Conformational flexibility of GRASP protein and its constituent PDZ subdomains reveals structural basis of its 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.

Significance Statement:

Golgi Reassembly and Stacking Proteins (GRASPs) play pivotal roles in the maintenance of Golgi structure as well as in unconventional protein secretion. Their broad network of interactions is mainly sustained by the two-PDZ domains located in the N-terminal portion of the protein. The asymmetry of the PDZ domains in terms of number and diversity of interacting partners has been long recognized, but the molecular determinants of that asymmetry remains largely unknown. The biophysical data presented here provide a firm basis for understanding why PDZ1 behaves differently to PDZ2 in solution, despite their similar 3D structures. Furthermore, we propose that PDZ2 assist ligand binding to PDZ1, by means of conformational stabilization.

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, which includes the Golgi Reassembly and Stacking Proteins (GRASPs), which participate in cisternae stacking and lateral linkage in vertebrates. GRASPs also have critical roles in other processes, with an unusual ability to interact with several different protein 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 the interactions between GRASPs and their binding 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.

Introduction

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

GRASPs were first observed to be essential factors for Golgi reassembly *in vitro* [6,7] and *in vivo* [8]. They participate in the lateral linkage of Golgi stacks in vertebrates (building the Golgi ribbon), and also in all the Golgi remodelling necessary in migrating cells [9]. Current models suggest that GRASP-mediated Golgi stacking occurs via GRASP dimerization in a *trans* orientation [10], whereby a monomeric GRASP on one cistern face interacts with a second GRASP monomer on an opposite cistern [10]. The mechanism of GRASP dimerization (and whether higher-order oligomers exist and might play any role) is still unclear and of intense debate [10,11,12].

Although first identified as stacking factors, other functions of GRASPs have also been observed. GRASP can interact directly with the C-terminus of the Transforming Growth Factor- α (TGF- α) and with several members of the p24 family, facilitating the conventional protein secretion pathway [13,14]. GRASPs also have a direct and essential role in the unconventional secretion of a large number of proteins, such as: (1) the soluble acyl-coenzyme A binding protein (ACBP) in *Dictyostelium* [15], *Saccharomyces cerevisiae* and *Pichia pastoris* [16,17], (2) the Golgi bypass of α PS1 integrin during the stage 10B of *Drosophila* embryogenesis [18], (3) the mutant Δ F508 of the cystic fibrosis transmembrane conductance regulator (CFTR) in an ER stress situation [19], (4) in the secretion of IL1- β [20], among others [21,22]. GRASP is apparently also required for the export of non-protein molecules. In *Cryptococcus neoformans*, deletion of GRASP (CnGRASP) results in inefficient secretion of its main virulence factor GXM, followed by reduced capsule formation and attenuated virulence [23].

Although many functions associated with GRASPs have been identified, their detailed molecular mechanisms are still unclear. The crystal structures of the N-terminal domain of human GRASP55 (24), *Rattus norvegicus* GRASP55 (25) and *Rattus norvegicus* GRASP65 (25) have only recently been reported. GRASPs are composed of two structurally similar PDZ domains (named after the three first PDZ-proteins found: PSD-95, DLG and ZO-1) with an atypical protozoa-like arrangement, which together comprise the so-called GRASP domain (DGRASP). This finding also suggests that a new sequence analysis criterion might reveal a host of unidentified eukaryotic PDZ domains [24]. A second region in the GRASP structure consists of a poorly conserved domain that is rich in serine and proline (SPR domain) and that has a regulatory function [26,27] and intrinsically disordered characteristics [28].

One of the most intriguing properties of GRASPs is their ability to interact with a large variety of different protein partners. In addition to the Golgins, mammalian GRASPs also interact with cargo receptors from the conventional secretory pathway [13] and form part of a complex with RAB2, essential for normal protein transport and Golgi structure [29]. GRASPs have also been shown to interact with TGF- α [14], the potassium channel tetramerization domain-containing protein 5 (KCTD5) [30], with SEC16A and Δ F508 CFTR in response to ER stress [31,19], with JAM-B and JAM-C during spermatogenesis [32], with LC3-II on the autophagosomes and LAMP2 on late endosomes/lysosomes [33], and also with itself during dimerization [34]. These interactions involve the PDZ1 domain rather than PDZ2, despite the structural similarity seen in the reported crystal structures [11,12,19,25,32]. PDZ1 was also observed in crystal structures to be sufficiently promiscuous to interact with a non-native truncated SPR C-terminus [25].

In a previous report, we showed that GRASPs from the fungal pathogen *C. neoformans* [28] and *S. cerevisiae* [35] present structural features usually observed in molten globule proteins, even in the absence of any mild denaturing conditions. As expected for a collapsed intrinsically disordered protein, many physicochemical perturbations induced by the crystal packing, (including local dehydration and changes in the dielectric constant, along with increased PEG concentration used in the crystallographic reservoir), promote disorder-to-order transitions in GRASPs [35,36]. Even though GRASPs appear to have a special affinity for divalene-containing C-terminal proteins [27], the overall sequences of their protein partners do not have a conserved sequence or structural feature in common (Figure S1), suggesting GRASPs might be highly promiscuous in terms of protein-protein interactions. Therefore, this apparent functional asymmetry of the PDZs and their contrasting promiscuity of interactions with binding partners, might benefit from detailed studies of GRASPs in solution.

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

Materials and Methods:

Protein expression and purification:

CnGRASP GRASP domain (DGRASP) and isolated PDZ samples were expressed and purified as described elsewhere [28]. ¹³C- and/or ¹⁵N-labeled DGRASP and PDZs were produced using a protocol reported previously [37]. The same protocol was used for purification of both labelled and unlabelled protein [28].

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 μ m and 50 μ m were mainly used. The parameters were the same as those used for the conventional CD experiments.

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 [38] and/or Bruker TopSpin 3.5 and spectra were analyzed using the CCPN software [39]. ^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 [40,41] using ca. 19,500 water molecules and Na^+ and Cl^- ions were added to render the system neutral. Simulations were performed with NAMD [42] using periodic boundary conditions, the CHARMM36 [43] force-field for the protein and peptides, and the TIP3P model [44] 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 [45]. Covalent bonds involving hydrogen atoms in the protein were constrained to their equilibrium distances using the SHAKE algorithm, while SETTLE [46] was used for water.

The systems were equilibrated as follows: 1) 1,000 steps of conjugate-gradient steps (CG) [47] 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 $\text{C}\alpha$ 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 [48,49] 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 (Figure S2). 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 $10 \text{ kcal.mol}^{-1}\text{Å}^2$ 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 [50] was used for visualization and figure preparation.

Results and Discussion

DGRASP as a molten globule-like protein

^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -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 1A). CnGRASP, like the orthologue GRASP from *S. cerevisiae* [35], has previously been identified as a molten globule-like protein in solution, and belongs to the family of collapsed intrinsically disordered proteins [28]. We have previously suggested that the isolated DGRASP domain of CnGRASP shares some of the molten globule features observed in the full-length structure [28]. Here we extended that analysis by showing that molten globule-like behavior is indeed observed for the isolated GRASP domain (DGRASP) (Figure 1). In the absence of urea, native DGRASP bound the fluorophore ANS with high affinity, compared to the fully unfolded protein (Figure 1B), 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 1C). This low cooperativity during unfolding was also observed in the loss of regular secondary structure, as monitored by CD (Figure 1D) and in the change of the tryptophan fluorescence anisotropy (Figure 1E). Our data indicate that DGRASP is a molten globule-like protein, and that its three-dimensional structure unfolds in a non-cooperative way.

Figure 1: 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.

Although intrinsically disordered proteins (IDPs) may have significant secondary structure content, their hydrophobic cores are not stable. IDPs therefore have a three-dimensional conformation that fluctuates over a timescale of μs -ms [51,52]. IDPs are frequently observed as important components of the cellular signaling machinery [53] and may function as central hubs [54], especially due to their unique structural plasticity [51]. 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 [55,56]. 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 [57,58].

PDZ asymmetry

It has been previously observed that PDZs in tandem might assist each other during folding [59,60]. For example, for the PDZ12 tandem of the GRIP1 PDZ1–3 cassette, the folding of PDZ1 strictly depends on the covalent attachment to PDZ2 [59]. In addition, PDZ5 of GRIP

is completely unstructured in solution but, when covalently bound to PDZ4, both PDZ domains become well folded and stable [60]. We explored the behavior of the PDZs of DGRASP in solution using NMR methods. The DGRASP ^1H - ^{15}N HSQC spectrum showed around 10% of the total expected resonance lines, as previously observed for other induced molten globule-proteins (Figure 2 – 0 M urea) [52,61]. Furthermore, the DGRASP ^1H - ^{15}N HSQC spectrum showed low proton dispersion (separation of less than 0.7 ppm), a strong indication that those peaks might be located in disordered, highly dynamic regions.

Figure 2: DGRASP urea titration monitored by ^1H - ^{15}N HSQC at 950 MHz at 20 °C. The sample was titrated from 0 to 9 M in 1 M steps. Representative spectra are presented.

The absence of resonance lines in the ^1H - ^{15}N HSQC spectrum of DGRASP, is consistent with the presence of a molten globule-like fluctuating structure in the characteristic timescale window of μs -ms [52]. To recover the missing peaks, the structure is disrupted by increasing urea concentration, to either increase the local dynamics or decrease the number of “conformational states” the protein visits. The strategy of using urea unfolding to gain structural information by NMR has been previously used to study the molten globule states of α -lactalbumin and p53 [62,63].

Urea was titrated from 0 to 9 M in 1 M steps and selected ^1H - ^{15}N HSQC spectra are shown in Figure 2. The resonance peaks start to appear during the titration indicating that the

Not all peaks were present at the very beginning of the urea titration and only started to appear as the urea concentration was raised (Fig. 2), (a common feature of molten globule states). The gradual appearance of the peaks is attributed to the unfolding of the corresponding region and all the analyses from here on rely on this assumption: if the resonance peak of a non-proline residue X was only observed in the ^1H - ^{15}N HSQC at Y M urea, this is the urea concentration where this residue is now assumed to be in a fully unfolded structure. This assumption is valid, since these peaks are all located in a very narrow chemical shift region close to 8 ppm in the proton dimension, typical of unfolded structures and intrinsically disordered proteins [64].

To further explore our data, we monitored every peak that could be unambiguously tracked during the titration, and the urea concentration where each peak appeared during the titration was plotted as a function of the amino acid sequence, giving the local unfolding plot (Figure 4A). Figure 4A clearly shows the differences in the unfolding patterns of the two PDZs in solution, in contrast to the available crystal structures, which suggest that they are structurally identical to each other [12,24]. In general, ~60% of the PDZ1 resonances appear at urea concentrations below 2 M, and the median urea concentration for the unfolding of each amino acid is 2 M (Figure 4A). The most stable part of this subdomain seems to be the region between the amino acids 70 and 95 with urea stability of ca. 4 M. PDZ1 unfolding is therefore weakly cooperative and at least two main unfolding transitions are expected.

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 [65], but GRASP PDZs have a circular permutation where the first two β -strands are located at the end of the sequence ($\beta_3\alpha_1\beta_4\beta_5\alpha_2\beta_6\beta_1\beta_2$) [12,24]. The PDZ binding groove is therefore formed by the final, rather than the second, β -strand within the fold [24]. PDZ binding grooves are structured into a short α -helix (α_2) coupled with the abovementioned β_2 and our NMR data therefore suggest that the PDZ1 binding pocket is formed by a stable β -strand and a more flexible and unstable α -helix (Figure 4B). Since the PDZ interaction mechanism involves the formation of a β -sheet between the β_2 and the disordered C-terminus of the protein partner [66], a more flexible α -helix might be used to control the overall accessibility to this domain.

The opposite behavior was observed for PDZ2, where the majority of the resonance peaks appeared at 5-7 M urea, in a much sharper transition (Figure 4A-B). The average urea concentration for PDZ2 unfolding is 5.7 M. It is clear that the GRASP domain is formed by the

much more stable PDZ2 and a PDZ1 with at least two regions of different stability, where the more unstable one is predominant (>60%) (Figure 4B).

Figure 4. 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 was visible in the ^1H - ^{15}N HSQC. The low-stability region inside the PDZ1 was highlighted in the graphic. B) DGRASP molecular model (build on Swiss model [67] 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 figures were built in Pymol [68].

The existing crystal structure of the human GRASP65 GRASP domain (DGRASP65) indicates that PDZ1 is more flexible than PDZ2 based on the B-factor values (Figure 4C). Instability during chemical unfolding may be closely related to structural flexibility, especially in the region close to and surrounding the binding pocket of PDZ1.

Dissecting the GRASP domain

Using the domain predictor Pfam (<http://pfam.xfam.org>, accessed in 2017), the disorder prediction published previously [28] and the information from our own solution NMR

experiments, we designed constructs for each individual PDZ domain. Residues 1-115 were assigned to “PDZ1” and 116-220 to “PDZ2” and the two constructs were expressed and purified by following the same protocol used for the full-length DGRASP.

SRCD data indicate that the PDZs have different secondary structure content (Figure 5A, Table 1). In solution, PDZ1 possesses lower helical content (approximately three times less than PDZ2) and more strand structures, in contrast to predictions based on the crystal structures, in which similar secondary structure content was observed [24,25]. It has previously been shown that GRASP structure is sensitive to changes in the physicochemical parameters of the medium, in particular the variation in the dielectric constant. [36]. High PEG concentrations have also been shown to induce disorder-to-order transitions in GRASP, so we could expect that more disordered regions might become more ordered in conditions used for crystallization [36].

Table 1: SRCD spectral deconvolution performed using the software CDSSTR and an appropriate database in the online server Dichroweb [69].

	Helix	Strand	Turns	Disord.	NRMSD
PDZ1	5%	37%	22%	34%	0.040
PDZ2	18%	30%	22%	30%	0.027

Figure 5: 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 [70]. 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) ^1H NMR of PDZ1 (black) and PDZ2 (red) with some resonance regions highlighted in the figure for illustration.

Changes in protein tertiary structure may be monitored using near-UV (250-350 nm) CD spectra [71]. If a protein has no well-defined 3D structure (e.g. IDPs or "molten-globule" structures), the signals in the near-UV region will be close to zero [72]. 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 [71].

PDZ1 (PDZ2) has 4 (6) tyrosines, 4 (3) phenylalanines, and 1 (3) tryptophans. Therefore, both PDZs have comparable numbers 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 5B). 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 1), individual PDZs show distinct unfolding transitions when compared to DGRASP and to each other (Figure 5C). 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 5C). This CD data is consistent with the NMR data, in which a significant part of the structure appears to be unfolded at the beginning of the urea titration. It is important to emphasize that CD and NMR provide complementary information: NMR allows us to analyse local denaturation, whereas, in the CD experiment, we observe global structural differences between the folded and the unfolded states. 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).

^1H NMR spectra of proteins are rarely informative because spectral crowding hinders assignment. However, by focusing first on the amide resonance regions, it is possible to see differences between the ^1H spectra of the individual PDZs (Figure 5D). PDZ2 has multiple peaks spread over the region 6-10.5 ppm, with most of them being very sharp, suggesting that PDZ2 behaves like a regular, ordered protein. A remarkably different behavior is observed for PDZ1 (Figure 5D). The resonances of the amide protons are distributed over a narrower region (~ 2 ppm) and the line superposition is much more severe. Even though PDZ1 has a poorer dispersion than PDZ2, this dispersion is still greater than those observed in extended intrinsically disordered proteins [73]. Of course, this was expected since neither CnGRASP nor its GRASP domain (DGRASP) is an extended IDP.

The NMR spectra suggest that DGRASP unfolds non-cooperatively and that PDZ2 is more stable, with PDZ1 having at least two regions with different stabilities. Urea titration and NMR experiments with the isolated PDZs validated the initial observations. PDZ1 shows the same pattern observed for DGRASP with fewer peaks at the very beginning of the titration but with a rapid appearance of the remaining lines as the urea concentration is increased (Figure 6A). Most of the peaks missing in the spectrum of the native structure, appeared at urea concentrations around 1-2 M, the same pattern observed in DGRASP (Figure 3B). This unfolding transition is associated with the part of the domain predicted to be the binding site for molecular partners, and it is interesting to note that its unfolding appears to be unaffected by the presence or absence of PDZ2. The remaining lines appeared around 3-4 M urea, suggesting that the most stable region of PDZ1 in the DGRASP construct decreases its stability in the absence of PDZ2 (from ~ 5 to 3-4 M urea). Interestingly, if we superimpose the PDZ1 and DGRASP spectra acquired in the absence of urea, it is clear that all resonances present for the DGRASP at the beginning of the titration correspond to PDZ1 residues (Figure S2). As discussed before, this suggests these regions are within disordered/flexible structures, and supports the transfer of the assignments from the 9 M urea condition.

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 6B). 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.

Figure 6: Urea titration experiments monitored by ^1H - ^{15}N HSQC in 950 MHz at 20°C. A) Isolated PDZ1. B) Isolated PDZ2.

The long-range coupled mechanism of action

Our NMR and SRCD data show that the PDZ domains of GRASPs have different structural stabilities and dynamics, but it is not clear how these properties are exploited to enable the protein to bind multiple 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 [74]. 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 [74,75].

To understand the cellular role of PDZ2, 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 7A). 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) is large, however when the protein partner is bound (holo) the profile is greatly reduced (Figure 7B). The data suggest that the protein partner significantly reduces the conformational space that DGRASPs can access by stabilizing one particular conformation.

Figure 7: (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) Conformational free energy profile of apo and holo DRASP55. The large accessible area of the apo free energy profile implies a greater conformational flexibility.

Previous crystallographic data have suggested that DGRASP55 transitions from an open to a closed conformation on binding JAM-C and JAM-B [74], however, this closure is not observed when DGRASP55 binds the C-terminus of Golgin45 [75], 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 8). 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 7B). 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 8). At physiological temperature the conformational flexibility of DGRASPs allows large movements in the relative orientations of the two domains, allowing different conformations to be adopted for different binding partners, thereby creating numerous potential secondary interactions via PDZ2. 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.

Figure 8: 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.

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 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.

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). 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.

References:

- 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) 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.
- 7) 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.
- 8) 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.
- 9) 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.
- 10) Rabouille C, Linstedt AD. GRASP: A Multitasking Tether. *Front Cell Dev Biol.* 2016 Jan 26;4:1.
- 11) 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.
- 12) 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.
- 13) 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.
- 14) 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.
15. Kinseth 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.
- 16) 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.
- 17) 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.
- 18) 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.

- 19) 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.
- 20) 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.
- 21) Giuliani F, Grieve A, Rabouille C. Unconventional secretion: a stress on GRASP. *Curr Opin Cell Biol*. 2011 Aug;23(4):498-504.
- 22) Rabouille C. Pathways of Unconventional Protein Secretion. *Trends Cell Biol*. 2017 Mar;27(3):230-240.
- 23) 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.
- 24) 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.
- 25) 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.
- 26) 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.
- 27) Vinke FP, Grieve AG, Rabouille C. The multiple facets of the Golgi reassembly stacking proteins. *Biochem J*. 2011 Feb 1;433(3):423-33.
- 28) 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.
- 29) 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.
- 30) 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.
- 31) 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.
- 32) 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.
- 33) 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.
- 34) 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.
- 35) 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.
- 36) 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.
- 37) 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).

- 38) 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.
- 39) 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.
- 40) 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.
- 41) 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.
- 42) 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.
- 43) 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.
- 44) 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.
- 45) 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.
- 46) 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.
- 47) 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.
- 48) 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.
- 49) 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.
- 50) Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996 Feb;14(1):33-8, 27-8.
- 51) Uversky VN. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta*. 2013 May;1834(5):932-51.
- 52) Redfield C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods*. 2004 Sep;34(1):121-32.
- 53) Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol*. 2015 Jan;16(1):18-29.
- 54) 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.
- 55) 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.
- 56) 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.
- 57) 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.
- 58) 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.
- 59) 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.
- 60) Zhang Q, Fan JS, Zhang M. (2001). Interdomain chaperoning between PSD-95, Dlg, and Zo-1 (PDZ) domains of glutamate receptor-interacting proteins. *J Biol Chem.* Nov 16;276(46):43216-20.
- 61) 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.
- 62) 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.
- 63) 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.
- 64) 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.
- 65) Lee HJ, Zheng JJ. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal.* 2010; 8: 8.
- 66) Sheng M, Sala C. PDZ domains and the organization of supramolecular complexes. *Annu Rev Neurosci.* 2001;24:1-29.
- 67) 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.
- 68) DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography.* 202. 40, 82-92.
- 69) Whitmore L, Wallace BA. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 2004 Jul 1;32(Web Server issue):W668-73.
- 70) Whitmore L, Wallace BA. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 2004 Jul 1;32(Web Server issue):W668-73.
- 71) Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochim Biophys Acta.* 2005 Aug 10;1751(2):119-39.
- 72) Vassilenko KS, Uversky VN. Native-like secondary structure of molten globules. *Biochim Biophys Acta.* 2002 Jan 31;1594(1):168-77.
- 73) Isabella C. Felli and Roberta Pierattelli Editors Intrinsically Disordered Proteins Studied by NMR Spectroscopy. *Advances in Experimental Medicine and Biology.* Volume 870.
- 74) 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.
- 75) 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

Supplementary Information

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

Figure S1: The C-terminus of different GRASP55/65 proteins partners were aligned using T-Coffee (<https://www.ebi.ac.uk/Tools/msa/tcoffee/> accessed in 2018).

NORMAL MODE ANALYSIS

REACTION COORDINATES

Figure S2: The reactions coordinates used in ABF simulations were chosen based on normal mode analysis (NMA). Briefly, normal modes describe the protein's natural motions, called normal modes and their associated frequencies. Biologically relevant modes, i.e, modes with low frequency are associated with large-scale motions. Additionally, if there are two known conformations of a protein, the normal modes that most contribute to the conformational change can be used to understand the pathway and to describe the conformational transitions in only a few degrees of freedom. To characterize GRASP collective motions, we performed ANM (Anisotropic Normal Mode) analysis using ProDy [Ref S1]. Visualization of the motion along the two modes with lower frequencies reveals opening/closing and twisting movements. Based on this analysis, the first reaction coordinates coordinate was defined as 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.

Ref S1: Bakan A, Meireles LM, Bahar I. ProDy: Protein Dynamics Inferred from Theory and Experiments Bioinformatics 2011 27(11):1575-1577.