

Identification of Amino Acid Conservation in The Curli Accessory Protein CsgF

Karen Guerrero¹, Emma Smith¹, Shruti Sunder Rajkumar¹, Zachary Cairo¹, Jonathan Adame¹, Renad Rawas¹, Ranim Rawas¹, Sajith Jayasinghe^{1S}

¹Chemistry and Biochemistry, California State University, San Marcos, San Marcos, California, United States

^STo whom correspondence should be addressed: sjayasin@csusm.edu

Abstract

The Curli-Specific gene product F (CsgF) plays an important role in the assembly of gram-negative bacterial cell surface filaments known as Curli. In order to evaluate amino acid conservation in the context of the solution and CsgG bound structures of CsgF we carried out a multiple sequence alignment of CsgF sequences from 35 gram-negative bacteria and correlated amino acid conservation to structural and functional importance. We identified conserved Pro and Gly residues within the N-terminal region of CsgF that may be required to adopt the loop confirmation observed in this region. Several conserved hydrophobic residues are found on the 3rd and 4th β -strands of the C-terminal β -sheet and extending to the C-terminal end, that may play a role in the reported observation that the C-terminus is needed for Curli formation. The importance of several conserved residues that were identified in this study has not yet been reported and investigating their impact on the structure and function of CsgF may add to our understanding of Curli assembly.

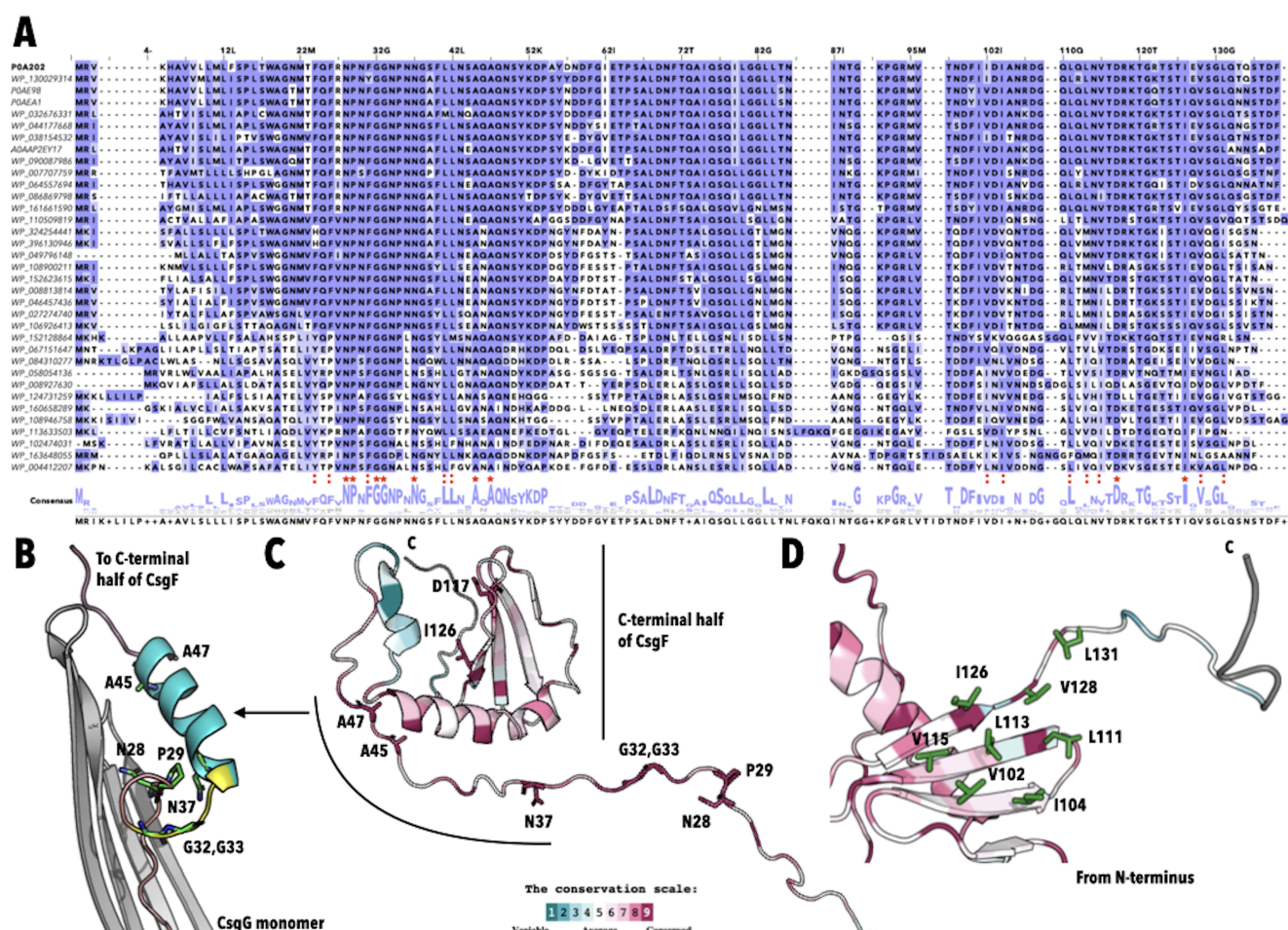


Figure 1. Amino acid conservation in CsgF:

A Multiple Sequence Analysis (MSA) of CsgF sequences from 35 bacterial species (see extended data for details). Based on the MSA we identified nine residues, N28, P29, G32, G33, N37, A45, A47, D117, and I126, that are conserved in all 35 sequences (red star "*"). Red ":" indicate conserved residues that correlate with published descriptions of amino acids or regions of CsgF that are implicated as being important for its structure or function. **B** Structure of the N-terminal residues of CsgF in association with the Curli secretion channel CsgG (PDBID 6L7C and 6LQH). Only one monomer of

CsgF and CsgG are shown. The N-terminal residues (20-31) of CsgF interact with the inner surface of the CsgG channel in an extended conformation ending in a loop (salmon). The loop extends between residues 29-36 of CsgF (yellow) and is followed by an α -helix that spans residues 37-46 (teal). Seven of the nine amino acids found in all 35 sequences, N28, P29, G32, G3, N37, A45, and A47, are found in this region (shown in stick representation). **C** Solution structure of CsgF (PDBID 5MIU) colored by evolutionary conservation as calculated by the ConSurf server where purple indicated high conservation and teal indicated low (see scale). All nine amino acids conserved in all 35 sequences are highlighted using stick representation. The region between positions 37-47 are in an extended conformation in the solution structure but adopt a helical structure in the CsgG associated form (Yan et al., 2020). **D** C-terminal β -sheet of the solution structure of CsgF colored by evolutionary conservation. The conserved positions containing amino acids with hydrophobic side chains are highlighted (green) and labeled.

Description

The outer surface of gram-negative bacteria contain filaments, known as Curli, that play an important role in cell-cell interactions and host-cell colonization (Collinson et al., 1997, Sukupolvi et al., 1997, Austin et al., 1998, Römling et al., 1998). Curli are composed of two Curli specific gene (Csg) proteins, CsgA and CsgB, with CsgA being the major component of the Curli filaments (Chapman et al., 2002, Hammar et al., 1996, Bian & Normark, 1997). The assembly of CsgA/CsgB into Curli involves four other proteins: CsgC, CsgE, CsgF, and CsgG. CsgC and CsgE are periplasmic proteins that are thought to prevent the intracellular aggregation of CsgA/CsgB and target these proteins to the outer membrane channel CsgG (Gibson et al., 2007, Taylor et al., 2011, Evans et al., 2015). CsgA and CsgB traverse through CsgG and are secreted to the bacterial outer surface where they associate to form Curli (Cao et al., 2014, Goyal et al., 2014). In contrast, CsgF is located on the extracellular surface of bacteria and is required for Curli assembly (Robinson et al., 2006, Nenninger et al., 2009, Chapman et al., 2002). Structural studies have shown that the N-terminus of CsgF interacts with the CsgG pore, and in-vivo complementation assays have shown that the C-terminal half of the protein is required for proper Curli assembly (Zhang et al., 2020, Yan et al., 2020, Swasthi et al., 2023). A majority, if not all, of the structure/function studies reported so far have concentrated on *E. Coli* CsgF. We carried out a multiple sequence alignment (MSA) of CsgF sequences from 35 gram-negative bacteria (see extended data) to determine amino acid conservations and evaluated these in the context of the three-dimensional structures of CsgF to correlate amino acid conservation to structural and functional importance.

The MSA shows significant conservation throughout the sequence (Figure 1A). There are nine amino acids that are conserved in all 35 sequences (numbered according to the position in the *Salmonella* sequence): N28, P29, G32, G33, N37, A45, A47, D117, and I126 (Figure 1A, red star "*"). We evaluated if these residues contributed to the interaction of CsgF with the CsgG channel using the cryo-EM structures of the CsgF-CsgG complexes available in the protein data bank (PDB ID 6L7C and 6LQH) (Zhang et al., 2020) (Yan et al., 2020). The N-terminal residues (20-31) of CsgF are observed to interact with the inner surface of the CsgG channel in an extended conformation ending in a loop (Figure 1B, salmon). The loop which extends between residues 29-36 of CsgF (Figure 1B, yellow) is followed by an α -helix that spans residues 37-46 (Figure 1B, teal). Seven of the nine amino acids found in all 35 sequences, N28, P29, G32, G3, N37, A45, and A47, are found in the region that contains these three structural elements (Figure 1B). N28 (100% conservation) and N30, which is conserved in 74% of the sequences included in this study, forms H-bonds with E216 of CsgG suggesting that they may play a role in anchoring CsgF to CsgG. Yan et. al report a reduced interaction of CsgF with CsgG for the N28A variant but not for the N30A variant (Yan et al., 2020). The loop region contains P29, G32, and G33 which are found in all 35 sequences. These Pro and Gly residues, both of which are found in structural motifs requiring tight turns or flexibility, may play an important role in this region's ability to adopt the loop observed in the CsgG bound form of CsgF. N37 which is conserved in all 35 sequences forms a H-bond with the backbone carbonyl of R27 and may also interact with the backbone carbonyl of P29 from an adjacent CsgF (which is 3.5 Å away), and therefore may also play a role the formation of the loop conformation. As far as we know the impact of variants of these four residues have not been reported but would help in determining the influence of the loop region on the function of CsgG bound CsgF. A45 and A47, which are also found in all 35 sequences, are found on the C-terminal end of the α -helix that associates with the upper part of the CsgG pore. Again, no variants of these two alanine residues have been investigated and therefore, it is not known if these residues play a role in the association of CsgF with CsgG or its function.

Residues that have been identified as helping anchor CsgF to CsgG via van der Waals or hydrophobic interactions F24 (60% F, 34% Y, F26 (66% F, 34% Y, F31 (97% F, 3% Y), L41 (97% L) and L42 (91% L 6% F), are also conserved (Figure 1A, red ":") (Zhang et al., 2020, Yan et al., 2020). Yan et. al report that the L41D and L42D variants of CsgF reduced its interaction with CsgG and that the F24D, F26D, and F31D, completely abolished the ability of CsgF to interact with CsgG (Yan et al., 2020). These observations correlate with the conservation observed for these five amino acids.

The cryo-EM structure of the CsgF-CsgG complex does not contain the C-terminal residues of CsgF. In order to visualize the amino acid conservation in the context of structure we used the ConSurf server to map evolutionary conservation onto the three-dimensional solution structure of CsgF (PDBID 5M1U, Figure 1C) (Yariv et al., 2023)(Schubeis et al., 2018) (Landau et al., 2005). We observe significant conservation throughout the solution structure of CsgF (Figure 1C, purple)

similar to what was observed via the MSA. In addition to conservation in the N-terminal half of the protein there is considerable conservation of residues in the α -helix and β -sheet found in the C-terminal half of the solution structure of CsgF. It has been observed that the β -sheet contains a hydrophobic surface formed from positions 102, 104, 113, and 115 on the 3rd and 4th β -strands, and it was speculated that this hydrophobic surface may facilitate protein interactions (Schubeis et al., 2018). The MSA and the evolutionary conservation predicted by ConSurf indicate that the hydrophobic surface is conserved and extends towards the unstructured C-terminal end to include position 126, 128 and 131 (Figure 1D). I126 is found in all 35 sequences investigated here and positions 102 (69% V, 23% I, 9% L), 104 (66% I, 31% V, 3% L), 111 97% L, 3% F), 113 (40% L, 21% V, 9% I), 115 (54% V, 46% I), 128 (94% V, 3% I, 3% L), 131 (77% L, 11% I, 3% V) are occupied by amino acids with hydrophobic side chains (Figure 1A red :). The conservation of these hydrophobic residues correlates well with the observation that this region is important for Curli assembly. In vivo complementation assays have shown that the C-terminal half of CsgF is needed for Curli assembly and that this region binds the nucleator protein CsgB (Swasthi et al., 2023, Yan et al., 2020, Schubeis et al., 2018). It appears that the 11 residues (between 128-138) at the C-terminal end of CsgF is required for interaction with CsgB and the proper assembly of Curli (Swasthi et al., 2023). D117 is conserved in all 35 sequences of CsgF used in this study however its importance to the structure or function of CsgF is not known.

Conclusion: Using a multiple sequence alignment of CsgF sequences from 35 bacterial species, as well as an evolutionary conservation analysis using the ConSurf server, we were able to identify significant amino acid conservation throughout the protein. We were able to correlate several of the conserved positions to published reports describing the regions important for the structure and function of CsgF. However there remains positions of amino acid conservation whose importance is not yet known. A systematic investigation of the conserved positions identified here may help add to our understanding of Curli assembly.

References

- Austin JW, Sanders G, Kay WW, Collinson SK. 1998. Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. FEMS Microbiology Letters 162: 295-301. DOI: [10.1111/j.1574-6968.1998.tb13012.x](https://doi.org/10.1111/j.1574-6968.1998.tb13012.x)
- Bian Z, Normark S. 1997. Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*. The EMBO Journal 16: 5827-5836. DOI: [10.1093/emboj/16.19.5827](https://doi.org/10.1093/emboj/16.19.5827)
- Cao B, Zhao Y, Kou Y, Ni D, Zhang XC, Huang Y. 2014. Structure of the nonameric bacterial amyloid secretion channel. Proceedings of the National Academy of Sciences 111: 10.1073/pnas.1411942111. DOI: [10.1073/pnas.1411942111](https://doi.org/10.1073/pnas.1411942111)
- Chapman MR, Robinson LS, Pinkner JS, Roth R, Heuser J, Hammar Mr, Normark S, Hultgren SJ. 2002. Role of *Escherichia coli* Curli Operons in Directing Amyloid Fiber Formation. Science 295: 851-855. DOI: [10.1126/science.1067484](https://doi.org/10.1126/science.1067484)
- Collinson SK, Clouthier SC, Doran JL, Baner PA, Kay WW. 1997. Characterization of the AgfBA Fimbrial Operon Encoding Thin Aggregative Fimbriae of *Salmonella Enteritidis*. Advances in Experimental Medicine and Biology, Mechanisms in the Pathogenesis of Enteric Diseases : 247-248. DOI: [10.1007/978-1-4899-1828-4_37](https://doi.org/10.1007/978-1-4899-1828-4_37)
- Evans ML, Chorell E, Taylor JD, Åden Jr, Götheson A, Li F, et al., Chapman. 2015. The Bacterial Curli System Possesses a Potent and Selective Inhibitor of Amyloid Formation. Molecular Cell 57: 445-455. DOI: [10.1016/j.molcel.2014.12.025](https://doi.org/10.1016/j.molcel.2014.12.025)
- Gibson DL, White AP, Rajotte CM, Kay WW. 2007. AgfC and AgfE facilitate extracellular thin aggregative fimbriae synthesis in *Salmonella Enteritidis*. Microbiology 153: 1131-1140. DOI: [10.1099/mic.0.2006/000935-0](https://doi.org/10.1099/mic.0.2006/000935-0)
- Goyal P, Krasteva PV, Van Gerven N, Gubellini F, Van den Broeck I, Troupiotis-Tsailaki A, et al., Remaut. 2014. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. Nature 516: 250-253. DOI: [10.1038/nature13768](https://doi.org/10.1038/nature13768)
- Hammar M, Bian Z, Normark S. 1996. Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proceedings of the National Academy of Sciences 93: 6562-6566. DOI: [10.1073/pnas.93.13.6562](https://doi.org/10.1073/pnas.93.13.6562)
- Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Research 33: W299-W302. DOI: [10.1093/nar/gki370](https://doi.org/10.1093/nar/gki370)
- Nenninger AA, Robinson LS, Hultgren SJ. 2009. Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. Proceedings of the National Academy of Sciences 106: 900-905. DOI: [10.1073/pnas.0812143106](https://doi.org/10.1073/pnas.0812143106)
- Robinson LS, Ashman EM, Hultgren SJ, Chapman MR. 2005. Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Molecular Microbiology 59: 870-881. DOI: [10.1111/j.1365-2958.2005.04997.x](https://doi.org/10.1111/j.1365-2958.2005.04997.x)
- Römling U, Bian Z, Hammar Mr, Sierralta WD, Normark S. 1998. Curli Fibers Are Highly Conserved between *Salmonella typhimurium* and *Escherichia coli* with Respect to Operon Structure and Regulation. Journal of Bacteriology

180: 722-731. DOI: [10.1128/jb.180.3.722-731.1998](https://doi.org/10.1128/jb.180.3.722-731.1998)

Schubeis T, Spehr J, Viereck J, Köpping L, Nagaraj M, Ahmed M, Ritter C. 2018. Structural and functional characterization of the Curli adaptor protein CsgF. FEBS Letters 592: 1020-1029. DOI: [10.1002/1873-3468.13002](https://doi.org/10.1002/1873-3468.13002)

Sukupolvi S, Lorenz RG, Gordon JI, Bian Z, Pfeifer JD, Normark SJ, Rhen M. 1997. Expression of thin aggregative fimbriae promotes interaction of Salmonella typhimurium SR-11 with mouse small intestinal epithelial cells. Infection and Immunity 65: 5320-5325. DOI: [10.1128/iai.65.12.5320-5325.1997](https://doi.org/10.1128/iai.65.12.5320-5325.1997)

Swasthi HM, Basalla JL, Dudley CE, Vecchiarelli AG, Chapman MR. 2023. Cell surface-localized CsgF condensate is a gatekeeper in bacterial curli subunit secretion. Nature Communications 14: 10.1038/s41467-023-38089-1. DOI: [10.1038/s41467-023-38089-1](https://doi.org/10.1038/s41467-023-38089-1)

Taylor JD, Zhou Y, Salgado PS, Patwardhan A, McGuffie M, Pape T, et al., Matthews. 2011. Atomic Resolution Insights into Curli Fiber Biogenesis. Structure 19: 1307-1316. DOI: [10.1016/j.str.2011.05.015](https://doi.org/10.1016/j.str.2011.05.015)

Yan Z, Yin M, Chen J, Li X. 2020. Assembly and substrate recognition of curli biogenesis system. Nature Communications 11: 10.1038/s41467-019-14145-7. DOI: [10.1038/s41467-019-14145-7](https://doi.org/10.1038/s41467-019-14145-7)

Yariv B, Yariv E, Kessel A, Masrati G, Chorin AB, Martz E, et al., Ben-Tal. 2023. Using evolutionary data to make sense of macromolecules with a “face-lifted” ConSurf. Protein Science 32: 10.1002/pro.4582. DOI: [10.1002/pro.4582](https://doi.org/10.1002/pro.4582)

Zhang M, Shi H, Zhang X, Zhang X, Huang Y. 2020. Cryo-EM structure of the nonameric CsgG-CsgF complex and its implications for controlling curli biogenesis in Enterobacteriaceae. PLOS Biology 18: e3000748. DOI: [10.1371/journal.pbio.3000748](https://doi.org/10.1371/journal.pbio.3000748)

Funding:

Supported by National Institute of General Medical Sciences (United States) 1R15GM123430-01 to Sajith Jayasinghe.

Author Contributions: Karen Guerrero: investigation, writing - review editing, formal analysis. Emma Smith: formal analysis. Shruti Sunder Rajkumar: formal analysis. Zachary Cairo: formal analysis. Jonathan Adame: formal analysis. Renad Rawas: formal analysis. Ranim Rawas: formal analysis. Sajith Jayasinghe: conceptualization, formal analysis, project administration, data curation.

Reviewed By: Irfana Muqbil

History: Received June 17, 2025 **Revision Received** July 16, 2025 **Accepted** July 8, 2025 **Published Online** July 16, 2025 **Indexed** July 30, 2025

Copyright: © 2025 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Guerrero K, Smith E, Sunder Rajkumar S, Cairo Z, Adame J, Rawas R, Rawas R, Jayasinghe S. 2025. Identification of Amino Acid Conservation in The Curli Accessory Protein CsgF. microPublication Biology. [10.17912/micropub.biology.001702](https://doi.org/10.17912/micropub.biology.001702)