

Reversal of protein chemical aging by enzymatic deglycation

Corresponding Author: Dr Aaron Cravens

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)

The study presents a highly novel and original approach by engineering CMLase, the first enzyme capable of reversing Nε-carboxymethyl-lysine (CML) modifications on proteins, through computational screening of over 44,000 sequences and directed evolution testing more than 500 million variants via a custom genetic selection system, demonstrating unprecedented repair of endogenous CML in aged human tissues ex vivo. The scientific logic is robust, progressing rationally from structural insights and promiscuous glycine oxidase activity to iterative improvements in substrate affinity and breadth, validated by assays like ELISA, proteomics, LC-MS/MS, and IHC, with strong controls confirming specificity. However, faults include methodological gaps such as limited off-target profiling, overreliance on ex vivo models without assessing in vivo accessibility or hydrogen peroxide byproducts' risks, quantitative inconsistencies (e.g., varying CML reduction metrics), and weak statistics with few samples and no blinding. Limitations encompass the absence of in vivo data, no demonstration of functional protein restoration or broader AGE reversal, substrate-dependent variability, and challenges in scalability for therapeutics. Moreover, the claim of "reversing protein aging" appears overstated, as while there is strong evidence that CML is a major marker of protein damage and is linked to aging and age-related diseases, direct evidence that it is a major cause of protein aging itself is less conclusive, as the accumulation of CML is often a consequence of oxidative stress and glycation, which are themselves contributors to aging; thus, enzymatic removal of CML (on protein surface) may not fully reverse protein aging, especially given correlative rather than definitive causal links in many studies.

Novelty

This work introduces a highly novel concept: the enzymatic reversal of Nε-carboxymethyl-lysine (CML) modifications, a type of advanced glycation endproduct (AGE) that accumulates on long-lived extracellular matrix (ECM) proteins during aging and is widely considered irreversible. Prior research on AGEs has focused primarily on preventing their formation (e.g., via small-molecule inhibitors that trap reactive carbonyls), but no previous studies have reported an enzyme capable of cleaving CML from peptides or proteins. The authors engineer "CMLase," a de novo enzyme derived from a glycine oxidase homolog, through computational screening of ~44,783 sequences and five rounds of directed evolution, testing over 500 million variants via a custom genetic selection system. This is the first demonstration of repairing endogenous CML in aged human tissues (lens, skin, aorta), reducing levels by 45–78% in lens and >55–70% in skin/aorta ex vivo. The approach establishes a platform for "deglycation" enzymes, potentially extending to other AGEs, which could transform interventions in aging, diabetes, and vascular diseases. While glycine oxidases have been studied for promiscuous activity on similar substrates, adapting one for protein-bound CML is unprecedented.

Originality

The study is highly original, blending computational biology, enzyme engineering, and gerontology in a unique way. The starting enzyme (from *Calidithermus roseus*) was identified through a novel structural screen for active-site accessibility, not just sequence homology. The genetic selection strategy—coupling CML cleavage to lysine auxotroph growth in *E. coli*, adapted for periplasmic peptidyl-CML—is inventive and enables ultra-high-throughput screening (>10⁸ variants), surpassing traditional methods. Key innovations include loop deletions to widen the active site, consensus mutations from homologs for affinity gains, and mixed-peptide substrates to ensure sequence-agnostic activity. The collaboration between Revel Pharmaceuticals, Calico Life Sciences, and academic labs lends credibility, but the core idea of "reversing protein aging" via engineered oxidases appears to originate from this team, with no direct precedents in the literature (e.g., a partial CML oxidase from *Cladosporium* was reported in 2004 but unsequenceable and inactive on peptides). The work avoids recycling established techniques, instead creating new tools like the peptidyl-CML selection.

Scientific Logic

The reasoning is logically sound and stepwise, grounded in biochemistry and evolutionary principles. CML structurally mimics glycine (carboxymethyl amine moiety), justifying glycine oxidase as a starting scaffold. Screening homologs for peptidyl activity addresses the challenge of substrate size, and structural modeling (AlphaFold, alignments) rationally guides mutations (e.g., deleting helix $\alpha 9$ for access, swapping residues for affinity). The genetic selection cleverly links catalysis to cell survival, with escalating stringency (lower substrate/IPTG, mixed peptides) to evolve broad activity. Results build progressively: from free CML (k_{cat}/K_M improvements from 0.0012 to 0.023 s⁻¹ mM⁻¹) to peptides (20% gain in final round) to proteins (52–97% CML removal by ELISA) to tissues (quantified by LC-MS/MS, ELISA, IHC). Controls are robust (inactive enzyme, unmodified proteins, antibody validation via competition). Proteomics confirms site-specific repair (30/33 sites reduced on BSA, independent of sequence/secondary structure), and age-dependent CML accumulation in tissues aligns with literature. Hydrogen peroxide production is acknowledged as a byproduct, consistent with flavoenzyme mechanics. Overall, the logic flows from hypothesis (promiscuous oxidase activity) to engineering to validation, with data supporting claims without overreach.

Limitations

1. Overstated Claim on Reversing Protein Aging: The claim of "reversing protein aging" appears overstated, as while there is strong evidence that CML is a major marker of protein damage and is linked to aging and age-related diseases, direct evidence that it is a major cause of protein aging itself is less conclusive, as the accumulation of CML is often a consequence of oxidative stress and glycation, which are themselves contributors to aging; thus, enzymatic removal of CML (maybe predominantly on protein surface) may not fully reverse protein aging, especially given correlative rather than definitive causal links in many studies.
2. Methodological Gaps in Specificity and Controls: While CMLase is claimed to be specific, there's limited data on off-target activity (e.g., does it oxidize native lysines or other AGEs like CEL?). The text mentions no activity on unmodified BSA, but broader substrate profiling (e.g., other amino acids, lipids) is absent. IHC antibody validation is good, but ELISA variability across proteins (52–97% reduction) isn't deeply probed—could reflect antibody epitope bias rather than true catalysis.
3. Overreliance on Ex Vivo Models: All tissue experiments are on homogenized or sectioned samples, potentially overestimating accessibility in intact ECM. No diffusion/permeability data for CMLase into dense tissues like aorta.
4. Quantitative Inconsistencies: CML reduction in lens differs between methods (45% by LC-MS/MS vs. 78% by ELISA), attributed to surface vs. buried residues, but not experimentally dissected (e.g., via denaturation). Proteomics on BSA shows 3 sites untouched, but no follow-up on why (steric hindrance?).
5. Potential Artifacts: Hydrogen peroxide assays assume all H₂O₂ comes from CML cleavage, but flavoenzymes can auto-oxidize; controls mitigate this, but long incubations (overnight) risk secondary reactions. IHC omits antigen retrieval to avoid de novo CML, but this might under-detect buried modifications.
6. Evolutionary Bias: Selections use poly-alanine peptides with one variable residue, potentially biasing toward simple contexts; real ECM proteins have complex sequences.
7. Lack of In Vivo Data: The study is purely in vitro/ex vivo; no animal models (e.g., diabetic mice with high CML) to assess pharmacokinetics, immunogenicity, or efficacy in living systems. H₂O₂ byproduct could cause oxidative stress in vivo, unaddressed here.
8. Incomplete Functional Restoration: While CML removal is shown, no assays test if it restores protein function (e.g., collagen elasticity, ECM signaling, RAGE binding inhibition). Pathogenic "hotspots" are mentioned but not prioritized or validated.
9. Scope of CML Accessibility: Enzyme works on surface-exposed sites (correlates weakly with B-factor), but buried or crosslinked CML in dense ECM may resist repair. Variability across proteins (e.g., 52% on eye TPE vs. 97% on casein) highlights substrate dependence.
10. Scalability and Therapeutics: Engineering focused on activity/stability, but no data on production yield, half-life, or delivery (e.g., to eyes/kidneys). As a large protein (~40 kDa inferred from homologs), systemic use faces barriers like renal clearance or immunogenicity.
11. Broader AGE Context: CML is one AGE; others (e.g., pentosidine crosslinks) aren't addressed, limiting impact on total AGE burden. Aging is multifactorial, so CML reversal alone may not "reverse aging."
12. Sample Limitations: Human tissues from limited donors/ages; no disease-specific samples (e.g., diabetic kidneys). Proteomics only on BSA, not tissues, missing real-world complexity.
13. Technical Constraints: Genetic selection relies on *E. coli*, potentially missing eukaryotic folding needs. IHC quantification (DAB intensity) is semi-quantitative, prone to variability.

Reviewer #2

(Remarks to the Author)

This manuscript describes the development of an engineered glycine oxidase that has new activity, enabling the enzymatic removal of CML from modified proteins. Overall, the computational screening and directed evolution is nicely done, and a CMLase enzyme would be a highly valued contribution to the field of glycation. However, there are major concerns about the characterization of CMLase activity, including the characterization of CML-modified substrates and the seemingly large potential for cross-reactivity, all of which should be addressed before this submission would become suitable for acceptance in *Nature Communications*. Detailed comments are below:

1. One issue throughout the manuscript related to a lack of clarity about how activity was assessed for the various CMLase enzymes. For example, minimal details were provided about the concentrations used for free vs peptide-CML in enzymatic assays. In later, critical experiments, I could not find details about the substrate concentration, identity or how the "loading" of CML on protein or peptide substrates was assessed. Greater clarity on these points would strengthen the conclusions of the manuscript.
2. There seems to be a major potential for cross reactivity with cleavage at C-terminal Gly residues. Did the authors observe any truncations at the C-terminus? How was the potential for other, potentially unexpected, cross-reactivity evaluated? Can it

work on N-terminal or Arg residues or modifications? I do not think I noticed any experiments to address the remaining (or potential for remaining) glycine oxidase activity, which could contribute to background H₂O₂ generation.

3. On a related note, additional characterization for CMLase-mediated CML loss from proteins would strengthen the contribution. As an example - A western blot showing in vitro CML modified proteins, treated with and without CMLase would show that the enzyme isn't digesting the whole protein (by Coomassie) and also decreases CML levels.

4. Additional characterization should also be provided for the extent of CML modification introduced into in vitro proteins. The glyoxylic acid reductive amination makes sense, but what is the overall degree of modification? The LC-MS data presented seems to focus on a %removal from a set of CML-modified sites. However, I find myself wondering if the initial extent of modification at each site (which was presumably not 100%, or was it?) affects the way one can interpret this data.

5. Given the large loading of CMLase to CML-modified protein (e.g. 5:1 CMLase to CML-BSA), it seems entirely plausible that some of the reduction observed by ELISA or immunohistochemistry could be due to blocking detection by Abs due to nonspecific (or even specific) binding to the substrate. This needs more careful controls throughout the paper.

6. Some discussion of the efficiency of the CMLase to other enzymes (and the native activity) should be included

7. Based on the data provided, I don't think that the following sentence is fully justified. "We show the new-to-nature CMLase repairs a majority of surface CML modifications on proteins and then mapped the specific sites of repair using proteomics." In particular "majority" seems inaccurate, as it definitely works better on some proteins than others as seen in figure 3.

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)

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 #5

(Remarks to the Author)

General:

This study demonstrated the development of CMLase, which reversed carboxymethyl-lysine in human tissue obtained from elderly donors. These findings show the first demonstration that the "irreversible" protein damage can be enzymatically repaired. The major conclusion from this study is that this newly developed enzyme and its effects could provide a platform to develop new enzymes that are able to reverse age-related molecular damage and repair tissue proteins compromised by aging and diseases.

Although this study may provide some interesting information, the experimental design has several potential major flaws, the data presented may require further verification with reproducibility, and the whole manuscript should be revised with proper and major format changes as a scientific article. Therefore, it is difficult to support the authors' conclusion based on the content of this manuscript. There are also several concerns, shortcomings, and deficits with this paper that I have listed below.

Specific Comments:

1. Overall, the style of writing in the article is more narrative than scientific, and does not provide deeper information regarding the role of CML in aging and diseases.

2. The Background section requires more information on the advanced glycation end products (AGE), aging, age-related functional changes, and age-related diseases.

3. In general, the Results section is a mixture of the description of the results including some of the methods and discussions, and should focus on just on the description of the results.

4. The first paragraph of the Results section needs to be removed or moved to the Discussion section.

5. The Results section should include sub-headlines to make it easier for the readers to follow.

6. The Figures and Figure Legend need to be separated from the body text.

7. The font in Figure 1B is too small and hard to see.

8. Although the authors stated that there are similar degrees of modification in various proteins, i.e., BSA, collagen, hemoglobin, casein, and eye total protein extract, using different glyoxylic acid concentrations, the degree of modification

seems to vary in various proteins (Figure 3A).

9. The experiments were conducted using different proteins and tissues, which seem to yield the data that support the authors' hypothesis. Unfortunately, it is a bit hard to reach the same conclusion as the authors because of the inconsistency of the usage of various proteins and tissues throughout the entire experiments.

10. The immunohistochemical analyses were conducted for the quantitative analysis of CML in human skin and the effects of CMLase treatment. The DAB intensity may not be the ideal way to quantitate the amount of antigen in the tissue. In addition, some of the immunostaining results possibly show over-staining and background staining, which could make quantitative analyses more difficult. Furthermore, the figure is too small to see the details as well.

11. It is not clear how the specificity of the antibodies and specific epitope of the antigen were determined. It is also not clear whether the authors have examined if the less (or no) staining after CMLase treatment was due to the direct effects or potential secondary effects by CMLase.

12. It is not clear whether the immunohistochemical analyses using formalin fixed paraffin embedded sections were ideal to test the enzyme activities on CML. Further confirmation may be required with frozen sections with/without fixation.

13. There is no clear statement regarding how many samples were used in many of the experiments, and it is not clear whether there was strong evidence of reproducibility for each experiment.

14. The Discussion section is too short, and there is not much information provided.

Reviewer #6

(Remarks to the Author)

REVIEW

The article "Reversal of protein aging by enzymatic deglycation" describes the development of a CMLase that removes carboxymethyl lysine modifications, a hallmark of aging. To achieve this, the authors identified the FAD-dependent glycine oxidase as a starting point for evolution. They tested 13 soluble GO homologs and identified 8 with activity on free CML, supporting promiscuity in this family. Then they used bioinformatics to screen over 44k enzymes in silico to find those with a more open and permissive active site. They identified *Caliditthermus roseus* (CrGO) as a promising starting point. They use a clever selection approach based on relief of lysine auxotrophy upon deglycation of CML peptides. They evolve for improved enzyme activity and then demonstrate deglycation on peptides, proteins, and tissue samples. Overall, the story is clearly written and quite interesting from a protein evolution standpoint. I am not an expert in protein aging, so I will leave it to the other reviewers to comment on the significance of the CML modification in the context of aging (how strong is the support that it leads to "aging" phenotypes that can be reverse by deglycation?).

Comments and suggestions below:

Glyoxylic acid is released in the deglycation process. This also seems to be the reagent used to generate CML proteins in the methods. From a quick read, it seems like the body should be able to metabolize glyoxylate and keep it at reasonable levels. The authors should comment on this- are there potential issues with freeing glyoxylic acid? Simply providing references for how it gets metabolized could suffice.

Panel 2E. The rate is slow ($\mu\text{M}/\text{hour}$). In the legend for Fig 2E, please state the concentration of enzyme used and state how the data are collected (are the plotted values averages? How many replicates?)

First round of evolution: error prone PCR. In the paragraph describing error prone PCR- how many rounds of error prone PCR and selection did the authors do? It seems like it was 1 round, but please clarify. 7 mutations is a lot for one round, but I see that the authors aimed for 4-8 mutations per gene, so I realize this could have been the hit from 1 round of selection. What was the estimated library size and coverage?

Exciting that this yielded a 5-fold improvement. Please list the mutations in this paragraph (I see the authors reference table Fig S2-1, but would be helpful to see it listed here, too).

How many colonies did the authors screen in each round of evolution? Please add this information to the methods or Figure S2-4 (to explain how you got to the estimated number of screening over 500 million protein variants as stated in the abstract).

Can you further explain the rationale for picking residues guided by CeGO to saturate the CrGO? Is it surprising that mutations that improve activity on free CML also improve activity on Peptidyl CML?

The results are very interesting and could be sold more explicitly in the discussion. Discuss the overall fold-improvement in enzyme activity in the discussion – this would strengthen the results of the paper.

When you tested activity on CML-BSA, why did you use low concentrations of CML-BSA ($1\ \mu\text{M}$) with $5\ \mu\text{M}$ CMLase? How many turnovers do you think the CMLase can do?

Apologies if I missed it- please provide information on how the selection strain was generated.

Can the authors provide more detailed methods for how the selection was done? What were the transformation efficiencies in each round, and what is the estimated coverage based on this? What temperature were the cells grown at during the selection (I see this information for the transformation)? When did the authors start to see colonies for the different rounds of selection?

Sequence context: I think it's great that Fig 3c highlights structural diversity of target sites and some sequence contexts. However, a more comprehensive depiction of sequence context would help a lot. Figure 3B- can the authors convert this information into sequence context information somehow (maybe logos plot?). This would help audience/end user to see if there are any limitations. A logos plot or some visual depiction of the sequence bias (or lack thereof) of enzymatic deglycation would be great.

Sequence context: When the authors did the error prone library and challenged for growth on mixtures of peptides- were cells selected on all peptides at once? What are the sequences of these peptides? In this case, do you worry that a particular variant could deglycate only a subset of the peptides? How do you modulate stringency? I think it would help if you added data for your final variant testing kinetics on a range of peptide contexts (At least the ones from the mixture you provided in the selection). In the end, you get variants with a 20% improvement on peptide, but you don't say which peptide. Also, Figure S2-6 doesn't say which peptide (sorry if I missed it).

Sequence context: Would be great to have sequence context for all residues from S4-3. Can the authors add that to the graph?

Fig 1c, Supplemental Figures S1-2 and 2-2. Please include methods for how the CML substrate was modeled in the active site. What docking algorithm or program did the authors use? Was this done manually?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I have carefully reviewed the authors' point-by-point response and the revised manuscript. Overall, the authors have responded logically, thoroughly, and satisfactorily to my comments. They addressed each concern with appropriate new data, clarifications, textual revisions, and explicit additions to the limitations section without dismissing any issues or weakening the core scientific claims. The revisions have markedly improved the rigor, clarity, and scientific tone of the paper.

Below is a concise evaluation of how the authors handled my specific comments:

1. Overstated claim on "reversing protein aging": Fully addressed. The title has been changed to "Reversal of Protein Chemical Aging by Enzymatic Deglycation," and the Discussion now clearly distinguishes chemical aging (accumulation of non-enzymatic modifications such as CML) from organismal/multifactorial aging. This is a precise and scientifically accurate revision.
2. Methodological gaps in specificity and controls (off-target activity, ELISA variability, epitope bias): Strongly addressed with new supporting data. The authors added broad substrate profiling (new Figures S2-6 and S2-7) demonstrating no oxidation of native lysine or other canonical amino acids, and they consistently used the catalytically inactive control (CeGO-763) in all antibody-based assays to rule out epitope masking. ELISA variability is now explicitly attributed to steric accessibility and is supported by orthogonal LC-MS/MS and proteomics data.
3. Overreliance on ex vivo models (no diffusion/permeability data): Acknowledged transparently. The authors added clear Discussion text framing macroscopic tissue penetration as a separate delivery challenge for future studies while defending the use of FFPE sections as a valid test of steric accessibility in a proof-of-concept context.
4. Quantitative inconsistencies (ELISA vs. LC-MS/MS differences; untouched sites): Well explained. The authors clarify that ELISA primarily detects surface-exposed CML (more enzyme-accessible), while LC-MS/MS measures the total pool including buried residues. Proteomics data confirm site-specific repair, and no correlation exists between initial modification level and repair efficiency. These orthogonal methods now strengthen the conclusions.
5. Potential artifacts (H₂O₂ auto-oxidation, IHC antigen retrieval): Clarified with controls and literature precedent. Peroxidase assays are limited to short incubations with strict controls; IHC omission of retrieval follows established protocols to avoid de novo CML formation, and LC-MS/MS on homogenates provides complementary data on total CML.
6. Evolutionary bias from poly-alanine peptides: Addressed through design rationale and counter-measures. The authors explain the need for baseline activity, the use of mixed-peptide libraries, and post-selection testing on complex CML-BSA substrates. Variants that did not translate to realistic proteins were discarded.
7. Lack of in vivo data (PK, immunogenicity, H₂O₂ risks): Properly framed as a study limitation. The Discussion now explicitly addresses these points, including H₂O₂/glyoxylate clearance and plans for future pharmacokinetic/immunogenicity evaluation.
8. Incomplete functional restoration (no biomechanics or RAGE assays): Correctly noted as outside the current scope and added as a limitation in the Discussion.

9–13. Scope of CML accessibility, scalability/therapeutics, broader AGE context, sample limitations, and technical constraints: All concerns have been explicitly acknowledged and incorporated into the revised Discussion. Key additions include surface-exposed preference with potential resistance of buried/cross-linked CML, immunogenicity and delivery challenges, CMLase as a proof-of-concept platform for other AGEs, rationale for the chosen protein/tissue panel, cross-method consistency for IHC, and suitability of the E. coli expression system.

The responses to the other reviewers (particularly #2 and #5) are equally constructive. New experimental details, supplementary figures (S2-6, S2-7, S2-9), expanded Background and Discussion sections, sub-headings, improved figure formatting, and clearer separation of results/methods/discussion have been implemented throughout the revised manuscript. The yellow-highlighted changes in the provided PDF align exactly with the promised revisions.

In summary, the authors have produced a high-quality revision that directly incorporates my feedback, adds necessary controls and data, and appropriately qualifies claims while preserving the novel and impactful nature of the work. The manuscript is now more concise, scientifically toned, transparent about limitations, and substantially stronger. I recommend acceptance with only minor copy-editing if needed.

Reviewer #2

(Remarks to the Author)

This manuscript describes the development of an engineered glycine oxidase that has new activity, enabling the enzymatic removal of CML from modified proteins. Overall, the edits to the manuscript have improved its quality, though they have also somewhat lowered the significance as the CMLase does not appear very active. Still, the authors have increased the clarity of their findings and addressed concerns about modification levels impacting enzymatic efficiency and epitope masking in the immunochemistry data, which is a big improvement. However, there are still concerns about the characterization of CMLase activity, including the potential for cross-reactivity to other AGE adducts and the lack of a western blot showing removal of CML. The latter point is more critical, and if addressed would make this submission suitable for acceptance in Nature Communications. Detailed comments are below:

1. The authors seemed to have misunderstood our prior question about activity on arginine residue modifications. Specifically, we wonder if the CMLase works on the equivalent arginine glycation adduct (CMA) and/or the equivalent modification on the N-terminus of peptides/proteins.
2. In figure S2-9 the treated CML-BSA lane seems to be closer to the fully modified than the unmodified BSA. Maybe using a substrate that gets closer to 100% recovery would be better to show in this data. As suggested before, we would also like to see a Western blot for the associated gel as an additional confirmation of reduction in CML modifications after incubation with the CMLase since the Coomassie doesn't show conclusive evidence of removal of adducts.

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)

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 #5

(Remarks to the Author)

General:

This revised manuscript describes the development of CMLase, which reversed carboxymethyl-lysine in human tissues obtained from elderly donors, and represents the first demonstration that “irreversible” protein damage can be enzymatically repaired. The authors conclude that this newly developed enzyme, along with its observed effects, could serve as a platform to develop new enzymes capable of reversing age-related molecular damage and repairing tissue proteins compromised by aging and disease.

Although this revised manuscript addressed the previous critiques, the experimental design still has several potential major flaws. Further verification with reproducibility may still be needed, and the manuscript may require more focus on a specific area of aging or age-related disease instead of covering many areas in age-related pathophysiology. This causes a loss of focus and makes the manuscript somewhat confusing. Therefore, it is difficult to support the authors' conclusion based on the content of this manuscript. There are still several concerns, shortcomings, and deficits with this paper that I have listed below.

Specific Comments:

1. Overall, this manuscript does not provide deeper scientific information regarding the role of CML in aging and diseases.
2. The Background section requires more focus on the role of CML accumulation in aging, age-related functional changes, or specific age-related disease. The current format addresses various aspects of age-related changes, making it difficult to clearly understand the purpose of this project.
3. The Results section still contains a mixture of the results description, and some of the methods and discussions, which should be revised to only focus on the description of the results.
4. Although the authors reported similar degrees of modification across various proteins, i.e., BSA, collagen, hemoglobin, casein, and eye total protein extract, it is not clear why the authors selected only these proteins for the study.
5. The experiments were conducted using different proteins, tissues, and species (including human samples), which yielded data that supports the authors' hypothesis. However, it is somewhat difficult to reach the same conclusion as the authors because of the inconsistencies in the use of various proteins and tissues across the experiments.
6. Immunohistochemical analyses were conducted for the quantitative analysis of CML in human arterial tissue and to evaluate the effects of CMLase treatment. DAB intensity may not be the ideal way to quantitate the amount of antigen in the tissue. In addition, some of the immunostaining results appear to have over-staining and high background, which could make quantitative analyses more difficult.
7. It is not clear whether the immunohistochemical analyses using formalin fixed paraffin embedded sections are ideal to test enzyme activity on CML. Further confirmation may be required using frozen sections with/without fixation.
8. There is no clear statement on the number of samples used in many experiments, and there is no strong evidence of reproducibility for each experiment.

Reviewer #6

(Remarks to the Author)

The authors have addressed my concerns regarding the clarity of the evolution process and the enzyme characterization. While the evolved enzyme is on the slow side, I think the engineering challenge and the improvement over the starting point is notable and interesting. The authors have also tempered their claims about the impact on aging, which is good. I am fine with this manuscript moving forward.

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors have addressed my remaining concerns and I feel that this manuscript is now acceptable for publication

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)

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.

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 Comments:

We would like to thank the reviewers for their insightful and constructive comments. We have made significant improvements based on these suggestions to strengthen the rigor and clarity of the work. Below are responses to each of the comments made by the reviewers.

Reviewer #1:

1. Overstated Claim on Reversing Protein Aging: The claim of "reversing protein aging" appears overstated, as while there is strong evidence that CML is a major marker of protein damage and is linked to aging and age-related diseases, direct evidence that it is a major cause of protein aging itself is less conclusive, as the accumulation of CML is often a consequence of oxidative stress and glycation, which are themselves contributors to aging; thus, enzymatic removal of CML (maybe predominantly on protein surface) may not fully reverse protein aging, especially given correlative rather than definitive causal links in many studies.

We thank the reviewer for this important distinction and agree that biological aging is a complex, multifactorial process. As suggested, we revised the title to "Reversal of Protein Chemical Aging by Enzymatic Deglycation". We believe this is scientifically precise because it distinguishes the specific molecular phenomenon of "chemical aging"—defined as the accumulation of non-enzymatic modifications like the Maillard reaction—from organismal aging. Since our enzyme chemically converts the pathogenic, aging-associated CML adduct back to its native lysine state, we are accurately reversing a specific chemical lesion that drives protein chemical aging. We've also updated the discussion to ensure accurate assessment of what we accomplished in the manuscript and limitations.

2. Methodological Gaps in Specificity and Controls: While CMLase is claimed to be specific, there's limited data on off-target activity (e.g., does it oxidize native lysines or other AGEs like CEL?). The text mentions no activity on unmodified BSA, but broader substrate profiling (e.g., other amino acids, lipids) is absent. IHC antibody validation is good, but ELISA variability across proteins (52–97% reduction) isn't deeply probed—could reflect antibody epitope bias rather than true catalysis.

We thank the reviewer for these rigorous methodological inquiries. We have addressed both the specificity and variability concerns with additional data and clarification:

Off-Target Activity & Specificity: To address the potential for off-target oxidation, we have expanded our characterization to include broad substrate profiling against a panel of native amino acids and peptides. These data, now included in Figures S2-6 and S2-7, demonstrate that CMLase retains high specificity for the carboxymethyl-amino moiety and does not oxidize native lysine or other canonical amino acids. Furthermore, the lack of activity on unmodified BSA (Fig. S2-9) confirms that the enzyme does not promiscuously degrade native protein structures.

ELISA Variability & True Catalysis: We acknowledge the variation in repair efficiency across different protein substrates (52–97%). We attribute this primarily to steric accessibility rather than antibody artifacts; the dense structural environment of certain substrates (e.g., Eye TPE) likely limits the physical access of the ~32 kDa enzyme to buried CML sites compared to more open scaffolds like casein. Inevitably with enzymes and antibodies binding to complex protein substrates there will be variation in binding depending on the local CML context. We believe this is rigorously explored in Figure 3C.

Ruling out Epitope Bias: To confirm that this reduction represents true catalysis and not just antibody epitope masking, we relied on our catalytically inactive control (CeGO-763). This variant yielded no reduction in signal, confirming that physical blocking of antibody binding sites is not the cause of the signal decrease.

Orthogonal Validation: Crucially, our site-specific LC-MS/MS data (Fig. 3B/C) independently confirms the chemical reversal of CML to lysine at varying efficiencies across different sites, corroborating the trends observed by ELISA without relying on immunological detection.

3. Overreliance on Ex Vivo Models: All tissue experiments are on homogenized or sectioned samples, potentially overestimating accessibility in intact ECM. No diffusion/permeability data for CMLase into dense tissues like aorta.

We thank the reviewer for this valid assessment. We acknowledge that sectioned samples reduce the diffusion barrier present in intact organs. However, our primary objective in this study was to validate the catalytic feasibility of reversing CML within the complex chemical environment of human tissue. We agree that FFPE sections, while physically thin, represent a rigorous test of steric accessibility due to extensive fixation-induced crosslinking. We view macroscopic tissue penetration as a delivery challenge distinct from the catalytic mechanism, to be addressed in future studies via formulation or administration routes. We have added text to the Discussion explicitly acknowledging challenges that remain for clinical translation. It's important to note that reversal of CML on extracellular proteins such as in plasma or in the ECM should prove more tractable than trying to remove intracellular proteins.

4. Quantitative Inconsistencies: CML reduction in lens differs between methods (45% by LC-MS/MS vs. 78% by ELISA), attributed to surface vs. buried residues, but not experimentally dissected (e.g., via denaturation). Proteomics on BSA shows 3 sites untouched, but no follow-up on why (steric hindrance?).

We thank the reviewer for this careful analysis. We employed three distinct analytical techniques—ELISA, bulk LC-MS/MS, and proteomics—specifically to provide robust, orthogonal confirmation of this unprecedented enzymatic reaction. We did not anticipate exact quantitative alignment given the fundamental differences in what each method measures:

Lens Tissue Discrepancy (ELISA vs. LC-MS/MS): The difference between the 78% reduction measured by ELISA and the 45% reduction measured by bulk LC-MS/MS is consistent with the accessibility of the substrate. The ELISA uses an antibody that primarily detects surface-exposed CML epitopes, which are also the most accessible to our enzyme. In contrast, the LC-MS/MS method involves total hydrolysis of the sample to individual amino acids, measuring the entire pool of CML, including residues buried within the protein core or macromolecular aggregates that are sterically shielded from the enzyme *ex vivo*. The fact that the "accessible" fraction (ELISA) shows higher clearance than the "total" fraction (LC-MS/MS) reinforces the conclusion that CMLase is effectively targeting the surfaceome where signaling interactions (e.g., RAGE binding) occur.

Untouched Sites in BSA: Regarding the three unmodified sites identified by proteomics, we attribute this to local steric hindrance or flexibility. Just as with the lens tissue, certain lysine residues may be sequestered by local secondary structure or protein folding, rendering them inaccessible to the enzyme's active site.

By showing consistent catalytic activity across these diverse contexts, we believe we have overcome the limitations inherent to any single method and provided compelling evidence of CMLase activity.

5. Potential Artifacts: Hydrogen peroxide assays assume all H₂O₂ comes from CML cleavage, but flavoenzymes can auto-oxidize; controls mitigate this, but long incubations (overnight) risk secondary

reactions. IHC omits antigen retrieval to avoid de novo CML, but this might under-detect buried modifications.

We thank the reviewer for thoughtfully pointing out potential artifacts in the hydrogen peroxide assay and IHC staining.

Regarding the peroxidase assays, we wish to clarify that these experiments are conducted on a timescale of six hours or less to minimize secondary hydrogen peroxide formation. We also include strict controls—utilizing either no substrate or unmodified substrate—to verify that no significant auto-oxidation occurs. Overnight incubations were utilized where noted in Figures 3 and 4 and only for experiments where hydrogen peroxide accumulation is not the readout.

Regarding immunohistochemistry, we deliberately omitted antigen retrieval steps based on established literature precedent (e.g., Schleicher et al., J Clin Invest 1997, ref. 29 in manuscript). These studies demonstrate that heat-induced epitope retrieval can generate de novo CML-like modifications, leading to potential false-positive staining. While we acknowledge that omitting retrieval may under-detect buried epitopes, we prioritized specificity to ensure that the staining observed—and its subsequent reduction by CMLase—reflected genuine, physiological CML accumulation rather than procedural artifacts. Furthermore, our LC-MS/MS data (Fig. 4A) serves as a complementary method to assess the total CML burden, including buried residues, in homogenized tissue.

6. Evolutionary Bias: Selections use poly-alanine peptides with one variable residue, potentially biasing toward simple contexts; real ECM proteins have complex sequences.

We thank the reviewer for pointing out potential bias in the peptide selection. A strong basis of activity on the selection peptide was needed to complement growth which is why we chose simpler targets that we knew had a baseline level of activity. To help counteract this, selections were performed with mixed peptides, which should favor enzymes with broad specificity. After selection, additional activity tests were performed on more complex sequences, and, most importantly, CML-BSA (Fig. 2E) which we confirmed by LC-MS has a diverse set of modified sites and structural contexts. We did discover some variants with improved activity on selection peptides that did not translate to more complex peptides and proteins and these were removed from consideration.

7. Lack of In Vivo Data: The study is purely in vitro/ex vivo; no animal models (e.g., diabetic mice with high CML) to assess pharmacokinetics, immunogenicity, or efficacy in living systems. H₂O₂ byproduct could cause oxidative stress in vivo, unaddressed here.

We thank the reviewer for the thoughtful comment regarding the in vitro and ex vivo models used in the paper. Our goal is to validate the novel enzymatic reaction and assess CMLase's chemical capabilities, prior to evaluating the *in vivo* efficacy. We have modified the discussion to address this as a limitation of the study that will be more closely examined in future work. We also address the question of potential hydrogen peroxide toxicity.

8. Incomplete Functional Restoration: While CML removal is shown, no assays test if it restores protein function(e.g., collagen elasticity, ECM signaling, RAGE binding inhibition). Pathogenic "hotspots" are mentioned but not prioritized or validated.

We thank the reviewer for bringing up this point about functional restoration of CML damaged proteins. We believe this is outside of the scope of the current work and have addressed this as a limitation of the study in the discussion.

9. Scope of CML Accessibility: Enzyme works on surface-exposed sites (correlates weakly with B-factor), but buried or crosslinked CML in dense ECM may resist repair. Variability across proteins (e.g., 52% on eye TPE vs. 97% on casein) highlights substrate dependence.

We thank the reviewer for this insightful observation regarding substrate dependence and steric accessibility. We agree that CMLase, like most enzymatic therapies, is primarily active on surface-exposed residues. We have revised the Discussion (paragraph 2) to explicitly acknowledge this limitation, stating: "efficacy varies across substrates (e.g., 97% on casein vs. 52% on eye TPE), suggesting that while surface-exposed CML is readily cleared, buried or crosslinked residues within dense ECM may resist enzymatic access."

10. Scalability and Therapeutics: Engineering focused on activity/stability, but no data on production yield, half-life, or delivery (e.g., to eyes/kidneys). As a large protein (~40 kDa inferred from homologs), systemic use faces barriers like renal clearance or immunogenicity.

We thank the reviewer for these important translational insights. We have revised the Discussion to explicitly acknowledge the challenges of therapeutic development, including immunogenicity considerations as well as potential product toxicity concerns.

11. Broader AGE Context: CML is one AGE; others (e.g., pentosidine crosslinks) aren't addressed, limiting impact on total AGE burden. Aging is multifactorial, so CML reversal alone may not "reverse aging."

We thank the reviewer for these comments. We agree that aging is a complex, multifactorial process and that CML represents only one component of the total AGE burden. We have revised the Discussion to explicitly state that CML reversal alone will not resolve the full aging phenotype. However, we emphasize that CMLase serves as a critical proof-of-concept, establishing a directed evolution platform that can now be adapted to target other pathogenic AGEs including both cross-linked AGEs (e.g. glucosepane, pentosidine) as well as non cross-linked AGEs (e.g. Argpyrimidine, carboxyethyllysine) to address the broader molecular landscape of aging.

12. Sample Limitations: Human tissues from limited donors/ages; no disease-specific samples (e.g., diabetic kidneys). Proteomics only on BSA, not tissues, missing real-world complexity.

We thank the reviewer for these comments on the limitations of selected samples. We agree that a more comprehensive assessment of activity across a variety of proteins and samples will be necessary to understand the full capabilities of the enzyme. As the basis of our study, we chose simple model proteins (BSA, Hemoglobin, Casein, and Collagen), as well as more complex targets (retinal total protein extract, human lens protein, aorta and skin samples). This sample variety, along with the diverse methods of assessing CML removal, make us confident in CMLase's ability to repair CML in a wide variety of contexts.

13. Technical Constraints: Genetic selection relies on E. coli, potentially missing eukaryotic folding needs. IHC quantification (DAB intensity) is semi-quantitative, prone to variability.

We thank the reviewer for raising these technical points.

Genetic Selection Host: We wish to clarify that the CMLase scaffold is derived from the bacterium *Calidithermus roseus* and does not require eukaryotic post-translational modifications (e.g., glycosylation) for activity. Therefore, *E. coli* serves as an appropriate host that accurately captures the enzyme's native folding requirements without the risk of missing eukaryotic-specific processing.

IHC Quantification: We acknowledge that DAB staining is semi-quantitative. To minimize variability, we processed all samples in parallel and analyzed multiple spatially aligned subregions. Crucially, the magnitude of CML reduction observed via IHC (~55–70%) is consistent with the fully quantitative reductions measured by our orthogonal methods (LC-MS/MS and ELISA) in homogenized samples. This cross-methodological consistency supports the validity of the histological data despite the inherent limitations of DAB.

Reviewer #2:

1. One issue throughout the manuscript related to a lack of clarity about how activity was assessed for the various CMLase enzymes. For example, minimal details were provided about the concentrations used for free vs peptide-CML in enzymatic assays. In later, critical experiemetns, I could not find details about the substrate concentration, identity or how the "loading" of CML on protein or peptide substrates was assessed. Greater clarity on these points would strengthen the conclusions of the manuscript

We thank the reviewer for pointing out the lack of clarity around enzyme and substrate concentrations in the manuscript. We have added more details about enzyme and substrate concentrations throughout the text and in the methods, especially "Peroxidase Assays" and "CML-modified protein ELISAs."

2. There seems to be a major potential for cross reactivity with cleavage at C-terminal Gly residues. Did the authors observe any truncations at the C-terminus? How was the potential for other, potentially unexpected, cross-reactivity evaluated? Can it work on N-terminal or Arg residues or modifications? I do not think I noticed any experiments to address the remaining (or potential for remaining) glycine oxidase activity, which could contribute to background H₂O₂ generation.

We thank the reviewer for pointing out potential cross reactivity with free amino acids and glycine terminal peptides. We added additional controls to quantify CMLase activity on natural free amino acids and several tripeptides, these data have been added to the manuscript as Fig S2-6 and S2-7. There is still residual glycine activity (Fig S2-3), however the K_M is significantly higher than the wildtype enzyme. At physiological concentrations, we expect this to be negligible.

3. On a related note, additional characterization for CMLase-mediated CML loss from proteins would strengthen the contribution. As an example - A western blot showing in vitro CML modified proteins, treated with and without CMLase would show that the enzyme isn't digesting the whole protein (by Coomasssie) and also decreases CML levels.

We thank the reviewer for bringing up the potential issue with protein substrates being digested, allowing the enzyme easier access. To verify true protein activity, we performed additional experiments to address the reviewer's concerns which have been included as Fig. S2-9 . Specifically, we ran an overnight reaction at identical conditions to the reactions analyzed by LC-MS and analyzed the products using SDS-PAGE to verify that the protein substrate was remaining intact.

4. Additional characterization should also be provided for the extent of CML modification introduced into in vitro proteins. The glyoxylic acid reductive amination makes sense, but what is the overall degree of modification? The LC-MS data presented seems to focus on a %removal from a set of CML-modified sites.

However, I find myself wondering if the initial extent of modification at each site (which was presumably not 100%, or was it?) affects the way one can interpret this data.

We thank the reviewer for this important question regarding the baseline modification levels. To clarify, the data in Figures 3B, 3C, and S3-3 are reported as the absolute percent abundance of the CML modification at each specific lysine residue. The samples treated with the catalytically inactive enzyme (CeGO-763) (represented by the lighter bars/points) serve as the control for the initial level of modification. As the reviewer suspected, the starting modification level is heterogeneous, ranging significantly from site to site rather than being uniformly 100%. Importantly, we analyzed the relationship between the initial modification burden and repair efficiency and found no significant correlation. CMLase demonstrated the ability to significantly reduce CML levels at both heavily modified and lightly modified sites, indicating that the initial degree of modification does not bias the interpretation of the enzyme's efficacy.

5. Given the large loading of CMLase to CML-modified protein (e.g. 5:1 CMLase to CML-BSA), it seems entirely plausible that some of the reduction observed by ELISA or immunohistochemistry could be due to blocking detection by Abs due to nonspecific (or even specific) binding to the substrate. This needs more careful controls throughout the paper.

We thank the reviewer for raising this critical control regarding potential epitope masking. To definitively rule out the possibility that the enzyme is simply blocking antibody access rather than catalyzing CML removal, we utilized a catalytically inactive enzyme variant (CeGO-763) as a negative control in all antibody-based assays. This inactive variant is structurally homologous to CMLase and was applied at the identical concentration (5 μ M) and incubation conditions. Crucially, the inactive control consistently showed no reduction in antibody binding signal compared to untreated samples. If steric blocking were the mechanism of signal loss, we would expect the inactive enzyme to produce a similar effect. Furthermore, our LC-MS/MS data (Figures 3B, 3C, and 4A) provides orthogonal validation, confirming that the loss of antibody signal corresponds to the physical chemical conversion of the CML adduct back to native lysine, independent of antibody binding dynamics.

6. Some discussion of the efficiency of the CMLase to other enzymes (and the native activity) should be included

We have revised the Discussion to explicitly contextualize the catalytic efficiency of CMLase. We note that while its rate is lower than that of naturally evolved metabolic enzymes, it compares favorably to other early-generation therapeutic enzymes engineered to act on non-native substrates. We also highlight the substantial fold-improvement achieved over the starting scaffold.

7. Based on the data provided, I don't think that the following sentence is fully justified. "We show the new-to-nature CMLase repairs a majority of surface CML modifications on proteins and then mapped the specific sites of repair using proteomics." In particular "majority" seems inaccurate, as it definitely works better on some proteins than others as seen in figure 3.

We thank the reviewer for the feedback surrounding certain claims made in the discussion. We have softened this statement to avoid overstating the efficacy of the enzyme. Additionally, we have expanded the discussion to further qualify the results and address shortcomings of the study so the reader can come to their own conclusions.

Reviewers #3 and #4:

We would like to thank reviewers #3 and #4 for co-reviewing the manuscript. We appreciate your time and consideration.

Reviewer #5:

1. Overall, the style of writing in the article is more narrative than scientific, and does not provide deeper information regarding the role of CML in aging and diseases.

We thank the reviewer for this feedback. We have revised the manuscript to ensure a more concise and objective scientific tone. Additionally, we have inserted subheadings within the main text to improve readability and structure. To address the request for deeper biological context, we have significantly expanded the Background and Discussion sections. These updates now provide a more thorough review of CML's pathogenic role—specifically detailing the CML-RAGE signaling axis and structural ECM damage—and contextualize how this study fits into the larger landscape of aging biology.

2. The Background section requires more information on the advanced glycation end products (AGE), aging, age-related functional changes, and age-related diseases.

We thank the reviewer for this suggestion. We agree that providing a broader context regarding the pathogenic role of AGEs in aging and disease strengthens the rationale for developing enzymatic reversal strategies.

We have revised the Background section (Paragraphs 1 and 2) to include a more comprehensive overview of AGE-mediated pathophysiology. Specifically, we have expanded the text to:

Distinguish between mechanical and signaling pathologies: We now explicitly differentiate between the structural stiffening caused by crosslinking AGEs (affecting vascular compliance and skin elasticity) and the inflammatory signaling caused by non-crosslinking adducts like CML.

List specific age-related diseases: We have added references to atherosclerosis, diabetic nephropathy, and neurodegenerative conditions to highlight the clinical stakes of AGE accumulation.

3. In general, the Results section is a mixture of the description of the results including some of the methods and discussions, and should focus on just on the description of the results.

We thank the reviewer for this feedback regarding the structure of the manuscript. We have carefully reviewed the text and revised the Results, Methods, and Discussion sections to better segregate these elements. Specifically, we have moved some technical protocols to the Methods where appropriate and interpretative commentary to the Discussion, ensuring the Results section remains focused on the description of experimental outcomes while maintaining the necessary logical flow of the engineering process.

4. The first paragraph of the Results section needs to be removed or moved to the Discussion section.

We appreciate the reviewer's suggestion to streamline the text. However, we respectfully prefer to retain this paragraph in the Results section because it describes specific experimental steps and negative data that were critical to our study design—specifically, the empirical testing and failure of the commercially available *Cladosporium* lysate to exhibit CML activity in our hands. These results were the decisive factor that necessitated the computational mining approach described in the subsequent text. We believe that reporting this experimental outcome is essential to the logical flow of the engineering narrative.

5. The Results section should include sub-headlines to make it easier for the readers to follow.

We thank the reviewer for the suggestion to include subheadlines to make the Results more clear. We have added these throughout the results section.

6. The Figures and Figure Legend need to be separated from the body text.

We thank the reviewer for the suggestion to update the format of the figures in the text. All figures and legends have been moved out of the main text to the end.

7. The font in Figure 1B is too small and hard to see.

We thank the reviewer for pointing out this formatting issue with Fig 1B. The figure has been updated with an increased font size.

8. Although the authors stated that there are similar degrees of modification in various proteins, i.e., BSA, collagen, hemoglobin, casein, and eye total protein extract, using different glyoxylic acid concentrations, the degree of modification seems to vary in various proteins (Figure 3A).

We thank the reviewer for bringing up their concern regarding different modification levels in the ELISA data. Modification conditions were tuned to get similar levels of signal in the ELISA, however it was challenging to get them identical, likely due to the difference in ease of glycation and potentially antibody bias. Figure S3-2 shows all modification conditions for each protein and we selected the modification level closest for each.

9. The experiments were conducted using different proteins and tissues, which seem to yield the data that support the authors' hypothesis. Unfortunately, it is a bit hard to reach the same conclusion as the authors because of the inconsistency of the usage of various proteins and tissues throughout the entire experiments.

We thank the reviewer for this important observation regarding the diversity of substrates employed in our study. We understand that the use of varying tissues and proteins may appear inconsistent; however, this was a deliberate design choice to progressively test CMLase on increasingly complex substrates. We believe that observing consistent activity across this spectrum—from peptides and recombinant proteins to human solid tissues—using orthogonal techniques (fluorescence assays, quantitative LC-MS/MS, proteomics, and IHC) strengthens the conclusion that CMLase is robust, independent of the model system. We have updated the main text and Discussion to clarify this rationale and address the limitations of specific samples.

10. The immunohistochemical analyses were conducted for the quantitative analysis of CML in human skin and the effects of CMLase treatment. The DAB intensity may not be the ideal way to quantitate the amount of antigen in the tissue. In addition, some of the immunostaining results possibly show over-staining and background staining, which could make quantitative analyses more difficult. Furthermore, the figure is too small to see the details as well.

We appreciate the reviewer's comment regarding the impact of tissue fixation on enzymatic activity. We selected FFPE sections specifically to leverage the availability of high-quality, age-stratified human tissue banks. However, we agree that fixation introduces steric complexity. To address the need for confirmation in non-fixed environments, we direct the reviewer to our experiments on human lens homogenates and soluble protein substrates (e.g., BSA, collagen, hemoglobin).

Confirmation in Non-Fixed Samples: As detailed in Figure 4A-B and Figure 3, we demonstrated robust CMLase activity on non-fixed, soluble samples using orthogonal detection methods (LC-MS/MS and ELISA). These experiments confirm that CMLase is highly active on CML modifications in their native state, independent of fixation-induced cross-linking.

FFPE as a Stringent Test: Formalin fixation induces extensive methylene bridge cross-linking, which sterically hinders enzyme access. Consequently, the FFPE experiments serve as a rigorous "stress test" for the enzyme. The fact that CMLase successfully reversed CML modifications within this highly cross-linked matrix demonstrates its ability to penetrate dense tissue structures.

Methodological Controls: To ensure the signal was not an artifact of sample preparation (e.g., heat-induced CML formation), we strictly omitted antigen retrieval steps in accordance with established protocols for detecting AGEs in fixed tissue (Ref. 28).

We believe the combination of quantitative LC-MS/MS data on non-fixed lens tissue and the high-resolution spatial data provided by FFPE histology provides a comprehensive validation of CMLase activity across diverse protein environments.

11. It is not clear how the specificity of the antibodies and specific epitope of the antigen were determined. It is also not clear whether the authors have examined if the less (or no) staining after CMLase treatment was due to the direct effects or potential secondary effects by CMLase.

We thank the reviewer for these rigorous controls. We have addressed both points experimentally:

Antibody Specificity: As described in the Results (and shown in Fig. 4C), we validated the specificity of the 6D12 monoclonal antibody using a competitive inhibition assay. Pre-incubation of the antibody with CML-BSA completely abolished tissue staining, whereas pre-incubation with unmodified BSA did not. This confirms that the antibody specifically recognizes the CML epitope rather than non-specific tissue components. In addition, the antibody has been extensively characterized by the manufacturer.

Ruling Out Secondary Effects: To confirm that the reduction in staining was caused by the direct enzymatic reversal of CML—rather than secondary effects such as epitope masking, steric hindrance by the enzyme, or non-specific proteolysis—we utilized a catalytically inactive enzyme control (CeGO-763). This variant was purified using the exact same protocol but lacks oxidative activity. In our IHC experiments, the inactive control did not reduce CML staining, whereas CMLase did. Furthermore, our LC-MS/MS data (Fig. 3B, 4A) provides definitive chemical proof that the CML adduct is converted back to native lysine, confirming that the loss of antibody signal corresponds to the physical removal of the antigen.

12. It is not clear whether the immunohistochemical analyses using formalin fixed paraffin embedded sections were ideal to test the enzyme activities on CML. Further confirmation may be required with frozen sections with/without fixation.

We thank the reviewer for raising this important technical point. We selected FFPE sections as our primary histological model because they represent the standard for human tissue banking and offer superior morphological preservation, which is critical for spatially resolved quantification. Regarding the comparison with frozen sections, we wish to highlight the following experimental validations:

Evaluation of Frozen Sections: We did attempt to perform these assays on fresh frozen tissue. While we observed qualitatively similar enzymatic reversal of CML, the fresh frozen samples showed significant tissue

fragmentation and loss of structural integrity. This rendered them unsuitable for the rigorous spatial quantification presented in the manuscript. The structural stability provided by FFPE was therefore necessary to maintain tissue architecture during the reaction.

Biological Validation (Young vs. Old): To confirm that the signal observed in FFPE was specific to age-accumulated CML and not a fixation artifact, we utilized tissue from young donors (20–25 years old) as biological controls. As shown in Figure S4, young FFPE sections displayed negligible CML staining compared to elderly donors, consistent with the known accumulation kinetics of CML. This confirms that our FFPE protocol preserves the antigen's biological specificity.

Confirmation on Non-Fixed Substrates: To address the concern about activity on non-fixed substrates, we point to our data on soluble proteins and human lens homogenates (Fig. 3 and 4A-B). These assays demonstrated robust CMLase activity on non-crosslinked, non-fixed samples using orthogonal detection methods (LC-MS/MS and ELISA), confirming that the enzyme is highly active on CML in its native state.

We believe the combination of the quantitative soluble data (non-fixed) and the spatially resolved FFPE data (fixed) provides the most comprehensive evaluation of CMLase activity.

13. There is no clear statement regarding how many samples were used in many of the experiments, and it is not clear whether there was strong evidence of reproducibility for each experiment.

We thank the reviewer for raising this important question about reproducibility. All experiments were done in triplicate and wherever possible all data points were shown (Fig 3C, 4A, B, and E). Some information regarding the number of replicates has been moved from the methods to the figure legends to clarify specific data (Fig 1D, E and 2E).

14. The Discussion section is too short, and there is not much information provided.

We thank the reviewer for this suggestion. We have significantly expanded the Discussion to provide a more comprehensive analysis of our findings and their context within the field. Specifically, we have added:

Enzymatic Context: A detailed comparison of CMLase efficiency relative to native metabolic enzymes and other therapeutic enzymes, explicitly highlighting the substantial fold-improvement achieved through directed evolution.

Translational Limitations: A dedicated assessment of the barriers to clinical translation, including the challenges of potential immunogenicity.

Safety & Metabolism: Clarification regarding the safety profile of the reaction, specifically citing the endogenous metabolic pathways responsible for clearing the reaction byproducts (glyoxylate and hydrogen peroxide).

We believe these additions provide the necessary depth and balance requested by the reviewer.

Reviewer #6:

Glyoxylic acid is released in the deglycation process. This also seems to be the reagent used to generate CML proteins in the methods. From a quick read, it seems like the body should be able to metabolize

glyoxylate and keep it at reasonable levels. The authors should comment on this- are there potential issues with freeing glyoxylic acid? Simply providing references for how it gets metabolized could suffice.

We thank the reviewer for raising this important safety consideration regarding reaction byproducts. We have revised the Discussion to explicitly address released glyoxylic acid. The text has been revised with references to note that glyoxylate is a naturally occurring metabolite efficiently cleared by endogenous enzymes (e.g., glyoxylate reductase/hydroxypyruvate reductase) and activity of CMLase would yield a negligible increase in glyoxylate compared to physiological levels.

Panel 2E. The rate is slow (uM/hour). In the legend for Fig 2E, please state the concentration of enzyme used and state how the data are collected (are the plotted values averages? How many replicates?)

We thank the reviewer for pointing out the lack of clarity in Fig. 2E. It has been updated to include concentrations of enzyme, substrates, and replicates.

First round of evolution: error prone PCR. In the paragraph describing error prone PCR- how many rounds of error prone PCR and selection did the authors do? It seems like it was 1 round, but please clarify. 7 mutations is a lot for one round, but I see that the authors aimed for 4-8 mutations per gene, so I realize this could have been the hit from 1 round of selection. What was the estimated library size and coverage?

We thank the reviewer for this comment regarding the error prone PCR in the first round of evolution. The first variant, CrGO-764, was found from a single aggressive library with 4-8 mutations per gene and the highest activity hit had 7 mutations. The text has been edited to clarify these points. The library size was approximately 5×10^7 and was plated with 2x coverage, hitting the upper limit of the selection 10^8 cells per plate. The methods have been updated to include this information.

Exciting that this yielded a 5-fold improvement. Please list the mutations in this paragraph (I see the authors reference table Fig S2-1, but would be helpful to see it listed here, too).

We thank the reviewer for the suggestion to include the mutations of CrGO-764 into the main text. We have added in the full list of mutations to improve clarity and consistency.

How many colonies did the authors screen in each round of evolution? Please add this information to the methods or Figure S2-4 (to explain how you got to the estimated number of screening over 500 million protein variants as stated in the abstract).

We thank the reviewer for requesting this clarification regarding our screening throughput. We have updated the Library Transformation section of the Methods to explicitly detail our library sizes. Specifically, we clarify that we generated libraries with an average diversity of $\sim 5 \times 10^7$ variants per round. Over the course of the project, we screened more than ten such libraries—including several exploratory, "dead-end" lineages not detailed in the main text—resulting in a cumulative screening total exceeding 5×10^8 variants.

Can you further explain the rationale for picking residues guided by CeGO to saturate the CrGO? Is it surprising that mutations that improve activity on free CML also improve activity on Peptidyl CML?

We appreciate the reviewer's question regarding the reasoning behind modelling a library after the active site of CeGO. We wanted to increase affinity for the CML side chain without affecting the more open active site, therefore we targeted residues inside the active site while leaving the surface residues untouched to retain peptide and protein activity. This has been expanded on in the text.

The results are very interesting and could be sold more explicitly in the discussion. Discuss the overall fold-improvement in enzyme activity in the discussion – this would strengthen the results of the paper.

We appreciate the reviewer's positive assessment of our results. We agree that emphasizing the magnitude of the engineering success strengthens the narrative. We have revised the Discussion to explicitly highlight the substantial fold-improvement in catalytic efficiency achieved through our five rounds of directed evolution, thereby emphasizing the efficacy of the engineering campaign.

When you tested activity on CML-BSA, why did you use low concentrations of CML-BSA (1 μ M) with 5 μ M CMLase? How many turnovers do you think the CMLase can do?

We thank the reviewer for this question regarding our assay conditions. We utilized 1 μ M BSA because the protein is heavily modified, resulting in an effective concentration of CML substrate estimated to be >10 μ M (due to multiple CML modifications per BSA molecule). We employed a relatively high concentration of CMLase (5 μ M) to ensure that the rate of hydrogen peroxide production was sufficient for detection in our coupled peroxidase assay within a short timeframe (hours), given that the reaction velocity on bulky protein substrates is slower than on free peptide. Regarding turnover capability, the enzyme is clearly catalytic rather than stoichiometric; on free CML substrates, we routinely observe more than 80 turnovers, indicating that the limiting factor in the CML-BSA assay is steric accessibility or diffusion rather than enzyme inactivation.

please provide information on how the selection strain was generated.

We thank the reviewer for pointing out this unintentional omission. The LysA knockout strain was purchased from Horizon (JW2806). The methods have been updated to include the supplier and catalog number.

Can the authors provide more detailed methods for how the selection was done? What were the transformation efficiencies in each round, and what is the estimated coverage based on this? What temperature were the cells grown at during the selection (I see this information for the transformation)? When did the authors start to see colonies for the different rounds of selection?

We thank the reviewer for these comments regarding the genetic selection methods. Selection plates were grown at 30 °C for up to one week. Colonies were picked as soon as they reached sufficient size, which was generally 1-7 days after plating, depending on the library and selection conditions. We have updated the "Genetic Selection" section of the methods to include more detailed information on growth temperature and time for selection plates. Additionally, we expanded on library transformation efficiency and estimated coverage in the "Library Transformation" section.

Sequence context: I think it's great that Fig 3c highlights structural diversity of target sites and some sequence contexts. However, a more comprehensive depiction of sequence context would help a lot. Figure 3B- can the authors convert this information into sequence context information somehow (maybe logos plot?). This would help audience/end user to see if there are any limitations. A logos plot or some visual depiction of the sequence bias (or lack thereof) of enzymatic deglycation would be great.

We thank the reviewer for requesting additional sequence context for the LC-MS data. It was originally omitted due to the complexity of the sequences and lack of a specific motif, however we have added the additional analysis as a sequence logos plot for additional information. This has been added as Fig S3-5. In brief, to evaluate the relationship between the local sequence environment and CMLase efficiency, the amino acid context surrounding CML-modified Lysine (CML-K) sites within BSA (± 5 aa) was analyzed. While no rigid

primary sequence motif was identified, local composition appears to influence accessibility. The most prominent differentiator was the distribution of Cysteine (C) residues; high-efficiency sites were frequently associated with flanking Cysteines, whereas Cysteine was entirely absent from the most resistant sites ($p < 0.001$, Fisher's Exact Test). This may reflect higher CMLase activity near more accessible domain boundaries or disulfide-linked regions. Beyond this, the resistant cohort indicated a potential shift toward a more hydrophobic microenvironment. We observed an increased density of hydrophobic residues and a subtle enrichment of Phenylalanine (F) on the downstream side of sites with lower CML reduction. These observations suggest that resistant modifications may reside within hydrophobic clusters or aromatic-rich regions that provide physical shielding from the enzyme. It remains unclear whether these residues directly obstruct enzymatic access, interfere with the catalytic process post-binding, or reflect BSA-specific structural constraints. While these trends highlight potential topographical barriers within the globular BSA fold, we remain cautious in generalizing these residue preferences to other protein architectures or to structural changes induced by the CML modification itself. A deeper exploration of these structural determinants across a wider set of proteins is beyond the scope of this study but remains an objective for future investigation.

We have updated the Results and Discussion sections to briefly include a more granular analysis of the local sequence environment.

Sequence context: When the authors did the error prone library and challenged for growth on mixtures of peptides- were cells selected on all peptides at once? What are the sequences of these peptides? In this case, do you worry that a particular variant could deglycate only a subset of the peptides? How do you modulate stringency? I think it would help if you added data for your final variant testing kinetics on a range of peptide contexts (At least the ones from the mixture you provided in the selection). In the end, you get variants with a 20% improvement on peptide, but you don't say which peptide. Also, Figure S2-6 doesn't say which peptide (sorry if I missed it).

We thank the reviewer for bringing up concerns regarding the peptide selection. The library was selected on a mix of all peptides at once (AAX-CML-AA-CML and AA-CML-XAA-CML, where X is every amino acid except R, K, or H) at a concentration such that enzyme activity needed to be significant on >50% of these sequences in order to complement growth. The enzyme was then assessed on a variety of peptide sequences, however our key indicator of both accommodation of protein substrate and diverse sequence context was the significant improvement in activity on CML-BSA (Fig. 2E) which would have been extremely unlikely if the enzyme had become hyper specific for specific simple sequences. In previous selections of the same library on single peptides, this effect was observed: significant improvements on the selection peptide with decreases on other peptides and a significant decrease in CML-BSA activity; these hits were removed from consideration. CMLase's broad specificity is further supported by the LC-MS data in Fig. 3C that demonstrates broad specificity with no trends in sequence preference. Stringency of the selection is modulated by changing total substrate concentration (keeping all peptides at equal concentrations) or IPTG concentration to alter the expression. Figure S2-4 has been updated to specify peptide sequences.

Sequence context: Would be great to have sequence context for all residues from S4-3. Can the authors add that to the graph?

We thank the reviewer for the suggestion to add additional sequence context to S3-3. We appended two sequence logos plots in S3-5 to address the need to present more sequence trends.

Fig 1c, Supplemental Figures S1-2 and 2-2. Please include methods for how the CML substrate was modeled in the active site. What docking algorithm or program did the authors use? Was this done manually?

We thank the reviewer for this question about docking methods and structural analysis. We have added a "Structural Analysis" section to the methods which includes information about the docking program that was used (Bioluminate).

Response to Reviewer Comments:

We would like to thank the reviewers and the editor for their continued engagement with our manuscript. We are pleased that Reviewers #1 and #6 found the revised manuscript satisfactory and recommended acceptance. Below we provide detailed point-by-point responses to the remaining comments from Reviewers #2 and #5. We have performed the two new experiments requested by Reviewer #2 and have addressed all tractable points raised by Reviewer #5, noting where the editor's framing of this work as a proof-of-concept study informs our scope.

Reviewer #1:

I have carefully reviewed the authors' point-by-point response and the revised manuscript. Overall, the authors have responded logically, thoroughly, and satisfactorily to my comments. They addressed each concern with appropriate new data, clarifications, textual revisions, and explicit additions to the limitations section without dismissing any issues or weakening the core scientific claims. The revisions have markedly improved the rigor, clarity, and scientific tone of the paper. [...] In summary, the authors have produced a high-quality revision that directly incorporates my feedback, adds necessary controls and data, and appropriately qualifies claims while preserving the novel and impactful nature of the work. The manuscript is now more concise, scientifically toned, transparent about limitations, and substantially stronger. I recommend acceptance with only minor copy-editing if needed.

We thank Reviewer #1 for their thorough re-evaluation and positive assessment of the revised manuscript. We are pleased that our revisions adequately addressed all of their concerns.

Reviewer #2:

This manuscript describes the development of an engineered glycine oxidase that has new activity, enabling the enzymatic removal of CML from modified proteins. Overall, the edits to the manuscript have improved its quality, though they have also somewhat lowered the significance as the CMLase does not appear very active. Still, the authors have increased the clarity of their findings and addressed concerns about modification levels impacting enzymatic efficiency and epitope masking in the immunochemistry data, which is a big improvement. However, there are still concerns about the characterization of CMLase activity, including the potential for cross-reactivity to other AGE adducts and the lack of a western blot showing removal of CML. The latter point is more critical, and if addressed would make this submission suitable for acceptance in Nature Communications. Detailed comments are below:

We thank the reviewer for their continued constructive engagement and for clearly identifying the two remaining items needed for acceptance. We have performed both requested experiments and describe the results in detail below.

1. The authors seemed to have misunderstood our prior question about activity on arginine residue modifications. Specifically, we wonder if the CMLase works on the equivalent arginine glycation adduct (CMA) and/or the equivalent modification on the N-terminus of peptides/proteins.

We apologize for the misunderstanding in our previous response. We now appreciate that the reviewer was specifically asking about N ω -carboxymethylarginine (CMA), the arginine-equivalent of CML, and we have directly tested CMLase activity on a CMA-modified protein substrate.

Activity on CMA: CMA-BSA was prepared by incubating BSA with glyoxylic acid and NaCNBH₃, then incubated overnight with active CMLase, inactive enzyme, or buffer alone. CMA levels were quantified by LC-MS/MS following enzymatic hydrolysis. CMLase treatment did not produce a statistically significant decrease in CMA compared to either the buffer-only or inactive enzyme controls, indicating that the carboxymethyl-guanidinium moiety of CMA is not a substrate for CMLase under conditions where the enzyme robustly oxidizes the chemically related carboxymethyl-amino moiety of CML. This finding reinforces that CMLase is selective for the lysine-derived CML adduct rather than acting broadly on carboxymethylated amino acids. We note that this selectivity may be advantageous in a therapeutic context, as it limits the spectrum of off-target modifications the enzyme could act upon. These data have been added as a new panel to Figure S2-8, and the Results and Methods sections have been updated accordingly.

Regarding the second part of the reviewer's question, on N-terminal carboxymethylation: we have not directly tested an α -amine carboxymethylated substrate in the present study. Physiologically relevant CML accumulation in long-lived ECM proteins occurs predominantly at internal lysine ϵ -amines rather than at protein N-termini, and our proteomic analysis (Fig. 3B) characterizes CMLase activity across thirty-three internal lysine sites, which we view as the most disease-relevant population. We agree that N-terminal activity is an interesting question for future characterization and have noted this in the Discussion.

2. In figure S2-9 the treated CML-BSA lane seems to be closer to the fully modified than the unmodified BSA. Maybe using a substrate that gets closer to 100% recovery would be better to show in this data. As suggested before, we would also like to see a Western blot for the associated gel as an additional confirmation of reduction in CML modifications after incubation with the CMLase since the Coomassie doesn't show conclusive evidence of removal of adducts.

We thank the reviewer for this important point and address both halves of the comment below.

First, we wish to clarify that the Coomassie-stained gel in Figure S2-9 was included specifically to demonstrate that CMLase does not fragment or degrade the BSA substrate — i.e., to confirm protein integrity rather than to quantify CML removal. As noted by the reviewer, CML adducts do not produce a highly obvious mass shift by SDS-PAGE, so the Coomassie gel was never intended as evidence of deglycation.

Regarding the suggestion to use a substrate that approaches 100% recovery: as shown in Figure 3A, CMLase achieves up to 97% reduction in CML ELISA signal on casein, indicating near-complete deglycation is achievable on certain substrates. The residual signal on BSA is consistent with our site-resolved proteomics analysis (Fig. 3B), which identified a small number of CMLase-resistant lysines (notably K131, K504, and K573) embedded in less accessible structural environments. We retained BSA as the substrate for Figure S2-9 because its well-characterized modification landscape and partial recalcitrance make it a more stringent test of enzyme activity, and because the Western blot described below directly addresses the reviewer's core question of whether CML is in fact being removed.

Western blot for CML removal: To demonstrate CML removal in an orthogonal assay to those described in our results, we performed the requested Western blot. CML-BSA was treated in duplicate with 1 μ M active CMLase or 1 μ M catalytically inactive control enzyme for 16 hours, then analyzed by automated capillary-based immunoassay (Jess Simple Western, Bio-Techne) using an anti-CML primary antibody (Abcam, ab27684). A CML-BSA standard curve (50, 125, 500, and 1500 pg, run in duplicate) was included to confirm linearity of detection across the assay's dynamic range, and unmodified BSA (50 and 500 pg, run in

duplicate) was included as a negative control to confirm antibody specificity for the CML adduct. We note that the standard curve was included as a linear-range reference for the immunoassay rather than as a direct quantitative anchor for the treated samples, since the standard curve and treatment samples were prepared via independent dilution routes prior to loading; the controlled quantitative comparison is between the active CMLase- and inactive-enzyme-treated lanes, both of which were prepared, incubated, and loaded under identical conditions.

Anti-CML signal was observed predominantly at the expected ~66 kDa molecular weight for the BSA monomer, with additional higher-molecular weight bands at ~120 and ~180 kDa consistent with glycation-induced BSA dimers and aggregates characteristic of CML-BSA preparations; no signal was detected in the unmodified BSA lanes. CMLase-treated lanes showed a clear and reproducible reduction in anti-CML signal at the BSA monomer band and a near-complete loss of the higher-molecular-weight species in both replicates compared to signal in lanes treated with inactive enzyme. Enzyme-only lanes (active and inactive enzymes without CML-BSA substrate) produced no anti-CML signal, confirming that the antibody does not cross-react with the enzyme. Together, these results directly confirm that the reduction in immunoreactivity observed by ELISA is due to enzymatic removal of the CML adduct. Full experimental details have been added to the Methods, the uncropped Western blot image has been added as new Supplemental Figure S2-11, and the raw data will be provided in a Supplementary Data File in accordance with Nature Communications guidelines. The Results section has been updated accordingly.

Reviewer #3:

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 participation in peer review as part of the Nature Communications early career researcher initiative.

Reviewer #4:

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 #4 for their participation in peer review as part of the Nature Communications early career researcher initiative.

Reviewer #5:

General: This revised manuscript describes the development of CMLase, which reversed carboxymethyl-lysine in human tissues obtained from elderly donors, and represents the first demonstration that "irreversible" protein damage can be enzymatically repaired. The authors conclude that this newly developed enzyme, along with its observed effects, could serve as a platform to develop new enzymes capable of reversing age-related molecular damage and repairing tissue proteins compromised by aging and disease. Although this revised manuscript addressed the previous critiques, the experimental design still has several potential major flaws. Further verification with reproducibility may still be needed, and the manuscript may require more focus on a specific area of aging or age-related disease instead of covering many areas in age-related pathophysiology.

This causes a loss of focus and makes the manuscript somewhat confusing. Therefore, it is difficult to support the authors' conclusion based on the content of this manuscript.

We thank Reviewer #5 for their continued engagement with our manuscript. We would like to note, respectfully, that many of the concerns raised in this second round were substantively addressed in our previous revision, with specific textual changes, new data, and expanded discussion provided in our prior point-by-point response. We direct the reviewer to those responses, which remain applicable. We also note the editor's characterization of this work as a "proof-of-concept, initial study, with guided enzyme evolution and enzymology insights" which informs the scope of our responses below. We have addressed each specific comment in turn.

1. Overall, this manuscript does not provide deeper scientific information regarding the role of CML in aging and diseases.

We respectfully note that the primary focus of this manuscript is the enzyme engineering and biochemical characterization of a novel deglycation activity, not a comprehensive review of CML biology in disease. As we described in our previous revision, we have expanded the Background and Discussion to provide appropriate context for CML as a pathogenic AGE modification linked to aging and age-related disease, with supporting references. The manuscript is not, by design, a clinical or pathological study of CML, and the editor has framed the work accordingly. We believe the current level of biological context is appropriate for the scope of the study and have made targeted further additions to the Introduction to clarify the significance of CML accumulation in aging.

2. The Background section requires more focus on the role of CML accumulation in aging, age-related functional changes, or specific age-related disease. The current format addresses various aspects of age-related changes, making it difficult to clearly understand the purpose of this project.

We thank the reviewer for this suggestion. In our previous revision, we expanded the Background to include additional context on CML accumulation in aging and its functional consequences. We acknowledge that the Background spans several aspects of AGE biology to motivate the broader relevance of our platform; however, we have now made targeted edits to sharpen the narrative arc of the Introduction, more explicitly connecting each element to the central purpose of the study: the development of a proof-of-concept enzymatic approach to reverse CML modification.

3. The Results section still contains a mixture of the results description, and some of the methods and discussions, which should be revised to only focus on the description of the results.

We thank the reviewer for this editorial point and have made an additional targeted pass through the Results section to remove interpretive language and methodological justification to the Discussion and Methods, respectively. We note that some contextual framing within the Results is standard practice in the field and aids reader comprehension, but we have minimized this where possible. The revised Results section now focuses more strictly on the presentation of experimental findings.

4. Although the authors reported similar degrees of modification across various proteins, i.e., BSA, collagen, hemoglobin, casein, and eye total protein extract, it is not clear why the authors selected only these proteins for the study.

We thank the reviewer for this comment; this was addressed in detail in our previous revision. Briefly, our protein panel was selected to span a wide range of structural complexity and

biological relevance: BSA and casein serve as accessible model substrates with well-characterized glycation chemistry; hemoglobin and collagen represent physiologically important proteins known to accumulate CML in aging; and retinal/eye total protein extract provides a complex, clinically relevant matrix. This rationale is now explicitly stated in the Results section (under the protein substrate characterization subsection). We believe this panel provides an appropriate foundation for a proof-of-concept study and note that the editor has framed the work in this context.

5. The experiments were conducted using different proteins, tissues, and species (including human samples), which yielded data that supports the authors' hypothesis. However, it is somewhat difficult to reach the same conclusion as the authors because of the inconsistencies in the use of various proteins and tissues across the experiments.

We thank the reviewer for this comment, which was similarly raised and addressed in our previous revision. The use of diverse substrates across experiments was a deliberate design choice to demonstrate that CMLase activity is robust across different protein scaffolds and tissue environments, rather than being an artifact of a single model system. The experimental logic is that each figure builds progressively from simple model substrates (free CML, peptides, recombinant proteins) to increasingly complex biological samples (homogenized lens, human FFPE tissue sections). Different proteins and tissues are used in different figures because each experiment is designed to answer a specific question at a specific level of complexity. We have updated the Results section to more explicitly state this progressive experimental logic at the beginning of each major subsection to help the reader follow the structure of the study.

6. Immunohistochemical analyses were conducted for the quantitative analysis of CML in human arterial tissue and to evaluate the effects of CMLase treatment. DAB intensity may not be the ideal way to quantitate the amount of antigen in the tissue. In addition, some of the immunostaining results appear to have over-staining and high background, which could make quantitative analyses more difficult.

We thank the reviewer for raising this point, which was also raised and addressed in our previous revision. As described therein, all IHC experiments were conducted with matched antibody concentrations, parallel processing, and analysis of multiple spatially defined subregions per section. DAB quantification was performed by densitometric analysis using ImageJ with background subtraction. The magnitude of CML reduction observed by IHC (~55–70%) is consistent with the quantitative reductions measured by orthogonal methods (LC-MS/MS and ELISA) on homogenized samples, providing cross-methodological validation. We acknowledge that DAB is semi-quantitative, and this limitation is explicitly stated in the Discussion. Regarding the concern about over-staining, representative images were selected to illustrate clearly stained tissue regions, and all images were acquired and processed under identical conditions.

7. It is not clear whether the immunohistochemical analyses using formalin fixed paraffin embedded sections are ideal to test enzyme activity on CML. Further confirmation may be required using frozen sections with/without fixation.

We thank the reviewer for this comment, which was addressed in our previous revision. As described therein, we did attempt experiments on fresh frozen tissue sections; however, these samples exhibited significant fragmentation and loss of structural integrity during the enzyme incubation step, rendering them unsuitable for rigorous spatial quantification. We therefore selected FFPE sections as the primary histological model. FFPE also serves as a stringent stress test of CMLase activity, given that formalin fixation introduces extensive methylene

bridge crosslinking that sterically restricts enzyme access. The fact that CMLase retains activity in this highly crosslinked matrix is, if anything, conservative evidence of its capability. Importantly, robust CMLase activity on non-fixed substrates is independently confirmed by our ELISA and LC-MS/MS data on soluble protein substrates and human lens homogenates (Figures 3 and 4A-B). We have added a brief note in the Methods to reiterate why FFPE was selected and to reference these complementary non-fixed experiments.

8. There is no clear statement on the number of samples used in many experiments, and there is no strong evidence of reproducibility for each experiment.

We thank the reviewer for this comment. As described in our previous revision, sample sizes (n values) are reported in the figure legends and Methods. We have done an additional systematic pass through all figure legends to confirm that n values, any error bar definitions as needed (mean $\pm$ SD or SEM as applicable), and statistical tests are explicitly stated for every panel. For human tissue experiments, the number of donors and sections analyzed per donor are also now explicitly stated. These additions address the reproducibility concern and meet the reporting standards required by Nature Communications.

Reviewer #6:

The authors have addressed my concerns regarding the clarity of the evolution process and the enzyme characterization. While the evolved enzyme is on the slow side, I think the engineering challenge and the improvement over the starting point is notable and interesting. The authors have also tempered their claims about the impact on aging, which is good. I am fine with this manuscript moving forward.

We thank Reviewer #6 for their positive re-evaluation and support for publication. We are pleased that our revisions addressed all of their concerns.

Response to Reviewer Comments (Revision #2)

We would like to thank the reviewers and the editor for their continued engagement with our manuscript. We are pleased that Reviewers #1 and #6 found the revised manuscript satisfactory and recommended acceptance. Below we provide detailed point-by-point responses to the remaining comments from Reviewers #2 and #5. We have performed the two new experiments requested by Reviewer #2 and have addressed all tractable points raised by Reviewer #5, noting where the editor's framing of this work as a proof-of-concept study informs our scope.

Reviewer #1:

I have carefully reviewed the authors' point-by-point response and the revised manuscript. Overall, the authors have responded logically, thoroughly, and satisfactorily to my comments. They addressed each concern with appropriate new data, clarifications, textual revisions, and explicit additions to the limitations section without dismissing any issues or weakening the core scientific claims. The revisions have markedly improved the rigor, clarity, and scientific tone of the paper. [...] In summary, the authors have produced a high-quality revision that directly incorporates my feedback, adds necessary controls and data, and appropriately qualifies claims while preserving the novel and impactful nature of the work. The manuscript is now more concise, scientifically toned, transparent about limitations, and substantially stronger. I recommend acceptance with only minor copy-editing if needed.

We thank Reviewer #1 for their thorough re-evaluation and positive assessment of the revised manuscript. We are pleased that our revisions adequately addressed all of their concerns.

Reviewer #2:

This manuscript describes the development of an engineered glycine oxidase that has new activity, enabling the enzymatic removal of CML from modified proteins. Overall, the edits to the manuscript have improved its quality, though they have also somewhat lowered the significance as the CMLase does not appear very active. Still, the authors have increased the clarity of their findings and addressed concerns about modification levels impacting enzymatic efficiency and epitope masking in the immunochemistry data, which is a big improvement. However, there are still concerns about the characterization of CMLase activity, including the potential for cross-reactivity to other AGE adducts and the lack of a western blot showing removal of CML. The latter point is more critical, and if addressed would make this submission suitable for acceptance in Nature Communications. Detailed comments are below:

We thank the reviewer for their continued constructive engagement and for clearly identifying the two remaining items needed for acceptance. We have performed both requested experiments and describe the results in detail below.

1. The authors seemed to have misunderstood our prior question about activity on arginine residue modifications. Specifically, we wonder if the CMLase works on the equivalent arginine glycation adduct (CMA) and/or the equivalent modification on the N-terminus of peptides/proteins.

We apologize for the misunderstanding in our previous response. We now appreciate that the reviewer was specifically asking about N ω -carboxymethylarginine (CMA), the arginine-equivalent of CML, and we have directly tested CMLase activity on a CMA-modified protein substrate.

Activity on CMA: CMA-BSA was prepared by incubating BSA with glyoxylic acid and NaCNBH₃, then incubated overnight with active CMLase, inactive enzyme, or buffer alone. CMA levels were quantified by LC-MS/MS following enzymatic hydrolysis. CMLase treatment did not produce a statistically significant decrease in CMA compared to either the buffer-only or inactive enzyme controls, indicating that the carboxymethyl-guanidinium moiety of CMA is not a substrate for CMLase under conditions where the enzyme robustly oxidizes the chemically related carboxymethyl-amino moiety of CML. This finding

reinforces that CMLase is selective for the lysine-derived CML adduct rather than acting broadly on carboxymethylated amino acids. We note that this selectivity may be advantageous in a therapeutic context, as it limits the spectrum of off-target modifications the enzyme could act upon. These data have been added as a new panel to Figure S2-8, and the Results and Methods sections have been updated accordingly.

Regarding the second part of the reviewer's question, on N-terminal carboxymethylation: we have not directly tested an α -amine carboxymethylated substrate in the present study. Physiologically relevant CML accumulation in long-lived ECM proteins occurs predominantly at internal lysine ϵ -amines rather than at protein N-termini, and our proteomic analysis (Fig. 3B) characterizes CMLase activity across thirty-three internal lysine sites, which we view as the most disease-relevant population. We agree that N-terminal activity is an interesting question for future characterization and have noted this in the Discussion.

2. In figure S2-9 the treated CML-BSA lane seems to be closer to the fully modified than the unmodified BSA. Maybe using a substrate that gets closer to 100% recovery would be better to show in this data. As suggested before, we would also like to see a Western blot for the associated gel as an additional confirmation of reduction in CML modifications after incubation with the CMLase since the Coomassie doesn't show conclusive evidence of removal of adducts.

We thank the reviewer for this important point and address both halves of the comment below.

First, we wish to clarify that the Coomassie-stained gel in Figure S2-9 was included specifically to demonstrate that CMLase does not fragment or degrade the BSA substrate — i.e., to confirm protein integrity rather than to quantify CML removal. As noted by the reviewer, CML adducts do not produce a highly obvious mass shift by SDS-PAGE, so the Coomassie gel was never intended as evidence of deglycation.

Regarding the suggestion to use a substrate that approaches 100% recovery: as shown in Figure 3A, CMLase achieves up to 97% reduction in CML ELISA signal on casein, indicating near-complete deglycation is achievable on certain substrates. The residual signal on BSA is consistent with our site-resolved proteomics analysis (Fig. 3B), which identified a small number of CMLase-resistant lysines (notably K131, K504, and K573) embedded in less accessible structural environments. We retained BSA as the substrate for Figure S2-9 because its well-characterized modification landscape and partial recalcitrance make it a more stringent test of enzyme activity, and because the Western blot described below directly addresses the reviewer's core question of whether CML is in fact being removed.

Western blot for CML removal: To demonstrate CML removal in an orthogonal assay to those described in our results, we performed the requested Western blot. CML-BSA was treated in duplicate with 1 μ M active CMLase or 1 μ M catalytically inactive control enzyme for 16 hours, then analyzed by automated capillary-based immunoassay (Jess Simple Western, Bio-Techne) using an anti-CML primary antibody (Abcam, ab27684). A CML-BSA standard curve (50, 125, 500, and 1500 pg, run in duplicate) was included to confirm linearity of detection across the assay's dynamic range, and unmodified BSA (50 and 500 pg, run in duplicate) was included as a negative control to confirm antibody specificity for the CML adduct. We note that the standard curve was included as a linear-range reference for the immunoassay rather than as a direct quantitative anchor for the treated samples, since the standard curve and treatment samples were prepared via independent dilution routes prior to loading; the controlled quantitative comparison is between the active CMLase- and inactive-enzyme-treated lanes, both of which were prepared, incubated, and loaded under identical conditions.

Anti-CML signal was observed predominantly at the expected ~66 kDa molecular weight for the BSA monomer, with additional higher-molecular weight bands at ~120 and ~180 kDa consistent with glycation-induced BSA dimers and aggregates characteristic of CML-BSA preparations; no signal was detected in the unmodified BSA lanes. CMLase-treated lanes showed a clear and reproducible reduction in anti-CML signal at the BSA monomer band and a near-complete loss of the higher-molecular-weight species in both replicates compared to signal in lanes treated with inactive enzyme. Enzyme-only lanes (active and inactive enzymes without CML-BSA substrate) produced no anti-CML signal, confirming that the antibody

does not cross-react with the enzyme. Together, these results directly confirm that the reduction in immunoreactivity observed by ELISA is due to enzymatic removal of the CML adduct. Full experimental details have been added to the Methods, the uncropped Western blot image has been added as new Supplemental Figure S2-11, and the raw data will be provided in a Supplementary Data File in accordance with Nature Communications guidelines. The Results section has been updated accordingly.

Reviewer #3:

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 participation in peer review as part of the Nature Communications early career researcher initiative.

Reviewer #4:

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 #4 for their participation in peer review as part of the Nature Communications early career researcher initiative.

Reviewer #5:

General: This revised manuscript describes the development of CMLase, which reversed carboxymethyl-lysine in human tissues obtained from elderly donors, and represents the first demonstration that "irreversible" protein damage can be enzymatically repaired. The authors conclude that this newly developed enzyme, along with its observed effects, could serve as a platform to develop new enzymes capable of reversing age-related molecular damage and repairing tissue proteins compromised by aging and disease. Although this revised manuscript addressed the previous critiques, the experimental design still has several potential major flaws. Further verification with reproducibility may still be needed, and the manuscript may require more focus on a specific area of aging or age-related disease instead of covering many areas in age-related pathophysiology. This causes a loss of focus and makes the manuscript somewhat confusing. Therefore, it is difficult to support the authors' conclusion based on the content of this manuscript.

We thank Reviewer #5 for their continued engagement with our manuscript. We would like to note, respectfully, that many of the concerns raised in this second round were substantively addressed in our previous revision, with specific textual changes, new data, and expanded discussion provided in our prior point-by-point response. We direct the reviewer to those responses, which remain applicable. We also note the editor's characterization of this work as a "proof-of-concept, initial study, with guided enzyme evolution and enzymology insights" which informs the scope of our responses below. We have addressed each specific comment in turn.

1. Overall, this manuscript does not provide deeper scientific information regarding the role of CML in aging and diseases.

We respectfully note that the primary focus of this manuscript is the enzyme engineering and biochemical characterization of a novel deglycation activity, not a comprehensive review of CML biology in disease. As we described in our previous revision, we have expanded the Background and Discussion to provide appropriate context for CML as a pathogenic AGE modification linked to aging and age-related disease, with supporting references. The manuscript is not, by design, a clinical or pathological study of CML, and the editor has framed the work accordingly. We believe the current level of biological context is appropriate for the scope of the study and have made targeted further additions to the Introduction to clarify the significance of CML accumulation in aging.

2. The Background section requires more focus on the role of CML accumulation in aging, age-related functional changes, or specific age-related disease. The current format addresses various aspects of age-related changes, making it difficult to clearly understand the purpose of this project.

We thank the reviewer for this suggestion. In our previous revision, we expanded the Background to include additional context on CML accumulation in aging and its functional consequences. We acknowledge that the Background spans several aspects of AGE biology to motivate the broader relevance of our platform; however, we have now made targeted edits to sharpen the narrative arc of the Introduction, more explicitly connecting each element to the central purpose of the study: the development of a proof-of-concept enzymatic approach to reverse CML modification.

3. The Results section still contains a mixture of the results description, and some of the methods and discussions, which should be revised to only focus on the description of the results.

We thank the reviewer for this editorial point and have made an additional targeted pass through the Results section to remove interpretive language and methodological justification to the Discussion and Methods, respectively. We note that some contextual framing within the Results is standard practice in the field and aids reader comprehension, but we have minimized this where possible. The revised Results section now focuses more strictly on the presentation of experimental findings.

4. Although the authors reported similar degrees of modification across various proteins, i.e., BSA, collagen, hemoglobin, casein, and eye total protein extract, it is not clear why the authors selected only these proteins for the study.

We thank the reviewer for this comment; this was addressed in detail in our previous revision. Briefly, our protein panel was selected to span a wide range of structural complexity and biological relevance: BSA and casein serve as accessible model substrates with well-characterized glycation chemistry; hemoglobin and collagen represent physiologically important proteins known to accumulate CML in aging; and retinal/eye total protein extract provides a complex, clinically relevant matrix. This rationale is now explicitly stated in the Results section (under the protein substrate characterization subsection). We believe this panel provides an appropriate foundation for a proof-of-concept study and note that the editor has framed the work in this context.

5. The experiments were conducted using different proteins, tissues, and species (including human samples), which yielded data that supports the authors' hypothesis. However, it is somewhat difficult to reach the same conclusion as the authors because of the inconsistencies in the use of various proteins and tissues across the experiments.

We thank the reviewer for this comment, which was similarly raised and addressed in our previous revision. The use of diverse substrates across experiments was a deliberate design choice to demonstrate that CMLase activity is robust across different protein scaffolds and tissue environments, rather than being an artifact of a single model system. The experimental logic is that each figure builds progressively from simple model substrates (free CML, peptides, recombinant proteins) to increasingly complex biological samples (homogenized lens, human FFPE tissue sections). Different proteins and tissues are used in different figures because each experiment is designed to answer a specific question at a specific level of complexity. We have updated the Results section to more explicitly state this progressive experimental logic at the beginning of each major subsection to help the reader follow the structure of the study.

6. Immunohistochemical analyses were conducted for the quantitative analysis of CML in human arterial tissue and to evaluate the effects of CMLase treatment. DAB intensity may not be the ideal way to quantitate the amount of antigen in the tissue. In addition, some of the immunostaining results appear to have over-staining and high background, which could make quantitative analyses more difficult.

We thank the reviewer for raising this point, which was also raised and addressed in our previous revision. As described therein, all IHC experiments were conducted with matched antibody concentrations, parallel processing, and analysis of multiple spatially defined subregions per section. DAB quantification was performed by densitometric analysis using ImageJ with background subtraction. The magnitude of CML reduction observed by IHC (~55–70%) is consistent with the quantitative reductions measured by orthogonal methods (LC-MS/MS and ELISA) on homogenized samples, providing cross-methodological validation. We acknowledge that DAB is semi-quantitative, and this limitation is explicitly stated in the Discussion. Regarding the concern about over-staining, representative images were selected to illustrate clearly stained tissue regions, and all images were acquired and processed under identical conditions.

7. It is not clear whether the immunohistochemical analyses using formalin fixed paraffin embedded sections are ideal to test enzyme activity on CML. Further confirmation may be required using frozen sections with/without fixation.

We thank the reviewer for this comment, which was addressed in our previous revision. As described therein, we did attempt experiments on fresh frozen tissue sections; however, these samples exhibited significant fragmentation and loss of structural integrity during the enzyme incubation step, rendering them unsuitable for rigorous spatial quantification. We therefore selected FFPE sections as the primary histological model. FFPE also serves as a stringent stress test of CMLase activity, given that formalin fixation introduces extensive methylene bridge crosslinking that sterically restricts enzyme access. The fact that CMLase retains activity in this highly crosslinked matrix is, if anything, conservative evidence of its capability. Importantly, robust CMLase activity on non-fixed substrates is independently confirmed by our ELISA and LC-MS/MS data on soluble protein substrates and human lens homogenates (Figures 3 and 4A-B). We have added a brief note in the Methods to reiterate why FFPE was selected and to reference these complementary non-fixed experiments.

8. There is no clear statement on the number of samples used in many experiments, and there is no strong evidence of reproducibility for each experiment.

We thank the reviewer for this comment. As described in our previous revision, sample sizes (n values) are reported in the figure legends and Methods. We have done an additional systematic pass through all figure legends to confirm that n values, any error bar definitions as needed (mean $\pm$ SD or SEM as applicable), and statistical tests are explicitly stated for every panel. For human tissue experiments, the number of donors and sections analyzed per donor are also now explicitly stated. These additions address the reproducibility concern and meet the reporting standards required by Nature Communications.

Reviewer #6:

The authors have addressed my concerns regarding the clarity of the evolution process and the enzyme characterization. While the evolved enzyme is on the slow side, I think the engineering challenge and the improvement over the starting point is notable and interesting. The authors have also tempered their claims about the impact on aging, which is good. I am fine with this manuscript moving forward.

We thank Reviewer #6 for their positive re-evaluation and support for publication. We are pleased that our revisions addressed all of their concerns.

Response to Reviewer Comments (Revision #1)

We would like to thank the reviewers for their insightful and constructive comments. We have made significant improvements based on these suggestions to strengthen the rigor and clarity of the work. Below are responses to each of the comments made by the reviewers.

Reviewer #1:

1. Overstated Claim on Reversing Protein Aging: The claim of "reversing protein aging" appears overstated, as while there is strong evidence that CML is a major marker of protein damage and is linked to aging and age-related diseases, direct evidence that it is a major cause of protein aging itself is less conclusive, as the accumulation of CML is often a consequence of oxidative stress and glycation, which are themselves contributors to aging; thus, enzymatic removal of CML (maybe predominantly on protein surface) may not fully reverse protein aging, especially given correlative rather than definitive causal links in many studies.

We thank the reviewer for this important distinction and agree that biological aging is a complex, multifactorial process. As suggested, we revised the title to "Reversal of Protein Chemical Aging by Enzymatic Deglycation". We believe this is scientifically precise because it distinguishes the specific molecular phenomenon of "chemical aging"—defined as the accumulation of non-enzymatic modifications like the Maillard reaction—from organismal aging. Since our enzyme chemically converts the pathogenic, aging-associated CML adduct back to its native lysine state, we are accurately reversing a specific chemical lesion that drives protein chemical aging. We've also updated the discussion to ensure accurate assessment of what we accomplished in the manuscript and limitations.

2. Methodological Gaps in Specificity and Controls: While CMLase is claimed to be specific, there's limited data on off-target activity (e.g., does it oxidize native lysines or other AGEs like CEL?). The text mentions no activity on unmodified BSA, but broader substrate profiling (e.g., other amino acids, lipids) is absent. IHC antibody validation is good, but ELISA variability across proteins (52–97% reduction) isn't deeply probed—could reflect antibody epitope bias rather than true catalysis.

We thank the reviewer for these rigorous methodological inquiries. We have addressed both the specificity and variability concerns with additional data and clarification:

Off-Target Activity & Specificity: To address the potential for off-target oxidation, we have expanded our characterization to include broad substrate profiling against a panel of native amino acids and peptides. These data, now included in Figures S2-6 and S2-7, demonstrate that CMLase retains high specificity for the carboxymethyl-amino moiety and does not oxidize native lysine or other canonical amino acids. Furthermore, the lack of activity on unmodified BSA (Fig. S2-9) confirms that the enzyme does not promiscuously degrade native protein structures.

ELISA Variability & True Catalysis: We acknowledge the variation in repair efficiency across different protein substrates (52–97%). We attribute this primarily to steric accessibility rather than antibody artifacts; the dense structural environment of certain substrates (e.g., Eye TPE) likely limits the physical access of the ~32 kDa enzyme to buried CML sites compared to more open scaffolds like casein. Inevitably with enzymes and antibodies binding to complex protein substrates there will be variation in binding depending on the local CML context. We believe this is rigorously explored in Figure 3C.

Ruling out Epitope Bias: To confirm that this reduction represents true catalysis and not just antibody epitope masking, we relied on our catalytically inactive control (CeGO-763). This variant yielded no reduction in signal, confirming that physical blocking of antibody binding sites is not the cause of the signal decrease.

Orthogonal Validation: Crucially, our site-specific LC-MS/MS data (Fig. 3B/C) independently confirms the chemical reversal of CML to lysine at varying efficiencies across different sites, corroborating the trends observed by ELISA without relying on immunological detection.

3. Overreliance on Ex Vivo Models: All tissue experiments are on homogenized or sectioned samples, potentially overestimating accessibility in intact ECM. No diffusion/permeability data for CMLase into dense tissues like aorta.

We thank the reviewer for this valid assessment. We acknowledge that sectioned samples reduce the diffusion barrier present in intact organs. However, our primary objective in this study was to validate the catalytic feasibility of reversing CML within the complex chemical environment of human tissue. We agree that FFPE sections, while physically thin, represent a rigorous test of steric accessibility due to extensive fixation-induced crosslinking. We view macroscopic tissue penetration as a delivery challenge distinct from the catalytic mechanism, to be addressed in future studies via formulation or administration routes. We have added text to the Discussion explicitly acknowledging challenges that remain for clinical translation. It's important to note that reversal of CML on extracellular proteins such as in plasma or in the ECM should prove more tractable than trying to remove intracellular proteins.

4. Quantitative Inconsistencies: CML reduction in lens differs between methods (45% by LC-MS/MS vs. 78% by ELISA), attributed to surface vs. buried residues, but not experimentally dissected (e.g., via denaturation). Proteomics on BSA shows 3 sites untouched, but no follow-up on why (steric hindrance?).

We thank the reviewer for this careful analysis. We employed three distinct analytical techniques—ELISA, bulk LC-MS/MS, and proteomics—specifically to provide robust, orthogonal confirmation of this unprecedented enzymatic reaction. We did not anticipate exact quantitative alignment given the fundamental differences in what each method measures:

Lens Tissue Discrepancy (ELISA vs. LC-MS/MS): The difference between the 78% reduction measured by ELISA and the 45% reduction measured by bulk LC-MS/MS is consistent with the accessibility of the substrate. The ELISA uses an antibody that primarily detects surface-exposed CML epitopes, which are also the most accessible to our enzyme. In contrast, the LC-MS/MS method involves total hydrolysis of the sample to individual amino acids, measuring the entire pool of CML, including residues buried within the protein core or macromolecular aggregates that are sterically shielded from the enzyme *ex vivo*. The fact that the "accessible" fraction (ELISA) shows higher clearance than the "total" fraction (LC-MS/MS) reinforces the conclusion that CMLase is effectively targeting the surfaceome where signaling interactions (e.g., RAGE binding) occur.

Untouched Sites in BSA: Regarding the three unmodified sites identified by proteomics, we attribute this to local steric hindrance or flexibility. Just as with the lens tissue, certain lysine residues may be sequestered by local secondary structure or protein folding, rendering them inaccessible to the enzyme's active site.

By showing consistent catalytic activity across these diverse contexts, we believe we have overcome the limitations inherent to any single method and provided compelling evidence of CMLase activity.

5. Potential Artifacts: Hydrogen peroxide assays assume all H₂O₂ comes from CML cleavage, but flavoenzymes can auto-oxidize; controls mitigate this, but long incubations (overnight) risk secondary reactions. IHC omits antigen retrieval to avoid de novo CML, but this might under-detect buried modifications.

We thank the reviewer for thoughtfully pointing out potential artifacts in the hydrogen peroxide assay and IHC staining.

Regarding the peroxidase assays, we wish to clarify that these experiments are conducted on a timescale of six hours or less to minimize secondary hydrogen peroxide formation. We also include strict controls—utilizing either no substrate or unmodified substrate—to verify that no significant auto-oxidation occurs. Overnight incubations were utilized where noted in Figures 3 and 4 and only for experiments where hydrogen peroxide accumulation is not the readout.

Regarding immunohistochemistry, we deliberately omitted antigen retrieval steps based on established literature precedent (e.g., Schleicher et al., J Clin Invest 1997, ref. 29 in manuscript). These studies demonstrate that heat-induced epitope retrieval can generate de novo CML-like modifications, leading to potential false-positive staining. While we acknowledge that omitting retrieval may under-detect buried epitopes, we prioritized specificity to ensure that the staining observed—and its subsequent reduction by CMLase—reflected genuine, physiological CML accumulation rather than procedural artifacts. Furthermore, our LC-MS/MS data (Fig. 4A) serves as a complementary method to assess the total CML burden, including buried residues, in homogenized tissue.

6. Evolutionary Bias: Selections use poly-alanine peptides with one variable residue, potentially biasing toward simple contexts; real ECM proteins have complex sequences.

We thank the reviewer for pointing out potential bias in the peptide selection. A strong basis of activity on the selection peptide was needed to complement growth which is why we chose simpler targets that we knew had a baseline level of activity. To help counteract this, selections were performed with mixed peptides, which should favor enzymes with broad specificity. After selection, additional activity tests were performed on more complex sequences, and, most importantly, CML-BSA (Fig. 2E) which we confirmed by LC-MS has a diverse set of modified sites and structural contexts. We did discover some variants with improved activity on selection peptides that did not translate to more complex peptides and proteins and these were removed from consideration.

7. Lack of In Vivo Data: The study is purely in vitro/ex vivo; no animal models (e.g., diabetic mice with high CML) to assess pharmacokinetics, immunogenicity, or efficacy in living systems. H₂O₂ byproduct could cause oxidative stress in vivo, unaddressed here.

We thank the reviewer for the thoughtful comment regarding the in vitro and ex vivo models used in the paper. Our goal is to validate the novel enzymatic reaction and assess CMLase's chemical capabilities, prior to evaluating the *in vivo* efficacy. We have modified the discussion to address this as a limitation of the study that will be more closely examined in future work. We also address the question of potential hydrogen peroxide toxicity.

8. Incomplete Functional Restoration: While CML removal is shown, no assays test if it restores protein function(e.g., collagen elasticity, ECM signaling, RAGE binding inhibition). Pathogenic "hotspots" are mentioned but not prioritized or validated.

We thank the reviewer for bringing up this point about functional restoration of CML damaged proteins. We believe this is outside of the scope of the current work and have addressed this as a limitation of the study in the discussion.

9. Scope of CML Accessibility: Enzyme works on surface-exposed sites (correlates weakly with B-factor), but buried or crosslinked CML in dense ECM may resist repair. Variability across proteins (e.g., 52% on eye TPE vs. 97% on casein) highlights substrate dependence.

We thank the reviewer for this insightful observation regarding substrate dependence and steric accessibility. We agree that CMLase, like most enzymatic therapies, is primarily active on surface-exposed residues. We have revised the Discussion (paragraph 2) to explicitly acknowledge this limitation, stating: "efficacy varies across substrates (e.g., 97% on casein vs. 52% on eye TPE), suggesting that while surface-exposed CML is readily cleared, buried or crosslinked residues within dense ECM may resist enzymatic access."

10. Scalability and Therapeutics: Engineering focused on activity/stability, but no data on production yield, half-life, or delivery (e.g., to eyes/kidneys). As a large protein (~40 kDa inferred from homologs), systemic use faces barriers like renal clearance or immunogenicity.

We thank the reviewer for these important translational insights. We have revised the Discussion to explicitly acknowledge the challenges of therapeutic development, including immunogenicity considerations as well as potential product toxicity concerns.

11. Broader AGE Context: CML is one AGE; others (e.g., pentosidine crosslinks) aren't addressed, limiting impact on total AGE burden. Aging is multifactorial, so CML reversal alone may not "reverse aging."

We thank the reviewer for these comments. We agree that aging is a complex, multifactorial process and that CML represents only one component of the total AGE burden. We have revised the Discussion to explicitly state that CML reversal alone will not resolve the full aging phenotype. However, we emphasize that CMLase serves as a critical proof-of-concept, establishing a directed evolution platform that can now be adapted to target other pathogenic AGEs including both cross-linked AGEs (e.g. glucosepane, pentosidine) as well as non cross-linked AGEs (e.g. Argpyrimidine, carboxyethyllysine) to address the broader molecular landscape of aging.

12. Sample Limitations: Human tissues from limited donors/ages; no disease-specific samples (e.g., diabetic kidneys). Proteomics only on BSA, not tissues, missing real-world complexity.

We thank the reviewer for these comments on the limitations of selected samples. We agree that a more comprehensive assessment of activity across a variety of proteins and samples will be necessary to understand the full capabilities of the enzyme. As the basis of our study, we chose simple model proteins (BSA, Hemoglobin, Casein, and Collagen), as well as more complex targets (retinal total protein extract, human lens protein, aorta and skin samples). This sample variety, along with the diverse methods of assessing CML removal, make us confident in CMLase's ability to repair CML in a wide variety of contexts.

13. Technical Constraints: Genetic selection relies on E. coli, potentially missing eukaryotic folding needs. IHC quantification (DAB intensity) is semi-quantitative, prone to variability.

We thank the reviewer for raising these technical points.

Genetic Selection Host: We wish to clarify that the CMLase scaffold is derived from the bacterium *Calidithermus roseus* and does not require eukaryotic post-translational modifications (e.g., glycosylation) for activity. Therefore, *E. coli* serves as an appropriate host that accurately captures the enzyme's native folding requirements without the risk of missing eukaryotic-specific processing.

IHC Quantification: We acknowledge that DAB staining is semi-quantitative. To minimize variability, we processed all samples in parallel and analyzed multiple spatially aligned subregions. Crucially, the magnitude of CML reduction observed via IHC (~55–70%) is consistent with the fully quantitative reductions measured by our orthogonal methods (LC-MS/MS and ELISA) in homogenized samples. This cross-methodological consistency supports the validity of the histological data despite the inherent limitations of DAB.

Reviewer #2:

1. One issue throughout the manuscript related to a lack of clarity about how activity was assessed for the various CMLase enzymes. For example, minimal details were provided about the concentrations used for free vs peptide-CML in enzymatic assays. In later, critical experiments, I could not find details about the substrate concentration, identity or how the "loading" of CML on protein or peptide substrates was assessed. Greater clarity on these points would strengthen the conclusions of the manuscript

We thank the reviewer for pointing out the lack of clarity around enzyme and substrate concentrations in the manuscript. We have added more details about enzyme and substrate concentrations throughout the text and in the methods, especially "Peroxidase Assays" and "CML-modified protein ELISAs."

2. There seems to be a major potential for cross reactivity with cleavage at C-terminal Gly residues. Did the authors observe any truncations at the C-terminus? How was the potential for other, potentially unexpected, cross-reactivity evaluated? Can it work on N-terminal or Arg residues or modifications? I do not think I noticed any experiments to address the remaining (or potential for remaining) glycine oxidase activity, which could contribute to background H₂O₂ generation.

We thank the reviewer for pointing out potential cross reactivity with free amino acids and glycine terminal peptides. We added additional controls to quantify CMLase activity on natural free amino acids and several tripeptides, these data have been added to the manuscript as Fig S2-6 and S2-7. There is still residual glycine activity (Fig S2-3), however the K_M is significantly higher than the wildtype enzyme. At physiological concentrations, we expect this to be negligible.

3. On a related note, additional characterization for CMLase-mediated CML loss from proteins would strengthen the contribution. As an example - A western blot showing in vitro CML modified proteins, treated with and without CMLase would show that the enzyme isn't digesting the whole protein (by Coomassie) and also decreases CML levels.

We thank the reviewer for bringing up the potential issue with protein substrates being digested, allowing the enzyme easier access. To verify true protein activity, we performed additional experiments to address the reviewer's concerns which have been included as Fig. S2-9. Specifically, we ran an overnight reaction at identical conditions to the reactions analyzed by LC-MS and analyzed the products using SDS-PAGE to verify that the protein substrate was remaining intact.

4. Additional characterization should also be provided for the extent of CML modification introduced into in vitro proteins. The glyoxylic acid reductive amination makes sense, but what is the overall degree of modification? The LC-MS data presented seems to focus on a %removal from a set of CML-modified sites. However, I find myself wondering if the initial extent of modification at each site (which was presumably not 100%, or was it?) affects the way one can interpret this data.

We thank the reviewer for this important question regarding the baseline modification levels. To clarify, the data in Figures 3B, 3C, and S3-3 are reported as the absolute percent abundance of the CML modification at each

specific lysine residue. The samples treated with the catalytically inactive enzyme (CeGO-763) (represented by the lighter bars/points) serve as the control for the initial level of modification. As the reviewer suspected, the starting modification level is heterogeneous, ranging significantly from site to site rather than being uniformly 100%. Importantly, we analyzed the relationship between the initial modification burden and repair efficiency and found no significant correlation. CMLase demonstrated the ability to significantly reduce CML levels at both heavily modified and lightly modified sites, indicating that the initial degree of modification does not bias the interpretation of the enzyme's efficacy.

5. Given the large loading of CMLase to CML-modified protein (e.g. 5:1 CMLase to CML-BSA), it seems entirely plausible that some of the reduction observed by ELISA or immunohistochemistry could be due to blocking detection by Abs due to nonspecific (or even specific) binding to the substrate. This needs more careful controls throughout the paper.

We thank the reviewer for raising this critical control regarding potential epitope masking. To definitively rule out the possibility that the enzyme is simply blocking antibody access rather than catalyzing CML removal, we utilized a catalytically inactive enzyme variant (CeGO-763) as a negative control in all antibody-based assays. This inactive variant is structurally homologous to CMLase and was applied at the identical concentration (5 μ M) and incubation conditions. Crucially, the inactive control consistently showed no reduction in antibody binding signal compared to untreated samples. If steric blocking were the mechanism of signal loss, we would expect the inactive enzyme to produce a similar effect. Furthermore, our LC-MS/MS data (Figures 3B, 3C, and 4A) provides orthogonal validation, confirming that the loss of antibody signal corresponds to the physical chemical conversion of the CML adduct back to native lysine, independent of antibody binding dynamics.

6. Some discussion of the efficiency of the CMLase to other enzymes (and the native activity) should be included

We have revised the Discussion to explicitly contextualize the catalytic efficiency of CMLase. We note that while its rate is lower than that of naturally evolved metabolic enzymes, it compares favorably to other early-generation therapeutic enzymes engineered to act on non-native substrates. We also highlight the substantial fold-improvement achieved over the starting scaffold.

7. Based on the data provided, I don't think that the following sentence is fully justified. "We show the new-to-nature CMLase repairs a majority of surface CML modifications on proteins and then mapped the specific sites of repair using proteomics." In particular "majority" seems inaccurate, as it definitely works better on some proteins than others as seen in figure 3.

We thank the reviewer for the feedback surrounding certain claims made in the discussion. We have softened this statement to avoid overstating the efficacy of the enzyme. Additionally, we have expanded the discussion to further qualify the results and address shortcomings of the study so the reader can come to their own conclusions.

Reviewers #3 and #4:

We would like to thank reviewers #3 and #4 for co-reviewing the manuscript. We appreciate your time and consideration.

Reviewer #5:

1. Overall, the style of writing in the article is more narrative than scientific, and does not provide deeper information regarding the role of CML in aging and diseases.

We thank the reviewer for this feedback. We have revised the manuscript to ensure a more concise and objective scientific tone. Additionally, we have inserted subheadings within the main text to improve readability and structure. To address the request for deeper biological context, we have significantly expanded the Background and Discussion sections. These updates now provide a more thorough review of CML's pathogenic role—specifically detailing the CML-RAGE signaling axis and structural ECM damage—and contextualize how this study fits into the larger landscape of aging biology.

2. The Background section requires more information on the advanced glycation end products (AGE), aging, age-related functional changes, and age-related diseases.

We thank the reviewer for this suggestion. We agree that providing a broader context regarding the pathogenic role of AGEs in aging and disease strengthens the rationale for developing enzymatic reversal strategies.

We have revised the Background section (Paragraphs 1 and 2) to include a more comprehensive overview of AGE-mediated pathophysiology. Specifically, we have expanded the text to:

Distinguish between mechanical and signaling pathologies: We now explicitly differentiate between the structural stiffening caused by crosslinking AGEs (affecting vascular compliance and skin elasticity) and the inflammatory signaling caused by non-crosslinking adducts like CML.

List specific age-related diseases: We have added references to atherosclerosis, diabetic nephropathy, and neurodegenerative conditions to highlight the clinical stakes of AGE accumulation.

3. In general, the Results section is a mixture of the description of the results including some of the methods and discussions, and should focus on just on the description of the results.

We thank the reviewer for this feedback regarding the structure of the manuscript. We have carefully reviewed the text and revised the Results, Methods, and Discussion sections to better segregate these elements. Specifically, we have moved some technical protocols to the Methods where appropriate and interpretative commentary to the Discussion, ensuring the Results section remains focused on the description of experimental outcomes while maintaining the necessary logical flow of the engineering process.

4. The first paragraph of the Results section needs to be removed or moved to the Discussion section.

We appreciate the reviewer's suggestion to streamline the text. However, we respectfully prefer to retain this paragraph in the Results section because it describes specific experimental steps and negative data that were critical to our study design—specifically, the empirical testing and failure of the commercially available *Cladosporium* lysate to exhibit CML activity in our hands. These results were the decisive factor that necessitated the computational mining approach described in the subsequent text. We believe that reporting this experimental outcome is essential to the logical flow of the engineering narrative.

5. The Results section should include sub-headlines to make it easier for the readers to follow.

We thank the reviewer for the suggestion to include subheadlines to make the Results more clear. We have added these throughout the results section.

6. The Figures and Figure Legend need to be separated from the body text.

We thank the reviewer for the suggestion to update the format of the figures in the text. All figures and legends have been moved out of the main text to the end.

7. The font in Figure 1B is too small and hard to see.

We thank the reviewer for pointing out this formatting issue with Fig 1B. The figure has been updated with an increased font size.

8. Although the authors stated that there are similar degrees of modification in various proteins, i.e., BSA, collagen, hemoglobin, casein, and eye total protein extract, using different glyoxylic acid concentrations, the degree of modification seems to vary in various proteins (Figure 3A).

We thank the reviewer for bringing up their concern regarding different modification levels in the ELISA data. Modification conditions were tuned to get similar levels of signal in the ELISA, however it was challenging to get them identical, likely due to the difference in ease of glycation and potentially antibody bias. Figure S3-2 shows all modification conditions for each protein and we selected the modification level closest for each.

9. The experiments were conducted using different proteins and tissues, which seem to yield the data that support the authors' hypothesis. Unfortunately, it is a bit hard to reach the same conclusion as the authors because of the inconsistency of the usage of various proteins and tissues throughout the entire experiments.

We thank the reviewer for this important observation regarding the diversity of substrates employed in our study. We understand that the use of varying tissues and proteins may appear inconsistent; however, this was a deliberate design choice to progressively test CMLase on increasingly complex substrates. We believe that observing consistent activity across this spectrum—from peptides and recombinant proteins to human solid tissues—using orthogonal techniques (fluorescence assays, quantitative LC-MS/MS, proteomics, and IHC) strengthens the conclusion that CMLase is robust, independent of the model system. We have updated the main text and Discussion to clarify this rationale and address the limitations of specific samples.

10. The immunohistochemical analyses were conducted for the quantitative analysis of CML in human skin and the effects of CMLase treatment. The DAB intensity may not be the ideal way to quantitate the amount of antigen in the tissue. In addition, some of the immunostaining results possibly show over-staining and background staining, which could make quantitative analyses more difficult. Furthermore, the figure is too small to see the details as well.

We appreciate the reviewer's comment regarding the impact of tissue fixation on enzymatic activity. We selected FFPE sections specifically to leverage the availability of high-quality, age-stratified human tissue banks. However, we agree that fixation introduces steric complexity. To address the need for confirmation in non-fixed environments, we direct the reviewer to our experiments on human lens homogenates and soluble protein substrates (e.g., BSA, collagen, hemoglobin).

Confirmation in Non-Fixed Samples: As detailed in Figure 4A-B and Figure 3, we demonstrated robust CMLase activity on non-fixed, soluble samples using orthogonal detection methods (LC-MS/MS and ELISA). These experiments confirm that CMLase is highly active on CML modifications in their native state, independent of fixation-induced cross-linking.

FFPE as a Stringent Test: Formalin fixation induces extensive methylene bridge cross-linking, which sterically hinders enzyme access. Consequently, the FFPE experiments serve as a rigorous "stress test" for the enzyme. The fact that CMLase successfully reversed CML modifications within this highly cross-linked matrix demonstrates its ability to penetrate dense tissue structures.

Methodological Controls: To ensure the signal was not an artifact of sample preparation (e.g., heat-induced CML formation), we strictly omitted antigen retrieval steps in accordance with established protocols for detecting AGEs in fixed tissue (Ref. 28).

We believe the combination of quantitative LC-MS/MS data on non-fixed lens tissue and the high-resolution spatial data provided by FFPE histology provides a comprehensive validation of CMLase activity across diverse protein environments.

11. It is not clear how the specificity of the antibodies and specific epitope of the antigen were determined. It is also not clear whether the authors have examined if the less (or no) staining after CMLase treatment was due to the direct effects or potential secondary effects by CMLase.

We thank the reviewer for these rigorous controls. We have addressed both points experimentally:

Antibody Specificity: As described in the Results (and shown in Fig. 4C), we validated the specificity of the 6D12 monoclonal antibody using a competitive inhibition assay. Pre-incubation of the antibody with CML-BSA completely abolished tissue staining, whereas pre-incubation with unmodified BSA did not. This confirms that the antibody specifically recognizes the CML epitope rather than non-specific tissue components. In addition, the antibody has been extensively characterized by the manufacturer.

Ruling Out Secondary Effects: To confirm that the reduction in staining was caused by the direct enzymatic reversal of CML—rather than secondary effects such as epitope masking, steric hindrance by the enzyme, or non-specific proteolysis—we utilized a catalytically inactive enzyme control (CeGO-763). This variant was purified using the exact same protocol but lacks oxidative activity. In our IHC experiments, the inactive control did not reduce CML staining, whereas CMLase did. Furthermore, our LC-MS/MS data (Fig. 3B, 4A) provides definitive chemical proof that the CML adduct is converted back to native lysine, confirming that the loss of antibody signal corresponds to the physical removal of the antigen.

12. It is not clear whether the immunohistochemical analyses using formalin fixed paraffin embedded sections were ideal to test the enzyme activities on CML. Further confirmation may be required with frozen sections with/without fixation.

We thank the reviewer for raising this important technical point. We selected FFPE sections as our primary histological model because they represent the standard for human tissue banking and offer superior morphological preservation, which is critical for spatially resolved quantification. Regarding the comparison with frozen sections, we wish to highlight the following experimental validations:

Evaluation of Frozen Sections: We did attempt to perform these assays on fresh frozen tissue. While we observed qualitatively similar enzymatic reversal of CML, the fresh frozen samples showed significant tissue fragmentation and loss of structural integrity. This rendered them unsuitable for the rigorous spatial quantification presented in the manuscript. The structural stability provided by FFPE was therefore necessary to maintain tissue architecture during the reaction.

Biological Validation (Young vs. Old): To confirm that the signal observed in FFPE was specific to age-accumulated CML and not a fixation artifact, we utilized tissue from young donors (20–25 years old) as biological controls. As shown in Figure S4, young FFPE sections displayed negligible CML staining compared to elderly donors, consistent with the known accumulation kinetics of CML. This confirms that our FFPE protocol preserves the antigen's biological specificity.

Confirmation on Non-Fixed Substrates: To address the concern about activity on non-fixed substrates, we point to our data on soluble proteins and human lens homogenates (Fig. 3 and 4A-B). These assays demonstrated robust CMLase activity on non-crosslinked, non-fixed samples using orthogonal detection methods (LC-MS/MS and ELISA), confirming that the enzyme is highly active on CML in its native state.

We believe the combination of the quantitative soluble data (non-fixed) and the spatially resolved FFPE data (fixed) provides the most comprehensive evaluation of CMLase activity.

13. There is no clear statement regarding how many samples were used in many of the experiments, and it is not clear whether there was strong evidence of reproducibility for each experiment.

We thank the reviewer for raising this important question about reproducibility. All experiments were done in triplicate and wherever possible all data points were shown (Fig 3C, 4A, B, and E). Some information regarding the number of replicates has been moved from the methods to the figure legends to clarify specific data (Fig 1D, E and 2E).

14. The Discussion section is too short, and there is not much information provided.

We thank the reviewer for this suggestion. We have significantly expanded the Discussion to provide a more comprehensive analysis of our findings and their context within the field. Specifically, we have added:

Enzymatic Context: A detailed comparison of CMLase efficiency relative to native metabolic enzymes and other therapeutic enzymes, explicitly highlighting the substantial fold-improvement achieved through directed evolution.

Translational Limitations: A dedicated assessment of the barriers to clinical translation, including the challenges of potential immunogenicity.

Safety & Metabolism: Clarification regarding the safety profile of the reaction, specifically citing the endogenous metabolic pathways responsible for clearing the reaction byproducts (glyoxylate and hydrogen peroxide).

We believe these additions provide the necessary depth and balance requested by the reviewer.

Reviewer #6:

Glyoxylic acid is released in the deglycation process. This also seems to be the reagent used to generate CML proteins in the methods. From a quick read, it seems like the body should be able to metabolize glyoxylate and keep it at reasonable levels. The authors should comment on this- are there potential issues with freeing glyoxylic acid? Simply providing references for how it gets metabolized could suffice.

We thank the reviewer for raising this important safety consideration regarding reaction byproducts. We have revised the Discussion to explicitly address released glyoxylic acid. The text has been revised with references to note that glyoxylate is a naturally occurring metabolite efficiently cleared by endogenous enzymes (e.g.,

glyoxylate reductase/hydroxypyruvate reductase) and activity of CMLase would yield a negligible increase in glyoxylate compared to physiological levels.

Panel 2E. The rate is slow (uM/hour). In the legend for Fig 2E, please state the concentration of enzyme used and state how the data are collected (are the plotted values averages? How many replicates?)

We thank the reviewer for pointing out the lack of clarity in Fig. 2E. It has been updated to include concentrations of enzyme, substrates, and replicates.

First round of evolution: error prone PCR. In the paragraph describing error prone PCR- how many rounds of error prone PCR and selection did the authors do? It seems like it was 1 round, but please clarify. 7 mutations is a lot for one round, but I see that the authors aimed for 4-8 mutations per gene, so I realize this could have been the hit from 1 round of selection. What was the estimated library size and coverage?

We thank the reviewer for this comment regarding the error prone PCR in the first round of evolution. The first variant, CrGO-764, was found from a single aggressive library with 4-8 mutations per gene and the highest activity hit had 7 mutations. The text has been edited to clarify these points. The library size was approximately 5×10^7 and was plated with 2x coverage, hitting the upper limit of the selection 10^8 cells per plate. The methods have been updated to include this information.

Exciting that this yielded a 5-fold improvement. Please list the mutations in this paragraph (I see the authors reference table Fig S2-1, but would be helpful to see it listed here, too).

We thank the reviewer for the suggestion to include the mutations of CrGO-764 into the main text. We have added in the full list of mutations to improve clarity and consistency.

How many colonies did the authors screen in each round of evolution? Please add this information to the methods or Figure S2-4 (to explain how you got to the estimated number of screening over 500 million protein variants as stated in the abstract).

We thank the reviewer for requesting this clarification regarding our screening throughput. We have updated the Library Transformation section of the Methods to explicitly detail our library sizes. Specifically, we clarify that we generated libraries with an average diversity of $\sim 5 \times 10^7$ variants per round. Over the course of the project, we screened more than ten such libraries—including several exploratory, "dead-end" lineages not detailed in the main text—resulting in a cumulative screening total exceeding 5×10^8 variants.

Can you further explain the rationale for picking residues guided by CeGO to saturate the CrGO? Is it surprising that mutations that improve activity on free CML also improve activity on Peptidyl CML?

We appreciate the reviewer's question regarding the reasoning behind modelling a library after the active site of CeGO. We wanted to increase affinity for the CML side chain without affecting the more open active site, therefore we targeted residues inside the active site while leaving the surface residues untouched to retain peptide and protein activity. This has been expanded on in the text.

The results are very interesting and could be sold more explicitly in the discussion. Discuss the overall fold-improvement in enzyme activity in the discussion – this would strengthen the results of the paper.

We appreciate the reviewer's positive assessment of our results. We agree that emphasizing the magnitude of the engineering success strengthens the narrative. We have revised the Discussion to explicitly highlight the

substantial fold-improvement in catalytic efficiency achieved through our five rounds of directed evolution, thereby emphasizing the efficacy of the engineering campaign.

When you tested activity on CML-BSA, why did you use low concentrations of CML-BSA (1 μ M) with 5 μ M CMLase? How many turnovers do you think the CMLase can do?

We thank the reviewer for this question regarding our assay conditions. We utilized 1 μ M BSA because the protein is heavily modified, resulting in an effective concentration of CML substrate estimated to be >10 μ M (due to multiple CML modifications per BSA molecule). We employed a relatively high concentration of CMLase (5 μ M) to ensure that the rate of hydrogen peroxide production was sufficient for detection in our coupled peroxidase assay within a short timeframe (hours), given that the reaction velocity on bulky protein substrates is slower than on free peptide. Regarding turnover capability, the enzyme is clearly catalytic rather than stoichiometric; on free CML substrates, we routinely observe more than 80 turnovers, indicating that the limiting factor in the CML-BSA assay is steric accessibility or diffusion rather than enzyme inactivation.

please provide information on how the selection strain was generated.

We thank the reviewer for pointing out this unintentional omission. The LysA knockout strain was purchased from Horizon (JW2806). The methods have been updated to include the supplier and catalog number.

Can the authors provide more detailed methods for how the selection was done? What were the transformation efficiencies in each round, and what is the estimated coverage based on this? What temperature were the cells grown at during the selection (I see this information for the transformation)? When did the authors start to see colonies for the different rounds of selection?

We thank the reviewer for these comments regarding the genetic selection methods. Selection plates were grown at 30 °C for up to one week. Colonies were picked as soon as they reached sufficient size, which was generally 1-7 days after plating, depending on the library and selection conditions. We have updated the “Genetic Selection” section of the methods to include more detailed information on growth temperature and time for selection plates. Additionally, we expanded on library transformation efficiency and estimated coverage in the “Library Transformation” section.

Sequence context: I think it's great that Fig 3c highlights structural diversity of target sites and some sequence contexts. However, a more comprehensive depiction of sequence context would help a lot. Figure 3B- can the authors convert this information into sequence context information somehow (maybe logos plot?). This would help audience/end user to see if there are any limitations. A logos plot or some visual depiction of the sequence bias (or lack thereof) of enzymatic deglycation would be great.

We thank the reviewer for requesting additional sequence context for the LC-MS data. It was originally omitted due to the complexity of the sequences and lack of a specific motif, however we have added the additional analysis as a sequence logos plot for additional information. This has been added as Fig S3-5. In brief, to evaluate the relationship between the local sequence environment and CMLase efficiency, the amino acid context surrounding CML-modified Lysine (CML-K) sites within BSA (± 5 aa) was analyzed. While no rigid primary sequence motif was identified, local composition appears to influence accessibility. The most prominent differentiator was the distribution of Cysteine (C) residues; high-efficiency sites were frequently associated with flanking Cysteines, whereas Cysteine was entirely absent from the most resistant sites ($p < 0.001$, Fisher's Exact Test). This may reflect higher CMLase activity near more accessible domain boundaries or disulfide-linked regions. Beyond this, the resistant cohort indicated a potential shift toward a more hydrophobic microenvironment. We observed an increased density of hydrophobic residues and a subtle enrichment of

Phenylalanine (F) on the downstream side of sites with lower CML reduction. These observations suggest that resistant modifications may reside within hydrophobic clusters or aromatic-rich regions that provide physical shielding from the enzyme. It remains unclear whether these residues directly obstruct enzymatic access, interfere with the catalytic process post-binding, or reflect BSA-specific structural constraints. While these trends highlight potential topographical barriers within the globular BSA fold, we remain cautious in generalizing these residue preferences to other protein architectures or to structural changes induced by the CML modification itself. A deeper exploration of these structural determinants across a wider set of proteins is beyond the scope of this study but remains an objective for future investigation.

We have updated the Results and Discussion sections to briefly include a more granular analysis of the local sequence environment.

Sequence context: When the authors did the error prone library and challenged for growth on mixtures of peptides- were cells selected on all peptides at once? What are the sequences of these peptides? In this case, do you worry that a particular variant could deglycate only a subset of the peptides? How do you modulate stringency? I think it would help if you added data for your final variant testing kinetics on a range of peptide contexts (At least the ones from the mixture you provided in the selection). In the end, you get variants with a 20% improvement on peptide, but you don't say which peptide. Also, Figure S2-6 doesn't say which peptide (sorry if I missed it).

We thank the reviewer for bringing up concerns regarding the peptide selection. The library was selected on a mix of all peptides at once (AAX-CML-AA-CML and AA-CML-XAA-CML, where X is every amino acid except R, K, or H) at a concentration such that enzyme activity needed to be significant on >50% of these sequences in order to complement growth. The enzyme was then assessed on a variety of peptide sequences, however our key indicator of both accommodation of protein substrate and diverse sequence context was the significant improvement in activity on CML-BSA (Fig. 2E) which would have been extremely unlikely if the enzyme had become hyper specific for specific simple sequences. In previous selections of the same library on single peptides, this effect was observed: significant improvements on the selection peptide with decreases on other peptides and a significant decrease in CML-BSA activity; these hits were removed from consideration. CMLase's broad specificity is further supported by the LC-MS data in Fig. 3C that demonstrates broad specificity with no trends in sequence preference. Stringency of the selection is modulated by changing total substrate concentration (keeping all peptides at equal concentrations) or IPTG concentration to alter the expression. Figure S2-4 has been updated to specify peptide sequences.

Sequence context: Would be great to have sequence context for all residues from S4-3. Can the authors add that to the graph?

We thank the reviewer for the suggestion to add additional sequence context to S3-3. We appended two sequence logos plots in S3-5 to address the need to present more sequence trends.

Fig 1c, Supplemental Figures S1-2 and 2-2. Please include methods for how the CML substrate was modeled in the active site. What docking algorithm or program did the authors use? Was this done manually?

We thank the reviewer for this question about docking methods and structural analysis. We have added a "Structural Analysis" section to the methods which includes information about the docking program that was used (Bioluminate).

During submission, some supplementary figure names were updated. The following table provides a corresponding list of old and new supplementary figure and table names.

Old Name	New Name
Supplementary Fig. 1-1	Supplementary Table 1
Supplementary Fig. 1-2	Supplementary Figure 1
Supplementary Fig. 2-1	Supplementary Table 2
Supplementary Fig. 2-2	Supplementary Figure 2
Supplementary Fig. 2-3	Supplementary Table 3
Supplementary Fig. 2-4	Supplementary Figure 3
Supplementary Fig. 2-5.	Supplementary Figure 4
Supplementary Fig. 2-6	Supplementary Figure 5
Supplementary Fig. 2-7	Supplementary Figure 6
Supplementary Fig. 2-8	Supplementary Figure 7
Supplementary Fig. 2-9	Supplementary Figure 8
Supplementary Fig. 2-10	Supplementary Figure 9
Supplementary Fig. 2-11	Supplementary Figure 10
Supplementary Fig. 3-1	Supplementary Figure 11
Supplementary Fig. 3-2	Supplementary Table 4
Supplementary Fig. 3-3	Supplementary Figure 12
Supplementary Fig. 3-4	Supplementary Figure 13
Supplementary Fig. 3-5	Supplementary Figure 14
Supplementary Fig. 4-1	Supplementary Figure 15
Supplementary Fig. 4-2	Supplementary Figure 16
Supplementary Fig. 4-3.	Supplementary Figure 17
Supplementary Fig. 4-4	Supplementary Figure 18
Supplementary Fig. 4-5	Supplementary Figure 19
Supplementary Fig. 4-6.	Supplementary Figure 20