Journal of the Indian Institute of Science
A Multidisciplinary Reviews Journal
ISSN: 0970-4140 Coden-JIISAD

REVIEWS

Structural and Functional Studies of Mammalian Peptidoglycan Recognition Protein, PGRP-S

*Pradeep Sharma, Divya Dube, Satya Prakash Yadav, Mau Sinha, Punit Kaur, Sujata Sharma and Tej P. Singh**

Abstract | Peptidoglycan recognition proteins belong to a broad family of innate immunity molecules. Mammals have four types of peptidoglycan recognition proteins designated as PGRP-S, PGRP-Iα, PGRP-Iβ and PGRP-L. PGRP-S is expressed in the granular polymorphonuclear leucocytes, PGRP-Iα is secreted from liver into blood and PGRP-Iβ, and PGRP-L are expressed in the skin, eyes, salivary glands, throat, tongue, esophagus, stomach and intestine. Peptidoglycan recognition proteins protect the host by carrying out early recognition of invading microorganisms. They contain a common domain known as peptidoglycan recognition domain whose lengths in various PGRPs vary from 165 to 175 residues. PGRP-S consists of a single peptidoglycan recognition domain while PGRP-Iα, PGRP-Iβ and PGRP-L have additional domains. Thus, PGRP-S represents the binding component of peptidoglycan recognition proteins and for understanding the mode of binding of these proteins, structural studies of PGRP-S are essential. So far, two structures of PGRP-S, one from human and another from *Camelus dromedarius are* available. The structure of human PGRP-S is found to be in monomeric state while the structure of camel PGRP-S consists of two distinct dimers in which dimeric interfaces involve opposite faces of the monomer. The observed monomeric and double dimeric structures of PGRP-S are well correlated to the differences in amino acid sequences of human and camel proteins. The binding sites in the dimers of camel PGRP-S are located at the contact sites of two molecules, whereas in human PGRP-S, it is supported by the single molecule. As a result, the binding clefts in camel protein are formed more efficiently as compared to the human protein. However, tertiary structures of both camel and human proteins are almost identical, with an average root mean squares shift of 1.2 Å for the backbone atoms. Since the ligand binding clefts in camel protein appear to have been evolved with better binding potencies than the human protein, the camel PGRP-S could be exploited for beneficial therapeutic applications against bacterial infections.

Keywords: *PGRP-S, CPGRP-S, HPGRP-S, Mycolic acid, Tuberculosis, LPS, PGN, LTA.*

Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India.
**tpsingh.aiims@gmail.com*

1 Introduction

Peptidoglycan recognition proteins (PGRPs) of the innate immune system provide the first line of defense against invading microbes.[1] These proteins recognize pathogen associated molecular patterns (PAMPS) by binding to them with high affinities, thus preventing the effects of infecting organisms.[2] The commonly occurring PAMPs include lipopolysaccharide (LPS) of Gram-negative bacteria,[3] lipoteichoic acid (LTA) from Gram-positive bacteria[4] and peptidoglycans (PGNs) of both Gram-positive and negative bacteria.[5] In the case of *Mycobacterium tuberculosis*, the prominent cell wall molecule is mycolic acid.[6] The mammalian PGRPs are broadly classified into three groups consisting of short PGRP (PGRP-S) with molecular mass ranging from 20 to 25 kDa, intermediate PGRPs (PGRP-Iα and PGRP-Iβ) with molecular mass of 40–45 kDa and long PGRP (PGRP-L) with molecular mass up to 90 kDa.[7] The expression of PGRP-S has been observed in the bone marrow.[8] It was found as a soluble protein in the granules of polymorphonuclear leucocytes (PMNs).[9] Its presence has also been confirmed in milk[10] as well as in intestinal M cells.[11] However, so far its significant presence has been reported only in the camel and porcine milks.[10] PGRP-L is expressed in intestinal follicle-associated epithelial cells.[12] In general, it is not expressed in healthy cells, its induction is caused by bacterial infections in keratinocytes.[13] PGRP-Iα and PGRP-Iβ are expressed in specialized epithelial cells.[14]

All the four PGRPs possess one common recognition domain known as PGRP domain. This domain consists of approximately 170 amino acid residues. PGRP domain is responsible for the recognition of bacterial cell wall molecules.[16] Since PGRP domain is common in all four mammalian PGRPs and is the main factor in the recognition of PAMPs, this review is concerned with the current status of structural studies on this protein. So far, the structure of PGRP-S is known only from two sources. The structure of camel protein (CPGRP-S)[17] was determined using the purified protein samples from the natural source of camel colostrum, while the structure of human protein (HPGRP-S) was obtained using the cloned protein.[18] The binding studies of CPGRP-S with various PAMPs[16,19–23] have shown high binding affinities and structure determinations of the complexes of CPGRP-S with different fragments of PAMPS[16,19–23] have revealed the site and mode of bindings. This review presents a brief comparison between monomeric structure of HPGRP-S and dimeric structure of CPGRP-S, and describes the therapeutic implications of the two oligomeric states of PGRP-S from two different sources.

2 Sequence Analysis

CPGRP-S and HPGRP-S share a sequence identity of 75% (Figure 1). The most important difference in the sequences of these two proteins concerns with an extra cysteine residue in HPGRP-S at position 8 (in HPGRP-S numbering scheme). The corresponding residue in CPGRP-S is Ala5 (in CPGRP-S numbering scheme). Other important differences in the sequences of CPGRP-S and HPGRP-S are indicated by giving the corresponding residues of HPGRP-S in parentheses which are listed here as Ser9 (Pro11), Arg23 (Ala26), Pro96 (His99), Ile101 (Met104), Asn126 (Gly129), Glu142 (Val145) and Pro151 (Arg153).

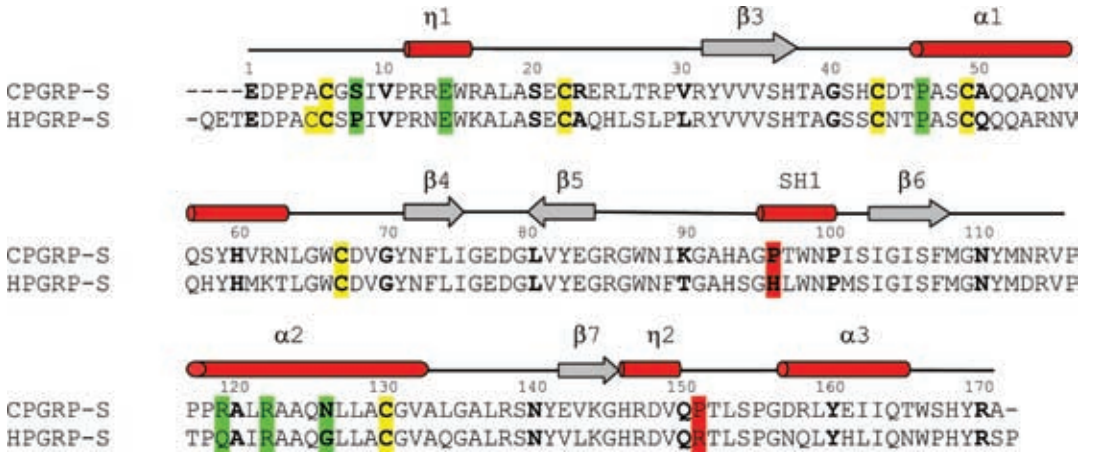

Figure 1: Sequence alignment for CPGRP-S and HPGRP-S showing important sequence differences. Cysteine residues are highlighted in yellow. In CPGRP-S, the important residues at the A-B interface are highlighted in green while those at C-D interface are highlighted in red. The secondary structure elements are indicated on the top.

The residues Pro96 and Pro151 are present on the one side while the remaining residues are part of the opposite face of the protein structure. It is well known that proline residues are reported to be favourable for dimerization.[24] The corresponding residues are histidine and arginine in HPGRP-S. On the other hand, the residues on the opposite face are better hydrogen bond formers in CPGRP-S than those in HPGRP-S.

3 Structural Studies

The structure determination of CPGRP-S indicated the presence of four crystallographically independent molecules in the asymmetric unit. The four molecules are designated as A, B, C and D (Figure 2). These molecules are associated in the form of two dimers named as dimer1 and dimer2. Dimer1 contains molecules A and B which are arranged in a face to face orientation with an approximate 2-fold rotation axis (Figure 3), while dimer2 is comprised molecules C and D with a back to back orientation in which molecules C and D are also related by an approximate two fold rotation (Figure 4). A calculation was carried using the PISA server[25] for ruling out other dimeric combinations. The results clearly indicated that the dimeric associations between molecule A, B, C, and D generated stable packing combinations of A-B (Figure 5) and C-D (Figure 6). However, because of face to face and back to back arrangements the overall structure of CPGRP-S seems to form polymeric arrangement (Figure 7). The overall structures of four crystallographically independent molecules, A, B, C and D were found to be identical with r.m.s shifts for the C^{α} atoms for any two molecules being less than 0.6 Å. Therefore, the subsequent description of molecular structure will be based on the structure of one molecule. The polypeptide chain folding of CPGRP-S contains a long non-repetitive flexible but well defined N-terminal segment (residues, 1–31). It includes a central β-sheet consisting of five β-strands, out of which four β-strands, β3 (residues, 31–38), β4 (residues, 71–76), β6 (residues, 103–108) and β7 (residues, 142–146) are parallel, while the fifth β5 strand (residues, 80–85) is antiparallel. The structure also contains three α-helices, α1 (residues, 46–64), α2 (residues, 118–132) and α3 (residues 157–164). In addition to these, it has two short one turn 3_{10}-helices, η1 (residues, 12–15) and η2 (residues, 146–149). The amino acid sequence of CPGRP-S contains six cysteine residues at positions 6, 22, 43, 49, 67 and 130. They form three

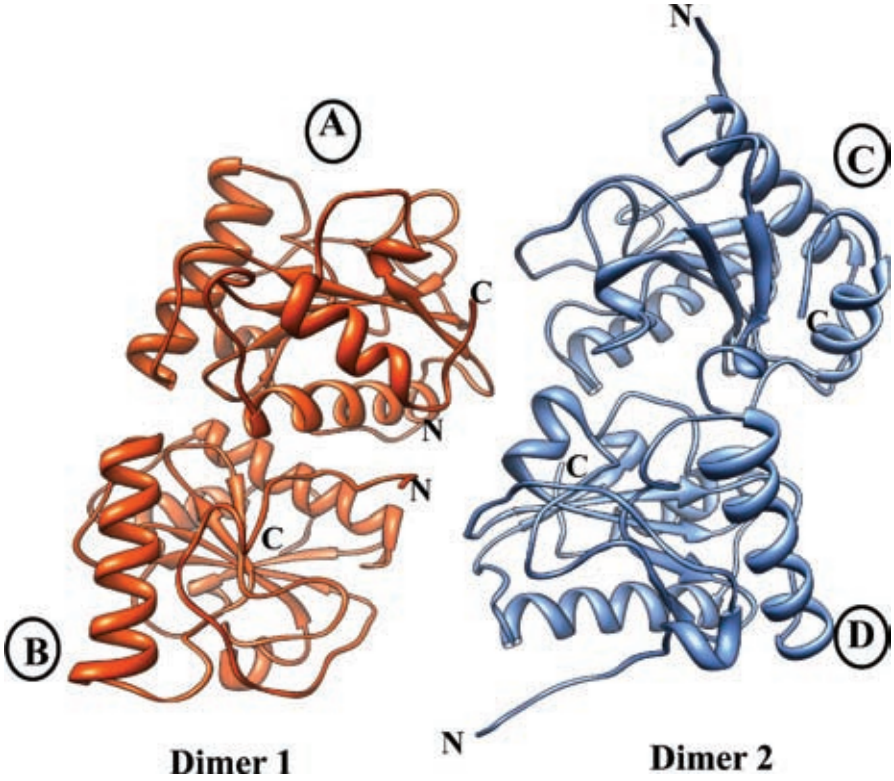

Figure 2: The ribbon representation of the structure of CPGRP-S. Molecules A and B form dimer 1 which also includes residues of N-terminal regions at the interface while molecules C and D form dimer 2 with Pro96 and Pro151 at the C-D interface.

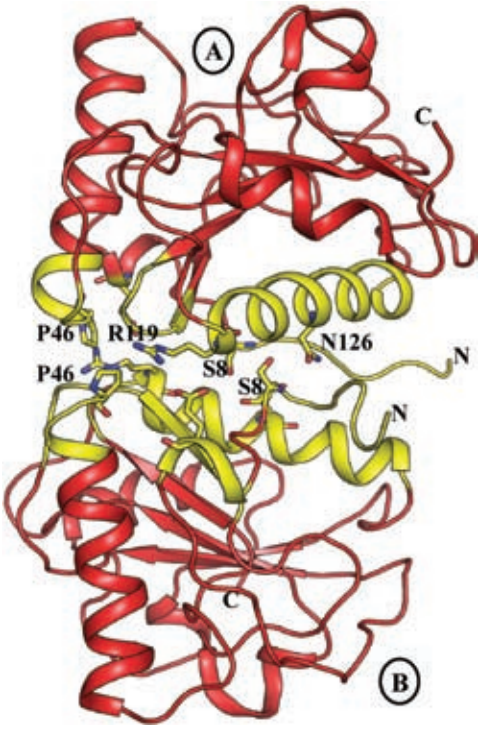

Figure 3: Dimer 1 has a buried surface area of 798 Å.^{2} The important residues at the interface are also indicated.

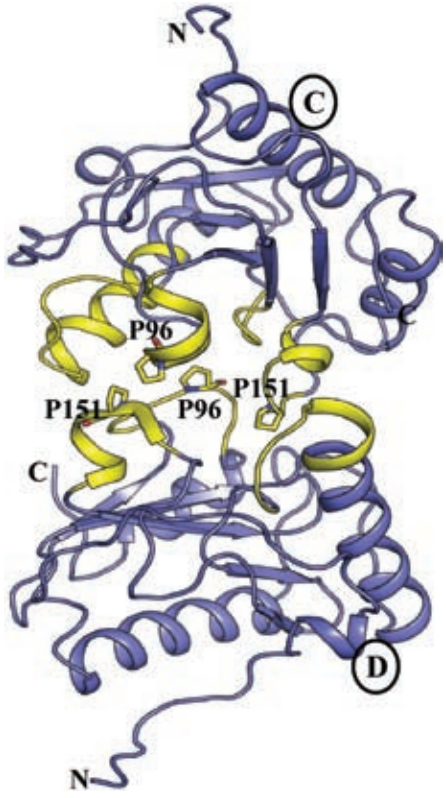

Figure 4: Dimer 2 is shown with molecules C and D with a buried surface area of 702 Å.^{2} The important residues in dimerization are indicated.

disulfide linkages involving cysteine residues at positions 6 and 130, 22 and 67 and 43 and 49. The former two disulfide bridges have Cys6 and Cys22 both from N-terminal segment (residues, 1–31), thereby stabilizing the N-terminal segment by tethering it to the longest helix α2 and a relatively short and tight loop α1–β4. The third disulfide bridge (Cys43-Cys49) connects a long loop β3–α1 to helix α1. All of these disulfide bonds have also been observed in the HPGRP-S. It may be noted that HPGRP-S contains seven cysteine residues with one unpaired cysteine side chain at position 8 (Figure 1). The crystal structure determination showed that HPGRP-S formed a monomeric structure. Overall foldings of both CPGRP-S and HPGRP-S are very similar to an average r.m.s shift of 1.2 Å for C^{α} atoms when the backbones of HPGRP-S is superimposed on the backbones of A, B, C and D molecules (Figure 8). It may also be noted here that residues from 1–8 were not observed in the structure of HPGRP-S. Hence, the structure of HPGRP-S contained the polypeptide chain with residues from 9 to 175. In contrast, in the structure of CPGRP-S, the N-terminus was well defined. In CPGRP-S, the N-terminus begins with Glu1 as obtained by determining the N-terminal sequencing of the first 10 residues using N-terminal protein sequencer. The first residue, Glu1 in CPGRP-S corresponds to the position 4 in HPGRP-S and all the N-terminal residues from Glu1 onwards were clearly defined in the electron density map. The N-terminal segment in the structure of CPGRP-S is an ordered motif which is stabilized by two disulfide bridges, Cys6-Cys130 and Cys22-Cys67; it is further rigidified by the presence of several proline residues in the sequence of N-terminal region. It is of significant interest to note that the N-terminal segment of CPGRP-S is involved in several intramolecular contacts, particularly with helix α2 with which it forms a relatively independent structure, indicating that this domain-like structure may be a favourable site for intermolecular binding as well.

The interface in dimer2 of CPGRP-S contains two proline residues, Pro96 and Pro151 in each monomer which promote a dimerization by bringing molecules C and D in contact through the interface containing these proline residues (Figure 5), while the corresponding residues are His99 and Arg154 in HPGRP-S, which clearly obstruct the dimer formation due to steric constraints (Figure 9). Similarly in dimer1 the A-B interface in CPGRP-S is induced by the formation of two hydrogen bonds involving Ser8 N and O^{γ} with Asn126 $O^{\delta}1$ and Arg122 NH1 with

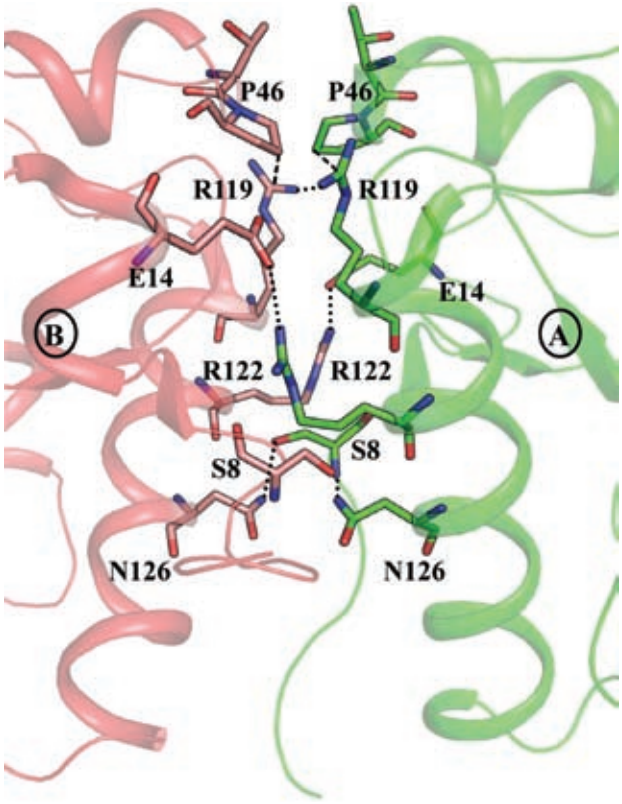

Figure 5: Showing interactions at the A-B interface. The critical interactions are formed between Ser8 and Asn126 and Arg122 and Glu14.

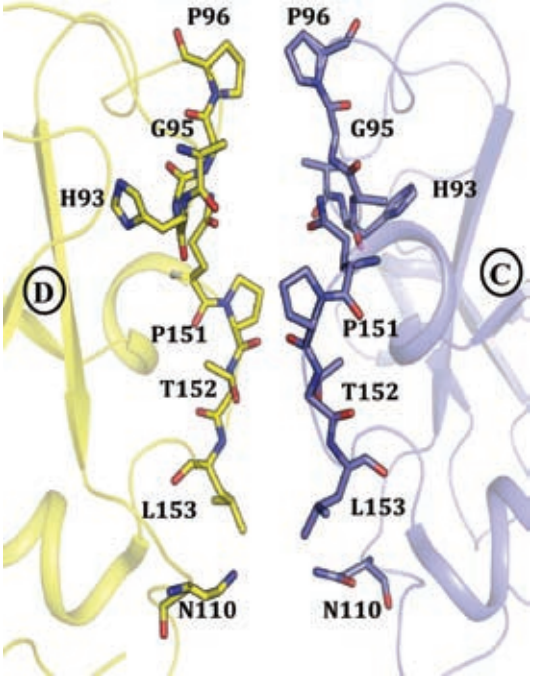

Figure 6: Showing C-D interface with Pro96 and Pro151 as critical residues for dimerization.

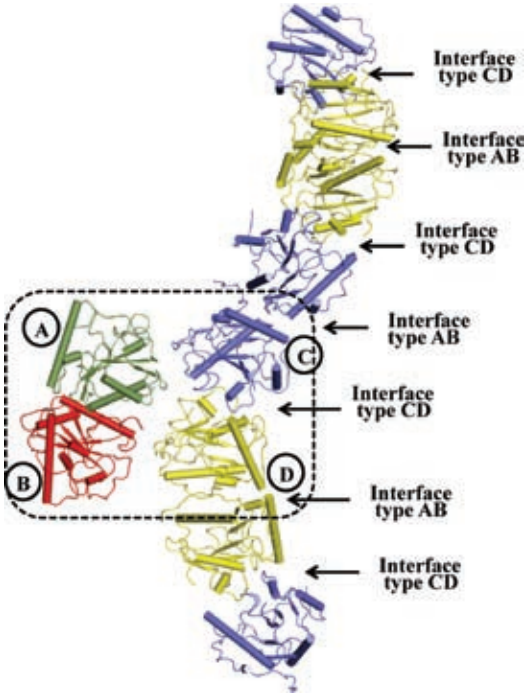

Figure 7: The A-B and C-D interfaces are on the opposite sides of CPGRP-S molecule indicating that the protein forms a polymeric arrangement with alternating A-B and C-D interfaces.

Glu14 (Figure 5). The corresponding residues in HPGRP-S are Pro11 and Gly129 which are incapable of forming hydrogen bonds, thus a dimer is not formed in HPGRP-S (Figure 10).

As shown by the structures of complexes of CPGRP-S, the binding sites are located at the

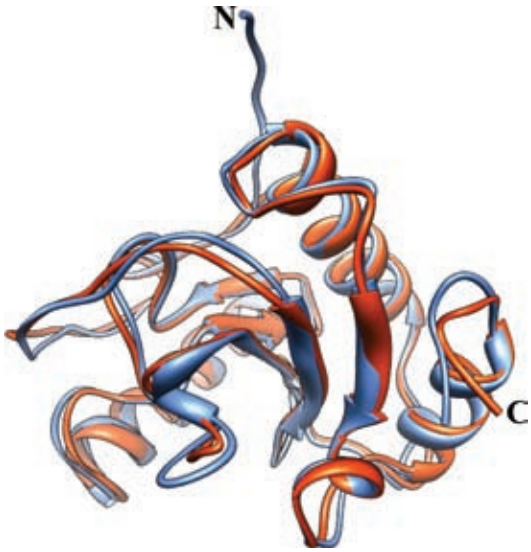

Figure 8: The superimposition of backbones of molecules of HPGRP-S and CPGRP-S shows that both proteins have similar folding.

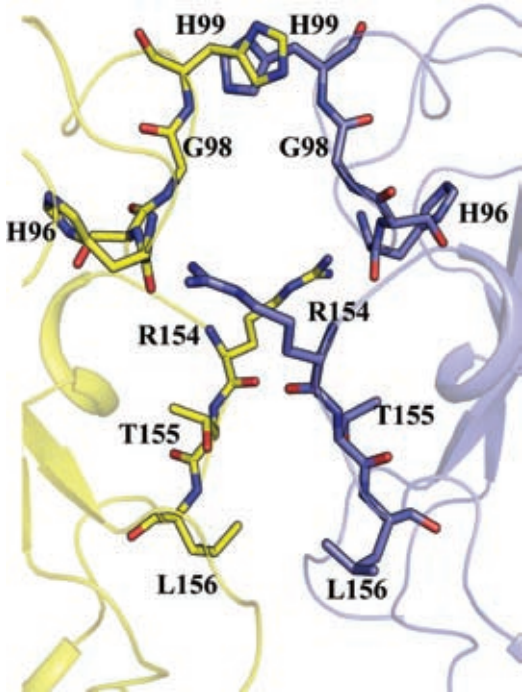

Figure 9: A stable interface involving molecules C and D can not be formed in HPGRP-S due to steric constraints caused by His99 and Arg154. The corresponding residues in CPGRP-S are Pro96 and Pro151 in CPGRP-S.

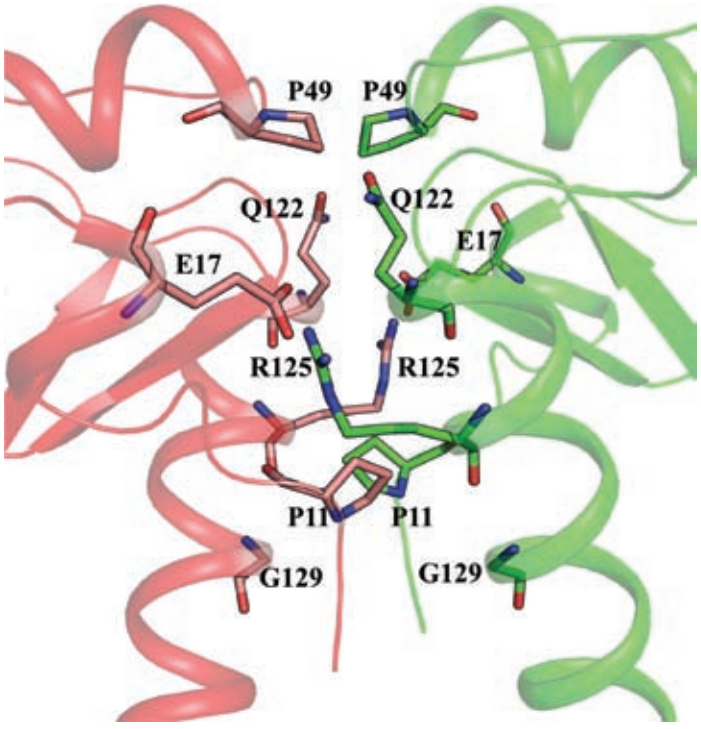

Figure 10: The A-B interface similar to that in CPGRP-S is not feasible in HPGRP-S because of the absence of hydrogen bonds between Pro11 and Gly129 as observed in CPGRP-S between the corresponding residues Ser8 and Asn126. Also Pro11 residues from two molecules produce incompatible constraints.

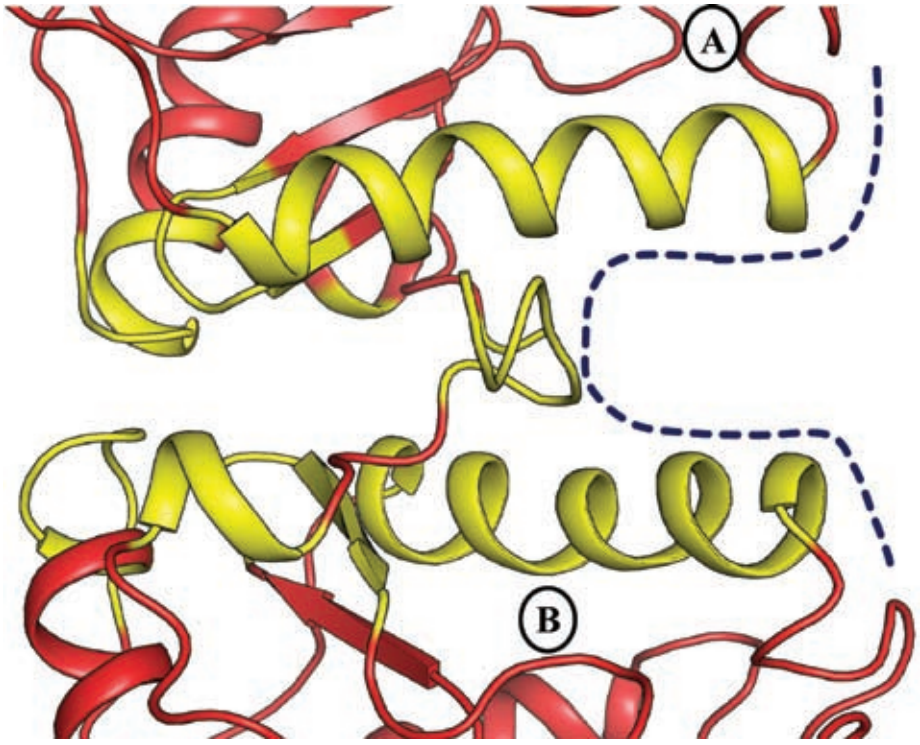

Figure 11: A deep cleft at the A-B contact.

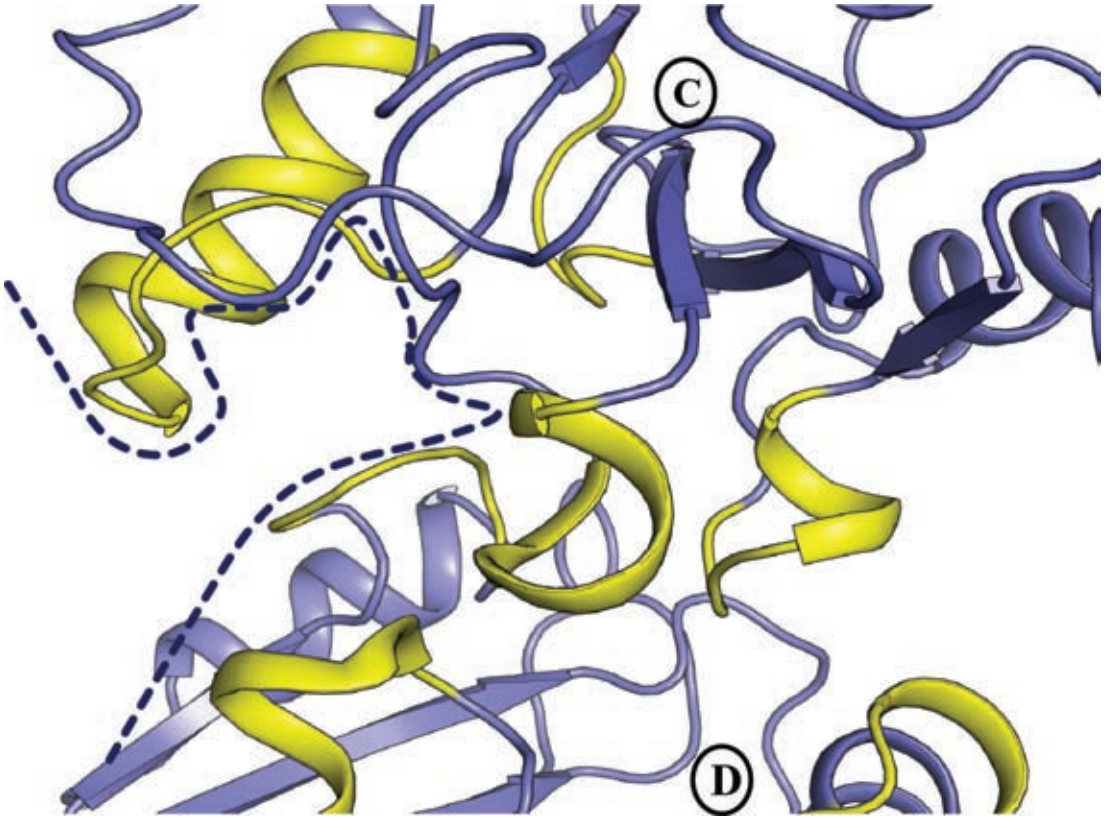

Figure 12: A deep cleft at the C-D contact.

contact points of two molecules in the dimeric structure of CPGRP-S resulting in the formation of deep clefts in the independent dimer 1 (Figure 11) and dimer2 (Figure 12) of CPGRP-S. The corresponding binding sites in the monomer of HPGRP-S are located on the shallow surface. The examination of binding sites in CPGRP-S and HPGRP-S shows that the clefts in CPGRP-S are capable of binding to PAMPs by involving larger interfaces with several intermolecular interactions resulting in higher affinities.[16] The binding contacts in the monomeric HPGRP-S through its shallow surface are expected to produce a limited number of interactions with PAMPs leading to low binding affinities.[26] This makes CPGRP-S a far more potent binding protein than HPGRP-S. Because of stronger and more versatile binding characteristics of CPGRP-S than those of HPGRP-S, CPGRP-S may be used as a therapeutic agent against bacterial infections.

Acknowledgements

The authors thank Department of Science and Technology (DST), Government of India for financial support. TPS thanks the Department of Biotechnology (DBT), New Delhi for the award of Distinguished Biotechnology Research Professorship to him. DD thanks the Council of Scientific and Industrial Research (CSIR), New Delhi and SPY. Thanks also to University Grants Commission (UGC) for the award of Fellowship.

Received 27 November 2013.

References

1. J. A. Hoffmann. The immune response of Drosophila. *Nature*, Vol. 426, pp. 33–38, 2003.
2. C. C. Tydell, J. Yuan, P. Tran, and M. E. Selsted. Bovine peptidoglycan recognition protein-S: Antimicrobial activity, localization, secretion, and binding properties. *J Immunol*, Vol. 176, pp. 1154–1162, 2006.
3. S. Lebeer, J. Vanderleyden, S. C. De Keersmaecker. Host interactions of probiotic bacterial surface molecules: Comparison with commensals and pathogens. *Nat Rev Microbiol*, Vol. 8, pp. 171–184, 2010.
4. I. Ginsburg. Role of lipoteichoic acid in infection and inflammation. *Lancet Infect Dis*, Vol. 2, pp. 171–179, 2002.
5. M. Xu, Z. Wang, R. M. Locksley. Innate immune responses in peptidoglycan recognition protein L-deficient mice. *Mol Cell Biol*, Vol. 24, pp. 7949–7957, 2004.
6. P. J. Brennan. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis,* Vol. 83, pp. 91–97, 2003.
7. X. Lu, M. Wang, J. Qi, H. Wang, X. Li, D. Gupta, R. Dziarski. Peptidoglycan recognition proteins are a new class of human bactericidal proteins. *J Biol Chem*, Vol. 28, pp. 5895–5907, 2006.
8. R. Dziarski, K. A. Platt, E. Gelius, H. Steiner, D. Gupta. Defect in neutrophil killing and increased susceptibility to infection with nonpathogenic gram-positive bacteria in peptidoglycan recognition protein-S (PGRP-S)-deficient mice. *Blood*, Vol. 102, pp. 689–697, 2003.
9. D. Kang, G. Liu, A. Lundstrom, E. Gelius, H. Steiner. A peptidoglycan recognition protein in innate immunity conserved from insects to humans. *Proc Natl Acad Sci U S A*, Vol. 95, pp. 10078–10082, 1998.
10. S. R. Kappeler, C. Heuberger, Z. Farah, Z. Puhan. Expression of the peptidoglycan recognition protein, PGRP, in the lactating mammary gland. *J Dairy Sci*, Vol. 87, pp. 2660–2668, 2004.
11. J. Wang, V. Gusti, A. Saraswati, D. D. Lo. Convergent and divergent development among M cell lineages in mouse mucosal epithelium. *J Immunol*, Vol. 187, pp. 5277–5285, 2011.
12. D. Lo et al. Peptidoglycan recognition protein expression in mouse Peyer's Patch follicle associated epithelium suggests functional specialization. *Cell Immunol*, Vol. 224, pp. 8–16, 2003.
13. H. Wang, D. Gupta, X. Li, R. Dziarski. Peptidoglycan recognition protein 2 (N-acetylmuramoyl-L-Ala amidase) is induced in keratinocytes by bacteria through the p38 kinase pathway. *Infect Immun,* Vol. 73, pp. 7216–7225, 2005.
14. C. Liu, Z. Xu, D. Gupta, R. Dziarski. Peptidoglycan recognition proteins: A novel family of four human innate immunity pattern recognition molecules. *J Biol Chem*, Vol. 276, pp. 34686–34694, 2001.
15. R. Dziarski R and D. Gupta. The peptidoglycan recognition proteins (PGRPs). *Genome Biol.* Vol. 7, pp. 232, 2006.
16. P. Sharma et al. Structural basis of recognition of pathogen-associated molecular patterns and inhibition of proinflammatory cytokines by camel peptidoglycan recognition protein. *J Biol Chem.* Vol. 286, pp. 16208–16217, 2011.
17. P. Sharma, N. Singh, M. Sinha et al. Crystal structure of the peptidoglycan recognition protein at 1.8 Å resolution reveals dual strategy to combat infection through two independent functional homodimers. *J Mol Biol*, Vol. 378, pp. 923–932, 2008.
18. R. Guan, Q. Wang, E. J. Sundberg, R. A. Mariuzza. Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 Å resolution. *J Mol Biol*, Vol. 347, pp. 683–691, 2005.
19. P. Sharma, D. Dube, M. Sinha et al. Multiligand Specificity of Pathogen-associated Molecular Pattern-binding Site in Peptidoglycan Recognition Protein. *J Biol Chem*, Vol. 286, pp. 31723–31730, 2011.
20. P. Sharma, D. Dube, M. Sinha et al. Structural basis of heparin binding to camel peptidoglycan recognition protein-S. *Int J Biochem Mol Biol*, Vol. 3, pp. 86–94, 2012.

21. P. Sharma, S. Yamini, D. Dube et al. Structural studies on molecular interactions between camel peptidoglycan recognition protein, CPGRP-S, and peptidoglycan moieties N-acetylglucosamine and N-acetylmuramic acid. *J Biol Chem*, Vol. 287, pp. 22153–22164, 2012.
22. P. Sharma et al. Structural basis of the binding of fatty acids to peptidoglycan recognition protein, PGRP-S through second binding site. *Arch Biochem Biophys*, Vol. 529, pp. 1–10, 2013.
23. P. Sharma, et al. Structural insights into the dual strategy of recognition by peptidoglycan recognition protein, PGRP-S: structure of the ternary complex of PGRP-S with lipopolysaccharide and stearic acid. *PLoS One*, Vol. 8, pp. e53756, 2013.
24. R. M. Kini, H. J. Evans. A hypothetical structural role for proline residues in the flanking segments of protein-protein interaction sites. *Biochem Biophys Res Commun*, Vol. 212, pp. 1115–1124, 1995.
25. E. Krissinel, K. Henrick. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, Vol. 372, pp. 774–797, 2007.
26. R. Guan, A. Roychowdhury, B. Ember, S. Kumar, G. J. Boons, R. A. Mariuzza. Structural basis for peptidoglycan binding by peptidoglycan recognition proteins. *Proc Natl Acad Sci U S A*, Vol. 101, pp. 17168–17173, 2004.

Pradeep Sharma is an INSPIRE faculty in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. He obtained his Ph.D. from All India Institute of Medical Sciences in 2012. His current research interests include structural and functional studies of peptidoglycan recognition protein from various species. His professional and research contributions have been recognized by awards such as Ludo Frevel Crystallography Scholarship, DST-INSPRE Faculty award, INSA Young Scientist Award.

Divya Dube is a post-doctoral scientist in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. She obtained her Ph.D. degree from Central Drug Research Institute, Lucknow in 2010. Her current research interests include structural and functional studies on peptidyl tRNA from *Acinetobactor baumanni* and ligand design against various other targets. Her professional and research contributions have been recognized by Fast Track projects from DST and Senior Research Associateship from CSIR.

Satya Prakash Yadav is a research scholar in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. He is registered for Ph.D. He obtained his M.Tech from Indian Institute of Technology, Kharagpur in 2010. He is working on structural and functional studies of porcine peptidoglycan recognition protein.

Mau Sinha is a senior scientist from DST. She is working in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. She obtained her Ph.D. degree from National Institute of Mental Health and Neurosciences, Bangalore in 2004. Her current research interests include structural and functional studies of bacterial and mammalian proteins. Her professional research contributions have been recognized by awards such as fast track projects from DST and Senior Research Associateship from CSIR.

Punit Kaur is a Professor in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. She obtained her Ph.D. from Department of Biophysics, AIIMS, New Delhi in 1991. Her research interests include X-ray crystallography of peptides and proteins. She has also been persuing molecular modeling and other computational calculations. Her professional research contributions have been recognized by Dr. Talekar award.

Sujata Sharma is an Additional Professor in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. She obtained her Ph.D. from Department of Biophysics, AIIMS, New Delhi in 1998. Her research interests include studies of proteins and drug discovery. Her professional research contributions have been recognized by awards such as National Bioscience Award (DBT), National Woman Bioscientist Award (DBT), Women Scientist of the Year 2007 (Biotech Research Society of India) and Dr. Talekar award of AIIMS.

T. P. Singh is currently working as DBT-Distinguished Biotechnology Research Professor in the Department of Biophysics at All India Institute of Medical Sciences, New Delhi. He obtained his Ph.D. degree from Indian Institute of Science, Bangalore in 1976. His research interests include protein structure determination, peptide design and drug discovery. His professional and research contributions have been recognized by some awards and Academy fellowships.