans. Therefore, both PDZs have comparable number of aromatic amino acids that are reasonably distributed along the primary structure and that can be used as good probes in near-UV CD experiments. The PDZ2 near-UV CD spectrum has multiple peaks in the region analyzed, suggesting that it has a more well-defined tertiary structure (Figure 6B). The same pattern is not observed for PDZ1, where only a bump is present in the tryptophan region, probably due to its single Trp being in a more collapsed region. The near-UV CD spectra therefore suggest that the isolated PDZ2 has a more well-defined tertiary structure than PDZ1.

Our data so far have revealed some important differences between PDZ1 and PDZ2 that could not be anticipated from the existing crystal structures. Although DGRASP unfolds in a non-cooperative (linear) manner during urea titration (Figure 2), individual PDZs show distinct unfolding transitions when compared to DGRASP and to each other (Figure 6C). PDZ1 starts to unfold very early during the titration (~1 M urea) and, at ~ 2-3 M urea, around 60% of the total population is already in the unfolded state (Figure 6C). These CD data are consistent with the NMR data, in which a significant part of the structure appears to be unfolded at the beginning of the urea titration. Although CD reports on the overall secondary structure content of a protein, only NMR chemical shifts and the level of their dispersion indicate the absence or presence of a defined three-dimensional conformation. Nevertheless, both NMR and CD data agree that PDZ1 has a less stable structure and with a weakly cooperative unfolding process. In contrast, PDZ2 shows a more strongly cooperative transition to the unfolded state with a urea concentration at half transition of 4.2 M, suggesting that, even though the PDZs have comparable amounts of regular secondary structure and disordered regions, PDZ2 has a more well-defined tertiary structure (consistent with its near-UV CD).

ENM (elastic network model) analysis of the DGRASP intrinsic dynamics reinforces the idea of a stronger cooperative unfolding process of PDZ2. The ability of a single residue to transmit information to other residues can be investigated by Markovian stochastic analysis coupled with ENM [66,67]. The commute time, $C(i,j)$, which is a measure of the time required for the information residing at residue i to be transmitted to residue j , was used to characterize the communication inside DGRASP. Figure 6D displays the commute time map of DGRASP. Lower values of $C(i,j)$ (red regions) correspond to higher communication abilities. As shown by the Commute time map, residues of PDZ2 are more correlated than in PDZ1, which means that

1 perturbations in one single residue of PDZ2 are transmitted more efficiently to other residues than
2 in PDZ1.

3
4 ^1H NMR spectra of proteins are rarely informative because spectral crowding hinders
5 assignment. However, by focusing first on the amide resonance regions, it is possible to see
6 differences between the ^1H spectra of the individual PDZs (Figure 6E). PDZ2 has multiple peaks
7 spread over the region 6-10.5 ppm, with most of them being very sharp, suggesting that PDZ2
8 behaves like a regular, ordered protein. A remarkably different behavior is observed for PDZ1
9 (Figure 6E). The resonances of the amide protons are distributed over a narrower region (~2 ppm)
10 and the line superposition is much more severe. DGRASP has also a similar poor dispersed NMR
11 spectrum in the amide ^1H dimension, just like those observed previously in molten globule states
12 [58,68] (Figure 7A). However, we can clearly see some peaks from DGRASP above 8.5 ppm and
13 below 0.6 ppm when those regions are amplified, and which are characteristic of folded structures
14 (Figure 7A). Although there is no perfect superposition, especially because the proton resonances
15 can change when the PDZs are connected in tandem, these characteristic peaks from DGRASP
16 have correlations with those observed in the isolated PDZ2 (Figure 7A). Thus, the ^1H NMR
17 spectrum of DGRASP seems to be the result of the combination of the molten globule-like
18 spectrum mainly from PDZ1 and from the PDZ1/2 interdomain dynamics, with the more well-
19 structured spectrum mainly from PDZ2. Interestingly, if we superimpose the PDZ1 and DGRASP
20 ^1H - ^{15}N HSQC spectra acquired in the absence of urea, almost all resonance lines present in
21 DGRASP spectrum at the beginning of the titration correspond to PDZ1 residues (Figure 7B).
22 Additional weak resonance lines present in the DGRASP HSQC spectrum came from the short
23 SPR portion, according to Figure 4A. As discussed before, this suggests those regions are within
24 disordered/flexible structures and supports the transfer of the assignments from the 9 M urea
25 condition.

26
27 In order to access the residue-specific unfold of the isolated PDZs, urea titration and NMR
28 experiments were once again performed. PDZ1 shows the same pattern observed for DGRASP
29 with fewer peaks at the very beginning of the titration but with a rapid appearance of the remaining
30 lines as the urea concentration is increased (Figure 8A). Most of the peaks missing in the spectrum
31 of the native structure, appeared at urea concentrations around 1-2 M, a similar pattern observed
32 in DGRASP (Figure 3). This unfolding transition is associated with the part of the domain
33 predicted to be the binding site for molecular partners, and it is interesting to note that its unfolding
34 appears to be unaffected by the presence or absence of PDZ2. The remaining lines appeared
35 around 3-4 M urea, suggesting that the most stable region of PDZ1 in the DGRASP construct
36 decreases its stability in the absence of PDZ2 (from ~5 to 3-4 M urea).

For PDZ2, the profile is consistent with a regular well-folded protein. The spectrum of the native structure has a reasonable spread of peaks in the proton dimension and, when the urea concentration reaches a specific value, the lines move to a very narrow region close to the ^1H chemical shift of 8 ppm (Figure 8B). The transition between these two states is therefore highly cooperative and occurs at a urea concentration of around 3-4 M. The cooperative pattern is the same observed previously for the DGRASP but the overall stability is significantly reduced (by at least 1 M urea). Surprisingly, PDZ2 is therefore even more stable when it is connected to PDZ1 in tandem.

The long-range coupled mechanism of action

Our NMR and SRCD data showed that the PDZ domains of GRASPs have different structure, structural stabilities and dynamics, but it is not clear how these properties are exploited to enable the protein to bind different partners. In the crystal structure of the DGRASP55 and JAM-C complex (PDB ID 5GMI), the C-terminus of JAM-C is buried in the PDZ1 binding pocket, however the remainder of JAM-C forms multiple additional interactions with both PDZ1 and PDZ2 [36]. In contrast, corresponding crystal structures of DGRASP55 in complex with Golgin45 and the JAM-B C-terminus (PDB IDs 5H3J and 5GMJ) show different molecular contacts with the PDZs [12,13].

To understand the role of PDZs during interaction with a partner, we performed MD simulations to sample the conformational flexibility of DGRASP55 in the presence (holo) and absence (apo) of a protein partner (the JAM-C C-terminus peptide) (Figure 9A-C). For the ABF analyses, two reaction coordinates are needed and, based on a normal mode analyses, we have chosen the distance between the center of mass of the two PDZs and the dihedral angle formed by the center of mass of the PDZs and the helix that connects the two domains (Figure 8B). Using Boltzmann statistics, for each increase of 1 kcal/mol in free energy at 310 K, a decrease of 80% probability for the observation of the higher energy states is expected. A free energy barrier of approximately 3 kcal/mol would represent an upper limit for the observable energy states at 310 K (with about 1% probability of occupancy). The conformational free-energy profile of DGRASP55 in the apo form at physiological temperature (310 K) exhibits a large region of accessible free-energy (Figure 9C). However, when the protein partner is bound (holo), the free-energy surface displays a much narrower low energy well (Figure 9C). The differences between the apo and holo free-energy profiles suggest that the binding of the protein partner happens through a conformational selection mechanism, although experimental kinetic data will be needed in order to fully differentiate this from an induced fit process. The wider free energy well of apo DGRASP55 indicates a greater conformational variability relative to the holo DGRASP55, with

open and closed conformations coexisting in equilibrium. The presence of the JAM-C C-terminus peptide significantly stabilizes closed conformations relative to all other conformations that are accessible in the absence of the protein partner.

Previous crystallographic data have suggested that DGRASP55 transitions from an open to a closed conformation upon binding to JAM-C and JAM-B [36Error! Bookmark not defined.]. However, such closure is not observed when DGRASP55 binds the C-terminus of Golgin45 [13], suggesting that the conformational change, if it really occurs in solution, may be partner dependent. The magnitude of the conformational change (a relative rotation of the PDZs) is sampled in the MD simulations to generate a free energy profile at physiological temperature (Figure 9). The MD simulations were performed using the crystal structures of holo DGRASP55+JAM-C, while the apo free energy landscapes were simulated from the holo structures by removing the peptide (Figure 9C). Under these conditions, the crystal structure conformation observed in the apo forms (PDB ID 4KFW and the homologue 3RLE) was accessible, as were the conformations observed in the complexes with Golgin45 (PDB ID 5H3J) and JAM-B (PDB ID 5GMJ) (Figure 10A). At physiological temperature the conformational flexibility of DGRASPs allows large movements in the relative orientations of the two domains, making different conformations accessible to different binding partners, thereby creating numerous potential secondary interactions via PDZ2. The crystallographic structures show that both JAM-B/C and Golgin45 C-terminus peptides are inserted into the canonical PDZ-peptide binding pocket in the PDZ1 domain and form multiple hydrogen bonds with the β 2 strand of DGRASP55. Additionally, Golgin45 C-terminus peptide also interacts with DGRASP55 through a cleft between PDZ1 and PDZ2, making contacts with both PDZs domains. Although the crystal structure of DGRASP55 bound to JAM-C peptide did not reveal any interaction between the peptide and PDZ2, our MD simulations show that small perturbations in the position of PDZ2 promote the formation of hydrogen bonds that might be responsible for the stabilization of the closed conformation when JAM-B peptide is present (Figure 10B). It seems that GRASP PDZs are structurally organized to provide a somewhat smoother energy landscape that allows the domain orientation to be flipped after a ligand binding. This effect, coupled to the promiscuity of the PDZ1 domain, enables DGRASPs to interact with multiple different protein partners, allowing it to fulfill a key role in cellular signaling.

Conclusion

GRASPs are formed by two asymmetric PDZs domains: PDZ1 is more unstable and flexible while PDZ2 is much more stable and structurally more well behaved. Many of the unstable

regions found in PDZ1 are located in the binding pocket, suggesting that structural promiscuity inside this domain allows the binding of multiple protein partners. In contrast, PDZ2 is more well-structured, and this correlates with the smaller number of associated proteins capable of anchoring with GRASPs using this domain. Our findings challenge the conclusions drawn from existing crystal structures, which suggest that both PDZs have a similar structure, even though their function is remarkably different. We have shown that DGRASPs exhibit a large conformational flexibility at physiological temperature, which reduces after binding to a specific protein partner. The coupled, coordinated movement between the two PDZs and the structural promiscuity of PDZ1 appears to enable GRASPs to interact with multiple partners, allowing it to act as a multitasking protein in the maintenance and remodeling of the Golgi body, as well as playing a key role in both conventional and unconventional secretory pathways.

Materials and Methods:

Protein expression and purification:

CnGRASP GRASP domain (DGRASP) and isolated PDZ samples were expressed and purified as described elsewhere [29]. ¹³C- and/or ¹⁵N-labeled DGRASP and PDZs were produced using a protocol reported previously [69]. The same protocol was used for purification of both labelled and unlabelled protein [29].

Circular Dichroism (CD):

CD measurements were performed using a Jasco J-815 Spectropolarimeter fitted with a Peltier temperature control unit. High-grade quartz cuvettes of 1 mm path length were used for data collection. The temperature was 20 °C. Further parameters were: data pitch 0.5 nm, D.I.T. of 1 sec, 1.00 nm bandwidth and a scanning speed of 50 nm/min. The sample preparation for the urea titration consisted of a protein dilution (less than 5% of the total final volume) directly into a solution containing 50 mM sodium phosphate, 10 mM NaCl and 1 mM β-mercaptoethanol, pH 7.4 with the corresponding urea concentration. Protein concentration was kept constant at 0.15 mg/mL.

Synchrotron Radiation Circular Dichroism (SRCD)

SRCD experiments were performed on the B23 Synchrotron Radiation CD beamline at Diamond Light Source, Oxfordshire, UK. Protein concentration was in the range of 5-10 mg/mL and demountable quartz cells of 20 μm and 50 μm were mainly used. The parameters were the same as those used for the conventional CD experiments. DICHROWEB was used in the analysis of protein secondary structures [70].

Fluorescence Spectroscopy

Steady-state fluorescence was monitored using a Hitachi F-7000 spectrofluorimeter equipped with a 150 W xenon arc lamp and polarized filters for anisotropy experiments. The excitation and emission monochromators were set at 5 nm slit width in all tryptophan experiments and 5 nm excitation with 10 nm emission for the 1-anilino-8-naphthalenesulfonic acid (ANS) experiments. For tryptophan fluorescence experiments, the selective tryptophan excitation wavelength was set at 295 nm and the emission spectra were measured from 310 up to 450 nm. For the anisotropy measurements, tryptophan was selectively excited at 300 nm and the mean anisotropy values were calculated as the mean average of the values from ± 10 nm (1 nm step acquisition) over the wavelength of maximum emission determined for each condition. For the ANS (250 μ M) fluorescence experiments, the excitation wavelength was set at 360 nm and the emission spectra were measured from 400 up to 650 nm. The urea unfolding experiments were performed using the same strategy described in the CD experiments. The first moment of the fluorescence spectrum, or spectral center of mass, is calculated according to the equation $CM = \sum_i \frac{F_i \lambda_i}{F_i}$ where F_i is the fluorescence intensity of the particular wavelength λ_i . All the experiments were performed at 25°C.

High-Field Solution Nuclear Magnetic Resonance (NMR)

All NMR experiments were carried out using spectrometers operating at ^1H frequencies of 600 MHz and 950 MHz, equipped with Oxford Instruments magnets and Bruker Avance III HD consoles. The 600 MHz and 950 MHz spectrometers were equipped with a 5mm TCI cryoprobe and a 5 mm room-temperature TXI probe, respectively. Data were processed using NMRPipe [71] and/or Bruker TopSpin 3.5 and spectra were analyzed using the CCPN software [72]. ^1H - ^{15}N HSQC spectra of DGRASP in 25 mM HEPES, 100 mM NaCl, 5 mM β -mercaptoethanol, pH 7.0 (95% H_2O /5% D_2O) and in a urea concentration gradient ranging from 0 to 9 M (1 M steps) were collected. The sample with 9 M urea was the one with the highest number of resonance peaks in the ^1H - ^{15}N HSQC spectrum and, therefore, chosen for assignment. CBCACONH, CBCANH, HNCO, HN(CA)CO, (H)CC(CO)NH, HBHA(CBCACO)NH, HSQC-TOCSY, HSQC-NOESY, HSQC-NOESY-HSQC were collected for assignment; these were transferred back to all other HSQC spectra collected with lower urea concentrations. The temperature was 20 °C in all experiments and the protein concentration varied from 150 to 400 μ M. NMR data collection for PDZ1 and PDZ2 followed the same protocols and buffers used with the DGRASP construction. Shigemi tubes were used in all experiments. The assignments have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under the accession number 27921.

Molecular Dynamics (MD) simulations

The crystal structure of DGRASP55 in complex with the JAM-C C-terminus peptide (PDB ID: 5GMI) was used to build models for holo-DGRASP55 (in which the peptide was present) and apo-DGRASP55 (from which the peptide was absent). The structures were solvated with Packmol [73,74] using ca. 19,500 water molecules and Na⁺ and Cl⁻ ions were added to render the system neutral. Simulations were performed with NAMD [75] using periodic boundary conditions, the CHARMM36 [76] force-field for the protein and peptides, and the TIP3P model [77] for water. Equilibration and simulations were performed in the NPT ensemble at 298.15 K and 1 atm. Temperature was controlled using Langevin dynamics with a friction coefficient of 10 ps⁻¹. The Nosé-Hoover algorithm was used for the pressure control, with a piston oscillation period of 200 fs and decay rate of 100 fs. A 2 fs time-step was used to integrate the equations of motion using the Verlet algorithm. A cutoff of 12 Å was used for Van der Waals interactions, whereas the long-range electrostatic interactions were handled by the Particle Mesh Ewald method [78]. Covalent bonds involving hydrogen atoms in the protein were constrained to their equilibrium distances using the SHAKE algorithm, while SETTLE [79] was used for water.

The systems were equilibrated as follows: 1) 1,000 steps of conjugate-gradient steps (CG) [80] followed by 200 ps MD keeping all protein and peptide atoms fixed, for solvent relaxation. 2) 500 CG minimization steps followed by 200 ps MD with fixed C α coordinates, allowing the side-chains of the protein to relax. 3) 2 ns MD without any restrictions. Five unbiased simulations with 50 ns were performed for each system. The final structures generated, were used for unbiased simulations and for Adaptive Biasing Force (ABF) simulations.

The ABF [81,82] simulations were performed using NAMD. The reaction coordinates were defined as the distance between the center of mass of the two PDZs and the dihedral angle formed by the center of mass of the PDZs and the helix, connecting the two domains. The first reaction coordinate was sampled by ABF simulations within 20 and 35 Å, with a precision of 0.1 Å, and the second reaction coordinate was sampled within -120° and 50°. To restrain the movements of the atoms inside the regions of interest, a harmonic boundary potential with 10 kcal.mol⁻¹Å² force constant was applied. The ABF force was applied only after 5000 samples of the mean force were generated in each bin. Ten independent ABF simulations with 100 ns each were performed for each system. The convergence of the free energy profiles was checked by computing the root mean square of the ABF forces. The profiles were considered converged if their root mean square was smaller than 0.2 kcal/mol. VMD [83] was used for visualization and figure preparation. DynOmics ENM Server was used to calculate the commute time map [84]. Swiss model was used to build the molecular model structures [85].

-
- 1) Marsh BJ, Howell KE. The mammalian Golgi--complex debates. *Nat Rev Mol Cell Biol.* 2002 Oct;3(10):789-95.
 - 2) Rothman JE. Mechanisms of intracellular protein transport. *Nature.* 1994;372: 55–63.
 - 3) Ladinsky MS, Mastronarde DN, McIntosh JR, Howell KE, Staehelin LA (1999). Golgi structure in three dimensions: functional insights from the normal rat kidney cell. *J Cell Biol* 144, 1135-1149.
 - 4) Barlow LD, Nývltová E, Aguilar M, Tachezy J, Dacks JB. A sophisticated, differentiated Golgi in the ancestor of eukaryotes. *BMC Biol.* 2018 Mar 7;16(1):27.
 - 5) Ramirez IB, Lowe M. Golgins and GRASPs: holding the Golgi together. *Semin Cell Dev Biol.* 2009 Sep;20(7):770-9.
 - 6) Mendes LFS, Fontana NA, Oliveira CG, Freire MCLC, Lopes JLS, Melo FA, Costa-Filho AJ. The GRASP domain in golgi reassembly and stacking proteins: differences and similarities between lower and higher Eukaryotes. *FEBS J.* 2019 Sep;286(17):3340-3358.
 - 7) Barr FA, Puype M, Vandekerckhove J, Warren G. GRASP65, a protein involved in the stacking of Golgi cisternae. *Cell.* 1997 Oct 17;91(2):253-62.
 - 8) Shorter J, Watson R, Giannakou ME, Clarke M, Warren G, Barr FA. GRASP55, a second mammalian GRASP protein involved in the stacking of Golgi cisternae in a cell-free system. *EMBO J.* 1999 Sep 15;18(18):4949-60.
 - 9) Bekier ME 2nd, Wang L, Li J, Huang H, Tang D, Zhang X, Wang Y. Knockout of the Golgi stacking proteins GRASP55 and GRASP65 impairs Golgi structure and function. *Mol Biol Cell.* 2017 Oct 15;28(21):2833-2842.
 - 10) Sutterlin, C., Polishchuk, R., Pecot, M. and Malhotra, V (2005). The Golgi-associated protein GRASP65 regulates spindle dynamics and is essential for cell division. *Mol. Biol. Cell.* 16, 3211–3222.
 - 11) Rabouille C, Linstedt AD. GRASP: A Multitasking Tether. *Front Cell Dev Biol.* 2016 Jan 26;4:1.
 - 12) Hu F, Shi X, Li B, Huang X, Morelli X, Shi N. Structural basis for the interaction between the Golgi reassembly-stacking protein GRASP65 and the Golgi matrix protein GM130. *J Biol Chem.* 2015 Oct 30;290(44):26373-82.
 - 13) Zhao J, Li B, Huang X, Morelli X, Shi N. Structural Basis for the Interaction between Golgi Reassembly-stacking Protein GRASP55 and Golgin45. *J Biol Chem.* 2017 Feb 17;292(7):2956-2965.
 - 14) Barr, F. A., Preisinger, C., Kopajtich, R. and Korner, R. (2001). Golgi matrix proteins interact with p24 cargo receptors and aid their efficient retention in the Golgi apparatus. *J. Cell Biol.* 155:885–891.
 - 15) Kuo A, Zhong C, Lane WS, Derynck R. Transmembrane transforming growth factor- α tethers to the PDZ domain-containing, Golgi membrane-associated protein p59/GRASP55. *EMBO J.* 2000 Dec 1;19(23):6427-39.
 - 16) Kineth MA, Anjard C, Fuller D, Guizzunti G, Loomis WF, Malhotra V. The Golgi-associated protein GRASP is required for unconventional protein secretion during development. *Cell.* 2007 Aug 10;130(3):524-34.
 - 17) Duran JM, Anjard C, Stefan C, Loomis WF, Malhotra V. Unconventional secretion of Acb1 is mediated by autophagosomes. *J Cell Biol.* 2010 Feb 22;188(4):527-36.
 - 18) Manjithaya R, Anjard C, Loomis WF, Subramani S. Unconventional secretion of *Pichia pastoris* Acb1 is dependent on GRASP protein, peroxisomal functions, and autophagosome formation. *J Cell Biol.* 2010 Feb 22;188(4):537-46.
 - 19) Schotman H, Karhinen L, Rabouille C. dGRASP-mediated noncanonical integrin secretion is required for *Drosophila* epithelial remodeling. *Dev Cell.* 2008 Feb;14(2):171-82.
 - 20) Gee HY, Noh SH, Tang BL, Kim KH, Lee MG. Rescue of $\Delta F508$ -CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell.* 2011 Sep 2;146(5):746-60.

-
- 21) Chiritoiu M, Brouwers N, Turacchio G, Pirozzi M, Malhotra V. GRASP55 and UPR Control Interleukin-1 β Aggregation and Secretion. *Dev Cell*. 2019 Mar 8. pii: S1534-5807(19)30102-9.
- 22) Giuliani F, Grieve A, Rabouille C. Unconventional secretion: a stress on GRASP. *Curr Opin Cell Biol*. 2011 Aug;23(4):498-504.
- 23) Rabouille C. Pathways of Unconventional Protein Secretion. *Trends Cell Biol*. 2017 Mar;27(3):230-240.
- 24) Kmetzsch L, Joffe LS, Staats CC, de Oliveira DL, Fonseca FL, Cordero RJ, Casadevall A, Nimrichter L, Schrank A, Vainstein MH, Rodrigues ML. Role for Golgi reassembly and stacking protein (GRASP) in polysaccharide secretion and fungal virulence. *Mol Microbiol*. 2011 Jul;81(1):206-18.
- 25) Truschel ST, Sengupta D, Foote A, Heroux A, Macbeth MR, Linstedt AD. Structure of the membrane-tethering GRASP domain reveals a unique PDZ ligand interaction that mediates Golgi biogenesis. *J Biol Chem*. 2011 Jun 10;286(23):20125-9.
- 26) Feng Y, Yu W, Li X, Lin S, Zhou Y, Hu J, Liu X. Structural insight into Golgi membrane stacking by GRASP65 and GRASP55 proteins. *J Biol Chem*. 2013 Sep 27;288(39):28418-27.
- 27) Wang Y, Satoh A, Warren G. Mapping the functional domains of the Golgi stacking factor GRASP65. *J Biol Chem*. 2005 Feb 11;280(6):4921-8.
- 28) Vinke FP, Grieve AG, Rabouille C. The multiple facets of the Golgi reassembly stacking proteins. *Biochem J*. 2011 Feb 1;433(3):423-33.
- 29) Mendes LF, Garcia AF, Kumagai PS, de Moraes FR, Melo FA, Kmetzsch L, Vainstein MH, Rodrigues ML, Costa-Filho AJ. New structural insights into Golgi Reassembly and Stacking Protein (GRASP) in solution. *Sci Rep*. 2016 Jul 20;6:29976.
- 30) Songyang Z, Fanning AS, Fu C, Xu J, Marfatia SM, Chishti AH, Crompton A, Chan AC, Anderson JM, Cantley LC. Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science*. 1997 Jan 3;275(5296):73-7.
- 31) Feng W, Zhang M. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat Rev Neurosci*. 2009 Feb;10(2):87-99.
- 32) Ye F, Zhang M. Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures. *Biochem J*. 2013 Oct 1;455(1):1-14.
- 33) Short B, Preisinger C, Körner R, Kopajtich R, Byron O, Barr FA. A GRASP55-rab2 effector complex linking Golgi structure to membrane traffic. *J Cell Biol*. 2001 Dec 10;155(6):877-83.
- 34) Dementieva IS, Tereshko V, McCrossan ZA, Solomaha E, Araki D, Xu C, Grigorieff N, Goldstein SA. Pentameric assembly of potassium channel tetramerization domain-containing protein 5. *J Mol Biol*. 2009 Mar 20;387(1):175-91.
- 35) Piao H, Kim J, Noh SH, Kweon HS, Kim JY, Lee MG. Sec16A is critical for both conventional and unconventional secretion of CFTR. *Sci Rep*. 2017 Jan 9;7:39887.
- 36) Cartier-Michaud A, Bailly AL, Betzi S, Shi X, Lissitzky JC, Zarubica A, Sergé A, Roche P, Lugari A, Hamon V, Bardin F, Derviaux C, Lembo F, Audebert S, Marchetto S, Durand B, Borg JP, Shi N, Morelli X, Aurrand-Lions M. Genetic, structural, and chemical insights into the dual function of GRASP55 in germ cell Golgi remodeling and JAM-C polarized localization during spermatogenesis. *PLoS Genet*. 2017 Jun 15;13(6):e1006803.
- 37) Zhang X, Wang L, Lak B, Li J, Jokitalo E, Wang Y. GRASP55 Senses Glucose Deprivation through O-GlcNAcylation to Promote Autophagosome-Lysosome Fusion. *Dev Cell*. 2018 Apr 23;45(2):245-261.e6.
- 38) Stein MF, Blume K, Heilingloh CS, Kummer M, Biesinger B, Sticht H, Steinkasserer A. CD83 and GRASP55 interact in human dendritic cells. *Biochem Biophys Res Commun*. 2015 Mar 27;459(1):42-8.
- 39) Roghi C, Jones L, Gratian M, English WR, Murphy G. Golgi reassembly stacking protein 55 interacts with membrane-type (MT) 1-matrix metalloprotease (MMP) and furin and plays a role in the activation of the MT1-MMP zymogen. *FEBS J*. 2010 Aug;277(15):3158-75.
- 40) Dicks M, Kock G, Kohl B, Zhong X, Pütz S, Heumann R, Erdmann KS, Stoll R. The binding affinity of PTPN13's tandem PDZ2/3 domain is allosterically modulated. *BMC Mol Cell Biol*. 2019 Jul 8;20(1):23.

- 41) Gianni S, Walma T, Arcovito A, Calosci N, Bellelli A, Engström A, Travaglini-Allocatelli C, Brunori M, Jemth P, Vuister GW. Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure*. 2006 Dec;14(12):1801-9.
- 42) van den Berk LC, Landi E, Walma T, Vuister GW, Dente L, Hendriks WJ. An allosteric intramolecular PDZ-PDZ interaction modulates PTP-BL PDZ2 binding specificity. *Biochemistry*. 2007 Nov 27;46(47):13629-37.
- 43) Kock G, Dicks M, Yip KT, Kohl B, Pütz S, Heumann R, Erdmann KS, Stoll R. Molecular Basis of Class III Ligand Recognition by PDZ3 in Murine Protein Tyrosine Phosphatase PTPN13. *J Mol Biol*. 2018 Oct 19;430(21):4275-4292.
- 44) McCann JJ, Zheng L, Chiantia S, Bowen ME. Domain orientation in the N-Terminal PDZ tandem from PSD-95 is maintained in the full-length protein. *Structure*. 2011 Jun 8;19(6):810-20.
- 45) Fontana NA, Fonseca-Maldonado R, Mendes LFS, Meleiro LP, Costa-Filho AJ. The yeast GRASP Grh1 displays a high polypeptide backbone mobility along with an amyloidogenic behavior. *Sci Rep*. 2018 Oct 24;8(1):15690.
- 46) Mendes LFS, Basso LGM, Kumagai PS, Fonseca-Maldonado R, Costa-Filho AJ. Disorder-to-order transitions in the molten globule-like Golgi Reassembly and Stacking Protein. *Biochim Biophys Acta Gen Subj*. 2018 Apr;1862(4):855-865.
- 47) Uversky VN. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta*. 2013 May;1834(5):932-51.
- 48) Redfield C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods*. 2004 Sep;34(1):121-32.
- 49) Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signaling and regulation. *Nat Rev Mol Cell Biol*. 2015 Jan;16(1):18-29.
- 50) Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J*. 2005 Oct;272(20):5129-48.
- 51) Alberti S, Halfmann R, King O, Kapila A, Lindquist S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell*. 2009 Apr 3;137(1):146-58.
- 52) Mukhopadhyay S, Krishnan R, Lemke EA, Lindquist S, Deniz AA. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc Natl Acad Sci U S A*. 2007 Feb 20;104(8):2649-54.
- 53) Silva JL, Cino EA, Soares IN, Ferreira VF, A P de Oliveira G. Targeting the Prion-like Aggregation of Mutant p53 to Combat Cancer. *Acc Chem Res*. 2018 Jan 16;51(1):181-190.
- 54) Pedrote MM, de Oliveira GAP, Felix AL, Mota MF, Marques MA, Soares IN, Iqbal A, Norberto DR, Gomes AMO, Gratton E, Cino EA, Silva JL. Aggregation-primed molten globule conformers of the p53 core domain provide potential tools for studying p53C aggregation in cancer. *J Biol Chem*. 2018 Jul 20;293(29):11374-11387.
- 55) Long J, Wei Z, Feng W, Yu C, Zhao YX, Zhang M (2008). Supramodular nature of GRIP1 revealed by the structure of its PDZ12 tandem in complex with the carboxyl tail of Frs1. *J Mol Biol*. Feb 1;375(5):1457-68.
- 56) Zhang Q, Fan JS, Zhang M. (2001). Interdomain chaperoning between PSD-95, Dlg, and Zonula Occludens-1 (PDZ) domains of glutamate receptor-interacting proteins. *J Biol Chem*. Nov 16;276(46):43216-20.
- 57) Bom AP, Freitas MS, Moreira FS, Ferraz D, Sanches D, Gomes AM, Valente AP, Cordeiro Y, Silva JL. The p53 core domain is a molten globule at low pH: functional implications of a partially unfolded structure. *J Biol Chem*. 2010 Jan 22;285(4):2857-66.
- 58) Redfield C. NMR studies of partially folded molten-globule states. *Methods Mol Biol*. 2004;278:233-54.
- 59) B.A. Schulman, P.S. Kim, C.M. Dobson, C. Redfield (1997). A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct.Biol.* 4, 630–634.
- 60) Park SJ, Borin BN, Martinez-Yamout MA, Dyson HJ. (2011). The client protein p53 adopts a molten globule-like state in the presence of Hsp90. *Nat Struct Mol Biol*. May;18(5):537-41.

-
- 61) Kurzbach D, Kontaxis G, Coudeville N, Konrat R. NMR Spectroscopic Studies of the Conformational Ensembles of Intrinsically Disordered Proteins. *Adv Exp Med Biol.* 2015;870:149-85.
- 62) Lee HJ, Zheng JJ. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal.* 2010; 8: 8.
- 63) Sheng M, Sala C. PDZ domains and the organization of supramolecular complexes. *Annu Rev Neurosci.* 2001;24:1-29.
- 64) Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochim Biophys Acta.* 2005 Aug 10;1751(2):119-39.
- 65) Vassilenko KS, Uversky VN. Native-like secondary structure of molten globules. *Biochim Biophys Acta.* 2002 Jan 31;1594(1):168-77.
- 66) Hu G, Di Paola L, Liang Z, Giuliani A. Comparative Study of Elastic Network Model and Protein Contact Network for Protein Complexes: The Hemoglobin Case. *Biomed Res Int.* 2017;2017:2483264.
- 67) Chennubhotla C, Bahar I. Signal propagation in proteins and relation to equilibrium fluctuations. *PLoS Comput Biol.* 2007 Sep;3(9):1716-26.
- 68) Mok KH, Nagashima T, Day IJ, Hore PJ, Dobson CM. Multiple subsets of side-chain packing in partially folded states of alpha-lactalbumins. *Proc Natl Acad Sci U S A.* 2005 Jun 21;102(25):8899-904.
- 69) Protocol from the NMR Structural Biology Facility from UConn Health webserver https://health.uconn.edu/structural-biology/wp-content/uploads/sites/177/2017/10/nmr_sample_preparation.pdf (accessed in 2019).
- 70) Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers.* 2008 May;89(5):392-400.
- 71) Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 6(3):277-93.
- 72) Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins.* Jun 1;59(4):687-96.
- 73) Martínez L, Andrade R, Birgin EG, Martínez JM. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J Comput Chem.* 2009 Oct;30(13):2157-64.
- 74) Martínez JM, Martínez L. Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking. *J Comput Chem.* 2003 May;24(7):819-25.
- 75) Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem.* 2005 Dec;26(16):1781-802.
- 76) MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FT, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B.* 1998 Apr 30;102(18):3586-616.
- 77) Jorgensen WL., Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of single potential functions for simulating liquid water. *J Chem Phys.* 1983, 79:926-935.
- 78) Darden T, York D, Pedersen L. Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* 1993, 98(12):10089-10092.
- 79) Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 1992, 13:952-962.
- 80) Darve E, Rodríguez-Gómez D, Pohorille A. Adaptive biasing force method for scalar and vector free energy calculations. *J Chem Phys.* 2008 Apr 14;128(14):144120.
- 81) Chipot C, Hénin J. Exploring the free-energy landscape of a short peptide using an average force. *J Chem Phys.* 2005 Dec 22;123(24):244906.

-
- 82) Comer J, Gumbart JC, Hénin J, Lelièvre T, Pohorille A, Chipot C. The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J Phys Chem B*. 2015 Jan 22;119(3):1129-51.
- 83) Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996 Feb;14(1):33-8, 27-8.
- 84) Li H, Chang YY, Lee JY, Bahar I, Yang LW. DynOmics: dynamics of structural proteome and beyond. *Nucleic Acids Res*. 2017 Jul 3;45(W1):W374-W380.
- 85) Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018 Jul 2;46(W1):W296-W303.

Table 1: Estimating the secondary structure content of DGRASP using SRCD. SRCD spectral deconvolution was performed using the software CDSSTR and an appropriate database in the online server Dichroweb [70]. Values for the DGRASP were taken from reference [6].

	Helix	Strand	Turns	Disord.	NRMSD
DGRASP	17%	29%	14%	40%	0.007
PDZ1	5%	37%	22%	34%	0.040
PDZ2	18%	30%	22%	30%	0.027

Figures

CFTR (P13569)	KPQ----IAAL-K-EETEEEVQDTR-L	} <i>Class I</i>
GM130 (Q08379)	PCI----PFFY-R-ADENDEVKITV-I	
KCTD5 (Q9NXV2)	GTA----SEPS-E-KAKILQERGSR-M	
TGF- α (P01135)	EKP----SALL-K-GRTACCHSETVV-	
Golgin45 (Q9H2G9)	RYE----NITF-N-CCNHCRGELIA-L	
JAM-B (P57087)	SSK----ATTM-S-ENDFKHTKSFI-I	} <i>Class II</i>
JAM-C (Q9BX67)	VNY----IRTD-E-EGDFRHKSSFV-I	
LC3-B (Q9GZQ8)	DGF----LYMV-Y-ASQETFGMKLS-V	
RAB2 (P61019)	ATN----ATHA-G-NQGGQAGGGC-C	
Sec16A (O15027)	ATS----GSSR---LGRIGQRKHLVLN	
TMED10 (P49755)	LAT----WQ-VFY-LRRFFKAKKLI-E	
TMED2 (Q15363)	AMT----LGQIYY-LKRFFEVRVV--V	
TMED9 (Q9BVK6)	AIG----VWQMRH-LKSFFEAKKL--V	
Truncated SPR GRASP55	YG----YLHRI-P-TRPFEEGKKIS-L	} <i>Class III</i>
Truncated SPR GRASP65	CGIGYGYLHRI-P-TQPSSQY-----K	
LAMP2 (P13473)	LVL----LAYF-IGLKHHH-AGYEQ-F	

Figure 1: The sequences of the C-terminus of different GRASP55/65 partners were aligned using T-Coffee (<https://www.ebi.ac.uk/Tools/msa/tcoffee/> accessed in 2018) and classified according to the different PDZ binding partner classes, as discussed in the main text.

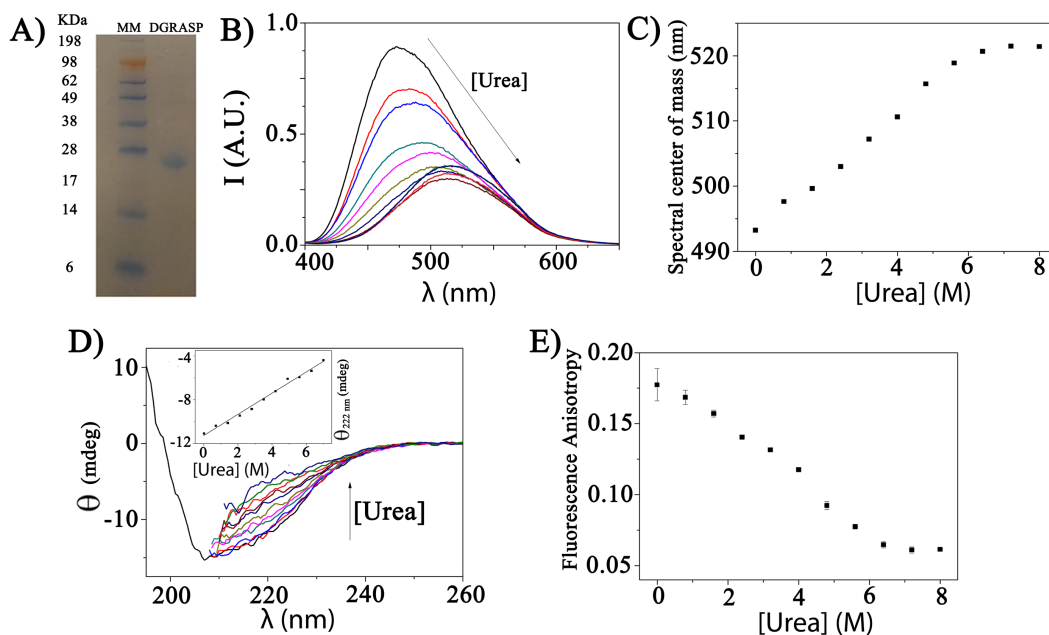


Figure 2: DGRASP as a molten globule-like protein: A) SDS-PAGE of purified $^{15}\text{N}/^{13}\text{C}$ -labeled DGRASP. Lane 1: molecular mass standards in kDa. Lane 2: purified DGRASP. B) ANS-based fluorescence. The arrow indicates the increasing urea concentration. C) The center of spectral mass is calculated from the ANS-based fluorescence spectra and plotted as a function of the urea concentration. D) DGRASP urea titration monitored by CD. Far-UV CD curves for each step of the titration (0.7 M urea from 0 to 7 M). The inset shows the CD intensity at 222 nm plotted as a function of the urea concentration. The dataset was fitted to a linear function. E) Variations in fluorescence anisotropy as a function of urea concentration. All the spectroscopy experiments were performed on duplicate with fresh new protein preparations. The figure was built using Adobe Fireworks CS6.

Figure 4: DGRASP resonance assignments: A) ^1H - ^{15}N HSQC spectrum of DGRASP in 9 M urea showing the HN resonances that were assigned successfully. Some regions were expanded for better reader visualization. In black, red and blue are the residues from PDZ1, PDZ2 and SPR, respectively. NMR experimental data for assignment are deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under the accession number 27921. The figure was built using Adobe Fireworks CS6.

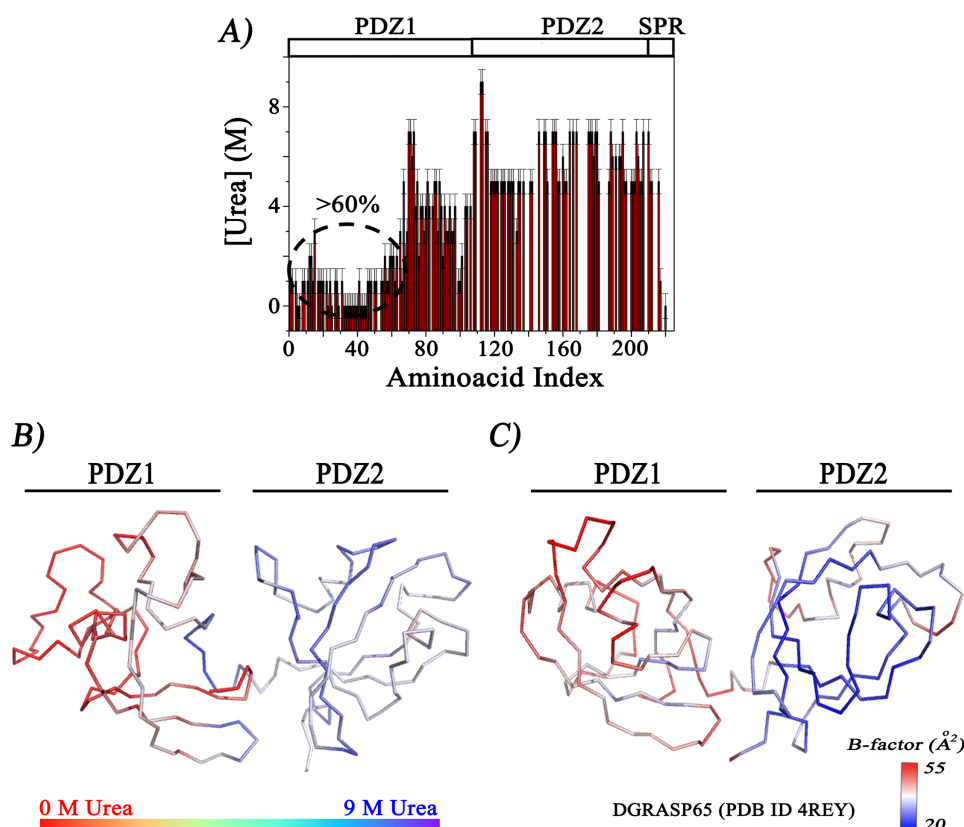


Figure 5. A) Chemical denaturation at residue-specific level. The graphic shows, for each assigned amino acid in the primary sequence, the urea concentration where the resonance line starts to be visible in the ^1H - ^{15}N HSQC operating in a 1H frequency of 950 MHz at 20 °C. We assumed that the urea concentration where the resonance line is now visible is the concentration necessary to unfold that specific protein region. The low-stability region inside the PDZ1 was highlighted in the graphic. B) DGRASP molecular model (build using Swiss model [85] and based on the structure of DGRASP65 – PDB I.D. 4REY) highlighted with a gradient of colors according to the urea stability presented in Figure 4A. We assume that the urea concentration necessary to unfold follows a smooth function of variation, so residues without an unambiguous observable urea concentration of unfolding, were given a number based on the simple average of the adjacent amino acids. C) Crystallographic structure of human DGRASP65 showing the experimental B-factors in a gradient of colors. The figure was built using Adobe Fireworks CS6.

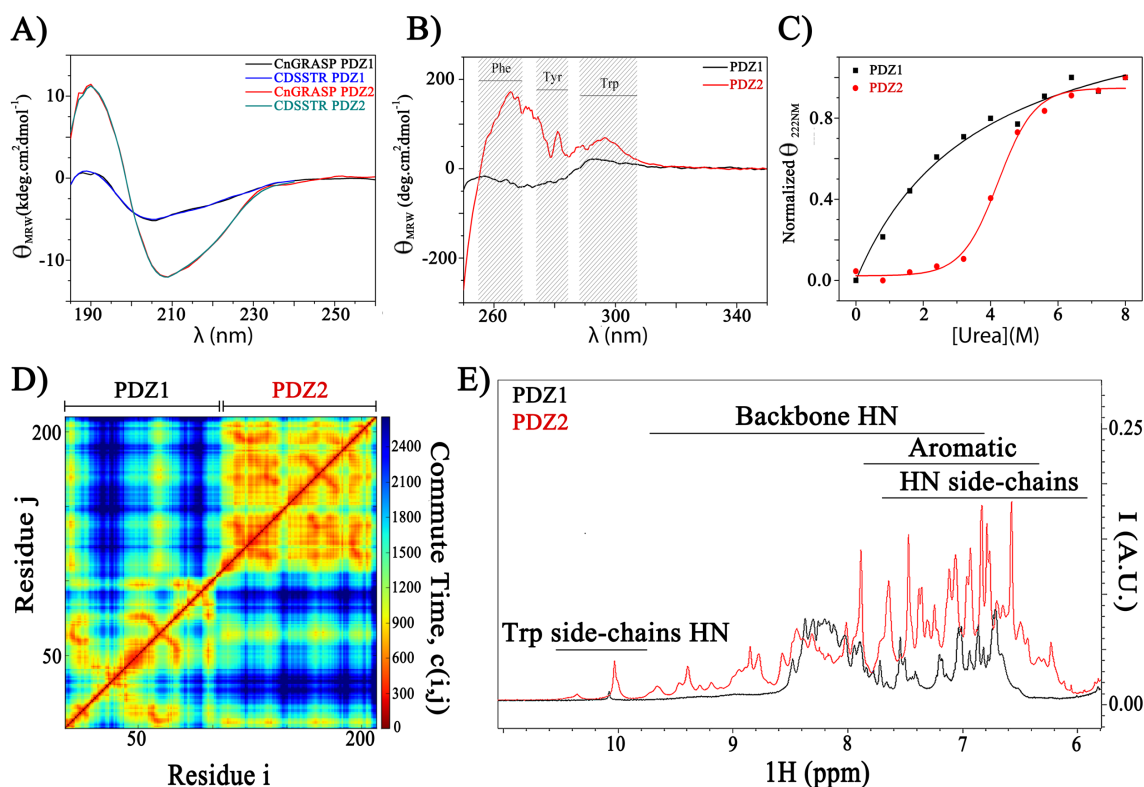


Figure 6: SRCD and ^1H NMR data for the separate PDZs of CnGRASP. A) SRCD data of PDZ1 (black) and PDZ2 (red) together with the deconvolution fit (blue and green, respectively) obtained using the software CDSSTR and an appropriate database in the online server Dichroweb. The elements of secondary structure are presented in table 1. B) Near-UV CD of PDZ1 (black) and PDZ2 (red) with the regions of Phe, Tyr and Trp resonances highlighted in grey. C) Urea titration monitored by the intensity at 222 nm using far-UV CD of PDZ1 (black) and PDZ2 (red) as a function of the urea concentration. The PDZ1 and PDZ2 data sets were fitted with a hyperbolic and a sigmoidal function, respectively. D) Commute Time, $C(i,j)$ for CnDGRASP. DynOmics ENM Server was used to calculate the commute time map [84]. E) ^1H NMR of PDZ1 (black) and PDZ2 (red) with some resonance regions highlighted in the figure for illustration. SRCD data was replicated using conventional CD. All the urea titration experiments were performed in duplicate at controlled temperature of 20 °C. The figure was built using Adobe Fireworks CS6.

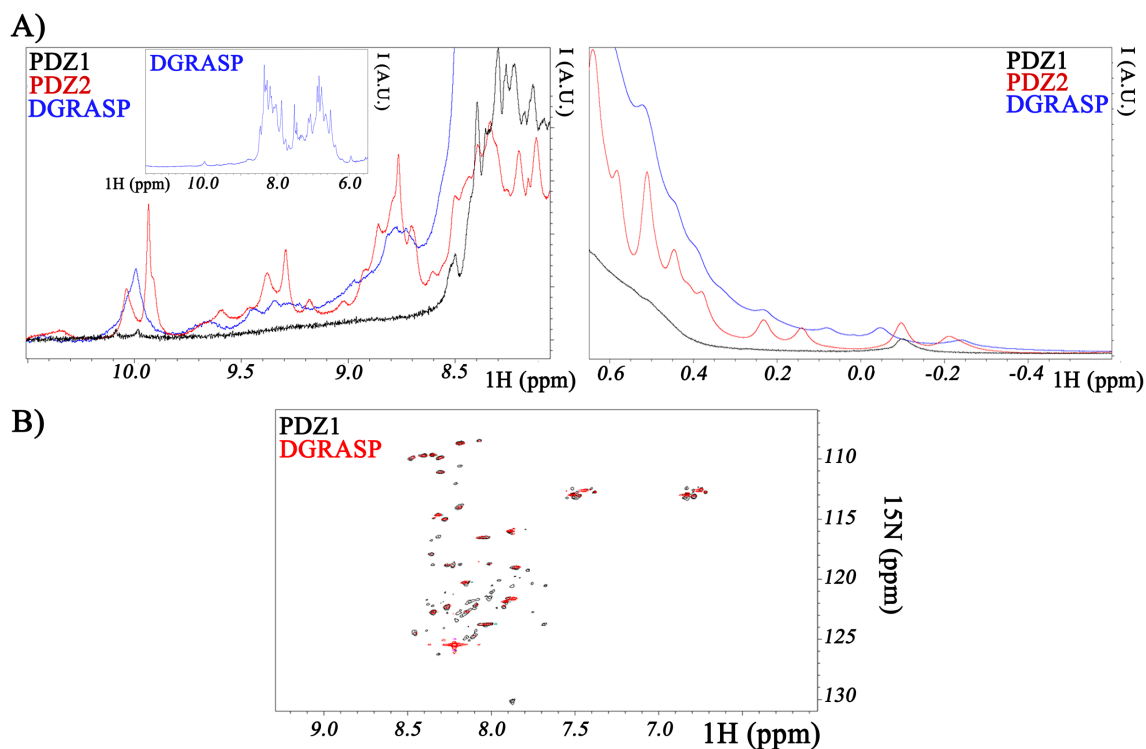


Figure 7: A) ^1H NMR of PDZ 1 (black), PDZ2 (red) and DGRASP (blue) highlighting the region of resonance between 0.6 and -0.5 ppm. In the inset is presented the ^1H NMR spectrum of DGRASP. The scales were adjusted to facilitate the analyze on the selected region. B) ^1H - ^{15}N HSQC spectra of PDZ1 (red) and DGRASP (black) in the absence of urea. The figure was built using Adobe Fireworks CS6.

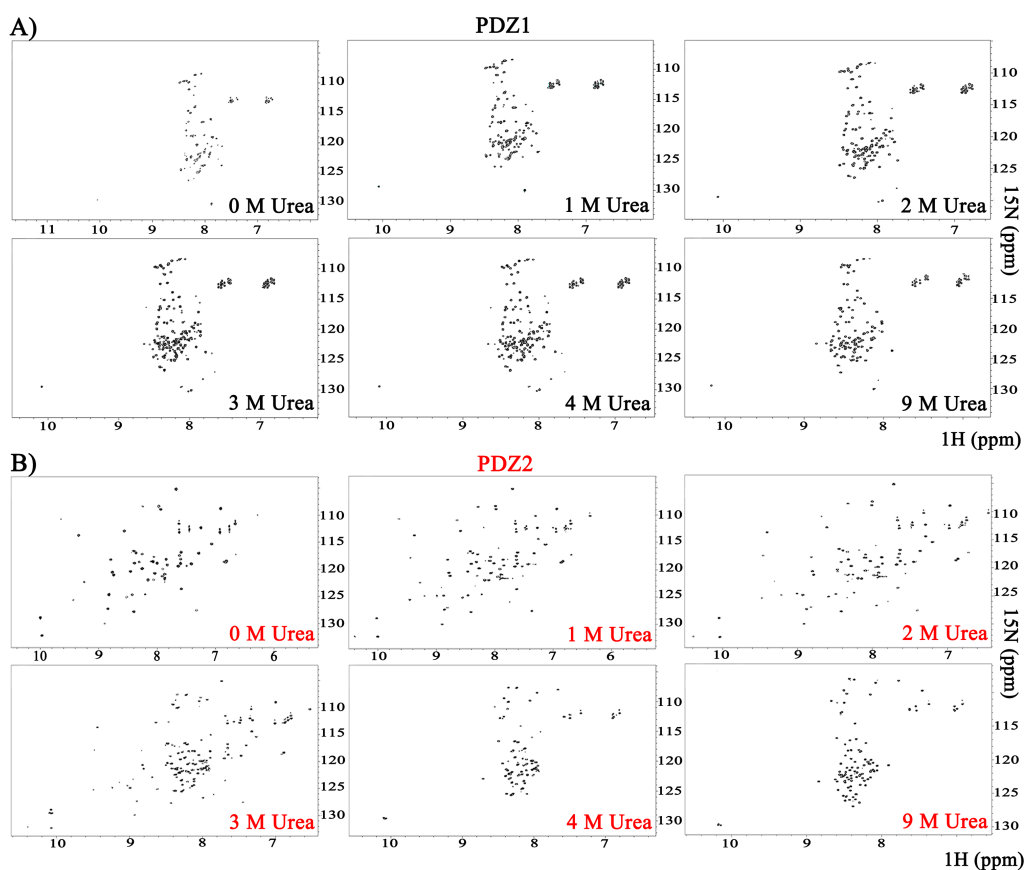


Figure 8: Urea titration experiments monitored by ^1H - ^{15}N HSQC in 950 MHz at 20°C. A) Isolated PDZ1. B) Isolated PDZ2. The figure was built using Adobe Fireworks CS6.

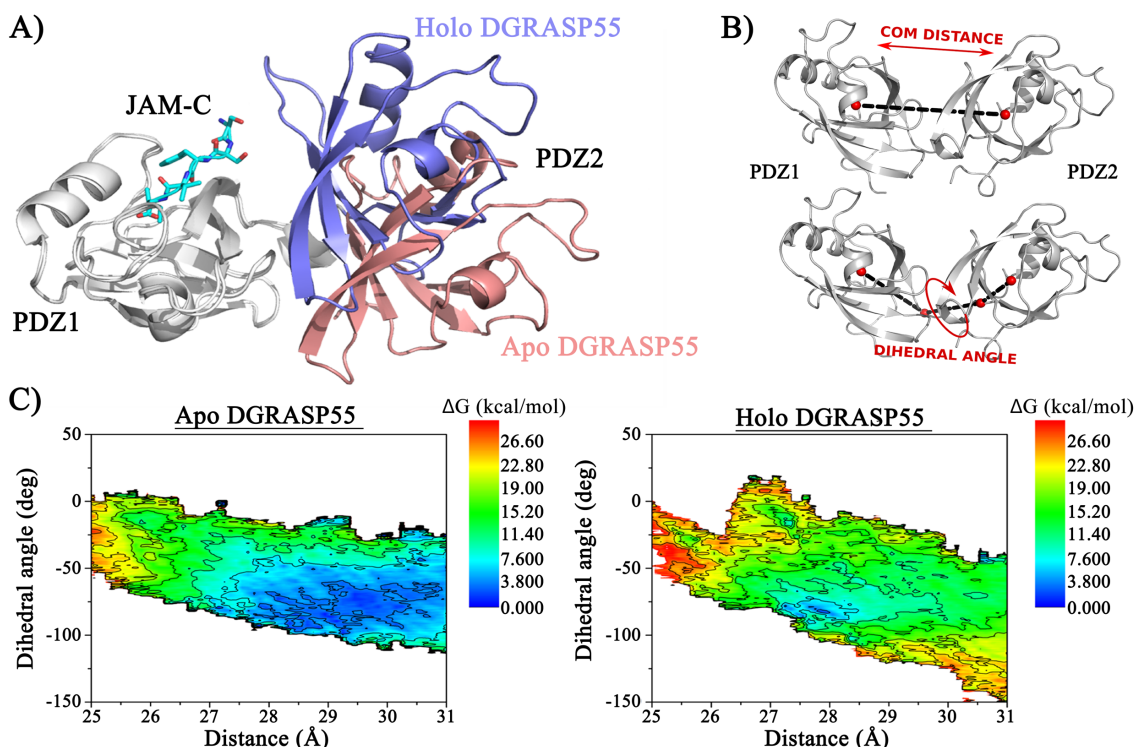


Figure 9: A) Superposition of crystallographic models of DGRASP55 in the presence (holo-DGRASP55) and in the absence (apo-DGRASP55) of JAM-C C-terminus peptide. The structures were aligned using only PDZ1 (represented in white). Holo-DGRASP55 exhibited a closed conformation characterized by a 33° rotation angle of the PDZ2 towards the PDZ1 in comparison with the apo structure. Apo and holo DGRASP55 PDZ2 are displayed in pink and blue, respectively. The peptide is represented as sticks and colored in cyan. B) The reactions coordinates used in ABF simulations. The first reaction coordinate was defined as the distance between the center of mass of the two PDZs to capture the opening/closing movements and the second reaction coordinate used in ABF simulations was set as the dihedral angle formed by the center of mass of the PDZs and the helix that connects the two domains to reproduce the twisting motions. C) Conformational free energy profile of apo and holo DRASP55. The large accessible area of the apo free energy profile implies a greater conformational flexibility. VMD [83] was used for visualization and figure preparation.

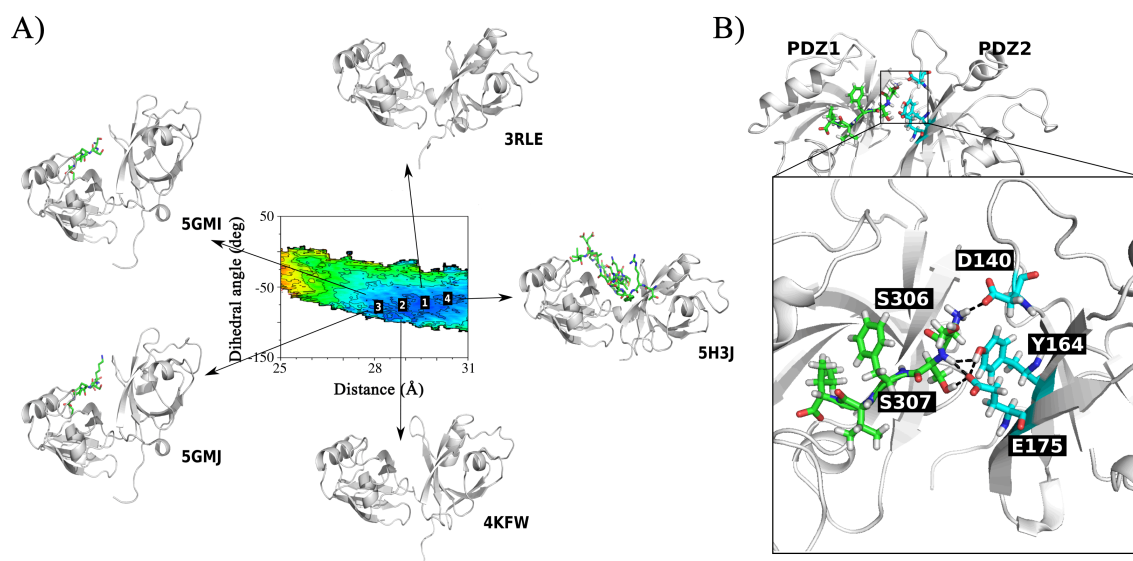


Figure 10: A) Available DGRASP55 crystallographic models and their coordinates in the free energy map: all the conformations observed in the experimental models are located at the accessible free energy area at physiological temperature for apo-DGRASP55. B) The molecular interactions between DGRASP55 and JAM-C C-terminus peptide. The residues involved in the PDZ2/JAM-C peptide interaction are shown with a stick model in which GRASP55 and JAM-C peptide backbones are colored cyan and green, respectively. VMD [83] was used for visualization and figure preparation.

Acknowledgments:

The authors thank the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the financial support through grants no. 2015/50366-7, 2012/20367-3, and 308380/2013-4. LFSM acknowledges FAPESP for the post-doctoral grants (Grant No. 2017/24669-8) and for the FAPESP Research Internships Abroad grant (Grant. No. 2016/09676-5). MRBB acknowledges FAPESP for the post-doctoral grant No. 2016/16328-3. AW acknowledges the UK BBSRC (research grant number BB/N006011/1). The authors acknowledge Monica Schroeder/Science Source for the Golgi cartoon used in the graphic abstract. We are grateful for the award of SRCD experimental time (SM16289) on the B23 Beamline at Diamond Light Source (UK) under the Membrane Protein Laboratory, funded by the Wellcome Trust grant No. 20289/Z16/Z and we

acknowledge the support of Dr Giuliano Siligardi, Dr Rohanah Hussein and Dr Charlotte Hughes. The Bruker console and probe upgrade to the 950 MHz NMR spectrometer was funded in Oxford with support from the Wellcome Institutional Strategy Support Fund, the John Fell Fund and the EPA Cephalosporin Fund.

Author contributions:

LFSM, AJCF, AW and CR designed research. LFSM collected all the experimental data. MRBB performed all the molecular dynamic simulations. LFSM, CR, AW, PJJ and AJCF performed the analyses. AJCF and AW conceived and coordinated the project. All authors contributed to the writing and reviewing of the results. All authors also approved the final version of the manuscript.

Authors' conflict of interest:

The authors declare no conflict of interest.