



A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

# CHEM **BIO** CHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

## Accepted Article

**Title:** Biocatalytic Reversal of Advanced Glycation End Product Modification

**Authors:** Nam Y Kim, Tyler N Goddard, Seungjung Sohn, David A Spiegel, and Jason Crawford

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

**To be cited as:** *ChemBioChem* 10.1002/cbic.201900158

**Link to VoR:** <http://dx.doi.org/10.1002/cbic.201900158>

WILEY-VCH

[www.chembiochem.org](http://www.chembiochem.org)

A Journal of



## FULL PAPER

# Biocatalytic Reversal of Advanced Glycation End Product Modification

Nam Y. Kim,<sup>[a, b]</sup> Tyler N. Goddard,<sup>[a, b]</sup> Seungjung Sohn,<sup>[a]</sup> David A. Spiegel,<sup>[a, c]</sup> and Jason M. Crawford\*<sup>[a, b, d]</sup>

**Abstract:** Advanced glycation end products (AGEs) are a heterogeneous group of molecules that emerge from the condensation of sugars and proteins via the Maillard reaction. Despite a significant number of studies showing strong associations between AGEs and the pathologies of aging-related illnesses, it has been a challenge to establish AGEs as causal agents primarily due to the lack of tools in reversing AGE modifications at the molecular level. Here, we show that MnmC, an enzyme involved in a bacterial tRNA-modification pathway, is capable of reversing the AGEs carboxyethyl-lysine (CEL) and carboxymethyl-lysine (CML) back to their native lysine structure. Combining structural homology analysis, site-directed mutagenesis, and protein domain dissection studies, we generated a variant of MnmC with improved catalytic properties against CEL in free amino acid form. We show that this enzyme variant is also active on a CEL-modified peptidomimetic and an AGE-containing peptide that has been established as an authentic ligand of the receptor for AGEs (RAGE). Our data demonstrate that MnmC variants are promising lead catalysts toward the development of AGE-reversal tools and a better understanding of AGE biology.

## Introduction

Advanced glycation end products (AGEs) are non-enzymatic post-translational modifications of proteins derived from the condensation of reducing sugars and nucleophilic amino acid residues, such as lysine and arginine.<sup>[1,2]</sup> Although AGEs are formed in the body as a part of normal metabolism, they can accumulate to high concentrations and contribute to the progressive decline of multiple organ systems.<sup>[3]</sup> This process is accelerated in diabetics, owing to their hyperglycemic conditions.<sup>[4]</sup> In addition to causing spontaneous damage by altering protein structure and function, AGEs also interact with the receptor for AGEs (RAGE), eliciting oxidative stress and activating the transcription factor NF- $\kappa$ B thought to be a major

contributor of AGE-associated chronic inflammation and cellular damage.<sup>[5]</sup> Elevated levels of AGEs are linked to the pathology of many metabolic and degenerative diseases of aging, such as diabetic complications, atherosclerosis, and Alzheimer's disease.<sup>[6-13]</sup> This association is manifested by age-dependent increases in cross-linking, browning, fluorescence, and AGE content in long-lived proteins such as collagens and lens crystallins.<sup>[14-17]</sup> Structural characterization and synthesis of some of the more prevalent AGEs (e.g., glucosepane) have allowed more focused investigations into their individual chemical properties and formation.<sup>[18]</sup> Indeed, chemical studies have shown strong correlations between specific AGEs and the development of age-related illnesses;<sup>[19,20]</sup> however, it has been difficult to unequivocally demonstrate that any AGEs are direct causal factors largely due to the lack of tools for investigating the reversal of mature AGE modifications at the molecular level.

Previous correlative strategies to probe the biological roles of AGEs have mainly focused on inhibiting their formation through chemical interventions. For example, aminoguanidine can intercept reactive  $\alpha$ -dicarbonyls, including the AGE substrates glyoxal (GO), methylglyoxal (MGO), and 3-deoxyglucosone (3DG).<sup>[21]</sup> Additionally, molecules such as ALT-711 were developed to "break" dicarbonyl protein crosslink intermediates *en route* to mature AGEs, which showed favorable therapeutic effects, including a reduction in collagen crosslinking and myocardial stiffness in animal models.<sup>[22]</sup> More recently, enzymes known as the Amadoriases and DJ-1 have been discovered to be able to "repair" glycated proteins and nucleic acids, respectively.<sup>[23,24]</sup> Although these enzymes are promising tools in AGE research, they only act on the early intermediates in glycation rather than mature AGEs. Currently, there are no known strategies available to repair mature AGE modifications and potentially ameliorate established AGE-associated pathologies.

N<sup>ε</sup>-(carboxyethyl)lysine (CEL) and N<sup>ε</sup>-(carboxymethyl)lysine (CML) are lysine-derived AGEs that have been extensively investigated in renal compartments, collagen, and lens.<sup>[25]</sup> Formation of CEL and CML involves the reaction between the  $\epsilon$ -amino group of lysine residues and MGO or GO, respectively. Alternatively, CML is also known to form from the oxidative cleavage of the Amadori product fructoselysine.<sup>[26]</sup> As two of the major physiological ligands of RAGE,<sup>[27,28]</sup> CEL and CML are associated with the progression of a host of diabetic complications and other neurodegenerative diseases. Studies have shown increased levels of CEL in serum samples of Alzheimer's disease patients, and reduction in permeability of the vitreous humor in the eye.<sup>[29,30]</sup> As one of the most abundant AGEs in renal compartments, CML was shown to be associated with a decline in renal function and chronic kidney disease, although it is still under debate whether or not CML is the causative agent<sup>[31-34]</sup> Here, we describe the identification and characterization of an enzyme family – MnmC – that is capable of cleaving CEL and to a lesser

[a] N. Y. Kim, T. N. Goddard, S. Sohn, Prof. D. A. Spiegel, Prof. J. M. Crawford

Department of Chemistry, Yale University  
225 Prospect Street, New Haven, CT 06511 (USA)

[b] N. Y. Kim, T. N. Goddard, Prof. J. M. Crawford  
Chemical Biology Institute, Yale University  
600 West Campus Drive, West Haven, CT 06516 (USA)

[c] Prof. D. A. Spiegel  
Department of Pharmacology, Yale School of Medicine  
333 Cedar Street, New Haven, CT 06520 (USA)

[d] Prof. J. M. Crawford  
Department of Microbial Pathogenesis, Yale School of Medicine  
333 Cedar Street, New Haven, CT 06536 (USA)  
E-mail: jason.crawford@yale.edu

Supporting information for this article is given via a link at the end of the document.

## FULL PAPER

extent CML modifications to restore the native lysine structure *in vitro*. MnmC naturally functions in macromolecular RNA modification as both an oxidase and a methyltransferase with  $K_m$  values of  $600 \pm 200$  nM and  $70 \pm 40$  nM for oxidation and methyltransferase activities, respectively.<sup>[35]</sup> However, here we show that the oxidase domain of the enzyme exhibits promiscuity for the AGE substrates at higher concentrations. Through homology analysis, site-directed mutagenesis, and protein domain dissection studies, we generated variants of the enzyme with an enhanced capacity to cleave CEL. We also demonstrate that the enzyme not only cleaves CEL in free amino acid form, but also cleaves CEL in both peptidomimetic and physiologically-relevant peptide contexts. To the best of our knowledge, this study provides the first biochemical characterization of a biocatalyst that is capable of reversing a mature AGE in a peptide context. MnmC variants could serve as directed evolution leads toward the future development of new molecular AGE tools with improved kinetic parameters for CEL and CML reversal.

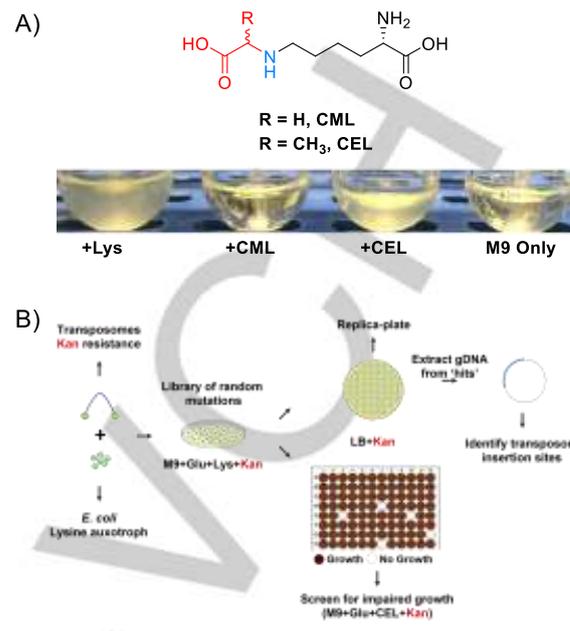
## Results

## Identification of a CEL-breaking biocatalyst

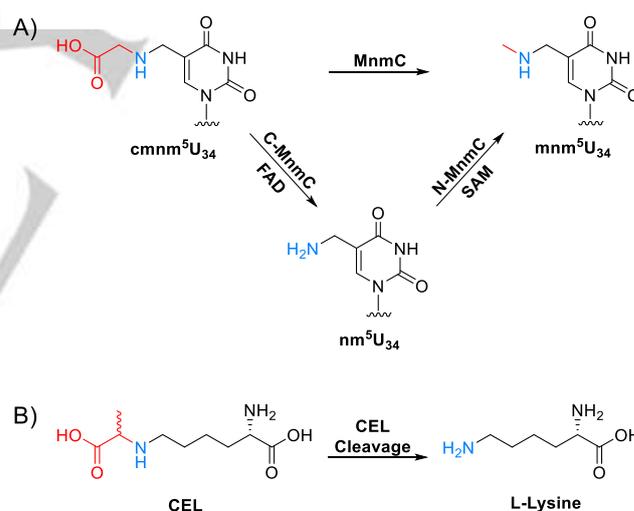
To assess whether *E. coli* was capable of cleaving the N<sup>ε</sup>-C bond in CEL or CML to restore Lys, we fed free CML or CEL as a sole Lys source to an *E. coli* Lys auxotroph ( $\Delta lysA$ ). *E. coli*  $\Delta lysA$  was cultivated in minimal M9 medium supplemented with 1 mg/mL Lys (positive control), CML, CEL or water vehicle (negative control) at 30 °C overnight under aerobic conditions. Strong growth was observed only for cultures supplemented with either Lys or CEL, indicating that *E. coli* can utilize CEL as a sole Lys source (Figure 1A). To identify the biocatalyst(s) responsible for CEL cleavage, we performed transposon mutagenesis in *E. coli*  $\Delta lysA$  and screened for mutants with impaired growth using CEL as a source of Lys (Figure 1B). One mutant from our non-saturating transposon mutant library (~3000 mutants were analyzed) exhibited impaired growth. The transposon insertion site was in *mnmG* (a.k.a., *gidA*), a gene located in the bacterial MnmEG pathway responsible for wobble U<sub>34</sub> tRNA modifications.<sup>[36,37]</sup> However, isolated MnmG was not able to cleave CEL *in vitro* (data not shown). Analysis of the broader pathway led us to further study MnmC, an enzyme that catalyzes a tRNA modification reaction analogous to CEL cleavage (Figure 2A, Figure 2B).<sup>[38]</sup>

## Characterization of MnmC and validation of its activity on CEL

MnmC is a bifunctional enzyme that is known to catalyze two steps in the biosynthesis of the 5-methylaminomethyl-2-thiouridine ( $mnm^5(S^2)U_{34}$ ) nucleoside in tRNA. The C-terminal domain (C-MnmC) is an FAD-dependent oxidase that catalyzes removal of the carboxymethyl functionality in  $cmnm^5U_{34}$  to generate a primary amine, and the N-terminal domain (N-MnmC) is a SAM-dependent methyltransferase that catalyzes methylation of this primary amine in  $nm^5U_{34}$  to yield  $mnm^5U_{34}$  (Figure 2A).



**Figure 1.** A) Growth of *E. coli* lysine auxotroph in M9 minimal medium supplemented with lysine, CML, CEL, or water vehicle control. B) Identification of genes required for CEL-cleavage through transposon mutagenesis.

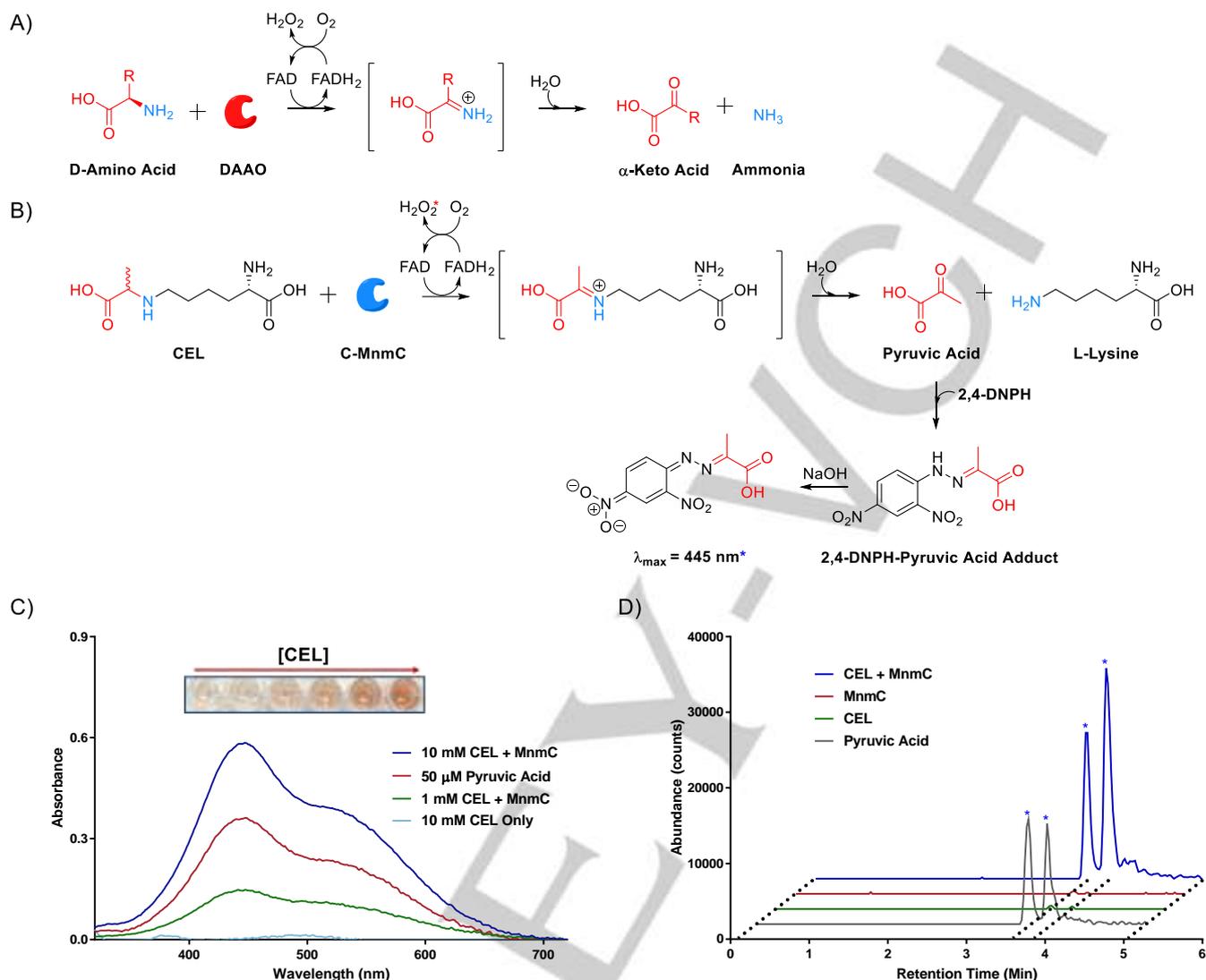


**Figure 2.** Analogy between the transformation of  $cmnm^5U_{34}$  to  $nm^5U_{34}$  by C-terminal domain of MnmC (C-MnmC) and the restoration of lysine from CEL. A) Two consecutive reactions catalyzed by the bifunctional tRNA-modifying enzyme MnmC in *E. coli*. C-MnmC catalyzes FAD-dependent removal of carboxymethyl group in  $cmnm^5U_{34}$  that generates a primary amine which is methylated by SAM-dependent N-MnmC. B) Lysine restoration from CEL cleavage. In both transformations, removal of carboxymethyl- or carboxyethyl-functionality generates a primary amine.

These domains are fused by an interdomain linker which, when cleaved, yields the two domains in isolation that still retain catalytic activity.<sup>[35,39]</sup> The C-MnmC domain-catalyzed removal of a carboxymethyl functionality to a primary amine in a

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER



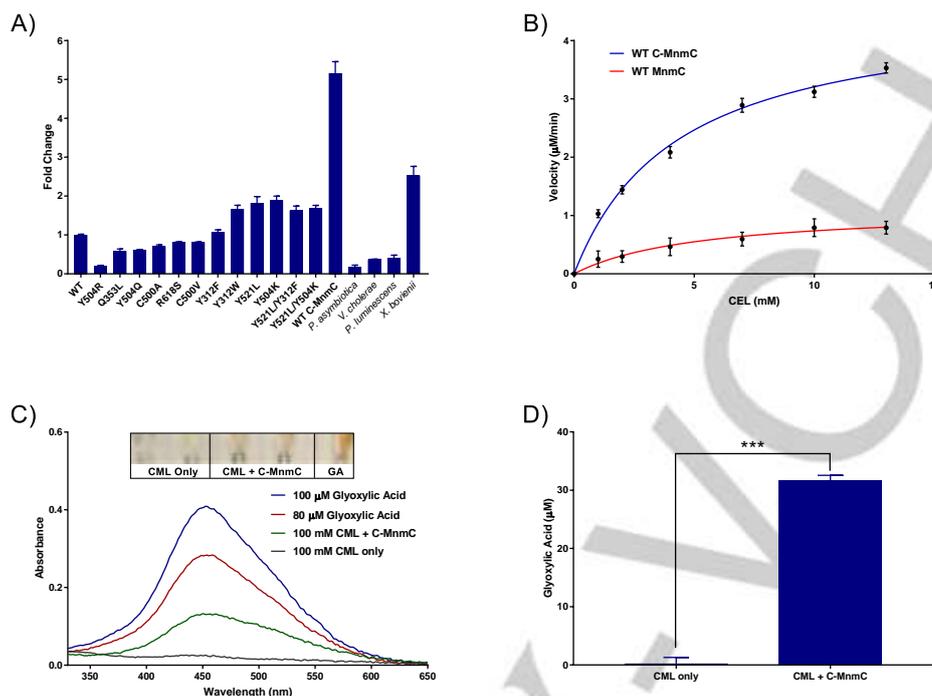
**Figure 3.** Validation of C-MnmC-mediated CEL cleavage using an  $\alpha$ -keto acid derivatization assay. A) Mechanism of deamination by D-amino acid oxidases (DAAO). B) Proposed mechanism of CEL cleavage and the formation of 2,4-DNPH-pyruvic acid hydrazone adduct. \*MnmC exhibits hydrogen peroxide scavenging activity. C) Detection of CEL cleavage-dependent hydrazone adduct formation spectrophotometrically. The adduct appears red-brown in color and its absorbance spectrum shows  $\lambda_{\max}$  at 445 nm, and a shoulder at 515 nm. D) LC-MS validation of MnmC-dependent generation of *syn*- and *anti*- isomers (\*) of the hydrazone adduct ( $m/z$  267, negative ion mode).

macromolecular target is analogous to the removal of a carboxyethyl (or carboxymethyl) functionality in the desired CEL (or CML) cleavage reaction (Figure 2B). D-amino acid oxidases (DAAOs) are the closest structural homologs of C-MnmC. Because DAAOs are FAD-dependent enzymes known to catalyze the oxidative deamination of various amines and neutral D-amino acids (e.g., glycine, sarcosine, D-alanine and D-proline) to give their corresponding  $\alpha$ -keto acids, ammonia, and hydrogen peroxide (Figure 3A), we first assessed MnmC's ability to turnover CEL using an established spectrophotometric hydrogen peroxide assay using horseradish peroxidase (HRP) and chromogenic substrates.<sup>[40]</sup> Like DAAOs, we expected that MnmC-mediated cleavage of CEL could involve oxidation of the N<sup>ε</sup>-C bond to

generate an imine, which is then hydrolyzed to yield pyruvic acid, L-Lys, and hydrogen peroxide (Figure 3B). However, we did not detect hydrogen peroxide production, as this product was consumed by MnmC (see Figure S1 for MnmC-mediated quenching of hydrogen peroxide). Thus, to demonstrate that MnmC exhibits activity toward CEL, we then adapted an  $\alpha$ -keto-acid derivatization assay which can be used to monitor pyruvic acid formation as a result of CEL cleavage (Figure 3B).<sup>[41]</sup> In short, 2,4-dinitrophenylhydrazine (2,4-DNPH) readily reacts with pyruvic acid to produce the 2,4-DNPH-pyruvic acid hydrazone adduct which has a distinctive red-brown color and absorbance maximum at 445 nm that can be monitored to measure the activity of MnmC on CEL (Figure 3C). At initial CEL concentrations of 1

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER



**Figure 4.** A) Relative activities of MnmC variants. Standalone WT C-MnmC shows about 5-fold improvement in CEL-cleaving activity. B) Steady-state kinetics of WT MnmC and WT C-MnmC showing concentration dependencies with CEL. C) Detection of CML cleavage-dependent 2,4-DNPH-glyoxylic acid hydrazone adduct formation spectrophotometrically. The adduct appears orange in color and its absorbance spectrum shows  $\lambda_{\max}$  at 455 nm. D) Quantitation of the amount of glyoxylic acid generated from CML cleavage by C-MnmC.

mm and 10 mM, we detected formation of 2,4-DNPH-pyruvic acid hydrazone adduct in an MnmC-dependent manner. Formation of 2,4-DNPH-pyruvic acid hydrazone products (*syn*- and *anti*-isomers) were further confirmed by LC-MS using authentic standards (Figure 3D).

In addition to CEL, we also tested whether MnmC is active on CML, which bears the carboxymethyl functionality similar to its native tRNA substrate. Despite the resemblance, we did not observe MnmC activity on CML at a 10 mM concentration, presumably due to differences in nonstandard substrate binding. This result is consistent with the observation that the *E. coli* Lys auxotroph does not grow in the presence of CML as a sole Lys source. Consistent with our overall findings, we did observe weak CML turnover using an MnmC variant at higher substrate concentrations (100 mM, see below).

### Structure-based engineering of MnmC

To enhance the activity of MnmC on CEL, we evaluated the X-ray crystal structure of *E. coli* MnmC (PDB: 3AWI). We first generated a  $\text{cmnm}^5\text{U}_{34}$  binding model of MnmC, and identified 10 residues around the bound substrate that could potentially be mutated to better accommodate the longer hydrophobic alkyl chain of CEL (Figure S2). Relative activities of site-directed mutants at these positions were then compared using the  $\alpha$ -keto acid derivatization assay (10 mM CEL, 5  $\mu\text{M}$  MnmC, 4 hours at 37 °C). Three single residue mutants (Y312W, Y521L, and Y504K) showed only about

a 2-fold increase in relative activities, and combinatorial double mutants did not show further improvements (Figure 4A). We then reasoned that since the reactions catalyzed by C-MnmC and N-MnmC are thought to be kinetically tuned,<sup>[35]</sup> dissection and analysis of the desired C-terminal monodomain could exhibit improved properties. Thus, we cleaved off the N-terminal domain using a cut site within the interdomain linker (between Leu250 and Pro251), allowing isolated C-MnmC to be expressed in a catalytically competent form. Compared to the previously isolated C-MnmC,<sup>[39]</sup> our C-MnmC variant contains an N-terminal His<sub>6</sub>-tag with a shorter linker. This isolated C-MnmC monodomain construct exhibited about a 5-fold enhancement in relative activity compared to the wildtype didomain MnmC (Figure 4A), suggesting that the N-terminal domain could participate in quality control. In addition to *E. coli* MnmC, we also tested the activities of didomain MnmC homologs from related *Gammaproteobacteria*, *Phototribadus asymbiotica*, *Vibrio cholerae*, *Phototribadus luminescens*, and *Xenorhabdus bovienii*, and their activities indicate that MnmC variants could more broadly catalyze CEL cleavage (Figure 4A).

Following the relative analyses of MnmC variants, we established the steady-state kinetics of wildtype didomain MnmC and monodomain C-MnmC for CEL using the  $\alpha$ -keto acid derivatization assay and a 2,4-DNPH-pyruvic acid standard curve (Figure 4B, Figure S3). The  $K_m$  values for MnmC and C-MnmC were comparable at  $5.2 \pm 1.3$  mM and  $4.3 \pm 0.6$  mM, respectively,

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER

**Table 1.** Kinetic parameters for CEL cleavage by MnmC and C-MnmC<sup>[a]</sup>

	$K_m$ (mM)	$k_{cat}$ (min <sup>-1</sup> )	$k_{cat} / K_m$ (M <sup>-1</sup> min <sup>-1</sup> )
MnmC	5.2 ± 1.3	0.04 ± 0.004	8 ± 1
C-MnmC	4.3 ± 0.6	0.18 ± 0.009	42 ± 6

[a] Reaction mixtures contained CEL (1–13 mM), MnmC or C-MnmC (25 μM), and FAD (5 μM) in 50 mM Tris buffer (pH 8.0). Kinetic parameters were calculated from GraphPad software (version 7.0) and are reported as mean ± SD. Experimental time points were run in duplicate.

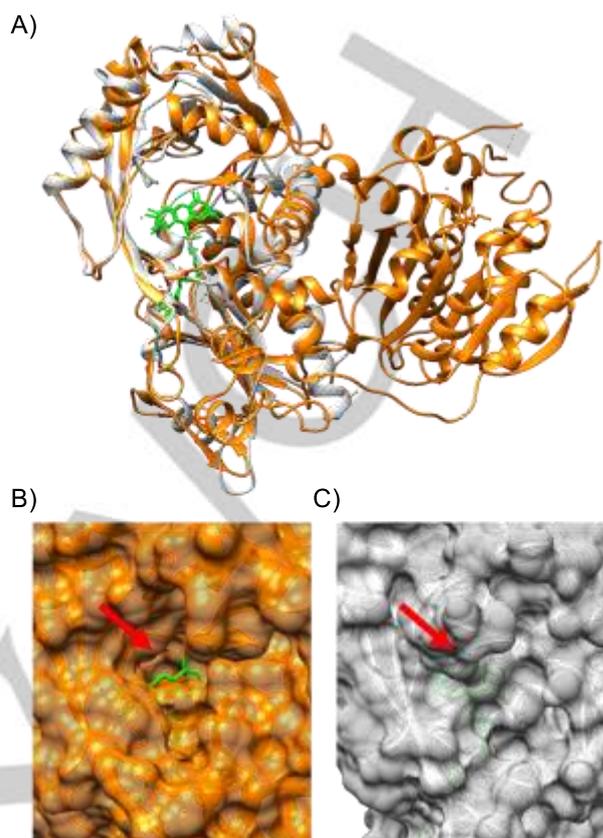
whereas  $k_{cat}$  values were 0.04 ± 0.004 min<sup>-1</sup> and 0.18 ± 0.009 min<sup>-1</sup>, respectively. These data indicate that the ~5.3-fold improvement of C-MnmC is largely a function of  $k_{cat}$ . Table 1 summarizes the kinetic parameters for CEL cleavage by MnmC and C-MnmC.

With the improved C-MnmC monodomain construct, we evaluated CML as a substrate. CML turnover was only observed at a higher substrate concentration (100 mM CML, 20 μM C-MnmC, 4 hours, 37 °C). C-MnmC-catalyzed CML cleavage leads to glyoxylic acid and lysine (Figure S4). Spectrophotometric detection of 2,4-DNPH-glyoxylic acid hydrazone adduct was monitored at 455 nm (Figure 4C) and quantified using a 2,4-DNPH-glyoxylic acid standard curve (Figure 4D, Figure S5).

### Substrate scope of C-MnmC

To have utility in probing the biological roles of CEL and CML AGEs, reversal biocatalysts must function on CEL/CML-modified peptide substrates. Encouraged by the fact that MnmC natively uses a macromolecular substrate like mature AGEs, we further evaluated the crystal structures of *E. coli* MnmC versus human DAAO. Superposition of the two crystal structures shows a high degree of structural similarity between the C-terminal domain of MnmC and DAAO (Figure 5A). However, when the surface representations of the active sites for the two enzymes are compared, it becomes apparent that C-MnmC exhibits a wide and exposed active site whereas DAAO has a closed active site (Figure 5B, Figure 5C). This difference likely arises from the fact that MnmC acts on a macromolecular tRNA substrate, whereas DAAO acts on small molecule substrates (e.g., D-amino acids). The open active site in C-MnmC suggests that the enzyme could also utilize CEL in peptide contexts.

To investigate the substrate scope of C-MnmC, we first synthesized diketopiperazine-CEL (DKP-CEL) and diketopiperazine-CML (DKP-CML) as peptidomimetics of mature CEL and CML AGEs. Diketopiperazines are cyclic dipeptides that mimic protein beta turns.<sup>[42]</sup> Consistent with weak turnover observed for CML in free amino acid form, under our experimental conditions C-MnmC was only able to turnover peptidomimetic DKP-CEL (Figure 6A) at a 10 mM substrate concentration (Figure 6B), suggesting that C-MnmC may cleave mature CEL residues in peptides. Activity for DKP-CEL was ~15% relative to free CEL.



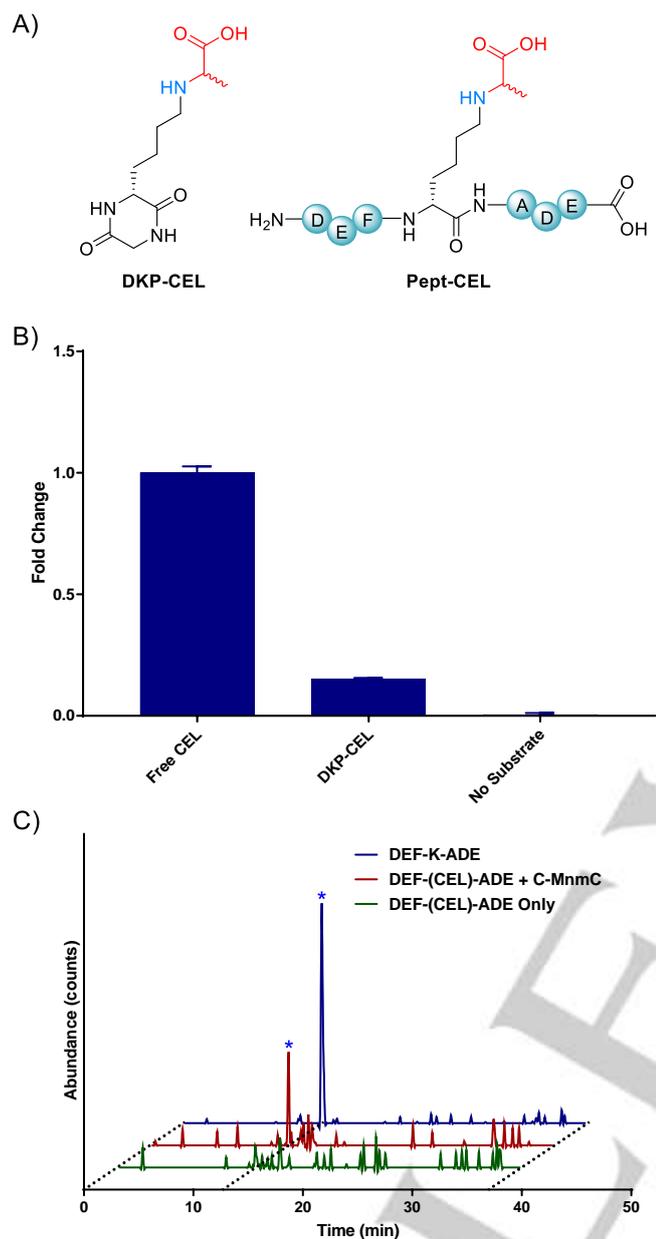
**Figure 5.** Comparison of the crystal structures of MnmC from *E. coli* (PDB: 3PS9, orange) and D-amino acid oxidase from human (PDB: 2DU8, grey). A) Superposition of MnmC and D-amino acid oxidase. B) Active site of the C-terminal domain of MnmC (indicated by red arrow). C) Active site of D-amino acid oxidase (indicated by red arrow). FAD is shown in green.

We then proceeded to test C-MnmC for its activity in a peptide context by synthesizing a short linear peptide substrate DEF-(CEL)-ADE (Pept-CEL). Pept-CEL is a mature CEL-modified peptide (Figure 6A) that is known to bind to the positively charged cavity of the V domain of RAGE largely through its interaction with the carboxyethyl functionality.<sup>[27,28]</sup> To detect activity of C-MnmC on Pept-CEL, an overnight enzymatic reaction containing 16 mM peptide substrate and 40 μM C-MnmC was analyzed by LC-MS. Pept-CEL was modestly converted to DEF-K-ADE in the presence of C-MnmC (Figure 6C). We performed HRMS/MS experiments to sequence the substrate and product peptides to validate C-MnmC-mediated CEL to Lys residue cleavage (Figure 7A, Figure 7B). These studies identified C-MnmC as the first catalyst capable of reversing a mature AGE modification and suggest that C-MnmC orthologs could serve as directed evolution leads for further development.

### Discussion

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER



**Figure 6.** Activity of WT C-MnmC on a peptidomimetic substrate DKP-CEL and a peptide substrate Pept-CEL. A) Structures of DKP-CEL and Pept-CEL. B) Activity of C-MnmC on DKP-CEL relative to its activity on free CEL. C) LC-MS detection of C-MnmC-mediated cleavage of Pept-CEL and the release of DEF-K-ADE (\*).

Despite numerous studies establishing AGEs as contributing factors to the progression of degenerative diseases, it has been a challenge to establish causal relationships between them.<sup>[20]</sup> One major challenge is an inability to reverse AGE modifications and evaluate potential disease amelioration. Biocatalysts that are capable of reversing specific mature AGEs could lead to a new arsenal of molecular tools for the study of AGEs, and if successful, to recombinant-enzyme therapies. However, no such tools currently exist. Here, we found that an *E. coli* lysine auxotroph can

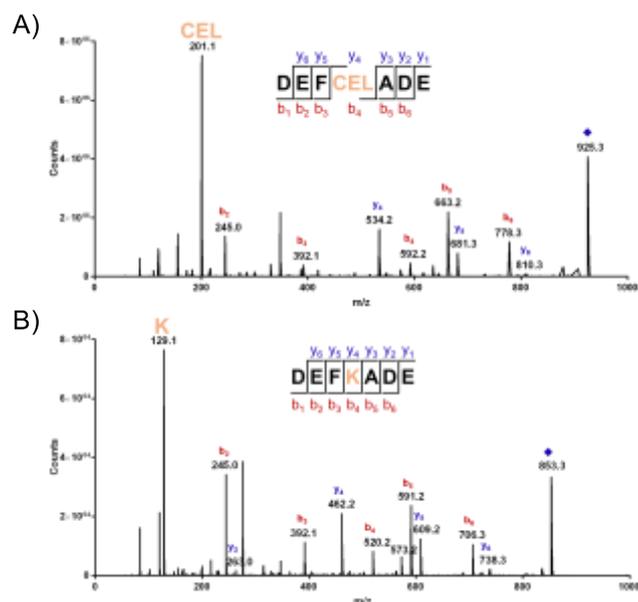
utilize CEL as a sole lysine source. Through a combination of transposon mutagenesis, screening, synthesis, and protein biochemical studies, we identify and characterize MnmC as an enzyme capable of CEL cleavage and Lys restoration. While our non-saturating transposon studies initially led to the identification of *mnmG*, we were unable to show MnmG-mediated CEL cleavage. This prompted us to examine the broader bacterial MnmEG pathway. Previous reports on the MnmEG pathway established that it is tightly controlled,<sup>[39]</sup> suggesting the possibility that *mnmG* disruption could affect other genes in the pathway. This led us to focus on MnmC, which catalyzes an analogous reaction to CEL/CML cleavage.

Previous homology searches on MnmC revealed that its C-terminal domain is structurally related to glycine oxidases (GOXs), sarcosine oxidases (SOXs), and DAAOs.<sup>[43,44]</sup> As FAD-dependent oxidases, these enzymes catalyze a similar enzymatic reaction where an N-C bond is oxidized to an imine which is subsequently hydrolyzed. Intriguingly, C-MnmC is known to catalyze a reaction in which imine generation followed by hydrolysis results in a decarboxymethylation to release a primary amine. Our homology modeling efforts showed that C-MnmC is structurally closest to the DAAOs. High structural homology between C-MnmC and DAAO could explain the fact that CEL, which is in effect an *N*-alkyl alanine, is a better substrate than CML, which is in effect an *N*-alkyl glycine. DAAOs are known to accept glycine as a substrate, albeit with much lower affinity than functionalized D-amino acids (Table S2).<sup>[45]</sup> This is consistent with our observation that C-MnmC, as a homolog of DAAOs, shows much weaker activity toward CML than CEL *in vitro* and in cell culture. Additionally, in a recent study, CML was reported to be converted to *N*-carboxymethylcadaverine (CM-CAD), *N*-carboxymethylaminopentanoic acid (CM-APA), and *N*-carboxymethyl- $\Delta^1$ -piperideinium ion by probiotic *E. coli* strains.<sup>[46]</sup> These data further support our cell culture observations and suggest that despite structural similarities between CEL and CML, their catabolism pathways in *E. coli* appear to be divergent.

The capability of C-MnmC to remove CEL from free, peptidomimetic, and mature peptide contexts can likely be attributed to the enzyme's open and large active site. Similarly, Amadoriases I and II, which are well-known FAD-dependent oxidases capable of oxidizing the early Amadori product (e.g., fructoselysine), are also known to be functional on peptide substrates with similar  $K_m$  values (4.2 mM and 5.6 mM for glycosylated-Lys and glycosylated-Lys-Phe, respectively).<sup>[47]</sup> Upon examination of the active sites of these Amadoriases, it becomes apparent that their active sites are also quite large, which could explain how these enzymes are able to function on peptide substrates unlike DAAOs and GOXs with small, buried active sites (Figure S6). One exception is the DAAO from the yeast *Rhodotorula gracilis* (RgDAAO). With an unusually large active site, the RgDAAO is known to accept cephalosporin C and convert it to glutaryl-7-aminocephalosporanic acid, a key intermediate in the enzyme-catalyzed synthesis of cephalosporin antibiotics.<sup>[48]</sup> CEL/CML cleavage and cephalosporin C inactivation mediated by these types of enzymes represent alternative activities that could be exploited for non-canonical functions.

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER



**Figure 7.** Representative tandem mass spectra for pre- and post-cleavage of CEL functionality by C-MnmC. A) HRMS/MS spectrum for pre-cleavage substrate DEF-(CEL)-ADE showing fragments that contain CEL. B) HRMS/MS spectrum for post-cleavage product DEF-K-ADE showing fragments that contain lysine instead of CEL. Addition of carboxyethyl functionality on a lysine side chain corresponds to an increase of 72 units in mass, and this difference is observed between fragments detected in pre- and post-cleavage tandem MS spectra.

## Conclusion

In summary, our study has shown that C-MnmC is a biocatalyst that can cleave mature CEL modifications to restore native lysine structures. To the best of our knowledge, this is the first biochemical demonstration of an enzyme that can reverse a mature AGE-functionalized peptide. While the kinetic parameters, which are similar to known Amadoriases, could be substantially improved, C-MnmC variants represent lead catalysts for further directed evolution and development. As MnmC natively acts on nucleic acids, glycosylated DNA (e.g., carboxyethyl/carboxymethyl-deoxyguanosine) may also be suitable substrates to test in future studies. Such improved AGE-reversal tools could in principle enable a better understanding of the biology of AGEs at the molecular level, elucidate their direct roles in the pathogenesis of age-related diseases, and serve as leads for recombinant enzyme therapies.

## Experimental Section

**General experimental procedures:** Low-resolution electrospray ionization-mass spectrometry (ESI-MS) data were collected on an Agilent 6120 Quadrupole LC/MS system equipped with a Phenomenex Kinetex C<sub>18</sub> (100 Å) 5- $\mu$ m (4.6 x 250 mm) column. High-resolution (HR) ESI-MS data were obtained using an Agilent iFunnel 6550 QTOF MS instrument

fitted with an electrospray ionization (ESI) source coupled to an Agilent 1290 Infinity HPLC system. Spectrophotometric measurements were performed using a Thermo Scientific NanoDrop™ 2000c Spectrophotometer. Protein modelling and active site comparisons were done using MOE and Chimera 1.12.<sup>[49,50]</sup>

**Materials:** N<sup>ε</sup>-(carboxyethyl)lysine (CEL) was purchased from Focus Synthesis, LLC. N<sup>ε</sup>-(carboxymethyl)lysine (CML) was purchased from Chem-Impex International. EZ-Tn5™ <KAN-2>TNP Transposome™ Kit was purchased from Epicentre. Ultra-micro (8.5 mm window height) cuvettes were purchased from BrandTech Scientific. Peptide Fmoc-D(otbu)E(otbu)FK(ivDDE)AD(otbu)E(otbu)-wang resin was synthesized by Biomatik.

**Strains:** *E. coli* lysine auxotroph ( $\Delta$ lysA) harboring kanamycin resistance gene (JW2806-1, Keio Collection) was obtained from Coli Genetic Stock Center at Yale University. In order to utilize transposomes in EZ-Tn5™ <KAN-2>TNP Transposome™ Kit which also contain the kanamycin resistance gene, we generated an *E. coli* lysine auxotroph strain sensitive to kanamycin. To this end, we excised the kanamycin resistance gene in JW2806-1 through FLP recombination using the plasmid pCP20 encoding the FLP recombinase, which yielded a kanamycin-sensitive lysine auxotroph (JW2806-1<sub>flpout</sub>) that was used as the host strain in the transposon mutagenesis studies.

**Sole lysine source growth studies, transposon mutagenesis and screening:** To evaluate whether *E. coli* can grow in presence of CEL or CML as sole lysine source, *E. coli* strain JW2806-1<sub>flpout</sub> was cultivated in 0.5 mL M9 minimal medium supplemented with 1 mg/mL lysine, 0.1% CEL, or 0.1% CML at 30 °C overnight under aerobic conditions. The same *E. coli* strain was then used as a host to generate a transposon mutant library using the EZ-Tn5™ <KAN-2>TNP Transposome™ Kit following manufacturer's protocol. Briefly, electrocompetent JW2806-1<sub>flpout</sub> cells were electroporated using 1  $\mu$ L of the EZ-Tn5 <KAN-2> TNP Transposome. 1 mL of SOC medium was then added to the cells and incubated on a 37 °C shaker for 1 hour to facilitate cell outgrowth. Dilutions of the aliquots of the recovered cells were then plated on M9 minimal medium agar plates supplemented with 50  $\mu$ g/mL kanamycin and 1 mg/mL lysine. Resulting colonies were then picked into 96-well plates containing 300  $\mu$ L of M9 minimal medium supplemented with 0.1% CEL and 50  $\mu$ g/mL kanamycin, and cultured at 30 °C overnight under aerobic conditions. Genomic DNA from transposon mutants that showed impaired growth was then extracted using the DNeasy Blood and Tissue Kit (Qiagen), which was used as template for arbitrarily primed PCR.<sup>[51]</sup> The amplified DNA segments near the transposon insertion sites were sequenced by Yale Keck Sequencing facility and subjected to BLAST analysis to identify the disrupted gene.

**MnmC expression and purification:** N-terminally His<sub>6</sub>-tagged MnmC was cloned and expressed as previously reported<sup>[35]</sup> with slight modifications. In brief, DNA encoding MnmC was amplified from *E. coli* BW25113 genomic DNA with primers MnmC-fw and MnmC-rv (Table S1), which contain an NdeI site at the 5'-end and a BglII site at the 3'-end. Upon gel extraction (NucleoSpin Gel and PCR Clean-up, Macherey Nagel), the amplified sequence was cloned into pET28a (Novagen) expression vector between NdeI and BamHI sites. N-terminally His<sub>6</sub>-tagged MnmC was then expressed in *E. coli* BL21 (DE3) (1L of TB media), grown to an OD<sub>600</sub> of 0.5 at 37 °C, induced with 0.4 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside, and further cultivated under aerobic conditions (250 rpm) overnight at 20 °C. Cells were lysed by sonication, and protein was purified from cell lysate using affinity chromatography (Ni-NTA agarose, Qiagen). Purified protein was then concentrated using an Amicon Ultra-15 centrifugal filter unit with 50 kDa cutoff, and prepared into glycerol stocks for storage at -20 °C. To clone N-terminally His<sub>6</sub>-tagged C-MnmC (Pro251 to stop codon), primers

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER

C-MnmC-fw and C-MnmC-rv (Table S1) were used to amplify DNA from *E. coli* BW25113 genomic DNA, and the protein was isolated using identical purification steps as in MnmC before concentrating with Amicon Ultra-15 centrifugal filter unit with 30 kDa cutoff.

**Site-directed mutagenesis:** MnmC mutants were generated via the QuickChange mutagenesis method with the corresponding pairs of primers shown in Table S1 in the Supporting Information. pET28a construct containing MnmC gene was used as a template. All mutations were validated by sequencing through Yale Keck Sequencing facility.

**Hydrogen peroxide consumption assay:** Hydrogen peroxide consumption by MnmC was monitored using a spectrophotometric hydrogen peroxide assay as previously described.<sup>140</sup> A typical experiment involved addition of H<sub>2</sub>O<sub>2</sub> (40 μM) to MnmC (10 μM or 20 μM) or C-MnmC (10 μM) in a total volume of 200 μL in phosphate buffer (0.2 M, pH 7.4) and incubation at 37 °C for 1 h. 40 μL of chromogenic solution (1 mM vanillic acid, 0.5 mM 4-aminoantipyrine, 4 U/mL peroxidase in phosphate buffer) was then added to the mixture, and incubated at 37 °C for 20 min for color development before measuring the absorbance at 490 nm.

**Cleavage assay for CEL and DKP-CEL:** In a typical cleavage assay, CEL or DKP-CEL (1 mM – 10 mM) was reacted with MnmC (1 μM – 5 μM) in PBS (pH 7.4) supplemented with FAD (5 μM) in a total volume of 100 μL for 2–4 h at 37 °C. 50 μL of 2,4-DNPH solution (1 mM in 1 N HCl) was then added to the mixture, incubated at 37 °C for 10 min, basified with 350 μL of 0.6 N NaOH for color development, and the absorbance of the red-brown mixture at 445 nm was measured. Calibration curve was obtained using pyruvate, which allowed determination of the amount of pyruvate generated after cleavage of CEL or DKP-CEL.

**Cleavage assay for Pept-CEL:** Pept-CEL (16 mM) was reacted with C-MnmC (40 μM) in Tris buffer (50 mM, pH 8.0) supplemented with FAD (20 μM) in a total volume of 100 μL for 18 h at rt. The reaction mixture was blow-dried with nitrogen gas, and then extracted using 200 μL of MeOH before being subjected to LC-MS and HRMS/MS analysis.

**Cleavage assay for CML:** CML (100 mM) was reacted with C-MnmC (25 μM) in Tris buffer (50 mM, pH 8.0) supplemented with FAD (20 μM) in a total volume of 50 μL for 4 h at 37 °C. 25 μL of 2,4-DNPH solution (1 mM in 1 N HCl) was then added to the mixture, incubated at 37 °C for 10 min, basified with 175 μL of 0.6 N NaOH for color development, and the absorbance of the orange mixture at 455 nm was measured. Background absorbance from the enzyme controls were subtracted from the experiments, and the amount of glyoxylic acid generated via cleavage of CML was determined using calibration curve obtained using glyoxylic acid.

**Enzyme kinetics:** The reaction mixtures (50 μL) consisted of 50 mM Tris buffer (pH 8.0), 25 μM enzyme, 5 μM FAD, and CEL at various concentrations (1 mM – 13 mM). All reactions were incubated at 37 °C, and stopped after 1 min – 21 min by adding 25 μL of 2,4-DNPH solution (1 mM in 1 N HCl). 175 μL of 0.6 N NaOH was then added to each reaction and the absorbance at 445 nm was measured.  $K_m$  and  $k_{cat}$  were determined from Michaelis-Menten plots using GraphPad Prism 7.

**Synthesis of DKP-CEL and Pept-CEL:** Detailed synthetic procedures are described in the supporting information.

## Acknowledgements

We thank the Yale Claude D. Pepper Older Americans Independence Center (Pilot Grant to J.M.C. and D.A.S.), the SENS Research Foundation (Grant to D.A.S.), the American Diabetes Association Pathway to Stop Diabetes (Grant 1-17-VSN-04 to D.A.S.), and Yale University for financial support.

**Keywords:** Enzyme catalysis • oxidoreductases • protein engineering • protein modifications

- [1] S. J. Cho, G. Roman, F. Yeboah, Y. Konishi, *Curr. Med. Chem.* **2007**, *14*, 1653-1671.
- [2] M. Hellwig, T. Henle, *Angew. Chem, Int, Ed, Engl.* **2014**, *53*, 10316-10329.
- [3] R. D. Semba, E. J. Nicklett, L. Ferrucci, *J. Gerontol. A. Biol. Sci. Med. Sci.* **2010**, *65*, 963-975.
- [4] S. Yamagishi, N. Nakamura, M. Suematsu, K. Kaseda, T. Matsui, *Mol. Med.* **2015**, *21 Suppl 1*, S32-40.
- [5] J. Xie, J. D. Mendez, V. Mendez-Valenzuela, M. M. Aguilar-Hernandez, *Cell Signal* **2013**, *25*, 2185-2197.
- [6] A. D. Bhatwadekar, J. V. Glenn, G. Li, T. M. Curtis, T. A. Gardiner, A. W. Stitt, *Invest. Ophthalmol. Vis. Sci.* **2008**, *49*, 1232-1241.
- [7] J. M. Forbes, M. E. Cooper, M. D. Oldfield, M. C. Thomas, *J. Am. Soc. Nephrol.* **2003**, *14*, S254-258.
- [8] A. Goldin, J. A. Beckman, A. M. Schmidt, M. A. Creager, *Circulation* **2006**, *114*, 597-605.
- [9] Z. Hegab, S. Gibbons, L. Neyses, M. A. Mamas, *World J. Cardiol.* **2012**, *4*, 90-102.
- [10] J. Li, D. Liu, L. Sun, Y. Lu, Z. Zhang, *J. Neurol. Sci.* **2012**, *317*, 1-5.
- [11] V. M. Monnier, G. T. Mustata, K. L. Biemel, O. Reihl, M. O. Lederer, D. Zhenyu, D. R. Sell, *Ann. N. Y. Acad. Sci.* **2005**, *1043*, 533-544.
- [12] V. M. Monnier, D. R. Sell, R. H. Nagaraj, S. Miyata, S. Grandhee, P. Odetti, S. A. Ibrahim, *Diabetes* **1992**, *41 Suppl 2*, 36-41.
- [13] R. D. Semba, K. Sun, A. V. Schwartz, R. Varadhan, T. B. Harris, S. Satterfield, M. Garcia, L. Ferrucci, A. B. Newman, Health ABC Study. *J. Hypertens.* **2015**, *33*, 797-803.
- [14] V. M. Monnier, R. R. Kohn, A. Cerami, *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 583-587.
- [15] D. G. Dyer, J. A. Dunn, S. R. Thorpe, K. E. Bailie, T. J. Lyons, D. R. McCance, J. W. Baynes, *J. Clin. Invest.* **1993**, *91*, 2463-2469.
- [16] A. W. Stitt, *Ann. N. Y. Acad. Sci.* **2005**, *1043*, 582-597.
- [17] R. H. Nagaraj, M. Linetsky, A. W. Stitt, *Amino Acids.* **2012**, *42*, 1205-1220.
- [18] C. Draghici, T. Wang, D. A. Spiegel, *Science.* **2015**, *350*, 294-298.
- [19] M. Fournet, F. Bonte, A. Desmouliere, *Aging Dis.* **2018**, *9*, 880-900.
- [20] J. Chaudhuri, Y. Bains, S. Guha, A. Kahn, D. Hall, N. Bose, A. Gugliucci, P. Kapahi, *Cell Metab.* **2018**, *28*, 337-352.
- [21] C. Henning, K. Liehr, M. Girndt, C. Ulrich, M. A. Glomb, *J. Biol. Chem.* **2014**, *289*, 28676-28688.
- [22] M. Asif, J. Egan, S. Vasan, G. N. Jyothirmayi, M. R. Masurekar, S. Lopez, C. Williams, R. L. Torres, D. Wagle, P. Ulrich, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 2809-2813.
- [23] G. Richarme, C. Liu, M. Mihoub, J. Abdallah, T. Leger, N. Joly, J. C. Liebart, U. V. Jurkunas, M. Nadal, P. Bouloc, J. Dairou, A. Lamouri, *Science* **2017**, *357*, 208-211.
- [24] S. Ferri, S. Kim, W. Tsugawa, K. Sode, *J. Diabetes Sci. Technol.* **2009**, *3*, 585-592.
- [25] C. Delgado-Andrade, *Food Funct.* **2016**, *7*, 46-57.
- [26] M. U. Ahmed, S. R. Thorpe, J. W. Baynes, *J. Biol. Chem.* **1986**, *261*, 4889-4894.
- [27] J. Xue, V. Rai, D. Singer, S. Chabierski, J. Xie, S. Reverdatto, D. S. Burz, A. M. Schmidt, R. Hoffmann, A. Shekhtman, *Structure* **2011**, *19*, 722-732.
- [28] J. Xie, S. Reverdatto, A. Frolov, R. Hoffmann, D. S. Burz, A. Shekhtman, *J. Biol. Chem.* **2008**, *283*, 27255-27269.

For internal use, please do not delete. Submitted\_Manuscript

## FULL PAPER

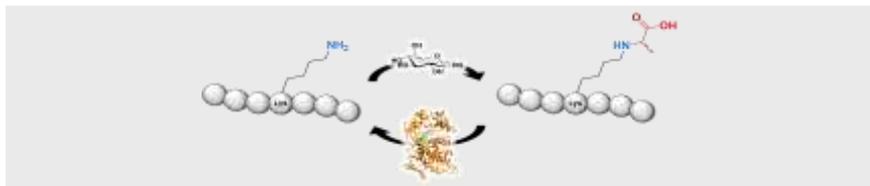
- [29] O. T. Lee, S. D. Good, R. Lamy, M. Kudisch, J. M. Stewart, *Invest. Ophthalmol. Vis. Sci.* **2015**, *56*, 2892-2897.
- [30] C. Y. Lin, J. J. Sheu, I. S. Tsai, S. T. Wang, L. Y. Yang, I. U. Hsu, H. W. Chang, H. M. Lee, S. H. Kao, C. K. Lee, C. H. Chen, Y. F. Lin, *Clin. Biochem.* **2018**, *56*, 75-82.
- [31] N. Rabbani, P. J. Thornalley, *Kidney Int.* **2018**, *93*, 803-813.
- [32] R. D. Semba, J. C. Fink, K. Sun, S. Bandinelli, J. M. Guralnik, L. Ferrucci, *Eur. J. Nutr.* **2009**, *48*, 38-44.
- [33] R. D. Semba, J. C. Fink, K. Sun, B. G. Windham, L. Ferrucci, *J. Ren. Nutr.* **2010**, *20*, 74-81.
- [34] S. B. Schwedler, T. Metzger, R. Schinzel, C. Wanner, *Kidney Int.* **2002**, *62*, 301-310.
- [35] D. Pearson, T. Carell, *Nucleic Acids Res.* **2011**, *39*, 4818-4826.
- [36] I. Moukadir, S. Prado, J. Piera, A. Velazquez-Campoy, G. R. Bjork, M. E. Armengod, *Nucleic Acids Res.* **2009**, *37*, 7177-7193.
- [37] R. Shi, M. Villarroya, R. Ruiz-Partida, Y. Li, A. Proteau, S. Prado, I. Moukadir, A. Benitez-Paez, R. Lomas, J. Wagner, A. Matte, A. Velazquez-Campoy, M. E. Armengod, M. Cygler, *J. Bacteriol.* **2009**, *191*, 7614-7619.
- [38] J. M. Bujnicki, Y. Oudjama, M. Roovers, S. Owczarek, J. Caillet, L. Droogmans, *RNA* **2004**, *10*, 1236-1242.
- [39] I. Moukadir, M. J. Garzon, G. R. Bjork, M. E. Armengod, *Nucleic Acids Res.* **2014**, *42*, 2602-2623.
- [40] K. Zhi, Z. Yang, J. Sheng, Z. Shu, Y. Shi, *Iran. J. Pharm. Res.* **2016**, *15*, 131-139.
- [41] V. Job, G. L. Marcone, M. S. Pilone, L. Pollegioni, *J. Biol. Chem.* **2002**, *277*, 6985-6993.
- [42] A. D. Borthwick, *Chem. Rev.* **2012**, *112*, 3641-3716.
- [43] A. Kitamura, T. Sengoku, M. Nishimoto, S. Yokoyama, Y. Bessho, *Protein Sci.* **2011**, *20*, 1105-1113.
- [44] J. Kim, S. C. Almo, *BMC Struct. Biol.* **2013**, *13*, 5.
- [45] G. Molla, S. Sacchi, M. Bernasconi, M. S. Pilone, K. Fukui, L. Polegioni, *FEBS Lett.* **2006**, *580*, 2358-2364.
- [46] M. Hellwig, C. Auerbach, N. Müller, P. Samuel, S. Kammann, F. Beer, F. Gunzer, T. Henle, *J. Agric. Food Chem.* **2019**, *67*, 1963-1972.
- [47] C. Mennella, R. C. Borrelli, F. Vinale, M. Ruocco, V. Fogliano, *Ann. N. Y. Acad. Sci.* **2005**, *1043*, 837-844.
- [48] M. S. Pilone, S. Buto, L. Pollegioni, *Biotechnol. Lett.* **1995**, *17*, 199-204.
- [49] Chemical Computing Group ULC. *Molecular Operating Environment (MOE)*, **2019**.
- [50] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605-1612.
- [51] G. A. O'Toole, L. A. Pratt, P. I. Watnick, D. K. Newman, V. B. Weaver, R. Kolter, *Methods Enzymol.* **1999**, *310*, 91-109.

## FULL PAPER

Entry for the Table of Contents (Please choose one layout)

Layout 2:

## FULL PAPER



Nam Y. Kim, Tyler N. Goddard,  
Seungjung Sohn, David A. Spiegel,  
Jason M. Crawford\*

Page No. – Page No.

**Biocatalytic Reversal of Advanced  
Glycation End Product Modification**

Advanced glycation end products (AGEs) are non-enzymatic post-translational modifications of proteins that can accumulate to high concentrations and contribute to the decline of body tissues. A lead biocatalyst has been discovered to reverse the AGE modifications *in vitro* which could be further developed to enable a better understanding of AGEs at the molecular level.

Accepted Manuscript