

Supplementary Information

Extensive Sampling of the Cavity of the GroEL Nanomachine by Protein Substrates Probed by Paramagnetic Relaxation Enhancement

Marielle A. Wälti, David S. Libich and G. Marius Clore*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases,
National Institutes of Health, Bethesda, MD 20892-0520

Experimental Methods and 4 Figures.

Experimental

Expression, purification and choice of Fyn SH3 variants. Wild type *Gallus gallus* Fyn SH3 (SH3^{WT}), the triple A39V/N53P/V55L mutant (SH3^{Mut}), and a three-residue C-terminal truncation of the wild type (SH3^{WT}Δ57) were expressed with a His₆ purification tag, uniformly ¹⁵N-labeled and purified as described previously.¹ The final gel filtration step was performed using Chelex-100 (Sigma Aldrich) treated 50 mM Tris buffer, pH 7.0 to ensure the removal of any paramagnetic metal ions. The final samples were concentrated with a 3,500 molecular weight cutoff centrifugal filter (Millipore) to a concentration of 0.4-0.5 mM, flash frozen with liquid nitrogen and stored at -80 °C. Tobacco Etch Mosaic virus protease used to cleave the N-terminal His tag was purified as described previously.²

At 10 °C, SH3^{WT} is entirely in the native, fully-folded state,³ SH3^{Mut} exists as a dynamic mixture between the major native species (~98%) and an excited, folding intermediate (~2%);³⁻⁴ and the folding intermediate mimetic SH3^{WT}Δ57 is entirely in the intermediate state.⁵ The intermediate state has essentially the same fold as the native state with the exception that the C-terminal β5 strand (residues 55-58) is disordered and no longer part of a β-sheet.⁵ In our previous relaxation-based work on the kinetics of the interaction of SH3 folding intermediate mimetics with apo GroEL, we also studied a very closely related construct to SH3^{WT}Δ57 comprising a four-residue C-terminal truncation of the triple mutant, SH3^{Mut}Δ56.⁵ The latter was not employed in the current work as, free in solution, it comprises a rapidly interconverting mixture between the major monomeric intermediate (~92%) with essentially the same structure as SH3^{WT}Δ57, a dimeric form of the intermediate (~4%) and a fully unfolded state (~4%)⁵ that could potentially complicate interpretation of the PRE data. Free SH3^{WT}Δ57 shows no ¹⁵N relaxation dispersion while large dispersions are observed for SH3^{Mut}Δ56; quantitative kinetic analysis of the modulation of ¹⁵N dispersions by GroEL, in conjunction with lifetime line broadening and dark state exchange saturation transfer data, led to the identification of two bound states: one immobilized by binding to the walls of the GroEL cavity; the other mobile, confined and enriched within the GroEL cavity.⁵ The current PRE data on SH3^{WT}Δ57 are fully consistent with the kinetic analysis of the relaxation-based data on the SH3^{Mut}Δ56-GroEL interaction (see main text).

GroEL mutations. The GroEL^{E315C} DNA sequence with all natural cysteines mutated to alanines and a cysteine introduced at position 315 (C138A, C458A, C519A, E315C) was synthesized by Genscript.⁶ We subsequently introduced a point mutation at position 315 (C315E) to obtain the Cys0 construct corresponding to the wild type sequence without any cysteines.⁶ The following mutations were introduced into the GroEL^{Cys0} construct using the QuickChange site-directed mutagenesis kit from Qiagen: S43C, A138C, R231C, R268C and G545C. The latter construct also contained an accidental deletion at Gly539; however, as the disordered C-terminal tail extends from residue 528 to 548, this deletion has no structural impact.

Expression and purification of GroEL. The initial purification in which GroEL is separated from other unbound proteins was carried out as described previously with 2 mM dithiothreitol (DTT) or 1 mM tris(2-carboxyethyl)phosphine) (TCEP) added to all buffers.⁶⁻⁷ Briefly, the plasmids for all mutants were transformed into BL21 StarTM (DE3) *E. coli* cells (Invitrogen), and expressed in Luria-Bertini broth (LB) at 37 °C. Induction was carried out at an OD₆₀₀ = 0.8 with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the cultures grown overnight at 20 °C. Cells were lysed and nucleic acids removed by streptomycin sulfate precipitation. The sample was centrifuged for 25 min at 50,000 x g and the supernatant applied to a self-packed 60 ml ion exchange column (Q Sepharose FF, GE Healthcare). GroEL was subjected to ammonium sulfate precipitation and then injected on to a HiPrep Sephacryl S-300 gel filtration column (GE Healthcare). The final clean-up procedure to remove proteins bound within the GroEL cavity was carried out either by acetone precipitation⁷

in the case of the S43C, A138C and G545C mutants, or using the recently described disassembly/reassembly protocol⁶ in the case of R231C, R268C and E315C mutants. The more time-consuming disassembly/reassembly procedure was only used in those cases where the acetone precipitation procedure either failed or resulted in very low yields. Assembly into a tetradecamer was confirmed by blue native polyacrylamide gel electrophoresis (BN-PAGE), and purity was assessed by the absence of tryptophan fluorescence⁸, sodium dodecylsulfate gel electrophoresis (SDS-PAGE) and mass spectrometry.

For the purposes of the current study, the effect of the various mutations used for spin-labeling on the ability of GroEL to undergo the various conformational changes accompanying ATP hydrolysis and binding of GroES is irrelevant, as the aim of the current work is directed at probing protein substrate sampling of the cavity of GroEL in the apo state. However, many of the sites of mutations used here have been used previously for functional studies with fluorescent labels, including S43C,⁹⁻¹⁰ R231C¹¹ and E315C.^{9,12} The A138C mutation simply reintroduces the native cysteine at position 128. The tail is intrinsically disordered and hence the G545C mutation is unlikely to have any significant effect on function, especially since GroEL with a 7-residue C-terminal deletion at position 541 is still functional.¹³ Although there is no data in the literature on the R268C mutant, the pattern of PREs observed for the spin-labeled R231C and R268C variants are very similar (see Fig. 2, main text).

Site-specific paramagnetic labeling. N-[S-(2-Pyridylthio)cysteaminy]ethylenediamine-N,N,N',N'-tetraacetic acid (Toronto Research Chemicals catalog no. P996250) was dissolved in methanol and incubated overnight with an excess of either Mn²⁺ or Ca²⁺. Conjugation of the cysteaminy paramagnetic (EDTA-Mn²⁺) and diamagnetic (EDTA-Ca²⁺) tags to surface engineered cysteines on GroEL was carried out prior to the final clean-up procedure (i.e. either before acetone precipitation, or before reassembly into the tetradecamer). Briefly, DTT and TCEP were removed by overnight dialysis into 50 mM Tris. In the case of the disassembly/reassembly protocol the buffer also included 3 M urea. The GroEL mutants were incubated at room temperature overnight with a 4-10 fold molar excess of the paramagnetic or diamagnetic label. The conjugation reaction was verified by mass spectrometry. The samples were further purified using the appropriate final clean-up protocol, stored at 4 °C, and used within 2 weeks after paramagnetic labeling.

Control of interaction between Fyn SH3 and the paramagnetic spin labels. To determine whether the PREs observed on the SH3 constructs with the various EDTA-Mn²⁺ paramagnetically-labeled GroEL constructs were not simply due to preferential interaction with the paramagnetic tag itself, we made use of two controls: soluble EDTA-Mn²⁺ in the form of hydroxylamine-EDTA-Mn²⁺ (hEDTA-Mn²⁺; Ricca Chemical Co.) which has the same charge as EDTA-Mn²⁺ conjugated to a protein, and a surface engineered cysteine mutant (E38C) of the completely unrelated maltose binding protein (MBP^{E32C}). The latter was expressed, purified and labeled with EDTA-Mn²⁺ (paramagnetic sample) or EDTA-Ca²⁺ (diamagnetic control) as described previously.¹⁴

NMR spectroscopy. All samples were buffer exchanged into chelex treated 50 mM Tris buffer, pH 7.0. Samples for NMR contained 100 μM ¹⁵N-labeled SH3 domain, and paramagnetically-labeled GroEL, MBP^{E38C} or hEDTA-Mn²⁺. For the diamagnetic controls, Mn²⁺ was replaced by Ca²⁺. In addition, all NMR samples contained 0.03% (wt/vol) NaN₃ and 5% (vol/vol) D₂O.

All NMR experiments were carried out at 10 °C on a Bruker 800 MHz NMR spectrometer equipped with a z gradient triple resonance cryoprobe. Assignments were taken from our previous work.^{3,5} ¹H_N-R₂ values were obtained using a ¹H-¹⁵N correlation-based experiment with a two-time point measurements recorded in an interleaved manner as described previously.¹⁵ The ¹H_N-Γ₂ values are given by the difference ¹H_N-R₂ transverse relaxation rates recorded for paramagnetic and diamagnetic samples and calculated as:

$$\Gamma_2 = 1/(T_a - T_b) \cdot \ln[I_{\text{dia}}(T_b) \cdot I_{\text{para}}(T_a) / I_{\text{dia}}(T_a) \cdot I_{\text{para}}(T_b)],$$

where T_a and T_b are the two time points, and I_{dia} and I_{para} are peak intensities (i.e. peak heights) for the diamagnetic (Ca^{2+}) and paramagnetic (Mn^{2+}) states, respectively.¹⁶

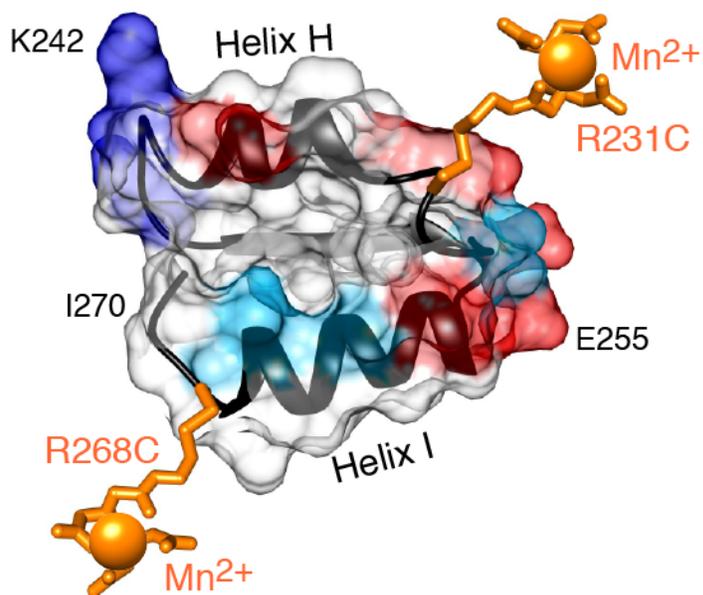


Figure S1. Molecular surface of a portion of the inner rim of the GroEL apical domain illustrating the location of the paramagnetic labels at R231C and R268C encompassing a hydrophobic patch formed by helices H and I. Color code: white, hydrophobic; cyan, hydrophilic; blue, positively charged; red, negatively charged. Helices H and I are shown as ribbons.

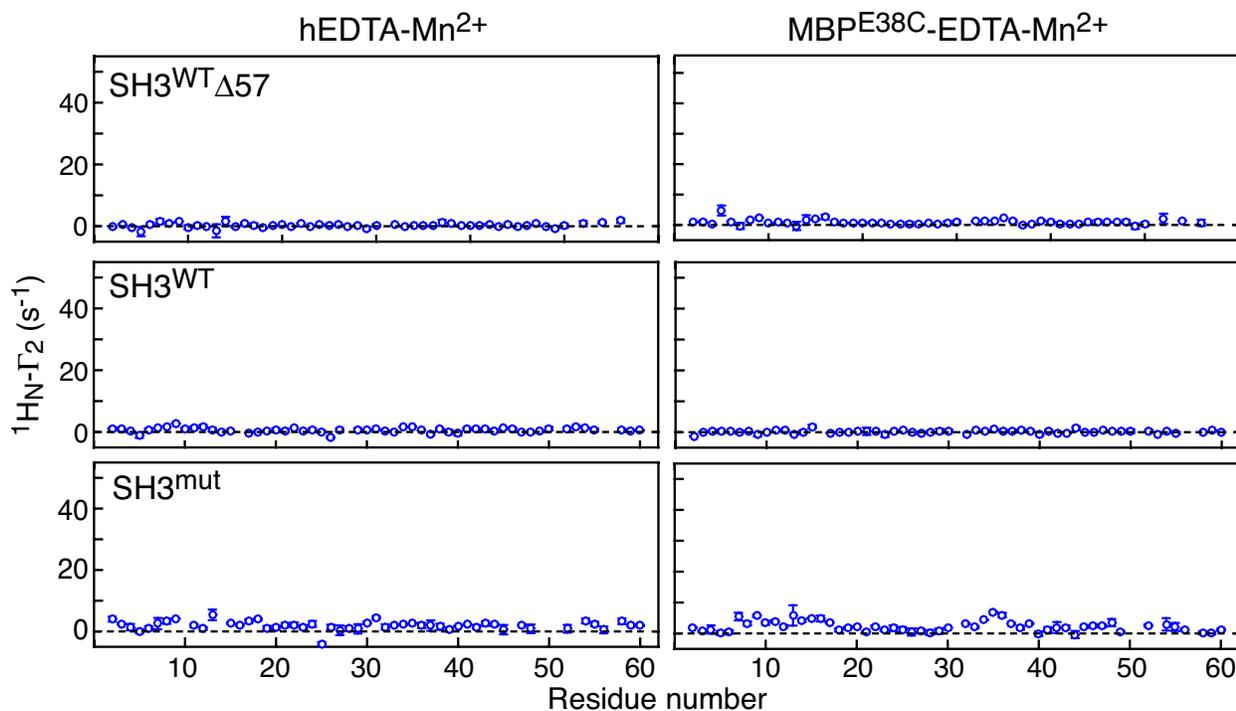


Figure S2. Control experiments demonstrating the absence of any significant intermolecular PREs observed on ¹⁵N-labeled SH3^{WT}Δ57 (top), SH3^{WT} (middle) or SH3^{Mut} (bottom) arising from the presence of either soluble Mn²⁺ (in the form of hEDTA-Mn²⁺) or the protein MBP^{E38C} paramagnetically-labeled with EDTA-Mn²⁺. ¹H_N-T₂ rates are measured as the difference in ¹H_N-R₂ rates between paramagnetic (Mn²⁺) and diamagnetic (Ca²⁺) samples. Data were recorded at 800 MHz and 10 °C on samples containing 100 μM SH3 domain and 100 μM paramagnetically-labeled component (hEDTA or MBP^{E38C}) in 50 mM Tris buffer pH 7.0, 5% (v/v) D₂O and 0.03% (w/v) NaN₃.

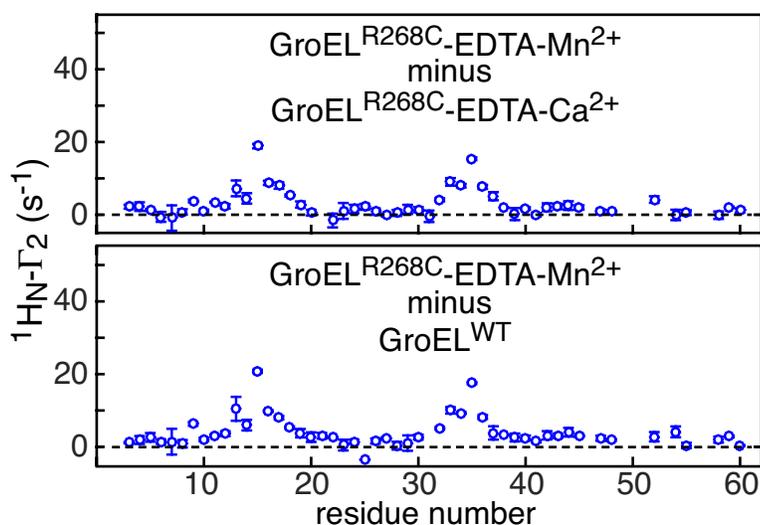


Figure S3. Control experiments demonstrating that the same intermolecular PREs for SH3^{Mut} are obtained using either GroEL^{R268C} conjugated to EDTA-Ca²⁺ (top) or wild type GroEL (bottom) as the diamagnetic reference sample. These data indicate that lifetime line broadening arising from interaction of SH3^{Mut} with diamagnetic GroEL is the same for wild type and R268C-EDTA-Ca²⁺ tagged GroEL, and that there is therefore no preferential binding to the EDTA-Ca²⁺ tag. These control experiments complement those reported in Fig. S1.

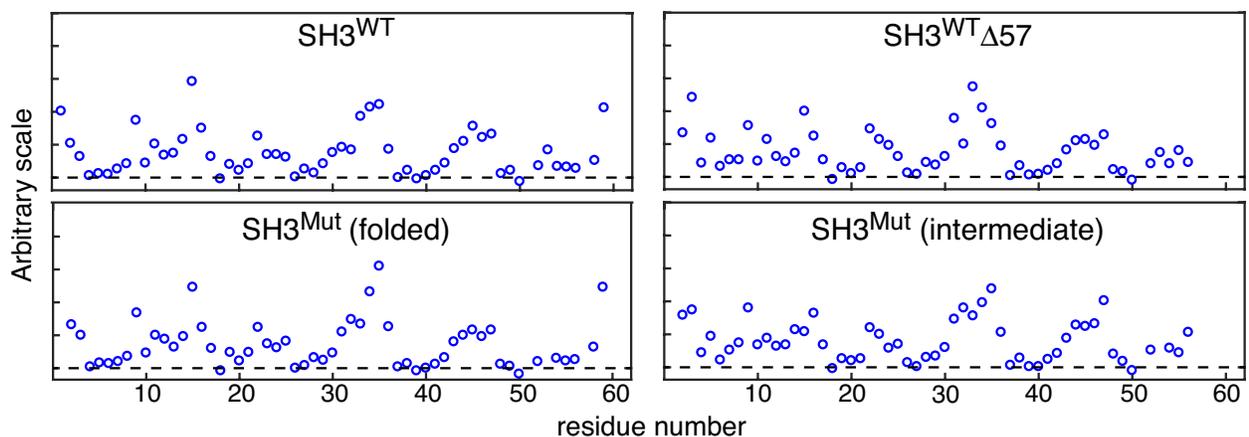


Figure S4. Predicted solvent PREs (arising from random collisions with the paramagnetic label) calculated on the basis of solvent accessibility of backbone amide protons using the program Xplor-NIH¹⁷⁻¹⁸ (as described by Iwahara et al.¹⁹) for SH3^{WT} (PDB 3UA6, with Val1 and Val5 substituted by Ser and Glu, respectively, to obtain the same sequence as found in *Gallus gallus*),²⁰ SH3^{WT}Δ57,⁵ SH3^{Mut} (native, fully folded; PDB 2LP5)⁴ and SH3^{Mut} (folding intermediate; PDB 2LP2)⁴. The only significant structural difference between the fully folded state (SH3^{WT} and SH3^{Mut} folded) versus the folding intermediate state (SH3^{Mut} intermediate and SH3^{WT}Δ57) lies in the presence or absence, respectively, of the C-terminal β5-strand.

Supplementary references

1. Maxwell, K. L.; Davidson, A. R., Mutagenesis of a Buried Polar Interaction in an SH3 Domain: Sequence Conservation Provides the Best prediction of Stability Effects. *Biochemistry* **1998**, *37*, 16172-16182.
2. Tropea, J. E.; Cherry, S.; Waugh, D. S., Expression and Purification of Soluble His6-Tagged TEV Protease. *Methods Mol. Biol.* **2009**, *498*, 297-307.
3. Libich, D. S.; Tugarinov, V.; Clore, G. M., Intrinsic Unfoldase/Foldase Activity of the Chaperonin GroEL Directly Demonstrated Using Multinuclear Relaxation-Based NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 8817-8823.
4. Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundstrom, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E., Structure of an Intermediate State in Protein Folding and Aggregation. *Science* **2012**, *336*, 362-366.
5. Libich, D. S.; Tugarinov, V.; Ghirlando, R.; Clore, G. M., Confinement and Stabilization of Fyn SH3 Folding Intermediate Mimetics Within the Cavity of the Chaperonin GroEL Demonstrated by Relaxation-Based NMR. *Biochemistry* **2017**, *56*, 903-906.
6. Walti, M. A.; Clore, G. M., Disassembly/Reassembly Strategy for the Production of Highly Pure GroEL, a Tetradecameric Supramolecular Machine, Suitable for Quantitative NMR, EPR and Mutational Studies. *Protein Expr. Purif.* **2018**, *142*, 8-15.
7. Grason, J. P.; Gresham, J. S.; Widjaja, L.; Wehri, S. C.; Lorimer, G. H., Setting the Chaperonin Timer: the Effects of K⁺ and Substrate Protein on ATP Hydrolysis. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17334-17338.
8. Todd, M. J.; Lorimer, G. H., Criteria for Assessing the Purity and Quality of GroEL. *Methods Enzymol* **1998**, *290*, 135-141.
9. Weaver, J.; Rye, H. S., The C-terminal Tails of the Bacterial Chaperonin GroEL Stimulate Protein Folding by Directly Altering the Conformation of a Substrate Protein. *J. Biol. Chem.* **2014**, *289*, 23219-23232.
10. Chen, D. H.; Madan, D.; Weaver, J.; Lin, Z.; Schroder, G. F.; Chiu, W.; Rye, H. S., Visualizing GroEL/ES in the Act of Encapsulating a Folding Protein. *Cell* **2013**, *153*, 1354-1365.
11. Motojima, F.; Yoshida, M., Polypeptide in the Chaperonin Cage Partly Protrudes Out and then Folds Inside or Escapes Outside. *EMBO J.* **2010**, *29*, 4008-4019.
12. Grason, J. P.; Gresham, J. S.; Lorimer, G. H., Setting the Chaperonin Timer: a Two-Stroke, Two-Speed, Protein Machine. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17339-17344.
13. Machida, K.; Kono-Okada, A.; Hongo, K.; Mizobata, T.; Kawata, Y., Hydrophilic residues ⁵²⁶KNDAAD⁵³¹ in the Flexible C-terminal Region of the Chaperonin GroEL are Critical for Substrate Protein Folding Within the Central Cavity. *J. Biol. Chem.* **2008**, *283*, 6886-6896.
14. Tang, C.; Ghirlando, R.; Clore, G. M., Visualization of Transient Ultra-Weak Protein Self-Association in Solution Using Paramagnetic Relaxation Enhancement. *J. Am. Chem. Soc.* **2008**, *130*, 4048-4056.
15. Iwahara, J.; Tang, C.; Clore, G. M., Practical Aspects of ¹H Transverse Paramagnetic Relaxation Enhancement Measurements on Macromolecules. *J. Magn. Reson.* **2007**, *184*, 185-195.
16. Iwahara, J.; Schwieters, C. D.; Clore, G. M., Ensemble Approach for NMR Structure Refinement Against ¹H Paramagnetic Relaxation Enhancement Data Arising From a Flexible Paramagnetic Group Attached to a Macromolecule. *J. Am. Chem. Soc.* **2004**, *126*, 5879-5896.
17. Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M., The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Reson.* **2003**, *160*, 65-73.
18. Schwieters, C. D.; Bermejo, G. A.; Clore, G. M., Xplor-NIH for Molecular Structure Determination From NMR and Other Data Sources. *Protein Sci.* **2018**, *27*, 26-40.
19. Iwahara, J.; Zweckstetter, M.; Clore, G. M., NMR Structural and Kinetic Characterization of a Homeodomain Diffusing and Hopping on Nonspecific DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15062-15067.
20. Martin-Garcia, J. M.; Luque, I.; Ruiz-Sanz, J.; Camara-Artigas, A., The promiscuous binding of the Fyn SH3 domain to a peptide from the NS5A protein. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 1030-1040.