nature portfolio

Peer Review File

Mechanistic conformational and substrate selectivity profiles emerging in the evolution of enzymes via parallel trajectories.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The manuscript reports the catalytic activity of a series of HsKYNase with different mutations. As each mutant exhibits different kinetics and bindings towards substrate KYN and OH-KYN, authors were able to conclude that the catalytic activities is highly dependent on the mutations, and further, genetically different evolutionary trajectories. Regarding on the HDX-MS experiments and results, here are some comments that needs authors' attention.

1. While performing the HDX experiments, authors mixed the substrate KYN and OH-KYN into the deuterated buffer. In other words, the H/D exchange reaction is competing with the catalytic reaction. Authors should comment on the differences in kinetics profile between H/D exchange reaction and catalytic reaction so that the conclusion is a true representation of catalytic dynamics observed via HDX.

2. In the experimental section, authors should revisit the description of LC conditions. If my understanding is correct, mobile phase A should be water with 0.1% formic acid whereas mobile phase B should be acetonitrile with 0.1% formic acid. The gradient should be a binary gradient rather than isocratic.

Reviewer #2 (Remarks to the Author):

In this very interesting paper, the authors follow on their previous studies on evolving specialist and generalist kynureninase enzymes (concerning two substrates, KYN and OH-KYN). This manuscript describes and extensive and intricate directed evolution campaign the resulted in an enzyme that, despite having the "canonical" KYN-specificity motif W102-T333, is a true generalist with high activity for both substrates. Using enzyme kinetics, structural models, and HDX-MS, they show that different evolutionary trajectories can lead to genetically distinct enzymes that rely on different mutations to reach a similar new activity by differently affecting protein dynamics and the catalytic mechanism. The study provides interesting insights on the relationship between conformational dynamics, substrate specificity, and catalytic activity of evolved enzymes and further illustrates the complexity of enzyme function.

To the best of my understanding, the manuscript is technically sound and sufficiently described so as to be reproduced. The findings will be of significant interest to the field of enzymology and enzyme evolution (both applied and basic science). The manuscript is clearly written. I have only minor comments.

1. Lines 108-114. In the first sentence I suggest the authors emphasize that HsKYNase_66 is a specialist by stating how selective it is for KYN over OH-KYN. Then in the second sentence make it very clear that other enzymes were created in the same or parallel experiments – not sure which - (HsKYNase_46 and BF-HsKYNase) that were generalists. Starting the second sentence with "in this process" makes it sound like the process for making HsKYNase_66 and the sentence is about HsKYNase_66, which is confusing until you only learn is not the case at the very end of the sentence with the parenthetical listing of the names of the two generalists. It took me a while to figure this out.
2. Main Table 2. The heading of the first column should be "kinetic parameter" not "enzyme." Also, missing space between units and number in the row for k1.
3. The supplementary tables provided as Excel sheets are not very large. I suggest they would be better incorporated in the supplementary document as normal tables, perhaps oriented sideways, if needed. The exception is the .csv file, which makes sense to keep in that format. Just a suggestion for ease of access.

Reviewer #3 (Remarks to the Author):

The authors report on the evolution of a KYNase generalist enzyme (HsKYNase_93D9) displaying equally high catalytic activities for both KYN and OH-KYN. They use pre-steady state kinetics, molecular docking and HDX-MS to deepen on the question that parallel evolutionary pathways can lead to enzyme spices with similar catalytic activities achieved by different conformational dynamics during catalysis by bypassing other sequence constraints.
In my opinion the work is of significance for the scientific community to the field and related fields since it provides rational understanding on the evolutionary trajectories of a system with a key engineering interest.
The overall conclusions are well supported and specific claims and methodology sound and well detailed. Nonetheless, the modelling part and related analysis included to support and rationalise their experimental observations is rather thin, regarding the expected standards of the journal. I would support publication after expanding on the modelling part of the article.
Docking is a valuable tool but has to be used with caution, specially with loopy proteins and structures build by homology modelling, where longer MD simulations (at least 50-100 ns and replicates) are mandatory to relay on the stability of the mentioned interactions.
In Figure 3 the authors show striking differences on the conformation of the substrate (KYN or OH-KYN) within the pocket of the different studied variants (HsKYNase, HsKYNase_93D9 and HsKYNase_66). From the table in Figure 1e it can be seen that the three variants are indeed catalytically active (as reported along the article) towards both KYN and OH-KYN, with the largest differences on KM, however in the text line 358 it is mentioned "...docking of OH-KYN to HsKYNase_66 suggested a totally unproductive mode of binding...". The authors should extend the MD simulations and report on the stability of the key catalytic distances, as well as the key distances identified for substrate stability in the binding pocket, otherwise they are discussing conformational dynamics on a static picture of the complex.

Reviewer #1 (Remarks to the Author):

The manuscript reports the catalytic activity of a series of HsKYNase with different mutations. As each mutant exhibits different kinetics and bindings towards substrate KYN and OH-KYN, authors were able to conclude that the catalytic activities is highly dependent on the mutations, and further, genetically different evolutionary trajectories. Regarding on the HDX-MS experiments and results, here are some comments that needs authors' attention.

1. While performing the HDX experiments, authors mixed the substrate KYN and OH-KYN into the deuterated buffer. In other words, the H/D exchange reaction is competing with the catalytic reaction. Authors should comment on the differences in kinetics profile between H/D exchange reaction and catalytic reaction so that the conclusion is a true representation of catalytic dynamics observed via HDX.

We thank the Reviewer very much for this comment. Within the context of our experimental setup the HDX is monitoring both the substrate binding and the catalytic reaction under steady-state turnover conditions. We performed steady-state kinetics of HsKYNase_93D9 against both substrates KYN & OH-KYN in deuterium oxide (D_2O) in order to obtain estimates for the k_{cat} and K_M parameters under HDX conditions. We show the values of those steady-state kinetic parameters (HsKYNase93_D9/KYN and HsKYNase93_D9/OH-KYN) below for the Reviewer's reference and also, we added them in the Supplemental Information as a new Supplementary Table 10.

Based on the estimated parameters in D_2O below, the total enzyme and substrate concentration that were used in the reaction (1.6 µM enzyme and 3 mM of each substrate as described in the Methods section) and the fast association constants for substrate binding (leading to fast equilibrium), HsKYNase_93D9 was actively turning over during the first 10 minutes of the reactions before either substrate would become limiting.

HsKYNase_93D9/KYN in D_2O
k_{cat}: 0.16 ± 0.012 s^{-1}
K_M: 250 ± 49 µM
k_{cat}/K_M: 640±134 $M^{-1}s^{-1}$

HsKYNase_93D9/OH-KYN in D_2O
k_{cat}: 0.135 ± 0.0065 s^{-1}
K_M: 50 ± 16 µM
k_{cat}/K_M: (2.7±0.45) × 10^3 $M^{-1}s^{-1}$

2. In the experimental section, authors should revisit the description of LC conditions. If my understanding is correct, mobile phase A should be water with 0.1% formic acid whereas mobile phase B should be acetonitrile with 0.1% formic acid. The gradient should be a binary gradient rather than isocratic.

We thank the Reviewer for catching this error. Corrections have been made.

Reviewer #2 (Remarks to the Author):

In this very interesting paper, the authors follow on their previous studies on evolving specialist and generalist kynureninase enzymes (concerning two substrates, KYN and OH-KYN). This manuscript describes and extensive and intricate directed evolution campaign the resulted in an enzyme that, despite having the "canonical" KYN-specificity motif W102-T333, is a true generalist with high activity for both substrates. Using enzyme kinetics, structural models, and HDX-MS, they show that different evolutionary trajectories can lead to genetically distinct enzymes that rely on different mutations to reach a similar new activity by differently affecting protein dynamics and the catalytic mechanism. The study provides interesting insights on the relationship between conformational dynamics, substrate specificity, and catalytic activity of evolved enzymes and further illustrates the complexity of enzyme function.

To the best of my understanding, the manuscript is technically sound and sufficiently described so as to be reproduced. The findings will be of significant interest to the field of enzymology and enzyme evolution (both applied and basic science). The manuscript is clearly written. I have only minor comments.

We would like to thank the Reviewer for his/her appreciation of our work and for recognizing the important insights it provides on the complex relationship between conformational dynamics, substrate specificity, and catalytic activity.

1. Lines 108-114. In the first sentence I suggest the authors emphasize that HsKYNase_66 is a specialist by stating how selective it is for KYN over OH-KYN. Then in the second sentence make it very clear that other enzymes were created in the same or parallel experiments – not sure which - (HsKYNase_46 and BF-HsKYNase) that were generalists. Starting the second sentence with "in this process" makes it sound like the process for making HsKYNase_66 and the sentence is about HsKYNase_66, which is confusing until you only learn is not the case at the very end of the sentence with the parenthetical listing of the names of the two generalists. It took me a while to figure this out.

We apologize to the Reviewer for this confusion. Indeed, the variant HsKYNase_46 was a OH-KYN-specialist and an evolutionary dead-end in terms of evolution of high KYN catalytic activity. BF-HsKYNase was isolated from a separate directed evolution campaign (and a different trajectory) that focused only on the mutagenesis and randomization of highly flexible (high B-factors) loops distal from the active site.
We edited the text accordingly to further clarify the point above. As proposed by the Reviewer, we explicitly stated the KYN specificity of the specialist HsKYNase_66 in the first sentence by adding in parenthesis the $(k_{cat}/K_M)_{KYN}/(k_{cat}/K_M)_{OH\text{-}KYN}$ ~160. In addition, in the same paragraph we clarified the features of HsKYNase_46 and BF-HsKYNase ("*Along these lines….similar catalytic properties as HsKYNase_46*").

2. Main Table 2. The heading of the first column should be “kinetic parameter” not “enzyme.” Also, missing space between units and number in the row for k1.

We thank the Reviewer for catching this typo and we have corrected it accordingly. We added the term “*Rate constant & steady-state kinetic parameter*” as the main heading of this column to cover both the estimated rate constants (k_1, k_2, k_3 etc) and the steady-stat kinetic parameters (k_{cat} and k_{cat}/K_M) that are shown in this column.

3. The supplementary tables provided as Excel sheets are not very large. I suggest they would be better incorporated in the supplementary document as normal tables, perhaps oriented sideways, if needed. The exception is the .csv file, which makes sense to keep in that format. Just a suggestion for ease of access.

We thank the Reviewer for this suggestion, and we also believe that such incorporation of the Tables into the main body of the Supplemental Information will indeed facilitate the reading of our manuscript. We transferred the Tables accordingly except for the last Table (Supplementary Table 12) that reports on the raw HDX-MS data (in agreement with Reviewer’s view).

Reviewer #3 (Remarks to the Author):

The authors report on the evolution of a KYNase generalist enzyme (HsKYNase_93D9) displaying equally high catalytic activities for both KYN and OH-KYN. They use pre-steady state kinetics, molecular docking and HDX-MS to deepen on the question that parallel evolutionary pathways can lead to enzyme spices with similar catalytic activities achieved by different conformational dynamics during catalysis by bypassing other sequence constraints.
In my opinion the work is of significance for the scientific community to the field and related fields since it provides rational understanding on the evolutionary trajectories of a system with a key engineering interest. The overall conclusions are well supported and specific claims and methodology sound and well detailed.

We are thankful to the Reviewer for his/her favorable acceptance of our work and for recognizing its value for the broader space of enzyme engineering.

Nonetheless, the modelling part and related analysis included to support and rationalise their experimental observations is rather thin, regarding the expected standards of the journal. I would support publication after expanding on the modelling part of the article.
Docking is a valuable tool but has to be used with caution, specially with loopy proteins and structures build by homology modelling, where longer MD simulations (at least 50-100 ns and replicates) are mandatory to relay on the stability of the mentioned interactions.
In Figure 3 the authors show striking differences on the conformation of the substrate (KYN or OH-KYN) within the pocket of the different studied variants (HsKYNase, HsKYNase_93D9 and HsKYNase_66). From the table in Figure 1e it can be seen that the three variants are indeed catalytically active (as reported along the article) towards both KYN and OH-KYN, with the

largest differences on KM, however in the text line 358 it is mentioned "...docking of OH-KYN to HsKYNase_66 suggested a totally unproductive mode of binding...". The authors should extend the MD simulations and report on the stability of the key catalytic distances, as well as the key distances identified for substrate stability in the binding pocket, otherwise they are discussing conformational dynamics on a static picture of the complex.

We appreciate the Reviewer's feedback on the modeling part. In response, we performed additional simulations and revised the molecular modeling as requested by the reviewer.

As the reviewer requested, we performed additional, longer MD-sim (50-nanoseconds) for all 6 complexes (3 enzymes x 2 substrates) and report this data in the revised version. In the manuscript we compared the energy minimized structures at t=0 against the time-averaged structures (instead of just 1 or 2 random frames from the trajectory) to obtain insights as to how key interactions between the substrates and active site residues dynamically progressed as well as to capture conformational changes of critical active site loops. This analysis is shown in the revised main Figure 3 and in Supplementary Figure 7 where we show RMSD values of representative complexes' trajectories. Our docking and MD simulations analysis provide relevant insights regarding the first step of substrate binding before the transaldimination reaction. The data are discussed in the main text in the respective modeling and MD-simulation section as well as in the discussion.

We would like to reiterate that the overarching goal of our work was to experimentally explore the impact of distinct evolutionary trajectories that achieve the same threshold of catalytic activity for a preferred substrate, yet do so with markedly different consequences regarding substrate selectivity, rate-determining step of the reaction and protein dynamics. We hope that the reviewer and Editor appreciate the very significant effort and experimental data already presented in the paper.

Performing further in-depth *in-silico* analyses in order to precisely define the specific chemical steps of the catalytic cycle for the 3 enzymes (HsKYNase, HsKYNase_93D9 & HsKYNase_66) with the two substrates (KYN and OH-KYN) is clearly outside the scope of the present study. Such work would require, at a minimum high-resolution crystal structures of the KYNase enzymes in complex with reaction intermediates coupled to extensive quantum mechanical/molecular mechanics (QM/MM) computations.

Additionally, it is important to keep in mind that the catalytic mechanism of PLP-dependent enzymes and of Kynureninases in particular consists of multiple steps including:
1) substrate binding; 2) trans-aldimination reaction (formation of Schiff base bond between the substrate and the co-factor pyridoxal-5'-phosphate, the so-called external aldimine); 3) hydrolysis reaction; & 4) release of the 2 products. Please see the following relevant references: i) Lima et al., "*Crystal structure of the H. sapiens Kynureninase-3-Hydroxyhippuric acid inhibitor…specificity*" J. Med. Chem. 2009, ii) Phillips RS, "*Structure and mechanism of kynureniase" Archives of Biochem. And Biophys. 2014*, iii) Karamitros CS et al., "*Leveraging intrinsic flexibility to engineer enhanced enzyme catalytic activity*" PNAS 2022)). Elucidating

interactions that are relevant for each of these steps would require an immense effort which, again is well beyond the scope of the present work.

REVIEWER COMMENTS

Reviewer #3 (Remarks to the Author):

I would like to recognize again the potential of the work performed by Karamitros et al. I understand that the overarching goal of their study was "experimentally"-driven, nonetheless the computational part of their study represents almost 1/3 of the results, and this is the reason why I would also like to reiterate the importance and the thoroughness of this part of the work.
The authors have now extended their simulationds to 50 ns as requested and show through RMSD that the complexes are stable within the time-scale used. However, it is not clear if they have performed replicates or not since in the methos they say "Three different simulations were performed for each spices and each simulation was initiated using different random seed number", but in Figure S7 there seem to be just one replicate for each system. This needs clarification somewhere.
I agree with the authors that a mechanistic analysis would be out of the scope of the present study and would require extensive QM/MM calculations, and none of my comments was actually suggesting this kind of request.
However, I would like to reiterate my request on a dynamical picture of the key interactions they highlight in Figure 3 and discuss along the manuscript. Although the comparison between the starting structures (t=0) and the time-averaged structures are an improvement, they do not mention how a time-averaged structure is obtained. Usually this kind of comparison should be made with the most representative structure along the simulation, namely the most populated cluster, since an average structure along the simulation could lead to distorted structures. The authors should define how this time-averaged structure is obtained. Nonetheless my request was simpler, as to show the dynamics of the interactions highlighted in Figure 3, just a table showing the % of the simulation time each interaction is maintained would be needed. It is a quite simple analysis to be done on the already existing MD simulations which I think is necessary instead of relaying in the comparison of two static pictures that do not give an idea of the stability of the discussed key interactions.

Reviewer #4 (Remarks to the Author):

See attached file.

[Editorial Note: This file is displayed over the next two pages]

First, I would like to thank the authors for submitting their study " Distinct mechanistic, conformational and substrate selectivity profiles emerging in the evolution of high catalytic activity enzymes via parallel trajectories" for consideration of publication by nature communications. This study follows up on earlier studies published by the authors (10.1073/pnas.2118979119, 10.1021/acschembio.0c00676, 10.1038/s41929-022-00856-6) and uses almost similar toolkit to address aspects of enzyme evolution in terms of their catalysis and dynamics. In this article, as per the author's claim, they generated a human kynureninase variant (HsKYNase_93D9) that exhibits catalytic activity for KYN comparable to that of HsKYNase_66 (a variant from one of their previous studies). They further show that it is a "generalist" enzyme that accepts 3'-hydroxykynurenine (OH-KYN) with the same proficiency. They used a combination of pre-steady state kinetics, HDX-MS and

1. It would be desirable for the authors to present data on the potential differences in the thermodynamic stability across the various enzymes using either chemical or heat mediated unfolding experiments (with specific focus on HsKYNase, HsKYNase_93D9 & HsKYNase_66). Interpretation of kinetics and dynamics is intricately related to the thermodynamic stability of a protein.
2. A selwyn's test assessment should be shown to understand that the claim for activity loss is not because of potential enzyme inactivation.
3. How are the authors assessing the active fraction of the enzyme (potential assessment of active fraction of enzyme using a tight-binding inhibitor would be helpful).
4. Supplemnetary table 2, 3, 5 and 7, I am surprised to see the values in the table listed without errors. The dispersion of the estimate about the mean is as important an estimate as the reported estimate itself. Errors will help me understand how reliable or conclusive comparisons are across several kcat/Km values. Also, it might benefit the readers to know how many technical and biological replicates resulted in the value reported.
5. For the sake of consistency and to avoid confusion, all Km values should preferably be reported in M (this is because all kcat/Km values have been reported as M-1 sec-1). In some tables, the Km values are reported in M, in others in mM and in yet others, they are uM.
6. I am bit confused about the values reported in supplementary table 7 and 8 for HsKYNase_93D9 with KYN. The steady state estimate indicates that the Km for the substrate is 0.000085 M. However, a crude calculation of Km from the numbers reported in table 8 shows a Km of 20 M for the substrate!!! Further, the kcat in suppl. Table 8 is 0.79 s-1 (with very narrow dispersion about the mean) while that in suppl. Table 7 is 1.45 s-1 (no dispersion provided). Which of these scattering of numbers are the authors using to compute their kcat/Km values that provides them crucial insights into evolutionary interpretations? Depending on which Km (from their steady state estimates or their pre-steady state dynamic simulations), the kcat/km will be approximately ~10^6 fold different placing the Fold-HsKYNase number anywhere you wish and interpreting the evolution how you would like to subjectively package it.
7. On similar lines, for data on HsKYNase_66 in supplemntary Table 9, the Km computation from the global fit analysis gives an estimate of 4.1 M. However, it is 0.000027 M as reported in the reference 10.1038/s41929-022-00856-6 by the same authors.
8. How do the authors explain the sudden drop of Km with D2O (supplementary table 10) of HsKYNase_93D9 against KYN to 250 uM those reported in point number 6 above.
9. Compounding the problem with respect to units and values (and demonstrating how casual the authors have been in showcasing this numbers without aking care to proofread them

appropriately), on page 6, line 267 of the main text, the authors state that the second order rate constant for binding for HsKYNase_66 is 5.9 µM-1 s -1. This is shown as 5.9 M-1 s -1 in suppl. Table 9 (once again a 10^6 difference!!!). On similar lines, the association rate constant for HsKYNase_93D9 is shown as 0.24 µM-1 s -1 (page 6, line 267) while the suplemnetary table 8 once again shows this as 0.24 M-1 s -1 ?? With the combined expertise of all the authors (some of whom are instrumental in laying the foundation of modern enzymology), such kind of errors is not only surprising but distressing.

10. In response to reviewer 3's comments, in their rebuttal letter the authors state that "Performing further in-depth in-silico analyses in order to precisely define the specific chemical steps of the catalytic cycle for the 3 enzymes (HsKYNase, HsKYNase_93D9 & HsKYNase_66) with the two substrates (KYN and OH-KYN) is clearly outside the scope of the present study". I disagree with this assertion in the light of such disparity of kinetic estimates reported as part of their current study and previous studies. This calls for the authors to look at all their numbers (reported in the present study and studies previously published) carefully in order to tidy up the interpretation.

Reviewer #3 (Remarks to the Author):

I would like to recognize again the potential of the work performed by Karamitros et al. I understand that the overarching goal of their study was "experimentally"-driven, nonetheless the computational part of their study represents almost 1/3 of the results, and this is the reason why I would also like to reiterate the importance and the thoroughness of this part of the work.
The authors have now extended their simulations to 50 ns as requested and show through RMSD that the complexes are stable within the time-scale used. However, it is not clear if they have performed replicates or not since in the methos they say "Three different simulations were performed for each species and each simulation was initiated using different random seed number", but in Figure S7 there seem to be just one replicate for each system. This needs clarification somewhere.
I agree with the authors that a mechanistic analysis would be out of the scope of the present study and would require extensive QM/MM calculations, and none of my comments was actually suggesting this kind of request.
However, I would like to reiterate my request on a dynamical picture of the key interactions they highlight in Figure 3 and discuss along the manuscript. Although the comparison between the starting structures (t=0) and the time-averaged structures are an improvement, they do not mention how a time-averaged structure is obtained. Usually, this kind of comparison should be made with the most representative structure along the simulation, namely the most populated cluster, since an average structure along the simulation could lead to distorted structures. The authors should define how this time-averaged structure is obtained. Nonetheless my request was simpler, as to show the dynamics of the interactions highlighted in Figure 3, just a table showing the % of the simulation time each interaction is maintained would be needed. It is a quite simple analysis to be done on the already existing MD simulations which I think is necessary instead of relaying in the comparison of two static pictures that do not give an idea of the stability of the discussed key interactions.

We would like to thank the Reviewer for his/her recognition of our work as well as the favorable acceptance of the first round of our revisions. Indeed, we performed three replicates of 50-nanosecond MD simulations for each complex that was studied in our present work. All three replicates yielded similar equilibration & stability profiles within the 50 nanoseconds (C_a RMSD), and this is why we showed representative RMSD plots in the first revision. Please note that raw data of all the replicates have been uploaded on the Zenodo database and will be published upon acceptance of our article. To further address the reviewer's request, we have included the RMSD plots for all the three replicates of each complex in the Supplementary Information (New SI Figures 7, 8 and 9).

Regarding the time-averaged structure, this was calculated by superposing each snapshot on the initial energy-minimized structure at time 0, and then averaging the resulting atom positions. Subsequently, the time-averaged structure was energy-minimized and structurally compared against the initial structure (we edited the main text accordingly).

We would like to thank the Reviewer for clarifying his/her suggestion regarding the inclusion of a Table showing the % of simulation time that each interaction was maintained. We did this analysis for the MD simulation trajectory that is shown in main Figure 3, and we present these data as a new main Table 3 now. Of note, the data of Table 3 correspond to the structural models shown in the main Figure 3 and

the respective RMSD plots of the trajectories are shown in the SI Figures 7a & 7d, SI Figures 8a & 8d and SI Figures 9a & 9d for HsKYNase, HsKYNase_93D9 and HsKYNase_66 respectively. We added this clarification in the main text as well as in the Figure/Table legends.

We adjusted the main text to accommodate the additional information that is provided in Table 3.

We hope that with these additional analyses, we have satisfied the Reviewer's requests and we would like to thank him/her for offering us the opportunity to further improve our manuscript.

Reviewer #4 (Remarks to the Author):

First, I would like to thank the authors for submitting their study "Distinct mechanistic, conformational and substrate selectivity profiles emerging in the evolution of high catalytic activity enzymes via parallel trajectories" for consideration of publication by nature communications. This study follows up on earlier studies published by the authors (10.1073/pnas.2118979119, 10.1021/acschembio.0c00676, 10.1038/s41929-022-00856-6) and uses almost similar toolkit to address aspects of enzyme evolution in terms of their catalysis and dynamics. In this article, as per the author's claim, they generated a human kynureninase variant (HsKYNase_93D9) that exhibits catalytic activity for KYN comparable to that of HsKYNase_66 (a variant from one of their previous studies). They further show that it is a "generalist" enzyme that accepts 3'-hydroxykynurenine (OH-KYN) with the same proficiency. They used a combination of pre-steady state kinetics, HDX-MS and

1. It would be desirable for the authors to present data on the potential differences in the thermodynamic stability across the various enzymes using either chemical or heat mediated unfolding experiments (with specific focus on HsKYNase, HsKYNase_93D9 & HsKYNase_66). Interpretation of kinetics and dynamics is intricately related to the thermodynamic stability of a protein.

We have determined the Tm of HsKYNase, HsKYNase_66 and HsKYNase_93D9 by differential scanning fluorimetry (DSF). All three enzymes are thermodynamically very stable and have comparable Tm values. We inserted the following in the main text: "*We note that HsKYNase, HsKYNase_93D9 and HsKYNase_66 are all thermostable with their respective Tm values measured by differential scanning fluorometry being ~73, ~76 and ~82 ℃, respectively*."

2. A selwyn's test assessment should be shown to understand that the claim for activity loss is not because of potential enzyme inactivation.

It is not entirely clear to us what activity loss the Reviewer is referring to. We did not observe any activity loss of the different enzyme variants that we characterized in this study. If the Reviewer is referring to the activity loss of the HsKYNase variants upon the incorporation of the H102W/N333T motif, this was not due to any inactivation events but attributed to the two amino acid substitutions which require a very well fine-tuned network of mutations to confer the expected substrate specificity advantage in favor to KYN based on the phylogenetic analysis as described in our previous publication (https://pubmed.ncbi.nlm.nih.gov/36465553/). Such H102W/N333T-dependent activity loss or decline was observed in the case of HsKYNase_93D9 scaffold as well (kinetic data shown in the manuscript).

3. How are the authors assessing the active fraction of the enzyme (potential assessment of active fraction of enzyme using a tight-binding inhibitor would be helpful).

We would like to note that the emphasis of this work is to demonstrate how a level of enzymatic activity can be attained by two completely different pathways which are characterized by distinct conformational dynamics and very different catalytic selectivity. Determining the active fraction using a tightly bound inhibitor is clearly beyond the scope of this work. Based on the calculation of certain burst amplitudes and measured product formation of HsKYNase, the enzyme exhibited ~50% site reactivity.

4. Supplemnetary table 2, 3, 5 and 7, I am surprised to see the values in the table listed without errors. The dispersion of the estimate about the mean is as important an estimate as the reported estimate itself. Errors will help me understand how reliable or conclusive comparisons are across several kcat/Km values. Also, it might benefit the readers to know how many technical and biological replicates resulted in the value reported.

In Supplementary Tables 2, 3, 5 and 7 we reported the kinetic features of over 50 evolutionary intermediate enzymes that were characterized in the course of the directed evolution campaign. We reported the kinetics of these intermediates solely for the shake of complete transparency. These evolutionary intermediates are peripheral to the story and have no material impact on our findings.
In any event to address the reviewer's concern we now added the errors of the different steady-state kinetic parameters for each variant as requested by the Reviewer. The purified evolutionary intermediates variants were kinetically characterized against KYN as described in the Methods section (each KYN concentration in 3 technical replicates). Of note, the final variants HsKYNase_93, HsKYNase_93D9 and HsKYNase_66 along with the wild-type HsKYNase were characterized several times from at least three biological replicates each.

5. For the sake of consistency and to avoid confusion, all Km values should preferably be reported in M (this is because all kcat/Km values have been reported as M-1 sec-1). In some tables, the Km values are reported in M, in others in mM and in yet others, they are uM.

We did not report any K_M values in M units in our manuscript. However, indeed, in the main Figure 1 the reported K_M values were shown in μM whereas in the SI Tables in mM. To avoid any confusion, we now

report all K_M values in µM (which makes easier the direct comparison) and all the k_{cat}/ K_M values in $M^{-1}s^{-1}$.

6. I am bit confused about the values reported in supplementary table 7 and 8 for HsKYNase_93D9 with KYN. The steady state estimate indicates that the Km for the substrate is 0.000085 M. However, a crude calculation of Km from the numbers reported in table 8 shows a Km of 20 M for the substrate!!!

We thank the Reviewer for pointing out a typo in 2 Supplementary Tables (8 & 9) where, the units for k_1 the second order rate constant were mistakenly shown as $M^{-1}s^{-1}$. Of course, the units are $\mu M^{-1}s^{-1}$ and this error has now been corrected. We also emphasize that the units for all the kinetic values for this and other variants in the main Figures and Tables are correct.

Further, the kcat in suppl. Table 8 is 0.79 s-1 (with very narrow dispersion about the mean) while that in suppl. Table 7 is 1.45 s-1 (no dispersion provided). Which of these scattering of numbers are the authors using to compute their kcat/Km values that provides them crucial insights into evolutionary interpretations?

We would like to emphasize that the difference between the 2 estimated values of 0.79 & 1.45 s^{-1} regarding the k_{cat} of HsKYNase_93D9 were estimated by performing entirely different kinetic methods covering totally distinct time regimes. That is, the steady-state kinetic value of 1.45 s^{-1} was estimated by carrying out steady-state kinetic experiments and measuring initial reaction rates within a few minutes of the reaction (measuring the absorbance of KYN disappearance at 365 nm). On the other hand, the value of 0.79 s^{-1} is an estimate that was obtained from pre-steady-state kinetic experiments in the milli-second range by monitoring the fluorescence of the released product anthranilate. In addition, the fitting of the data that yielded those 2 values was different (analytic fit of the Michaelis-Menten model in case of steady-state values vs numerical integration of the rate equations in case of the pre-steady-state value). Taking into account all these significantly different aspects of experimental execution and data analysis, the 2 values differ by less than 2-fold (1.45/0.79 ~1.83-fold) which is rather remarkable and we believe speaks to the rigor of our analysis and experimental techniques.

Depending on which Km (from their steady state estimates or their pre-steady state dynamic simulations), the kcat/km will be approximately ~10^6 fold different placing the Fold-HsKYNase number anywhere you wish and interpreting the evolution how you would like to subjectively package it.

We think the comment above stemmed from the confusion created by the typo in the Supplementary Tables (to reiterate there was NO TYPO in the main text).

We would also like to note that the main findings from our pre-steady-state kinetic analysis are: 1) primarily the different rate-determining steps of the two evolved variants HsKYNase_93D9 and HsKYNase_66 despite their similar steady-state kinetic profiles against KYN. Our pre-steady-state assay measures directly the rate of chemistry in the active site of the enzymes as it monitors the formation and release of the first product anthranilate (2^{nd} being Alanine), thereby enabling the observation of either the presence or the lack of a burst (burst informs on the nature of the rate-limiting steps of the

reaction). 2) Secondarily, through our KYN titration experiments and fitting the data globally to a minimal kinetic model we found that KYN binds significantly faster to HsKYNase_66 relative to HsKYNase_93D9 as evidenced by the respective k_1 values (5.9 $\mu M^{-1}s^{-1}$ vs 0.24 $\mu M^{-1}s^{-1}$ for HsKYNase_66 & HsKYNase_93D9 respectively). We note that, since this pre-steady-state assay is primarily measuring the rate of chemistry rather than directly the rate of substrate binding, additional experiments would be needed as next steps to estimate & resolve the k_1 more accurately. However, we experimentally compared the two evolved variants in parallel under the same conditions and the obtained values provide a very good relative estimate for the two species.

7. On similar lines, for data on HsKYNase_66 in supplemntary Table 9, the Km computation from the global fit analysis gives an estimate of 4.1 M. However, it is 0.000027 M as reported in the reference 10.1038/s41929-022-00856-6 by the same authors.

Please see response in point #6 above as well as the data that we report in the main text. In our Nature Catalysis publication that the author refers to, the steady-state k_{cat} and K_M values of HsKYNase_66 against KYN are 1.54 s^{-1} and 27 µM and in our present study the respective values are 1.68 s^{-1} and 41 µM as shown in Figure 1e. These values are in good agreement considering that different scientists performed the kinetic characterization using different protein purification lots.

8. How do the authors explain the sudden drop of Km with D2O (supplementary table 10) of HsKYNase_93D9 against KYN to 250 uM those reported in point number 6 above.

In fact, the K_M of HsKYNase_93D9 against KYN did not drop in D_2O but it increased (from ~85 µM as shown in main Figure 1e to 250 µM in the SI Table 10). In contrast, the k_{cat} decreased 9-fold (from 1.45 s^{-1} to 0.16 s^{-1}). Of note, our kinetic analysis of all the KYNase species in D_2O against both substrates KYN and OH-KYN have shown that the k_{cat} decreases within a range of 2-20-fold as we have reported in our different publications. For example, as described in https://www.nature.com/articles/s41929-022-00856-6 for HsKYNase_66: k_{cat}^{KYN} ~1.54 s^{-1} and 0.24 s^{-1} in H_2O and D_2O respectively and $k_{cat}^{OH\text{-}KYN}$ ~0.44 s^{-1} and 0.23 s^{-1} in H_2O and D_2O respectively. Same trends were observed for HsKYNase as described here https://pubs.acs.org/doi/10.1021/acschembio.0c00676. There are several factors that contribute to the observed apparent K_M increases or decreases and one of those might be linked to the higher viscosity of D_2O relative to that of H_2O. The higher viscosity can affect the way that the substrate diffuses into the active site and how it interacts with the enzyme. In addition, the substrate solubility can be affected by the higher viscosity. As shown by Cioni et al., (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1302113/) D_2O affects the flexibility of the polypeptide chain, making enzymes more rigid, thereby affecting their kinetic properties. Within the context of our KYNase species, we have seen that for the KYN reaction, the K_M values generally increase in D_2O, whereas for the OH-KYN hydrolysis either remain at similar levels (HsKYNase_66) or they decrease by a factor of ~3-fold (HsKYNase and HsKYNase_93D9). Overall, additional experiments would be needed to further resolve in detail the effect of D_2O on the kinetic parameters of the different variants including pH- and temperature-dependent as well as stability assessments.

9. Compounding the problem with respect to units and values (and demonstrating how casual the authors have been in showcasing this numbers without aking care to proofread them appropriately), on page 6, line 267 of the main text, the authors state that the second order rate constant for binding for HsKYNase_66 is 5.9 µM-1 s -1. This is shown as 5.9 M-1 s -1 in suppl. Table 9 (once again a 10^6 difference!!!). On similar lines, the association rate constant for HsKYNase_93D9 is shown as 0.24 µM-1 s -1 (page 6, line 267) while the suplemnetary table 8 once again shows this as 0.24 M-1 s -1 ?? With the combined expertise of all the authors (some of whom are instrumental in laying the foundation of modern enzymology), such kind of errors is not only surprising but distressing.

We addressed this typo issue in the point #6 above. We feel that the reviewer is being overly harsh for a typo in one value out of many 100s of kinetic parameters reported in our study.

10. In response to reviewer 3's comments, in their rebuttal letter the authors state that "Performing further in-depth in-silico analyses in order to precisely define the specific chemical steps of the catalytic cycle for the 3 enzymes (HsKYNase, HsKYNase_93D9 & HsKYNase_66) with the two substrates (KYN and OH-KYN) is clearly outside the scope of the present study". I disagree with this assertion in the light of such disparity of kinetic estimates reported as part of their current study and previous studies. This calls for the authors to look at all their numbers (reported in the present study and studies previously published) carefully in order to tidy up the interpretation.

It seems that this comment is in reaction to the one typo regarding units discussed in #6.

Please note that reviewer #3 agreed with our response in our initial rebuttal letter regarding the need for a detailed assessment of the different chemical steps by *in-silico* means which would clearly be outside the scope of the present study. We addressed the additional requests of Reviewer 3# regarding our *in-silico* work in our revised version of the manuscript and we will defer to him/her to comment as to whether this is satisfactory or not.

REVIEWER COMMENTS

Reviewer #3 (Remarks to the Author):

New table 3 is very interesting and I think that the results are aligned with their experimental results, in where the stability of some interactions, like for example R434 among others, follows the trend according to the experimental results. Nonetheless, when they mention the key distances both in the discussion of the results and in the figures, they should use average values (and ideally standard deviations) at least through the last few ns (once the complexes are stabilised) of the three replicates. If I understood it properly, they have used only one replicate to build the table while the main goal of doing replicates is to have more sampling.

I see that one of the problems the authors seem to face is that some of the complexes are not stable at 50ns for example SI figures 7d,f; 8c,f and 9c,f. This is not commented anywhere and it is of special relevance in comparing between binding modes, specially in one system where the Ca and BB RMSD seems to be stable but not the all-atoms one, meaning probably that the ligands is unbinding. In may opinion these should be analysed and discussed elsewhere to clarify what is happening and if the simulations should be extended or some kind of restraints could be necessary.

Reviewer #4 (Remarks to the Author):

Thank you very much for your revised manuscript addressing all the comments. Though I still believe that an estimation of the active fraction of the enzyme might be an essential requisite for studies of this nature, I believe your "response to reviewer" has addressed and allayed the most pressing concerns I had about the manuscript, which principally stemmed from the typographical error in the supplemental tables. I once again thank the authors for systematically and rationally addressing the concerns raised and would like to congratulate them on the work.

Response to Reviewer #3 (NCOMMS-23-30683B)

We would like to emphasize once more that for the purpose of our work, the MD simulation was employed to simply build a homology model of our evolved enzyme species HsKYNase_93D9 and to perform ligand docking assessments. This is a minor point in the overall story and one tangential to all the key messages of our work which addresses parallel pathways of molecular evolution and how conformational flexibility (determined by experimental HDX-MS studies) relates to enzyme substrate specificity.

Nonetheless, following Reviewer's #3 comments from both revisions, we have significantly expanded the MD simulation part by: 1) extending the simulations from 10 to 50 ns for all the 6 complexes in triplicates, 2) addressing his/her request to further analyze the dynamical picture of key interactions over the course of the simulations and report the % of the simulation time that each interaction is maintained (New Table 3) and 3) extensively revising the main Figure 3 and adding new Supplemental Figures 7, 8, & 9. Yet, after having addressed all the major and minor points from Reviewer #3, we feel that he/she continuously raises additional points, which are not in any way relevant to the major conclusions of the work, not on the interpretation of our data.

More specifically, regarding Reviewer's #3 latest points:

"New table 3 is very interesting and I think that the results are aligned with their experimental results, in where the stability of some interactions, like for example R434 among others, follows the trend according to the experimental results. Nonetheless, when they mention the key distances both in the discussion of the results and in the figures, they should use average values (and ideally standard deviations) at least through the last few ns (once the complexes are stabilized) of the three replicates. If I understood it properly, they have used only one replicate to build the table while the main goal of doing replicates is to have more sampling."

RESPONSE *=> Given that the MD simulation triplicates for each complex showed similar stabilization trends, minor differences in the averaged interaction distances from the MD calculations, likely by 10-20% are not relevant. The replicates were very consistent, so what is the rationale for this request, especially in the context of this study? Further, expanding on this point would actually be a distraction because it would add further weight to the MD simulations at the expense of the other much more important facets of the work.*

"I see that one of the problems the authors seem to face is that some of the complexes are not stable at 50ns for example SI figures 7d,f; 8c,f and 9c,f. This is not commented anywhere and it is of special relevance in comparing between binding modes, specially in one system where the Ca and BB RMSD seems to be stable but not the all-atoms one, meaning probably that the ligands is unbinding. In my opinion these should be analysed and discussed elsewhere to clarify what is happening and if the simulations should be extended or some kind of restraints could be necessary."

RESPONSE *=> The only RMSD plot that appears to be indeed unstable is the one shown in the Supplemental Figure 8f and this is happening only in 1 out of the 3 replicates. Regarding the other plots that were mentioned i.e. SI Figures 7d,f, 8c & 9c,f there is only a suspicion of very slight destabilization trend at around 50 ns just for 2 of them (SI Figures 7d,f) whereas the other 3 (8c & 9c,f) show stabilization.*

It is exceedingly challenging to interpret what is happening in case of the unstable replicate shown in SI Figure 8f as different reasons may be causing this local instability. It could well be that the ligand diffuses away from the active site and significantly longer simulation times would be needed to assess this hypothesis. More importantly, note that, the instability that is shown in Figure 8f may not be relevant because the ligand binds to the active site, followed by formation of the external aldimine with the cofactor PLP and finally catalysis. The latter events are not and CANNOT be captured at all by the MD simulations. Also note that SI Fig. 8f refers to OH-KYN which is of secondary importance to the story. HsKYNase93D9 was evolved for high activity to the Kyn substrate and the lack of selectivity towards OH-Kynurenine was an unexpected finding.

*In response and for additional clarity around this observation, we added the following text in the respective Results section: "*Of note, the all-heavy-atom RMSDs analysis of one out of three simulation replicates of HsKYNase_93D9/OH-KYN complex showed some evidence of destabilization during the second half of the simulation (**Supplemental Fig. 8f**). This observation may suggest that the ligand gradually diffuses away from the active site. However, extensive studies well beyond the scope of the present study would be needed to further investigate such a hypothesis and more broadly their implication, if any, on such fine mechanistic questions to the consequences of divergent evolutionary trajectories, which are the focus of this work".

In short, extending the simulations beyond 50 ns and further analyzing various specific aspects for each of the complexes in additional detail is outside the scope of this study, whose main message is that from 2 parallel evolutionary trajectories, one can obtain 2 genetically distinct enzyme species with similar catalytic activity for the evolved substrate but different substrate specificities, rate-determining steps and conformational dynamics.

We would like to emphasize that to implement all the changes requested by Reviewer #3, we were forced to expand the MD simulations section of the text to double that of the initial version. While we find this undesirable and distracting from the key points of the work and the message of the manuscript, we reluctantly changed the flow and the size of the text to satisfy the reviewer. We cannot see any way to further expand or edit the MD simulation portion without seriously negatively impacting the paper.

Last but not least, we would like to remind Reviewer #3 that he/she had already acknowledged from the beginning that: "In my opinion the work is of significance for the scientific community to the field and related fields since it provides rational understanding on the evolutionary trajectories of a system with a key engineering interest. The overall conclusions are well supported and specific claims and methodology sound and well detailed".