Published on Dec 20, 2023 by Arcadia Science **DOI**: 10.57844/arcadia-6ec1-0c09

How can we identify the common molecular signatures underlying polyphosphate accumulation?

Since releasing our pub on polyphosphate-forming proteins in bacteria, we've noticed the community has similar problems studying this process in diverse organisms. We're actively seeking feedback with a focus on advancing basic discoveries and useful tools in this space!

Contributors (A-Z)

Feridun Mert Celebi, Michael Downey, Megan L. Hochstrasser, Elizabeth A. McDaniel , Jasmine Neal

Version 2 · Mar 31, 2025

Purpose

We recently performed a comparative analysis of protein sequence and structural similarities to identify signatures of bacterial polyphosphate accumulation phenotypes. Since releasing this work, new questions have emerged that we've compiled to encourage the broader community to discuss where research in this area should go next. We’re particularly interested in further computational exploration by connecting gene neighborhoods and regulatory motifs with our structural results. While we don’t know if we’ll pursue this research ourselves, we hope that by starting an open discussion, others will be inspired to follow up.

- This is a follow-up to work described in a **prior pub**, “Discovering shared protein structure signatures connected to polyphosphate accumulation in diverse bacteria.” Visit that pub for complete background info and context.

Background on the original pub

We recently shared a pub comparing the structural similarity of phosphate-accumulating PPK1 enzymes from diverse microbes **[1]**. We were interested in understanding why certain bacteria in wastewater treatment plants are particularly good at accumulating polyphosphate and others are not. If we could better predict the molecular mechanisms contributing to microbial polyphosphate accumulation, this could enable rational engineering approaches to make wastewater treatment processes more reliable. Additionally, polyphosphate accumulation is important for virulence and biofilm formation in bacterial pathogens and plays a role in many processes in mammalian cells (albeit through unknown pathways).

We hypothesized that, despite great sequence divergence, protein structure could be a better indicator of function and explain why some bacteria accumulate large amounts of polyphosphate. We found through a comparative analysis of 28,000 PPK1 proteins that this can sometimes be the case, but doesn’t fully explain things **[1]**.

New questions emerge

In an attempt to identify structural or genetic signatures of effective polyphosphate accumulation, a commenter suggested that we consider whether the presence of phosphatase enzymes would reflect an organism’s ability to accumulate polyphosphate, as an organism should possess the ability to mobilize the polyphosphate in some capacity.

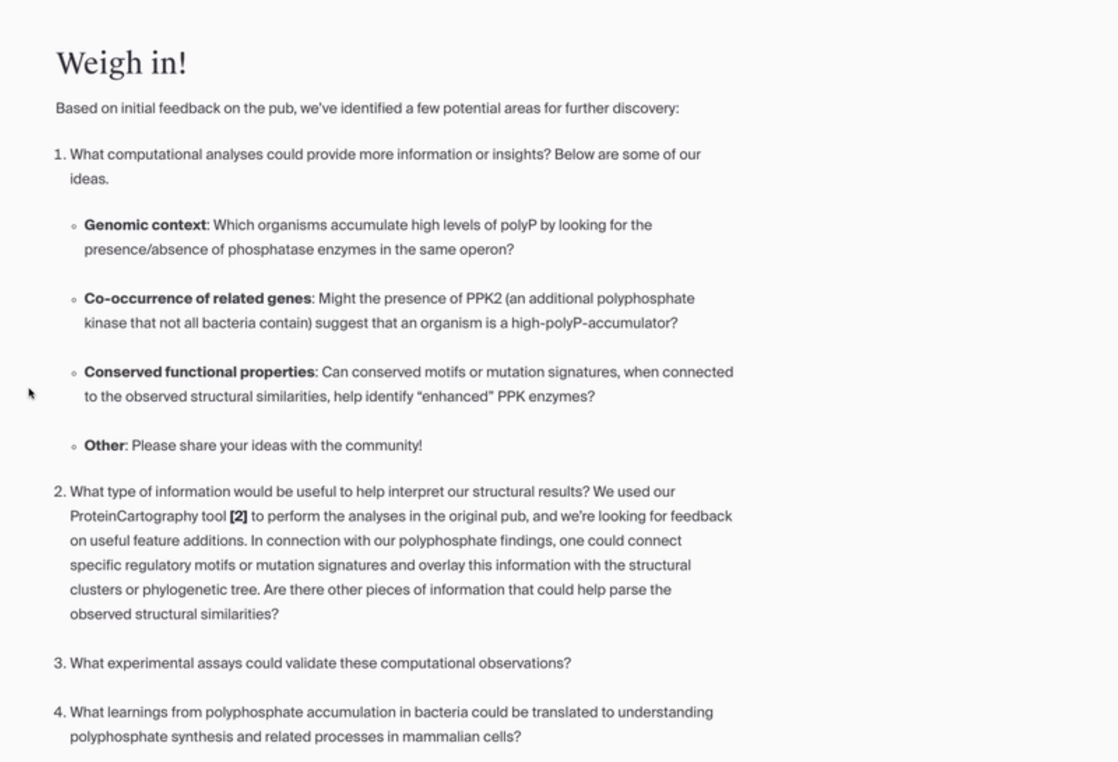
With this in mind, we're hoping to spark discussion on potential next steps. We see the most promise in the following two areas, and invite others to pursue these questions on their own:

- Computational analyses that provide more insights into our structural similarity results, such as looking at gene neighborhoods and mutation signatures
- Comparing *in vitro* and *in vivo* approaches to test the function and activity of PPK1 homologs from understudied organisms that are hypothesized to have enhanced polyphosphate accumulation activity, specifically those that are similar to Accumulibacter

We'd love for others to chime in and talk through some key questions. Our goal is to start a public conversation on this research topic, as we believe polyphosphate accumulation is of broad interest to many research communities. Arcadia isn't pursuing climate technologies in the near term since our focus is on therapeutics, but there are clear implications in this space that we hope other groups pursue.

How can I weigh in?

We hope you'll respond publicly to our questions below by selecting/highlighting the question you'd like to answer, clicking the comment icon, and typing in your thoughts (as shown in the GIF below)! You'll need a PubPub account to do this, but it's free and quick to make one. Here's a quick tutorial on all of this.



Weigh in!

Based on initial feedback on the pub, we've identified a few potential areas for further discovery:

1. What computational analyses could provide more information or insights? Below are some of our ideas.
    - **Genomic context**: Which organisms accumulate high levels of polyP by looking for the presence/absence of phosphatase enzymes in the same operon?
    - **Co-occurrence of related genes**: Might the presence of PPK2 (an additional polyphosphate kinase that not all bacteria contain) suggest that an organism is a high-polyP-accumulator?
    - **Conserved functional properties**: Can conserved motifs or mutation signatures, when connected to the observed structural similarities, help identify "enhanced" PPK enzymes?
    - **Other**: Please share your ideas with the community!
2. What type of information would be useful to help interpret our structural results? We used our ProteinCartography tool **[2]** to perform the analyses in the original pub, and we're looking for feedback on useful feature additions. In connection with our polyphosphate findings, one could connect specific regulatory motifs or mutation signatures and overlay this information with the structural clusters or phylogenetic tree. Are there other pieces of information that could help parse the observed structural similarities?
3. What experimental assays could validate these computational observations?
4. What learnings from polyphosphate accumulation in bacteria could be translated to understanding polyphosphate synthesis and related processes in mammalian cells?

Weigh in!

Based on initial feedback on the pub, we've identified a few potential areas for further discovery:

1. What computational analyses could provide more information or insights? Below are some of our ideas.
    - **Genomic context**: Which organisms accumulate high levels of polyP by looking for the presence/absence of phosphatase enzymes in the same operon?
    - **Co-occurrence of related genes**: Might the presence of PPK2 (an additional polyphosphate kinase that not all bacteria contain) suggest that an organism is

   a high-polyP-accumulator?
   - **Conserved functional properties**: Can conserved motifs or mutation signatures, when connected to the observed structural similarities, help identify “enhanced” PPK enzymes?
   - **Other**: Please share your ideas with the community!
2. What type of information would be useful to help interpret our structural results? We used our ProteinCartography tool **[2]** to perform the analyses in the original pub, and we’re looking for feedback on useful feature additions. In connection with our polyphosphate findings, one could connect specific regulatory motifs or mutation signatures and overlay this information with the structural clusters or phylogenetic tree. Are there other pieces of information that could help parse the observed structural similarities?
3. What experimental assays could validate these computational observations?
4. What learnings from polyphosphate accumulation in bacteria could be translated to understanding polyphosphate synthesis and related processes in mammalian cells?

We appreciate any public feedback on the above questions or anything else that came up while reading the pub! We feel that an open conversation on this project will benefit the polyphosphate community as a whole, and we hope others will build on the ideas presented here and in future comments.

References

1 Avasthi P, Celebi FM, McDaniel EA. (2024). Discovering shared protein structure signatures connected to polyphosphate accumulation in diverse bacteria. https://doi.org/10.57844/ARCADIA-AC10-23E7

2 Avasthi P, Bigge BM, Celebi FM, Cheveralls K, Gehring J, McGeever E, Mishne G, Radkov A, Sun DA. (2024). ProteinCartography: Comparing proteins with structure-based maps for interactive exploration. https://doi.org/10.57844/ARCADIA-A5A6-1068