

# Structural basis of undecaprenyl phosphate glycosylation leading to polymyxin resistance in Gram-negative bacteria

Corresponding Author: Dr Vasileios Petrou

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This manuscript reports the CryoEM structure of ArnC - the glycosyl transferase that transfers aminoarabinose from the corresponding UDP-sugar to undecaprenol phosphate.

ArnC is an interesting target due to its significance in bacterial antibiotic resistance. The work is generally executed well, however, I find some “disconnects” certain important aspects of the work. These issues are addressed in the following comments and require attention.

The structure is determined in MSP-based nanodiscs that are prepared by supramolecular assembly using the DDM purified protein followed by treatment with hydrophobic beads and POPG/MSP1E3D1 NDs. SEC data look good, but there would be some ambiguity from either the DDM or the NDs

The phospholipid (PL) composition is 100% POPG in the ND assembly, however, as bacterial inner membranes generally include a mixture of PLs (typical gram-negative bacteria ~65 mol% phosphatidyl ethanolamine, ~23 mol% phosphatidyl glycerol, and 10 mol% cardiolipin), it is unclear why a non-native lipid composition lacking the zwitterionic PE is used. Fig. S1B,C is called out suggesting evidence for optimization however, this is not shown in the SI fig. The authors should comment as these choices could impact activity and oligomerization.

ArnC is proposed to be a native tetramer, however, the PfDPMS, another characterized Pren-P GT is a monomer. Additionally, as the PL composition of the ND is non-native, there is not direct evidence supporting the physiological tetramerization state. The authors should consider directly extraction of ArnC from a bacterial membrane preparation into SMALP Nanodiscs using the affinity tags and perform cross-linking to support a native tetrameric state. Also, in this context the authors should comment on specific interactions in the soluble and membrane domains, which might promote tetramerization that are absent in the PfDPMS structure? More corroborating evidence would be significant.

The discussion of oligomerization state is also important because there is discussion of significant structural changes between Apo and UDP bound with C4 symmetry imposed. However, there is a concern that more conformations might be obscured, which would be visualized (and of functional importance) in the absence of symmetry.

Another disconnect appears to be in the simulations of ArnC and analysis of lipid binding – It is not really transparent whether all the simulations and comparisons with PfDPMS are done with an extracted monomer of ArnC from the structural data? If so, can the authors comment on any changes in lipid binding and access to binding sites that might result from ArnC tetramerization?

Can the authors comment on the use of Mn(II) concentrations at 1 mM in the various analyses. This concentration is considerably above reported levels in bacterial cells (1-10  $\mu$ M (micromolar) range). Also, a comparison is made with Mg(II) at 1 mM, however Mg(II) concentrations are in the 10-20 mM range. There is no doubt that the variation of UDP affinity trends that are presented are valid in the presented experimental condition – it is more an issue of the physiological relevance. A better understanding of the metal ion dependence at physiological concentrations of metal ions is critical for defining mechanistic hypotheses.

NOTE: 1. Ref 36 is from 2007 thus the information is out of date – the authors should revert to the CAZY database for the current numbers of GT families.

#### Reviewer #2

(Remarks to the Author)

In this manuscript, Ashraf et al. present the cryo-EM structures of ArnC from *S. enterica* in apo and UDP bound states. ArnC is a tetrameric, membrane-bound glycosyltransferase that transfers L-Ara4N to undecaprenyl phosphate. For building the UDP-bound state, the authors use microscale thermophoresis experiments to demonstrate that manganese ion is the favoured divalent cation coordinating the DXD motif. The authors also present coarse-grained and atomistic molecular dynamics simulations to gain insight into substrate binding and catalysis.

This pathway contributes to polymyxin resistance in Gram-negative bacteria. Given that polymyxin is an antibiotic of last resort and in view of the rising levels of antibiotic resistance, this an important area of research. The main conclusions of the paper are in line with earlier proposals and predictions and are generally supported by the data. However, there is no functional validation of the structural data or of the proposed mechanism, nor is there evidence that the purified protein is active. Functional assays would provide much needed support of the models and suggested mechanisms.

As the authors mention, ArnC is a potential antimicrobial target making structural insight potentially useful for drug development. However, as no ternary complex structure is presented here, the biological insight of the structural results is somewhat limited with respect to the usefulness of the models for rational drug design. While the GT domain shifts observed upon binding of UDP are intriguing, without observing both substrates the explanations of this movement are speculative.

The comparison of 200 and 300 keV cryo-EM data as well as discussion of the relationship between particle number and resolution does not convey biological insight and contributes only limited technical insight. It is unlikely that a difference of 0.04 Å in resolution is meaningful, in particular since the authors note that there were no differences observed in the maps. Given that the data comparison contains only data from a single protein, the observations are somewhat anecdotal. Perhaps this section could be reduced or removed as it does not contribute to the main story.

#### Reviewer #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

#### Reviewer #4

(Remarks to the Author)

General comments

The work provides important insights that guide the elucidation of mechanistic aspects of the glycosyltransferase ArnC. The theme of antibiotic resistance is of paramount importance for global issues and for scientific community. The work is well written and its major strength is to bring apo and UDP-bound structures of ArnC by Cryo-EM. However, some concerns should be addressed to enable the publication.

1- The authors cite the structures of ArnC already solved by crystallography. However, it is not cited the sequence identity between those proteins neither structural comparisons such as RMSDs between each ArnC previously solved and the current structures.

2- It seems there is a structure published in Jan. 2025, reporting a *Salmonella typhimurium* ArnC (Patel et al., doi:10.1002/pro.70037) with similar features. Is this a homologous enzyme? If yes, this should be discussed and the differences between them evaluated.

Other major points:

3- Is it possible that a catalytic likely proficient configuration was not reached by the states generated by course grained simulations due to the level of theory used?

4- I think it would be necessary to enlarge atomistic simulations production time (100 ns for each simulation is too short to see major structural rearrangement). Perhaps this would also allow the authors to see better coordination generated by state 1 and state 2 with D100, which seem to be an important residue.

5- In the sentence in the discussion "Our atomistic simulations have shown that the acceptor phosphate of UndP is coordinated in the proximity of D100...", I agree that, for simulation started from model 3, the coordination with D100 is maintained (Fig. 6c). However, is it not a consequence of the starting model? Which were the basis for its construction?

6- It is not clear for me the real contribution of the CG simulations for the mechanism elucidation.

7- Overall, I think that the authors could consider more simulations (or methods such as for enhancing sampling) for the conclusions made.

8- Discussion line 461. If there is a study already showing the importance of the first aspartic (and mutation?), this could be discussed.

Minor points:

- CAZY database currently has more than 130 families of transferases. Introduction line 84/85 should be updated and also

the reference (I suggest to bring the GTs families description reference and also the most recent up-to-date CAZy database reference).

- Introduction line 99. "several structures". I think that there are just a few, based on the works discussion. Moreover, to anticipate the information of the novelty brought from the apo structure could strengthen the findings.
- Fig 1. I suggest to bring the scheme in "a" in a separate figure. Also, the order of pannels a,b,c,d,e is bit confusing the way it is currently presented.
- Fig. 6 legend. RMSDs of the protein were based in which atoms? Backbone?
- Line 371-372. "... indicates that the presence of the lipid is not required for UDPA binding". Could it be due to the sampling?

Version 1:

Reviewer comments:

Reviewer #4

(Remarks to the Author)

The authors of the manuscript addressed the questions raised during the reviewing process and reformulated key point of the manuscript. I appreciate the changes made, specially regarding structural comparisons and MD clarifications. I only have few observations:

- The phrase "During the final stages of preparation for this article, we became aware of a competing publication reporting a cryo-EM structure of ArnC from *S. enterica* in apo (2.75Å) and UDPbound conformations (3.8Å)" could be replaced by "A recent publication highlighted a cryo-EM structure of ArnC from *S. enterica* in apo (2.75Å) and UDPbound conformations (3.8Å)". If significant changes were observed (and clarifications made in this new version) I do not see the publication as competing but complementary.
- Fig 6c: plots are too small
- Line 383 "Three repeats of unrestrained 500 ns MD simulations were performed for each system" Replace to: "three independent simulations of unrestrained 500 ns ...." (if 3 independent simulations started from different random velocities).

Reviewer #5

(Remarks to the Author)

I provided comments in the attached document.

Reviewer #6

(Remarks to the Author)

The authors have supplemented experiments and provided comprehensive responses to Reviewer #1's concerns.

The reviewer #2 mainly raised two questions:

Q1: The reviewer #2 emphasized that establishing a purified ArnC enzymatic activity assay would be necessary to fully elucidate the mechanism.

In the response, the authors pointed out that no commercial substrate is currently available, making it impossible to perform in vitro enzymatic activity assays. However, in fact, the substrate UDP-L-Ara4FN can be obtained using the method described in Reference 27 (J. Biol. Chem. 280, 14154–14167) cited by the authors. Given that the substrate is radiolabeled with C14, the authors can supplement bacterial genetic mutations in vivo assays to validate the complete mechanistic understanding of ArnC alternatively.

Q2: Another major concern raised by Reviewer #2 was the lack of a ternary complex structure (donor-acceptor-enzyme), which may limit the mechanistic understanding of ArnC catalysis and its potential for guiding drug design.

In this regard, the authors employed molecular dynamics (MD) simulations to validate donor substrate binding and dynamics. Combined with their previously resolved structural data, the authors have addressed question 2 from Reviewer #2.

**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

## Response to Reviewer's comments (Responses in Blue)

We thank the reviewers for their insightful comments. In response we have incorporated a number of changes to the manuscript. We provide responses and describe the changes made below.

Reviewer #1 (Remarks to the Author):

*This manuscript reports the CryoEM structure of ArnC - the glycosyl transferase that transfers aminoarabinose from the corresponding UDP-sugar to undecaprenol phosphate.*

*ArnC is an interesting target due to its significance in bacterial antibiotic resistance. The work is generally executed well, however, I find some "disconnects" certain important aspects of the work. These issues are addressed in the following comments and require attention.*

*The structure is determined in MSP-based nanodiscs that are prepared by supramolecular assembly using the DDM purified protein followed by treatment with hydrophobic beads and POPG/MSP1E3D1 NDs. SEC data look good, but there would be some ambiguity from either the DDM or the NDs*

We thank reviewer #1 for their positive feedback on the significance and quality of our work.

*The phospholipid (PL) composition is 100% POPG in the ND assembly, however, as bacterial inner membranes generally include a mixture of PLs (typical gram-negative bacteria ~65 mol% phosphatidyl ethanolamine, ~23 mol% phosphatidyl glycerol, and 10 mol% cardiolipin), it is unclear why a non-native lipid composition lacking the zwitterionic PE is used. Fig. S1B,C is called out suggesting evidence for optimization however, this is not shown in the SI fig. The authors should comment as these choices could impact activity and oligomerization.*

We thank the reviewer for the request to discuss the reasoning behind selecting POPG for the nanodisc reconstitution. We have now included a nanodisc incorporation trial in Supplementary Fig. 2C. Based on the monodispersity of the nanodisc peaks in SEC, a mixture of *E. coli* polar lipids (PL) is suboptimal for the stability of ArnC. PL leads to a severely diminished peak, which is also not monodisperse. On the contrary, peaks for nanodisc containing either POPC or POPG produce monodisperse peak of approximately equal magnitude. POPG was selected for the final sample because it seems to produce somewhat reduced aggregation compared to POPC. It is conceivable that in the native membrane, ArnC is making additional interactions with other proteins that enable it to be stable within the native lipid environment, but in the context of nanodisc-reconstituted protein, PLs do not help stabilize ArnC. Additional text was added in lines 120-122 describing the new figure panel.

*ArnC is proposed to be a native tetramer, however, the PfDPMS, another characterized Pren-P GT is a monomer. Additionally, as the PL composition of the ND is non-native, there is not direct evidence supporting the physiological tetramerization state. The authors should consider directly extraction of ArnC from a bacterial membrane preparation into SMALP Nanodiscs using the affinity tags and perform cross-linking to support a native tetrameric state. Also, in this context the authors should comment on specific interactions in the soluble and membrane domains, which might promote tetramerization that are absent in the PfDPMS structure? More corroborating evidence would be significant.*

We thank the reviewer for their suggested experiment. We have now incorporated in Supplementary Fig. 2B an experiment where ArnC is extracted with cubipol polymer. Cubipol is similar to SMA but uses a different backbone. ArnC protein extracted with cubipol was crosslinked using glutaraldehyde treatment. Two major species are observed after crosslinking, a likely dimeric and tetrameric species. The tetrameric band is increasing in abundance with longer treatment, while the dimeric band decreases in the same timeframe. This suggests that the protein is tetrameric, and that shorter treatment is mainly able to crosslink adjacent protomers, while longer treatment captures the full tetramer. Additional text was added in lines 122-126 describing the new figure panel, and in lines 551-563 describing the new methodology.

In addition, we discussed extensively in the manuscript that ArnC features a similar fold with the bacterial protein GtrB, which catalyzes the transfer of glucose from a UDP nucleotide to undecaprenyl

phosphate. GtrB is a tetramer as is shown in Supplementary Fig. 8. PfDPMS is an archaeal protein and indeed crystallized as a monomer, but the most likely oligomeric state is dimeric (Gandini et al. 2017)<sup>1</sup>. The reason for the dimeric state for PfDPMS is that it contains an additional two transmembrane segments located almost perpendicular to the vertical TM helices, which almost certainly prevent more than two protomers to assemble. Supplementary figures 8 and 9 show that the transmembrane domain of ArnC is very similar to the one from GtrB and not as similar to the one from PfDPMS. Thus, a tetrameric configuration is also supported by the structural features of the transmembrane domain.

Finally, in lines 310-331, we have added extensive discussion about recently published structures of ArnC that have been determined in amphipol. Despite the distortions in the amphipol structures that we discuss extensively, the oligomeric state observed for ArnC by a different group after reconstitution into amphipol is also tetrameric.

Thus, considering all available evidence, we believe that a tetrameric configuration for ArnC is well supported.

*The discussion of oligomerization state is also important because there is discussion of significant structural changes between Apo and UDP bound with C4 symmetry imposed. However, there is a concern that more conformations might be obscured, which would be visualized (and of functional importance) in the absence of symmetry.*

We thank the reviewer for bringing up this concern. To address this, we have created Supplementary Fig. 12. In this figure, we show reconstructions of the final particle stacks without imposing any symmetry (C1). A new ab initio (without symmetry) and a new non-uniform refinement (in C1) were performed using the final particle stack in each dataset. Supplementary Fig. 12A shows top views for both *apo* datasets. And in 12B, we show a reconstruction in C1 for the UDP-bound dataset in three orthogonal views. In all these cases it is obvious that the reconstructions exhibit fourfold symmetry, even when it is not imposed during the reconstructions. In addition, as can be seen in the local resolution maps presented in Supplementary Fig. 3F, 4F and 10F, only the top of the transmembrane domain and the JM helices exhibit reduced resolution, which could be interpreted as higher conformational fluidity. Thus, it is unlikely that any additional conformations in our dataset were missed or obscured, since averaged conformations would have led to significantly worsened resolution for parts our reconstructions, which is not observed. Finally, in Supplementary Fig. 12C we show the density of UDP in each protomer of ArnC. The densities for all protomers have similar appearance/quality even without imposing fourfold symmetry, arguing that all protomers are saturated with UDP. We believe that in the present manuscript we have captured two stable states of ArnC, apo and fully saturated UDP-bound conformation. In the future, it would be interesting to test whether lower concentrations of UDP induce states where only some of the GT-A domains bind UDP and complete the conformational transition into the UDP-bound state, which would be expected if the four GT-A modules in each molecule operate independently. We have included new text in lines 137-139, and 282-286 describing the findings using the C1 reconstructions.

*Another disconnect appears to be in the simulations of ArnC and analysis of lipid binding – It is not really transparent whether all the simulations and comparisons with PfDPMS are done with an extracted monomer of ArnC from the structural data? If so, can the authors comment on any changes in lipid binding and access to binding sites that might result from ArnC tetramerization?*

All simulations of ArnC, both coarse-grained and atomistic, were performed using the full tetrameric assembly of ArnC as resolved in the experimental structure. The coarse-grained simulation captured one event of an undecaprenyl phosphate lipid going through the JM helices and entering the active site of the GT-A domain of one protomer. The initial 100ns atomistic simulations were initiated from the top ranked CG pose, and observed the interactions of the undecaprenyl lipid within one protomer of the ArnC tetramer. In the longer 500ns atomistic simulations, both substrates (UDP-L-Ara4FN and undecaprenyl phosphate) were added to each protomer of the tetrameric ArnC in the configurations described in the manuscript. We hope that this clarifies our workflows, we have added some text on lines 352, 375-376, and 383 to further clarify this in text.

Based on current available evidence, we believe that the four GT-A domains act independently to bind the UDP-L-Ara4FN substrate. The reaction in each protomer will be completed once an undecaprenyl phosphate lipid enters each active site and is positioned within it productively for catalysis to occur.

Can the authors comment on the use of Mn(II) concentrations at 1 mM in the various analyses. This concentration is considerably above reported levels in bacterial cells (1-10  $\mu$ M (micromolar) range). Also, a comparison is made with Mg(II) at 1 mM, however Mg(II) concentrations are in the 10-20 mM range. There is no doubt that the variation of UDP affinity trends that are presented are valid in the presented experimental condition – it is more an issue of the physiological relevance. A better understanding of the metal ion dependence at physiological concentrations of metal ions is critical for defining mechanistic hypotheses.

We appreciate the comment by the reviewer, and understand the concern raised here regarding physiological relevance for Mn(II) given the concentration Mn(II) vs Mg(II) within the cell. Yet, the goal of our MST experiment, which is also reflected on the experimental design, was to determine the optimal ion for structural studies based on the binding affinity of the enzyme towards UDP in the presence of each metal. For this reason, we chose to use a saturating concentration of each metal (1mM), to fully occupy available binding sites with the metal and be able to perform a direct comparison in the binding affinity of UDP. As a result, we only claim that Mn(II) is optimal for our structural studies, not necessarily that it is the physiologically relevant metal ion for ArnC. It may be the case that despite the lower affinity towards UDP(sugar) in the presence of Mg(II), the enzyme is still able to catalyze its reaction. In this case the high concentration of Mg(II) in the cell would help to quickly fill all binding sites on each enzyme. Alternatively, it may be the case that only Mn(II) is able to stably coordinate UDP(sugar) and the reaction must wait for the “correct” metal to arrive in the binding site. Additional experiments in the form of a reconstituted catalytic assay would be needed to determine which of these possibilities is taking place, but because the donor substrate is not commercially available, such an assay is beyond the scope of the present work. To make clear that we are not making a case for Mn(II) as the physiological metal for ArnC we added text in lines 273-274 indicating that the metal ion used by ArnC in the cellular environment may depend on the relative abundance of metal ions near each metal-binding site.

**NOTE:** 1. Ref 36 is from 2007 thus the information is out of date – the authors should revert to the CAZY database for the current numbers of GT families.

We thank the reviewer for their suggestion, which was also brought up by reviewer #4. The text in line 84/85 has been updated with the current number of GT families and references for the GT family description and the CAZy database (latest one) have been added.

Reviewer #2 (Remarks to the Author):

*In this manuscript, Ashraf et al. present the cryo-EM structures of ArnC from S. enterica in apo and UDP bound states. ArnC is a tetrameric, membrane-bound glycosyltransferase that transfers L-Ara4N to undecaprenyl phosphate. For building the UDP-bound state, the authors use microscale thermophoresis experiments to demonstrate that manganese ion is the favoured divalent cation coordinating the DXD motif. The authors also present coarse-grained and atomistic molecular dynamics simulations to gain insight into substrate binding and catalysis.*

*This pathway contributes to polymyxin resistance in Gram-negative bacteria. Given that polymyxin is an antibiotic of last resort and in view of the rising levels of antibiotic resistance, this an important area of research. The main conclusions of the paper are in line with earlier proposals and predictions and are generally supported by the data.*

We thank reviewer #2 for highlighting the significance of our work and concurring that our main conclusions are supported by the data presented.

*However, there is no functional validation of the structural data or of the proposed mechanism, nor is there evidence that the purified protein is active. Functional assays would provide much needed support of the models and suggested mechanisms.*

We fully understand the comment by the reviewer regarding lack of a functional/catalytic assay in our



work. Developing a functional assay for ArnC is one of our highest priorities for our future experimentation. The reason that such an assay is not included in the present work is that the donor substrate for the reaction catalyzed is not commercially available and must be purified by reconstituting the aminoarabinose pathway to synthesize the soluble UDP-L-Ara4FN substrate. As a comparison, the other Pren-P GT enzymes we discuss in our manuscript utilize commercially available substrates: GtrB (UDP-glucose) and PfDPMS (GDP-mannose). If we could readily access a commercially available substrate that is compatible with our enzyme, reconstituting a UDP-Glc glycosyltransferase assay would be rather straightforward; but because we do not, a catalytic assay is not within scope of the present work. In addition, the reviewer mentions that there is no evidence presented that the purified protein is active. While this is technically true given that we don't currently have the ability to test the catalytic activity of the purified protein directly, we would like to highlight that the protein is able to bind its partial donor substrate UDP (MST, cryo-EM experiments), and it is fully ordered (UDP-bound structure is fully traced, and apo structure is only missing a small number of residues on one loop due to flexibility). Moreover, both GtrB and PfDPMS purified with similar approaches to ours yield catalytically active proteins. Thus, there is little reason to argue that ArnC specifically will be the exception. In fact, our ability to incorporate ArnC into lipid nanodiscs that provide a more native membrane environment increases the chance that our sample will provide a better approximation of the catalytic activity of ArnC compared to detergent, once we have successfully established an activity assay. We hope that the reviewer will acknowledge the additional hurdle that the absence of substrate availability poses and accept our commitment to characterize the catalytic activity of the enzyme in subsequent studies.

*As the authors mention, ArnC is a potential antimicrobial target making structural insight potentially useful for drug development. However, as no ternary complex structure is presented here, the biological insight of the structural results is somewhat limited with respect to the usefulness of the models for rational drug design. While the GT domain shifts observed upon binding of UDP are intriguing, without observing both substrates the explanations of this movement are speculative.*

The comment by the reviewer has two parts that need to be addressed. The first part claims that since an experimental ternary complex (with both substrates present) was not achieved, the biological insight gained as a result is limited, and in addition this also limits the usefulness of the models for rational drug design. First, a ternary complex was in fact determined experimentally, with a partial substrate and metal that completes the coordination sphere of the substrate. This confirms beyond sequence similarity that this glycosyltransferase enzyme functions like other similar enzymes. Second, a novel conformational transition was identified that has not been previously described in the family because all previous structures were in fact ternary complexes with ligands, but a true apo conformation was never determined for the Pren-P GT subfamily of glycosyltransferases before (due to constraints of X-ray crystallography methods). Third, we have used well-established computational methods to verify that movements of the lipidic substrate, which have been either proposed (GtrB) or observed (PfDPMS) in similar enzymes, are likely taking place within the chemical environment of the new experimentally-derived structure. We also use these computational methods to extensively characterize the potential coordination of both substrates in several positions within the active site, insight which will be difficult to capture using experimental methods, especially given the special donor substrate that this enzyme utilizes. And we propose a catalytic mechanism that is likely to operate in the entire family of Pren-P GTs, given the conservation of the DXD motif within the family. Finally, we disagree with the claim that the present work has limited usefulness for rational drug design targeting the ArnC enzyme. In fact, the donor substrate-binding groove, which is thoroughly characterized based on our highest-resolution structure available for the UDP-bound conformation for ArnC, represents the best possible target for designing a competitive inhibitor. The donor substrate is necessary for the reaction to occur, so displacing it should readily inhibit the enzyme. The binding groove is also rigid and can be readily targeted for rational drug design.

The second part of the comment by the reviewer is rather puzzling. The reviewer claims that a conformational transition that takes place with (and thus is triggered by) only the partial donor substrate UDP, must be characterized in the presence of both substrates, so as to not be speculative. We think it is important to note that in most enzymatic reactions, substrates will bind sequentially, and the binding of each substrate may have different kinetic parameters. Here, we provide experimental characterization of a conformational transition that happens with the binding of the (partial) donor substrate, and we use computational methods to predict how the acceptor substrate may be coordinated within the active site for catalysis to occur. Based on the available evidence that we have, we believe that the donor site and the acceptor site are liganded



independently, and represent two different kinetic components: the lipid snaking through JM helices is slow, while a diffusible substrate binding on a groove and triggering a conformational transition is fast. We believe that our claims are well supported by the entire body of evidence presented here and the literature for previously characterized enzymes in the Pren-P GT family. At the same time, our study builds on and extends previous findings with both experimental and computational evidence. We hope that the reviewer will recognize the value of our contribution, which expands the literature in a GT family that has been understudied.

*The comparison of 200 and 300 keV cryo-EM data as well as discussion of the relationship between particle number and resolution does not convey biological insight and contributes only limited technical insight. It is unlikely that a difference of 0.04 Å in resolution is meaningful, in particular since the authors note that there were no differences observed in the maps. Given that the data comparison contains only data from a single protein, the observations are somewhat anecdotal. Perhaps this section could be reduced or removed as it does not contribute to the main story.*

We concur with reviewer #2 that the 0.04Å difference is not biologically significant. This is, in fact, crucial to our technical point: that the voltage effect is insignificant compared to the specimen-intrinsic contributors to the B-factor.

Because this was unclear, we revised the text to make it more explicit that this is providing evidence for the absence of a significant effect of voltage and other imaging factors. The original draft read, "We conclude that biochemical and/or sample preparation factors more significant limiting factors for ArnC structure determination than electron microscope voltage." We deleted this sentence and inserted the following text summarizing our findings: "The B-factor and resolution are governed by sample-extrinsic factors such as the accelerating voltage of the imaging system and the detective quantum efficiency, as well as sample-intrinsic factors such as the flexibility of the molecules and the background scattering. In the case of ArnC, we show that (within these two high-quality imaging systems) extrinsic factors such as accelerating voltage had a negligible effect on resolution, but sample-intrinsic factors imposed definitive limits on resolution." (lines 213-217)

We disagree that the technical finding conveys limited technical insight. Although several recent papers or preprints have solved structures at different voltages (e.g. McMullan et al 2023, Karia et al 2024)<sup>2,3</sup>, the experiments are typically not sufficiently controlled to make direct head-to-head comparisons, but moreover signal analysis is not performed. Also, there are examples of cases for meaningful resolution improvements when moving from a 200kV to a 300kV microscope (e.g. Cushing et al., 2024)<sup>4</sup>, but our results show that gains will be sample-dependent. We hope the reviewer will concur that our observations are significant to providing a more rounded view of voltage effects in the literature.

#### Reviewer #3 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We thank reviewer #3 for their service, and we hope that our manuscript provided a good review experience.

#### Reviewer #4 (Remarks to the Author):

##### General comments

The work provides important insights that guide the elucidation of mechanistic aspects of the glycosyltransferase ArnC. The theme of antibiotic resistance is of paramount importance for global issues and for scientific community. The work is well written and its major strength is to bring apo and UDP-bound structures of ArnC by Cryo-EM. However, some concerns should be addressed to enable the publication.

We thank reviewer #4 for their kind evaluation of our work and their constructive suggestions for improving our manuscript.

*1- The authors cite the structures of ArnC already solved by crystallography. However, it is not cited the sequence identity between those proteins neither structural comparisons such as RMSDs between each ArnC previously solved and the current structures.*

We should clarify that in this manuscript we have discussed only structures from two other enzymes in the same glycosyltransferase subclass (Pren-P GTs), GtrB and PfDPMS, which have been previously determined using X-ray crystallography. There are no structures of ArnC that have been determined by crystallography, though other enzymes in the aminoarabinose pathway have been previously determined by X-ray crystallography (i.e. ArnA, ArnB, ArnD, and ArnT). The sequence identity between ArnC and GtrB is 30.2%, while the sequence identity between ArnC and PfDPMS is 24%. RMSDs between such divergent sequences or structures may not be very informative. Instead, we have provided an extensive fold comparison between ArnC and the other two proteins in Supplementary Fig. 8 and 9. In addition, a sequence alignment between the three proteins, showcasing the most conserved elements, is included in Supplementary Fig. 15B.

*2- It seems there is a structure published in Jan. 2025, reporting a Salmonella typhimurium ArnC (Patel et al., doi:10.1002/pro.70037) with similar features. Is this a homologous enzyme? If yes, this should be discussed and the differences between them evaluated.*

We confirm that a competing publication reporting structures of *Salmonella* ArnC (same enzyme) was published around the time that we were finalizing our manuscript for submission. This publication includes an *apo* reconstruction with similar reported resolution as ours, and a much lower-resolution UDP-bound reconstruction (nominal 3.8Å). These structures were determined in amphipol. The reconstructions made in this publication were only able to support C2 symmetry, likely because of distortions in the transmembrane domain resulting from being embedded in a non-membrane environment or the quality of the preparations used. We believe that the distortions observed in both the TM and JM helices are non-physiological. We use RMSD to quantify the divergence of the amphipol structures from our nanodisc-embedded structures, showing a very significant RMSD of 6.5Å in the case of the *apo* structure, and a more modest 3.3Å RMSD in the case of the UDP-bound protein. We have added an extensive comparison between our structures and the amphipol structures in lines 309-330.

*Other major points:*

*3- Is it possible that a catalytic likely proficient configuration was not reached by the states generated by course grained simulations due to the level of theory used?*

Yes, it is indeed possible (even likely) that a catalytically proficient configuration was not fully reached in the coarse-grained (CG) simulations. CG simulations are useful for sampling large conformational spaces and identifying general lipid binding modes and sites, however they simplify atomic details and therefore we switch to atomistic level of detail to capture specific interactions crucial for catalysis, such as hydrogen bonding, metal coordination, or electrostatics within the active site.

In our study, CG simulations were employed to capture the likely entry pathway and potential lipid binding sites, guiding subsequent atomistic simulations. Atomistic simulations initiated from the top-ranked CG pose revealed a stable coordination of UndP by the key residues R131 and R137, along with hydrophobic interactions between the lipid tail and the JM helices.

Our multi-scale simulation strategy supports a model where UndP initially engages with the JM helices and enters the GT-A domain (as suggested by CG simulations), then transitions into a catalytically competent configuration in the presence of the second substrate and the divalent ion (as proposed in State 3). Thus, while CG simulations alone likely do not capture a fully catalytically favorable state, they served as a valuable tool to generate starting configurations for higher-resolution atomistic simulations, which were used to assess the catalytic relevance in greater detail.

We have modified text on lines 352, 375-376, and 383 to provide better distinction between CG and the two levels of atomistic simulations performed in this work.

*4- I think it would be necessary to enlarge atomistic simulations production time (100 ns for each simulation is too short to see major structural rearrangement). Perhaps this would also allow the authors to see better coordination generated by state 1 and state 2 with D100, which seem to be an important residue.*

We thank reviewer #4 for this observation. We would like to clarify that only the initial atomistic simulations of ArnC bound to UndP alone were run for 100 ns. These shorter simulations were designed specifically to evaluate the stability and plausibility of the UndP binding pose derived from the coarse-grained (CG) simulations.

For the mechanistic analysis involving ArnC in complex with both substrates (UndP and UDP-A) and the divalent cation, all atomistic simulations (States 1, 2, and 3) were performed in triplicate and run for 500 ns each. These longer simulations were essential to evaluate substrate coordination, relative positioning, and catalytically relevant interactions over a more appropriate timescale. Notably, interactions involving D100 were consistently observed in the simulations of State 3, supporting its proposed role in catalysis. We have attempted to make these details clearer in the manuscript, on lines 352, 375-376, and 383.

*5- In the sentence in the discussion “Our atomistic simulations have shown that the acceptor phosphate of UndP is coordinated in the proximity of D100...), I agree that, for simulation started from model 3, the coordination with D100 is maintained (Fig. 6c). However, is it not a consequence of the starting model? Which were the basis for its construction?*

We thank reviewer #4 for this observation. Indeed, in the atomistic simulations initiated from model 3, the proximity of the acceptor phosphate of UndP to D100 is influenced by the initial configuration. However, the construction of this model was not arbitrary. Model 3 was designed specifically to test a potential catalytically competent pose, in which the UndP phosphate is placed deeper within the catalytic pocket and positioned for a SN<sub>2</sub>-like nucleophilic attack on the C1 atom of the sugar. The configuration was guided by mechanistic considerations, particularly the requirement for a catalytically viable geometry where the acceptor phosphate lies near the C1 carbon of the sugar, enabling nucleophilic attack. This arrangement also positions D100 as a potential catalytic base, analogous to D89 in PfDPMS, which has been shown to be essential for activity. Additionally, our simulations show that although the starting pose was optimized for catalysis, the coordination between the UndP phosphate and D100 remains stable over 500 ns of simulation. This persistent interaction supports the functional relevance of the model and suggests that the observed coordination is not merely an artifact of the starting structure, but reflects a stable, catalytically favorable arrangement. We have added further details to the manuscript to reflect on your point on lines 456-457.

*6- It is not clear for me the real contribution of the CG simulations for the mechanism elucidation.*

The primary contribution of the CG simulations was to identify potential lipid (UndP) entry pathways and binding sites on ArnC. Due to their efficiency in sampling large conformational spaces, CG simulations allowed us to capture the spontaneous association of UndP with the transmembrane domain and its progression toward the GT-A domain through the juxtamembrane (JM) helices. This supports our hypothesis that the JM helices facilitate substrate access, similar to the mechanism proposed for GtrB and observed in PfDPMS.

While the CG approach does not provide atomic-level resolution of catalytic interactions, it was employed as a first step to generate a plausible initial binding pose for UndP. This CG-derived pose was subsequently used as starting points for higher-resolution atomistic simulations, where the second substrate (UDP-A) and the divalent cation were included to explore potential catalytically favorable configurations in detail. Thus, the CG simulations were fundamental in capturing early-stage substrate recruitment and positioning, which, in combination with atomistic simulations, enabled us to propose a more complete mechanistic model that includes both substrates.

We cover these details in the manuscript alongside point 3, raised earlier, on lines 352-356. Indeed this section explains the rationale for the CG simulations.

*7- Overall, I think that the authors could consider more simulations (or methods such as for enhancing sampling) for the conclusions made.*

We thank reviewer #4 for this suggestion. We agree that additional simulations or the application of enhanced sampling methods can provide deeper insights into the conformational landscape and catalytic mechanism of ArnC. In our study, we tested three distinct configurations for the catalytic complex of ArnC, which were carefully designed based on insights from our coarse-grained (CG) simulations and structural data from related enzymes (*PfDPMS* and *GtrB*). These configurations were modelled to reflect potential catalytic configurations and explore different orientations of the substrates within the active site. The main goal of our atomistic simulations was to assess the stability and compatibility of these proposed configurations with catalysis. To this end, we performed **triplicate 500 ns** atomistic simulations for each of the three proposed models. These simulations allowed us to evaluate substrate positioning, coordination with key catalytic residues (including D100), and the dynamic behavior of the active site for each model. While enhanced sampling techniques could offer complementary insights, we believe that the current simulation strategy provides a strong basis for the mechanistic proposals presented.

*8- Discussion line 461. If there is a study already showing the importance of the first aspartic (and mutation?), this could be discussed.*

As we discuss in lines 457-460, mutation of D89 in *PfDPMS* (the residue equivalent to D100 in ArnC) has been shown to significantly reduce the catalytic activity of *PfDPMS*, while mutation of the second aspartate of the DXD motif (D91A) does not lead to as severe reduction of catalytic activity (Gandini et al)<sup>44</sup>. This argues that D89 likely operates as a catalytic base in *PfDPMS*, similar to ArnC. Moreover, the complete conservation of the DXD motif within the family of Pren-P GTs argues that our proposed mechanism is likely operative throughout the family. In the future, we aim to reconstitute ArnC in a catalytic assay, using purified donor substrate, and we intend to further characterize the contributions of these residue and adjacent residues to the catalytic activity of ArnC experimentally.

*Minor points:*

*- CAZy database currently has more than 130 families of transferases. Introduction line 84/85 should be updated and also the reference (I suggest to bring the GTs families description reference and also the most recent up-to-date CAZy database reference).*

We thank the reviewer for the suggestion. The text in line 84/85 has been updated and the suggested references for the GT family description and the CAZy database have been added.

*- Introduction line 99. “several structures”. I think that there are just a few, based on the works discussion. Moreover, to anticipate the information of the novelty brought from the apo structure could strengthen the findings.*

We have modified our statement in line 99 to more accurately describe the fact that structures for only two Pren-P glycosyltransferases were available prior to this work.

*- Fig 1. I suggest to bring the scheme in “a” in a separate figure. Also, the order of pannels a,b,c,d,e is bit confusing the way it is currently presented.*

We thank the reviewer for the suggestion. The scheme from Fig. 1a has been moved to the supplement as Supplementary Fig. 1. Also, the panels in Fig. 1 have been rearranged to have a more natural flow.

*- Fig. 6 legend. RMSDs of the protein were based in which atoms? Backbone?*

The RMSD values shown in Fig. 6 were calculated based on the backbone atoms of the protein. The legend has been revised to clarify this point.

*- Line 371-372. “ ... indicates that the presence of the lipid is not required for UDPA binding”. Could it be due to the sampling?*

We know from MST and cryo-EM experiments that UDP binds to the ArnC GT-A domain (Fig. 3B, Supplementary Figs. 11, 12C). The statement referenced above is supported by our atomistic simulations of the full UDP-L-Ara4FN substrate, which indicate that UDPA remains stably coordinated within the GT-A domain regardless of the presence or positioning of the undecaprenyl phosphate (UndP) lipid. In our simulations of states 2 and 3, UndP was modelled in distinct conformations within the catalytic site. In both cases, UDPA remained stably bound throughout the 500 ns simulations, suggesting that its binding is not significantly affected by the positioning of the lipid. Furthermore, in state 1—where UDPA was initially modelled in a different pose—we observed a spontaneous rearrangement during the simulation, with the sugar moiety adopting a conformation closely resembling that observed in states 2 and 3. This convergence across independently generated models supports the hypothesis that UDPA can bind and attain a stable, catalytically competent configuration regardless of the initial presence or conformation of UndP. We have updated lines 416–417 to include a reference to the experimental UDP-binding data to further clarify this point.

## References

1. Gandini, R., Reichenbach, T., Tan, T.-C. & Divne, C. Structural basis for dolichylphosphate mannose biosynthesis. *Nat. Commun.* **8**, 120 (2017).
2. McMullan, G. *et al.* Structure determination by cryoEM at 100 keV. *Proc. Natl. Acad. Sci.* **120**, e2312905120 (2023).
3. Karia, D. *et al.* Sub-3 Å resolution protein structure determination by single-particle cryo-EM at 100 keV. 2024.09.05.611417 Preprint at <https://doi.org/10.1101/2024.09.05.611417> (2024).
4. Cushing, V. I. *et al.* High-resolution cryo-EM of the human CDK-activating kinase for structure-based drug design. *Nat. Commun.* **15**, 2265 (2024).



## Response to Additional Reviewer's comments (Responses in Blue)

We thank the reviewers for their additional comments. We provide responses and describe the changes made below. The changes made to the manuscript for this revision are highlighted in cyan (previous revision changes are highlighted in yellow).

A new supplementary figure has been added in this revised version that shows time traces of different distances described in the manuscript over the course of the MD simulations (Supplementary Fig. 14).

### Reviewer #1 (Remarks to the Author):

*The authors of the manuscript addressed the questions raised during the reviewing process and reformulated key point of the manuscript. I appreciate the changes made, specially regarding structural comparisons and MD clarifications. I only have few observations:*

*- The phrase "During the final stages of preparation for this article, we became aware of a competing publication reporting a cryo-EM structure of ArnC from S. enterica in apo (2.75Å) and UDPbound conformations (3.8Å)" could be replaced by "A recent publication highlighted a cryo-EM structure of ArnC from S. enterica in apo (2.75Å) and UDPbound conformations (3.8Å)". If significant changes were observed (and clarifications made in this new version) I do not see the publication as competing but complementary.*

We thank reviewer #1 for the positive feedback and this excellent suggestion. We have changed the referenced sentence to the following: "During the final stages of preparation of this article, a complementary publication also reported cryo-EM structures of ArnC from S. enterica in apo (2.75Å) and UDP-bound conformations (3.8Å)"

*- Fig 6c: plots are too small*

The size of plots in Fig. 6c has been adjusted for better legibility.

*- Line 383 "Three repeats of unrestrained 500 ns MD simulations were performed for each system" Replace to: "three independent simulations of unrestrained 500 ns ...." (if 3 independent simulations started from different random velocities).*

The simulations were indeed started from different random velocities. The referenced sentence has been modified to emphasize that the three repeats of MD simulations were independent. It now reads: "Three independent unrestrained 500 ns MD simulations were performed for each system."

### Reviewer #2 (Remarks to the Author):

*The manuscript presents data and interpretation in two contexts:*

- 1. Structural and mechanistic analysis: Examination of three cryo-EM structures of ArnC from Salmonella enterica, supported by simulations and thermophoresis data, to propose a putative mechanism of action for ArnC.*
- 2. Methodological evaluation: Comparison of cryo-EM data collected with 200 kV and 300 kV instruments, leading to the conclusion that, for ArnC, the limiting factor is the biochemical quality of the purified protein rather than the microscope accelerating voltage.*

*I was not involved in the initial review. In my assessment, the authors have addressed the prior reviewers' comments adequately, and the manuscript in its current form is suitable for publication. However, to strengthen both the authors' credibility and the clarity of information provided to readers, I recommend the following revisions:*

We truly appreciate the highly detailed assessment of our observations and conclusions by the reviewer. This review highlighted two areas that required some changes in the presentation of our findings to the reader,

and as a result enhanced our ability to convey our findings.

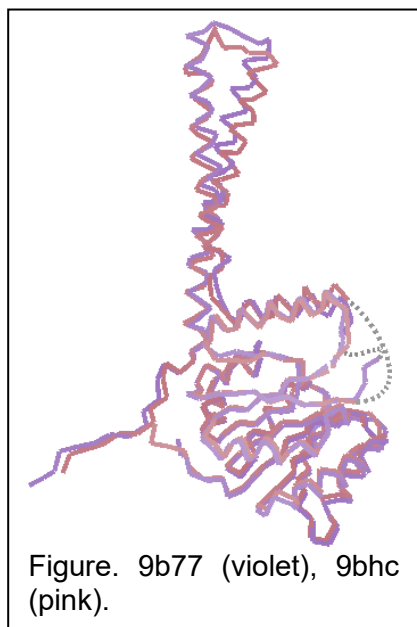


Figure. 9b77 (violet), 9bhc (pink).

**Ad [1] ArnC mechanism:** The section discussing ArnC structures from *Salmonella enterica* as reported by Patel et al. should be rewritten. The current text describes these structures as being significantly different from those presented in the manuscript under review, which is misleading. The difference is at the quaternary level: Patel et al. determined ArnC in C2 rather than C4 symmetry. At the monomer level, the functional unit of the enzyme according to Ashraf et al., the structures are nearly identical. Comparing C2 and C4 assemblies without clarification obscures the fact that the reported RMSD reflects symmetry differences, not meaningful changes in tertiary structure. Readers unfamiliar with the system could easily misinterpret this. For instance, my own comparison of chain A from 9B77 with chain A of 9BHC results in an RMSD of 1.3 Å, and with chain C of 9BHC, 1.6 Å. These structures are virtually identical (Figure). A more domain-specific comparison would likely reduce RMSD values even further.

We agree with the reviewer that the differences between our structures determined in lipid nanodiscs and the ones by Patel et al. determined in amphipol are mainly at the quaternary level, though differences at the top

of the transmembrane domain are also evident in the apo structures upon closer examination. We concur with the argument that presenting RMSD differences between the full tetramers partially obscures the fact that when monomers are aligned independently of their positions in the tetrameric structure, they are largely similar (though not identical). At the same time, we hope that the reviewer will recognize that presenting an RMSD between the isolated protomers alone would equally obscure the fact that the quaternary structure of the molecules presented in our paper and Patel et al. (9BHC) are in fact significantly different, especially in the apo state. This difference in the quaternary arrangement is significantly less when comparing the UDP-bound structures, with the UDP-bound structure presented by Patel et al. (9BHE) moving significantly closer to a square arrangement of TM helices compared to the rhomboid arrangement seen in the apo state. It is possible that the apo state is uniquely affected by embedding in amphipol, and that upon triggering of the conformational transition in the presence of nucleotide the quaternary arrangement of ArnC moves closer to the C4 configuration. We will also discuss this further in the next part about amphipol vs membrane environment.

To respond to this valid concern by the reviewer we have significantly re-arranged, and in part rewritten the paragraph comparing the structures, and included language to differentiate between the monomeric fold and quaternary structure (line 313-332).

*The work of Ashraf et al. is valuable not only for providing a higher-resolution UDP-bound ArnC structure but also precisely because it presents alternative quaternary arrangement. This difference deserves discussion. It is possible that Patel et al.'s use of amphipols restricted assembly to C2 symmetry, but it is also plausible that symmetry breaking is functionally relevant. Please check whether your simulations allowed for symmetry breakage in the tetramer and if so whether you observed such changes.*

We thank the reviewer for their assessment of our work and their suggestion to examine whether the change in symmetry observed in the amphipol structure may be functionally relevant. Based on our current understanding of the system, there is no indication that symmetry change is functionally relevant or physiological, especially because the transition away from C4 symmetry in the amphipol structure represents a compaction of the TM bundle (in the amphipol apo structure). In many cases, when a transition away from C4 symmetry is observed in tetrameric proteins, the observed state is typically a less-compact non-symmetrical (or lower symmetry) arrangement. Moreover, we did not observe symmetry breakage in our simulations. As shown in Supplementary Fig. 14, the distances and RMSDs calculated for each subunit display similar behavior across the trajectory. Although small differences are present, they do not indicate distinct behavior among the subunits. The simplest explanation of the observed differences between the apo



conformations is that amphipol (a polymer that only marginally mimics the physicochemical properties of a lipid bilayer) is imposing a conformation that is not observed in the context of a lipid bilayer and may represent a distorted arrangement due to the unique properties of the polymer. We further posit that the UDP-bound state in amphipol more closely resembling the square arrangement of the TM observed in lipid nanodiscs provides circumstantial evidence supporting an effect of the polymer on the structure of ArnC. Based on the structures and our current understanding of the mechanism of ArnC, this change in the conformation of the TM domain in amphipol can only be attributed to direct effects of the polymer environment.

We believe that the comparison of the structures as presented in this revised version is now well-balanced and appropriately conveys the observed differences between the structures.

*In addition, the argument about different JM helix positions in membrane is not compelling. The authors themselves note “better” ArnC behavior in a non-physiological lipid environment in their purifications, and helices rearrangements are a common feature of membrane-mediated functions.*

*I strongly recommend reframing the discussion to acknowledge that quaternary structure differences may reflect an intrinsic aspect of ArnC’s mechanism. This point should be clearly addressed when citing Patel et al. and expanded upon in the discussion.*

In regard with this statement by the reviewer, we must disagree, because of two considerations:

1) The lipid nanodiscs in our study do not have a physiological lipid composition due to experimental constraints, but this does not negate the fact that the lipid nanodisc provides a lipid bilayer environment that closely approximates the physicochemical properties of the native membrane environment, whereas a polymer (like amphipol) only marginally does so by providing a substitute (non-membrane) environment that stabilizes the protein in solution. The physicochemical properties of the two are completely distinct, and in addition amphipol is not known to preserve annular lipids associated with transmembrane proteins like other polymers (e.g. cubipol, SMA, etc.).

2) Helix rearrangements, defined in the broad sense, are indeed a common feature of membrane-mediated functions, but that is typical for helices already embedded within the membrane (i.e. transmembrane helices). Amphipathic helices on the other hand are known to intercalate into membranes in some cases, but not to traverse the plane of the membrane. Juxtamembrane helices, in particular, are known to associate with the membrane to stabilize the positioning of protein domains (i.e. cytoplasmic domain). Based on our work and that of others (Ardiccioni et al. 2016, Gandini et al. 2017), the two juxtamembrane helices (per protomer) in this family of proteins (ArnC, GtrB, DPMS) are positioned perpendicular to the membrane adjacent or partially embedded into the lipid bilayer to channel the acceptor substrate into the active site of the GT domain. The amphipol apo structure has the juxtamembrane helices of at least one protomer positioned at the same height as the halfway point of the adjacent TM helice(s), meaning approximately halfway into the lipid bilayer. There is no indication that this is a physiological arrangement, and certainly the fact that it is observed in a polymer environment is not a strong argument towards its physiological relevance.

Despite our disagreement, we concur that the statement included in text (“In addition, in the amphipol structures, the juxtamembrane helices have been modeled at different vertical positions in adjacent protomers. Given that the JM helices are located directly adjacent or partially embedded in membrane, this arrangement of the JM helices is highly unlikely to be accurate in a membrane environment.”) may not be easily interpretable by some readers without the additional context provided above. Thus, we have removed this statement from our comparison, and we will rely on the scientific community to further weigh in on the differences observed between the structures.

**Ad [2] 300 kV vs. 200 kV analysis:** *I concur with the previous reviewer: this analysis should not be included. It detracts from the main discussion on ArnC, and more importantly, the data were collected in a way that does not support the conclusions the authors attempt to draw. Two experimental factors likely compromised the 300 kV dataset, making it impossible to determine whether it could have yielded superior results. First, in Supplementary Table 1, the reported dose rate is 19.9685 (BTW, too many significant digits). It is unclear whether CDS mode was used, but even without CDS, Sun et al. [J Struct Biol. 2021;213(3):107745]*

*showed that coincidence loss occurs at such high dose rates for K3. A rate of  $\sim 20 \text{ e}^-/\text{px/s}$  is too high for optimal data collection, leading to lower DQE and potentially reduced resolution. Second, the dose fractionation strategy was problematic. The 300 kV dataset had a total dose of  $57.42 \text{ e}^-/\text{\AA}^2$  distributed over 40 frames, meaning each frame carried a very high per-frame dose ( $\sim 1.5 \text{ e}^-/\text{\AA}^2$  that corresponds to  $\sim 6 \text{ MGy}$ ). Consequently, even the earliest frames already suffered significant radiation damage, and upweighting could not recover the lost signal. Standard practice is to use roughly half that per-frame dose, if not lower. In contrast, the Talos/K2 dataset was collected with well- optimized parameters, balancing radiation damage and signal recovery effectively.*

We concur that a lower dose rate would have been optimal for ArnC collection on the Krios. The work of Sun et al. is clear that this dataset could have been collected better with CDS mode and at lower dose, but this is not because the lower dose rate would have given more similar detector performance to the performance obtained on a K2. In fact, the conditions we used were comparable, and a low dose rate with CDS mode would be expected to increase DQE beyond what is possible with a K2 camera.

First, in the absence of correlated double-sampling, the DQE as a fraction of physical Nyquist frequency is comparable between K2 and K3 cameras. D. Morado has suggested that the physical pixel design is actually the same between the two cameras—whether or not this is true, the vendor-reported DQE curves are virtually identical (except when CDS mode is used). At the fraction of physical Nyquist used here, the 300kV dataset's DQE should be  $\sim 62\%$  and the 200kV dataset's DQE should be  $\sim 45\%$  because the finer sampling of the 300kV dataset dominates over Nyquist effects.

Second, when correlated double-sampling is not used, the utilized dose does fall within the linear response rate with minimal coincidence loss. The coincidence loss of a K3 at  $20 \text{ e}^-/\text{px/s}$  is comparable to the coincidence loss of a K2 at  $6 \text{ e}^-/\text{px/s}$ . The K2 incident electron recovery rate for  $6 \text{ e}^-/\text{px/s}$  is around 80% (Li et al., Nat Meth 2013) and the K3 incident electron recovery rate at  $20 \text{ e}^-/\text{px/s}$  is around 75% according to Sun et al. or around 80% according to vendor documents.

FIGURE REDACTED DUE TO COPYRIGHT



FIGURE REDACTED.

In summary, using a more optimized dose rate as suggested by the reviewer could have made the structure even better, but does not compromise the comparison across voltages. To address the importance of the dose rate raised by the reviewer, we appended a sentence after “The most reasonable explanations for this difference include: effect of higher magnification ( $0.646 \text{\AA}/\text{px}$  vs.  $1.038 \text{\AA}/\text{px}$ ), effect of detector performance

(Gatan K3 vs. Gatan K2 direct electron detector), effect of microscope model and voltage (300kV Titan Krios vs. 200 kV Talos Arctica), effect of the experimenter/site, or other uncontrolled variation between experiments (e.g. intrabatch variation in grid quality, environmental interference, day-to-day variation in instrument performance etc.),” to read, “It is possible that an improvement at 300 kV or from another factor was masked by a countervailing variable: for instance, perhaps the higher dose rate used at 300 kV could have worsened data quality to approximately the same extent that the higher magnification improved data quality, resulting in similar-resolution structures.”

Finally, the conclusion that in this case voltage was not dispositive would be logically supported even if a non-voltage factor explained the similarity. Dose rate or motion could be described as such non-voltage factors, though we believe that based on the comparisons provided (above and below), they are unlikely to be the limiting factor instead of sample limitations. We hope that in subsequent studies outside the scope of this manuscript we will be able to focus exclusively on isolating and analyzing resolution-limiting factors and follow up on this point.

*Finally, particle movement is an additional, poorly controlled variable when comparing instruments. Even with the same grid, movement patterns can differ between microscopes due to both sample-specific and instrument-specific factors. While software can partially correct these effects, its ability is limited. A proper comparison would require repeating the 300 kV experiment across multiple 300 kV microscopes.*

*In sum, while I appreciate the authors’ effort, the combination of excessive dose and coarse dose slicing in the Krios dataset prevents their conclusions from being supported.*

Motion blur is a resolution-limiting factor and can be a major factor in some experiments. Because this blur can be quantified, we are able to rule this out as the key resolution-limiting factor in this particular experiment. The total motion was 2.89Å (median value; 90% of utilized micrographs were between 2.03Å and 17.7Å) for the 300kV dataset and 6.75Å (3.89–22.34Å) for the 200kV dataset. Given the higher framerate used at 300kV, this corresponds to 0.07Å (0.05–0.44Å) of average motion per frame at 300kV and 0.19Å (0.11–0.64Å) of motion per frame at 200kV. We would be Shannon-Nyquist limited to a resolution of, for example, 2.5Å if the motion were more than 1.25Å in the first frame. Although the first-frame motion is higher to estimate than the whole-movie motion, it would have to be at least 17-fold higher in the first frame than the average frame to cause enough blur in this dataset that it would be Nyquist-limited.

The reviewer was right to point this out as a potential limiting factor. The information readers needed to know that in this case it is not was not present in the un-revised text, so we added an additional line to Table 1 for “median per-frame motion (Å).”

*Minor point: SI units require a space between the number and the unit (e.g., 10 mM, not 10mM). This convention should be applied consistently. Similarly, the number of significant digits should be reported reasonably and consistently.*

We thank the reviewer for pointing out the inconsistencies in how units were reported in the previous version. This has now been corrected throughout the manuscript.

### **Reviewer #3 (Remarks to the Author):**

*The authors have supplemented experiments and provided comprehensive responses to Reviewer #1’s concerns.*

*The reviewer #2 mainly raised two questions:*

*Q1: The reviewer #2 emphasized that establishing a purified ArnC enzymatic activity assay would be necessary to fully elucidate the mechanism.*

*In the response, the authors pointed out that no commercial substrate is currently available, making it impossible to perform in vitro enzymatic activity assays. However, in fact, the substrate UDP-L-Ara4FN can*

*be obtained using the method described in Reference 27 (J. Biol. Chem. 280, 14154–14167) cited by the authors. Given that the substrate is radiolabeled with C14, the authors can supplement bacterial genetic mutations in vivo assays to validate the complete mechanistic understanding of ArnC alternatively.*

We thank the reviewer for this suggestion. We are aware of the published method to isolate the full donor substrate and we intend to use this exact method in our future studies to isolate the donor substrate and use it in both functional assays and structural analyses. Because no commercial substrate exists, and we currently don't have this method implemented in our laboratory, the suggested experiments are beyond the scope of the current manuscript. We hope that the reviewer will accept our commitment to address this suggestion in future studies of ArnC.

*Q2: Another major concern raised by Reviewer #2 was the lack of a ternary complex structure (donor-acceptor-enzyme), which may limit the mechanistic understanding of ArnC catalysis and its potential for guiding drug design.*

*In this regard, the authors employed molecular dynamics (MD) simulations to validate donor substrate binding and dynamics. Combined with their previously resolved structural data, the authors have addressed question 2 from Reviewer #2.*

We appreciate the concurrence by the reviewer that we have addressed the previous question 2 by Reviewer #2 (previous round) in a satisfactory manner.