# Supporting Information

## SANS investigation of fungal loosenins reveal substrate dependent impacts of protein action on inter-fibril distance and packing order of cellulosic substrates

Deepika Dahiya[1], Zsuzsanna Péter-Szabó[2], Manjula Senanayake[4], Sai Venkatesh Pingali[4], Wellington Leite[4], James Byrnes[5], Garry W. Buchko[6], Pramod Sivan[2], Francisco Vilaplana[2,3], Emma R. Master[1,7*], Hugh M. O'Neill[4*]

[1]Department of Bioproducts and Biosystems, Aalto University, Kemistintie 1, 02150 Espoo, Finland.

[2]Division of Glycoscience, Department of Chemistry, KTH Royal Institute of Technology, Stockholm, SE-106 91, Sweden.

[3]Wallenberg Wood Science Centre, KTH Royal Institute of Technology, Stockholm, SE-100 44, Sweden.

[4]Neutron Scattering Division and Center for Structural Molecular Biology, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA.

[5]Brookhaven National Laboratory, National Synchrotron Light Source II, Bldg. 745, P.O. Box 5000, Upton, NY 11973-5000

[6]Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99354, United States of America; School of Molecular Biosciences, Washington State University, Pullman, WA, 99164, United States of America

[7]Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College Street, M5S 3E5, Toronto, Ontario, Canada

* co-corresponding

**Table S1.** Chemical shifts (ppm) for the $^{13}C^{b}$ carbons of the four cysteine residues in oxidized *Pca*LOOL12.

| | **C26** | **C50** | **C73** | **C76** |
|---|---|---|---|---|
| ***Pca*LOOL12** | 39.4 | 40.1 | --- | 35.7 |

**Table S2.** Summary of the statistics for the NMR solution structures of *Pca*LOOL12.[a]

| Protein Name | *Pca*LOOL12 |
|---|---|
| PDB ID | 9CE9 |
| BMRB ID | 31180 |
| **Restraints for Structure Calculations** | |
| Total NOEs | 1300 |
| Intraresidue NOEs | 290 |
| Sequential (i, i + 1) NOEs | 327 |
| Medium-range (i, i + j; 1 < j ≤ 4) NOEs | 121 |
| Long-range (i, i + j; j > 4) NOEs | 492 |
| Phi (□) angle restraints | 75 |
| Psi (□) angle restraints | 75 |
| Hydrogen bond restraints (2 restraints/bond) | 68 |
| Disulfide bond restraints (3 restraints/bond) | 6 |
| **Structure Calculations** | |
| Number of structures calculated | 100 |
| Number of structures used in ensemble | 20 |
| **RMSD to Mean (Å) for PSVS Ordered Residues[b]** | |
| Backbone N-C□-C=O Atoms | 0.50 ± 0.07 Å |
| All Heavy Atoms | 0.84 ± 0.06 Å |
| **Ramachandran Plot Summary (from Procheck)** | |
| Most favored regions | 90.7 % |
| Additionally allowed regions | 9.3 % |
| Generously favored regions | 0.0 % |
| Disallowed | 0.0 % |
| **Global Quality Scores (Z-score (raw))[c]** | |
| Procheck (all) | -3.31(-0.56) |
| Procheck (□□□□) | -2.20 (-0.64) |
| MolProbity clash score | 0.30 (7.15) |

[a]All statistics are for the ensemble deposited in the Protein Data Bank.
[b]Ordered residues: T10 -T21, L23-P74, G77-T104.

**Table S3.** Monosaccharide and lignin content of untreated native wood (Eucalyptus and Spruce) and holocellulose samples after 6 hours delignification with $NaOCl_2$.

| Samples | Native eucalyptus (mg/g sample) | Eucalyptus holocellulose (6 h) (mg/g sample) | Native spruce (mg/g sample) | Spruce holocellulose (6 h) (mg/g sample) |
|---|---|---|---|---|
| Fucose | 0.4± 0.2 | 0.6± 0.3 | 0.3± 0.0 | 0.3± 0.1 |
| Rhamnose | 0.2± 0.3 | 3.7± 2.2 | 1.8± 2.0 | 1.6± 1.4 |
| Arabinose | 2.8± 0.1 | 4.1± 4.6 | 15.7± 0.4 | 15.0± 0.3 |
| Galactose | 13.2± 2.2 | 14.4± 10.5 | 22.0± 3.0 | 22.1± 0.8 |
| Glucose | 506.0± 77.5 | 575.8± 58.8 | 513.3± 61.2 | 593.9± 54.7 |
| Xylose | 129.9± 15.3 | 168.2± 23.3 | 71.1± 6.3 | 83.9± 7.2 |
| Mannose | 7.5± 0.9 | 9.5± 7.3 | 128.5± 12.2 | 130.3± 19.8 |
| 4O-MeGlcA | 1.4± 0.4 | 0.8± 0.8 | 8.2± 3.4 | 7.9± 3.9 |
| GalA | 0.6± 0.3 | 3.3± 3.0 | 2.6± 1.0 | 3.3± 1.9 |
| GlcA | 6.4± 0.8 | 4.1± 3.6 | 7.4± 1.9 | 4.0± 1.4 |
| Lignin | 336.6± 7.0 | 43.8± 16.5 | 302.1± 2.9 | 74.9± 3.5 |
| Total | 1005.1 | 828.5 | 1073.0 | 937.2 |

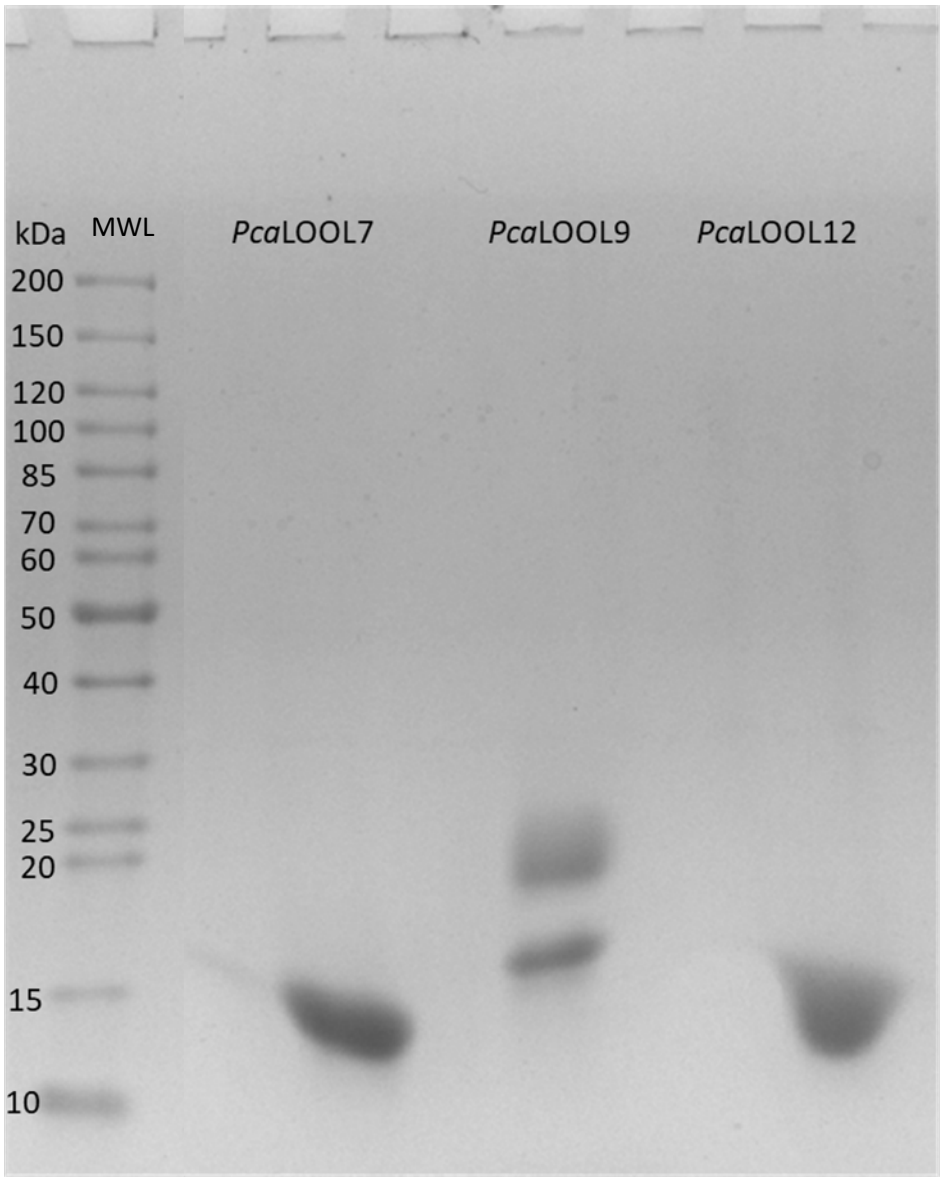


**Figure S1. Confirmation of *Pca*LOOL purity following recombinant production in *Pichia pastoris*.** Affinity-purified proteins were separated on a 4 - 20% Mini-PROTEAN® TGX™ gel and stained with Coomassie G-250 using the PageBlue™ Protein Staining Solution. MWL: PageRuler™ Prestained Protein Ladder

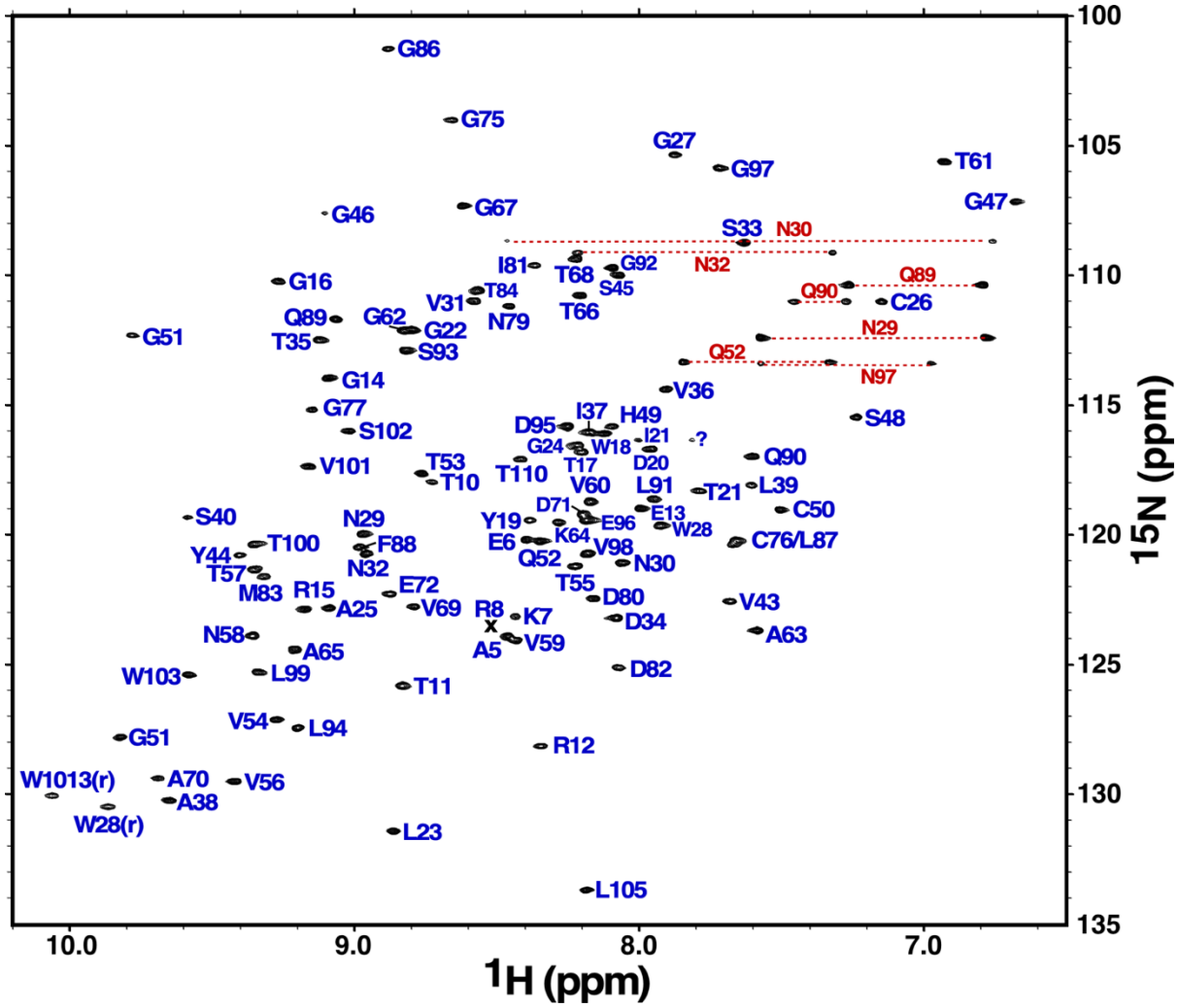


**Figure S2**. The $^{1}$H-$^{15}$N HSQC spectrum of $^{13}$C- and $^{15}$N-labelled *Pca*LOOL12 (~0.8 mM; recombinantly expressed in *Escherichia coli*) collected at a $^{1}$H resonance frequency of 600 MHz at 20 ºC in 100 mM NaCl, 20 mM Tris, pH 7 with the assigned amide cross peaks labelled (blue). The red dash lines connect side chain amide cross peak pairs, an "x" represents a cross peak below the contour threshold, and a "?" represents an unassigned resonance.

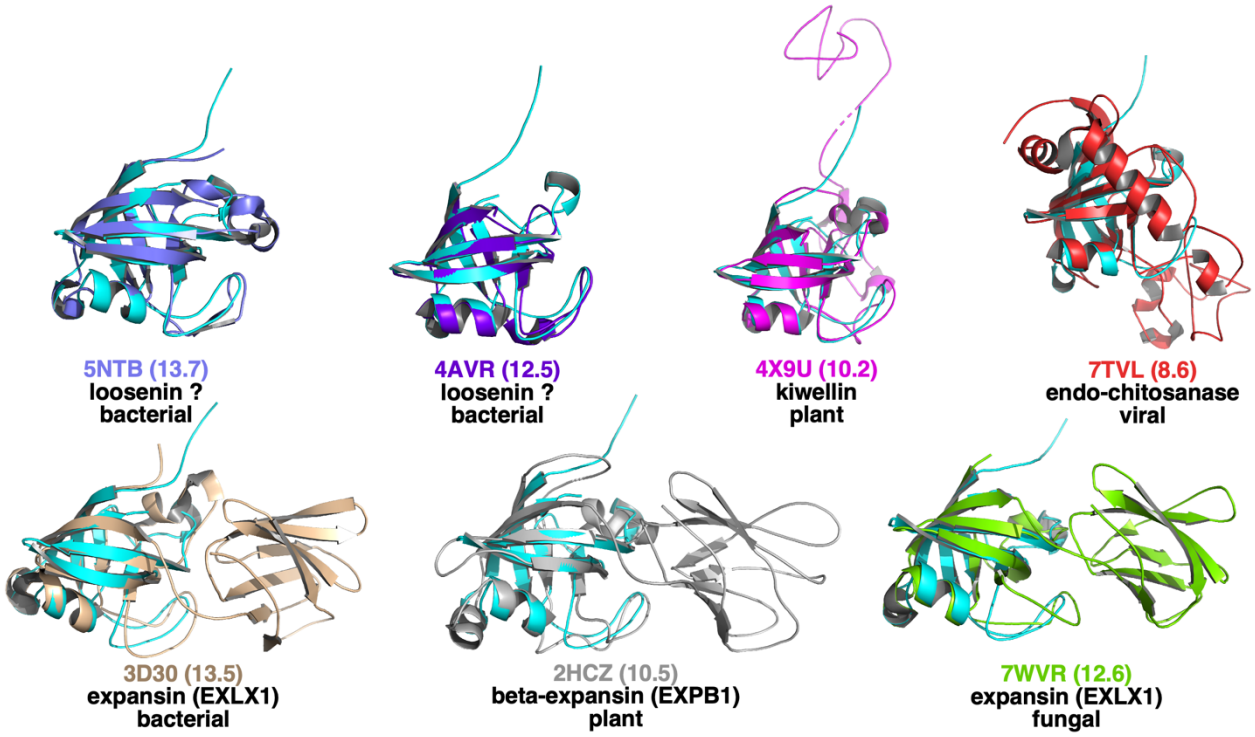


**Figure S3.** Alignment of the structure of *Pca*LOOL12 (9CE9) (https://www.rcsb.org/alignment) onto the top seven similar structures identified using the Dali server (http://ekhidna2.biocenter.helsinki.fi/dali/). . The structure of *Pca*LOOL12 is colored cyan in all the comparisons with the Z-score indicated in brackets next to the PDB ID. 5NTB - *Streptomyces mobaraensis* papain inhibitor; 4AVR - *Pseudomonas aeruginosa* Pa4485; 4X9U – *Actinidia chinensis* kiwellin; 7TVL – soil phage auxiliary metabolic gene product with endo-chitosanase activity; 3D30 – *Bacillus subtilis* expansin EXLX1; 2HCZ – *Zea mays* expansin EXPB1; 7WVR – *Talaromyces leycettanus* expansin EXLX1.

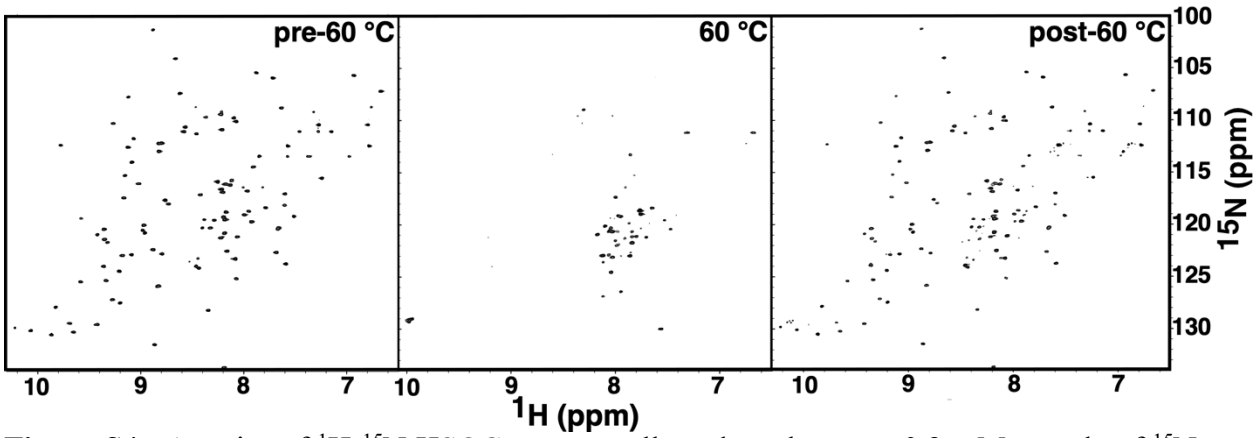


**Figure S4.** A series of $^{1}$H-$^{15}$N HSQC spectra collected on the same 0.2 mM sample of $^{15}$N-labelled *Pca*LOOL12 at a $^{1}$H resonance frequency of 600 MHz at 20 ºC pre-heating, 60 ºC, and 20 ºC post-heating.

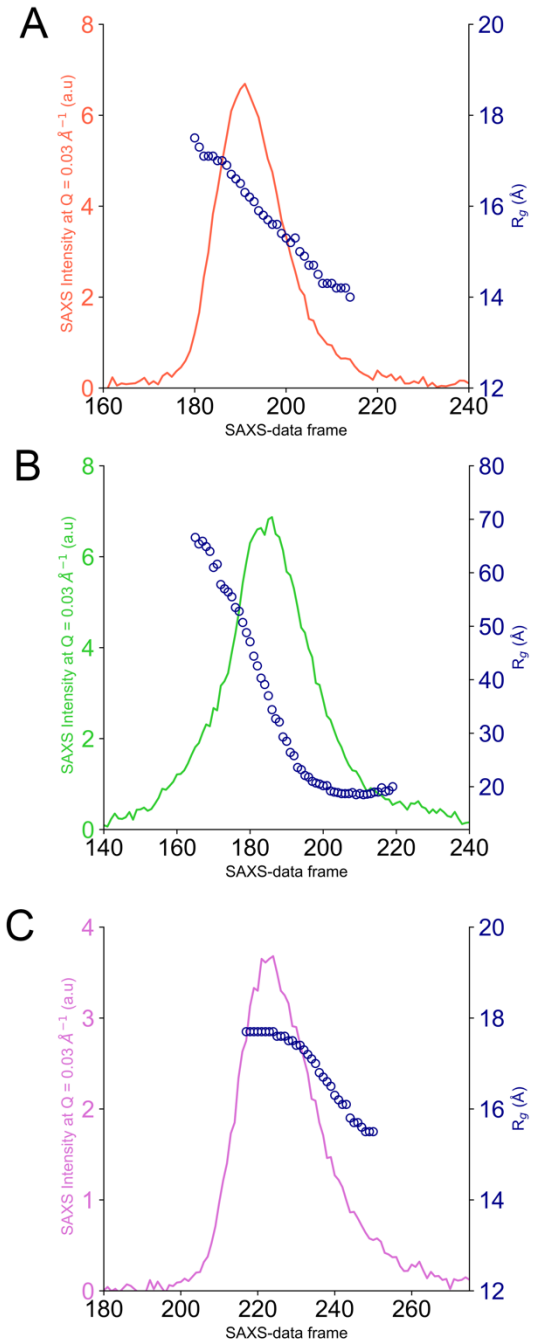


**Figure S5.** SEC-SAXS analysis of *Pca*LOOL proteins. Each plot show the SAXS intensity at Q = 0.03 Å$^{-1}$ for each SAXS frame vs. SAXS frame number (pink line). The radius of gyration (Rg) for each frame is represented by open blue circles. Panel A is *Pca*LOOL7, injected onto the SEC column at a concentration of 13 mg/mL. Panels B and C are for *Pca*LOOL9 and *Pca*LOOL12 injected at a concentration of 23 mg/mL and 13 mg/mL, respectively. The running buffer was 20 mM sodium acetate, 150 mM NaCl, pH 5.5.

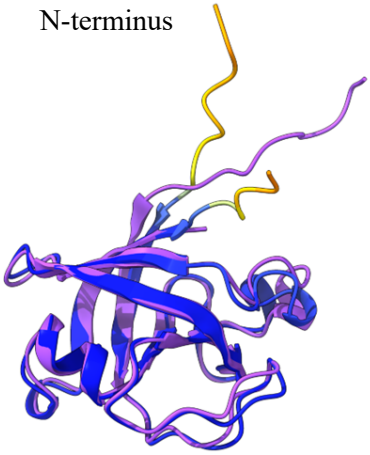


**Figure S6.** AlphaFold2 model of *Pca*LOOL12 (color code based on pLDDT values) overlaid with NMR solution structure (purple). The additional residues at the C-terminus of the AlphaFold2 model is associated with the $His_6$-tag present in the samples measured by SEC-SAXS.

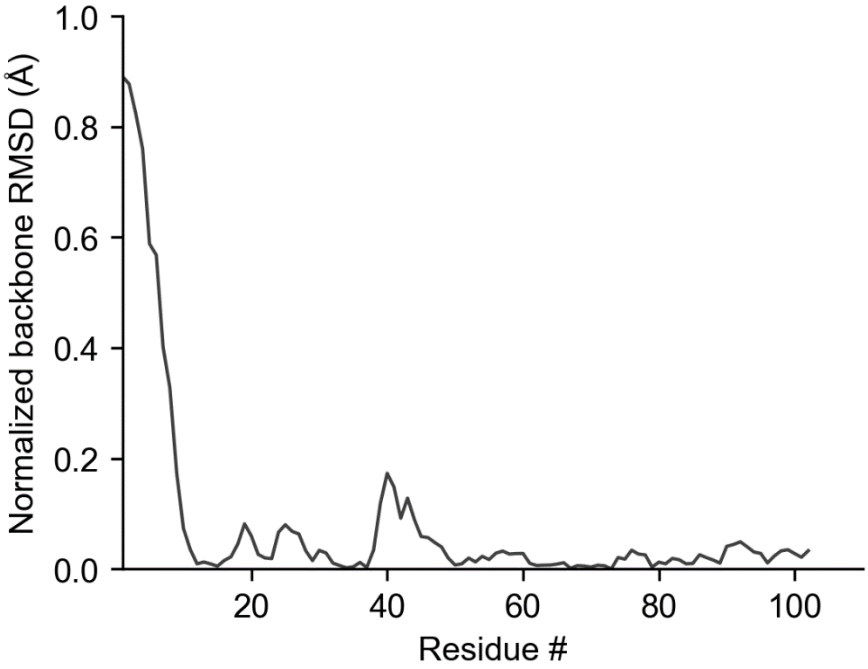


**Figure S7.** Normalized backbone RMSD between *Pca*LOOL12 model obtained using Normal Mode Analysis and NMR structure.

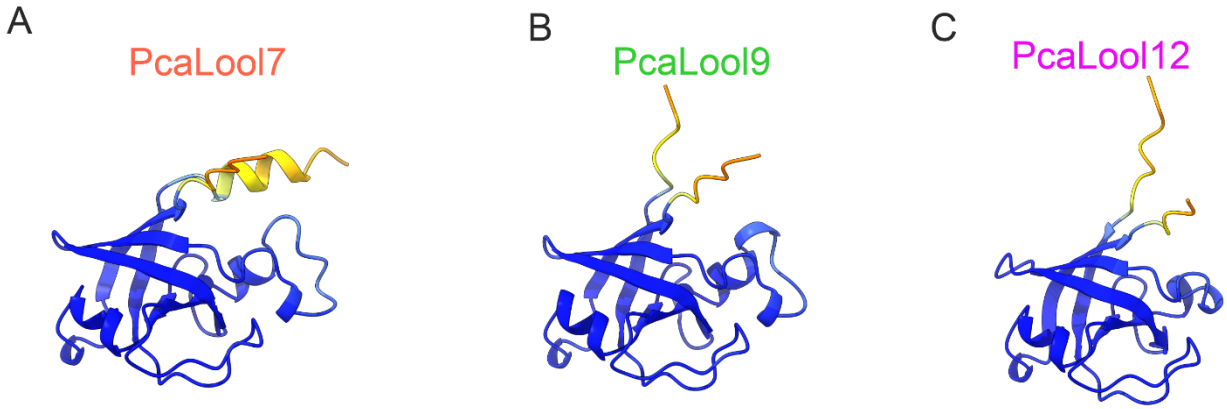


**Figure S8**. **Alphafold2 models of all *Pca*LOOL proteins**. AlphaFold2 models for *Pca*LOOL7 (A), *Pca*LOOL9 (B), and *Pca*LOOL12 (C). All AlphaFold2 models color-coded based on pLDDT scores (0-100): dark blue (>90), light blue (90-70), yellow (70-50), and red (<50).