

Conserved catalytic motifs encode enzyme-like supramolecular peptide assemblies

Ivan Korendovych

`ivan_korendovych@baylor.edu`

Baylor University <https://orcid.org/0000-0001-8144-783X>

Liam Marshall

Baylor University <https://orcid.org/0000-0002-5698-4841>

Sagar Bhattacharya

University of California, San Francisco <https://orcid.org/0000-0002-8020-9255>

Leonardo Serafim

University of Miami

Parth Rathee

University of Miami

Malitha Widanage

Michigan State University

Prerana Dash

Baylor University

Yang Li

UT Southwestern Medical Center <https://orcid.org/0000-0001-9601-4885>

Ho Yee Joyce Fung

UT Southwestern <https://orcid.org/0000-0002-0502-1957>

Souvik Panda

Baylor University

Marianne Jensen

University of Copenhagen

Lars Hemmingsen

Københavns Universitet

Tuo Wang

Michigan State University <https://orcid.org/0000-0002-1801-924X>

Rajeev Prabhakar

University of Miami

Keywords:

Posted Date: August 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10744902/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Conserved catalytic motifs encode enzyme-like supramolecular peptide assemblies

Liam R. Marshall,¹ Sagar Bhattacharya,¹ Leonardo F. Serafim,² Parth Rathee,² Malitha C. Dickwella Widanage,³ Prerana Dash,¹ Yang Li,⁴ Ho Yee Joyce Fung,⁴ Souvik Panda Mahapatra,¹ Marianne L. Jensen,⁵ Lars Hemmingsen,⁵ Tuo Wang,³ Rajeev Prabhakar,² Ivan V. Korendovych^{1*}

¹Department of Chemistry and Biochemistry; Baylor University, Waco, TX, USA

²Department of Chemistry, University of Miami; Coral Gable, FL, USA

³Department of Chemistry, Michigan State University; East Lansing, MI, USA

⁴Department of Biophysics, University of Texas Southwestern Medical Center; Dallas, TX, USA

⁵Department of Chemistry, University of Copenhagen; Copenhagen, Denmark

*Corresponding author. ivan_korendovych@baylor.edu

Abstract A seven-residue fragment derived from the active-site region of carbonic anhydrase spontaneously forms Zn²⁺-binding amyloid fibrils that catalyze carbon dioxide hydration with catalytic efficiencies surpassing all previous carbonic anhydrase mimics and approaching those of natural enzymes. Cryo-electron microscopy at 2.2 Å resolution, supported by perturbed angular correlation spectroscopy, solid-state NMR, molecular dynamics simulations and QM/MM calculations, reveals a supramolecular active site that recapitulates key structural and mechanistic features of the enzyme despite arising from a fundamentally different protein fold. Guided by structure–activity relationships, minimal sequence modifications followed for further optimization to reach carbon dioxide hydration activity of $1.3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$. Extending this strategy to a conserved motif from superoxide dismutase yielded copper-binding catalytic amyloids that promote superoxide dismutation near the diffusion limit. Together, these findings establish conserved catalytic motifs as a rich source of functional catalysts and demonstrate that they encode sufficient information to construct highly active supramolecular active sites independently of the globular protein fold. More broadly, they demonstrate that complex catalytic function can emerge from remarkably simple self-assembling peptide architectures, providing an experimentally tractable framework for investigating the fundamental principles that underlie enzymatic catalysis.

Modern enzymes achieve extraordinary catalysis through precisely organized active sites within complex protein folds, but early functional peptides likely lacked such structural sophistication.^{1,2} This raises the question of how efficient catalysis emerged before folded proteins evolved, and whether self-assembly provided an accessible route to organized, functional structures.³⁻⁵ Short β -sheet peptide assemblies provide an attractive system for addressing this

question because very short peptides spontaneously assemble into diverse, ordered supramolecular architectures, including catalytic amyloids.⁶⁻¹¹ The recent discovery of high catalytic activities in model reactions exerted by amyloid-like assemblies further supports the connection between peptide assemblies and early enzymes.^{8,10,12-14} However, it remains unclear whether catalytic amyloids can recapitulate the structural and mechanistic features that underlie the extraordinary catalytic efficiencies of modern enzymes.

Carbonic anhydrases that facilitate the equilibrium between carbon dioxide and water are absolutely critical for carbon-based life and therefore are present in all known organisms, with a wealth of functional and structural information available for different classes of these enzymes that are presumed to have evolved in a convergent fashion.¹⁵ CA's rely on a concerted action of a single metal ion and an intricate arrangement of amino acid side chains to increase the nucleophilicity of metal bound water and to promote CO₂ hydration.¹⁶ The dominant structural classes vary between the different kingdoms of life, but the metal binding motif of the α -class of CAs, that are found in all organisms, has a consensus sequence QFHFH remarkably conserved across all kingdoms of life (Fig. 1, Fig. S1, Supplementary Information). The consensus sequence is even longer when CA's of different kingdoms are considered separately, e.g. QFHFHWG is found in the most characterized *Animalia* enzymes (Fig. 1). This fragment shows a classic alternating hydrophilic/hydrophobic residue pattern, prompting us to ask whether these conserved catalytic motifs retain sufficient information to direct the assembly of functional supramolecular active sites outside the context of the native protein fold.

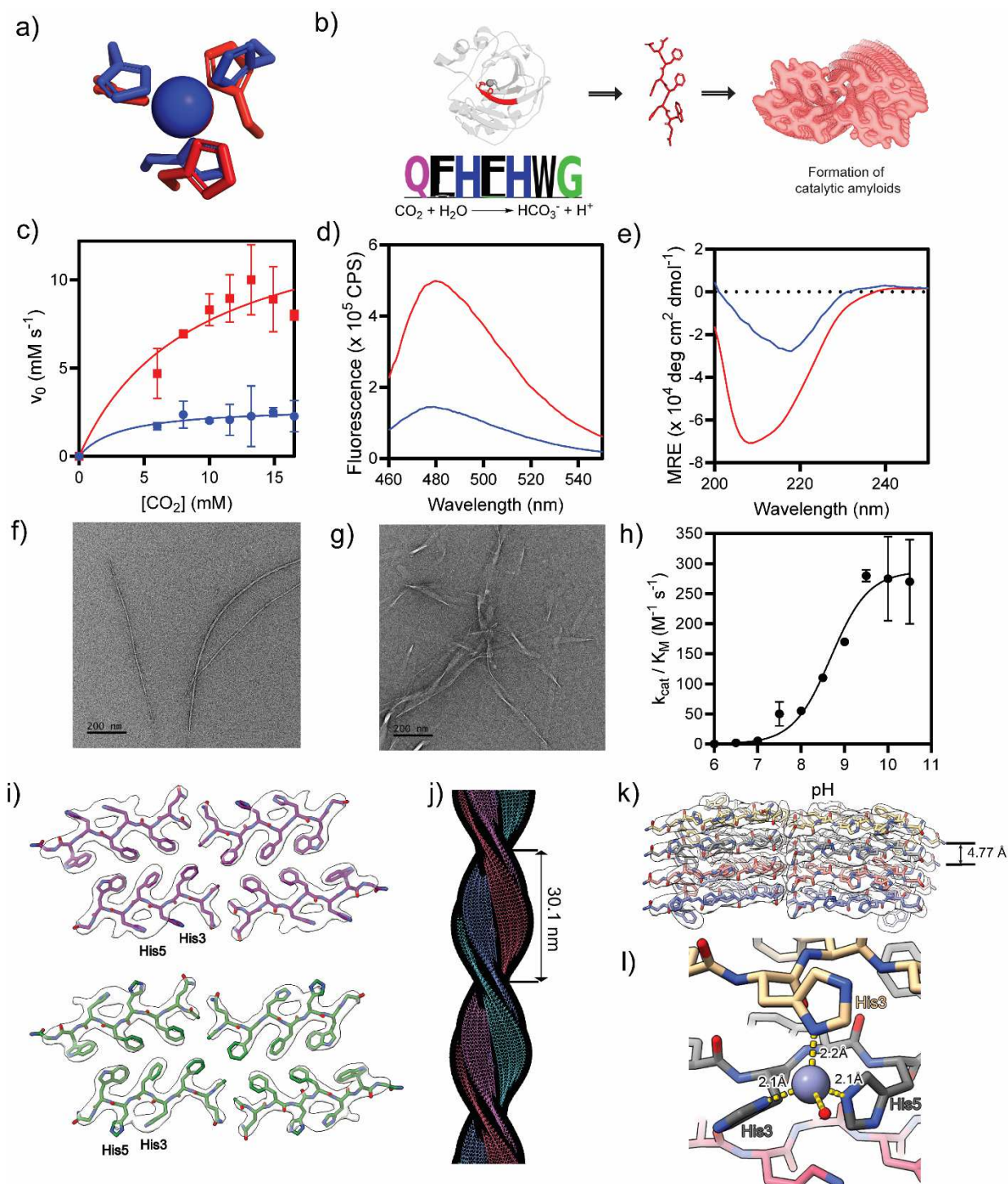


Fig. 1 | Carbonic anhydrase-like activity of amyloid designs a) Overlaid zinc binding site of carbonic anhydrase (PDB 1CA2) and Ac-IHVHLQI-NH₂ (PDB 5UGK). b) An amyloid-like sequence can be identified within the active site of carbonic anhydrase, which self-assembles to form catalytic amyloids. c) Kinetic plots showing CO₂ hydration activity of the Ac-QFHFHWG-NH₂ fragment (blue) and the evolved Ac-FYFHFHFQF-NH₂ (red). 2.5 μM peptide,

2.5 μM Zn^{2+} , 50 mM CHES, pH 9.5 d) Thioflavin T fluorescence assay of the peptides Ac-QFHFHWG-NH₂ (blue) Ac-FYFHFHFQF-NH₂ (red) e) Circular dichroism of the peptides Ac-QFHFHWG-NH₂ (blue) Ac-FYFHFHFQF-NH₂ (red) 100 μM peptide, 5 mM HEPES, pH 8 .100 μM peptide, 25 μM ThT, 25 mM Tris, pH 8. Transmission electron micrographs of f) Ac-QFHFHWG-NH₂ or g) Ac-FYFHFHFQF-NH₂ (100 μM peptide, 25 mM Tris, pH 8). h) pNPA hydrolysis efficiency of Ac-FYFHFHFQF-NH₂ as a function of pH. 25 μM Ac-FYFHFHFQF-NH₂, 25 mM buffer. i) Single layer of the Ac-QFHFHWG-NH₂ fibril model fitted into the cryo-EM density maps of conformation 1 (purple) and conformation 2 (green). His3 and His5, which adopt alternative conformations, are indicated. j) A synthetic four-start fibril reconstructed from the cryo-EM structure model. K) Side view of a four-layer fibril model fitted into the cryo-EM density map. l) Putative Zn^{2+} coordination site located between two layers, adopting a tetrahedral geometry. Zn is shown as a gray sphere, and water as a small red sphere.

Synthetically prepared Ac-QFHFHWG-NH₂ (**1**) self-assembles at near neutral pH to form well defined β -sheet rich fibrils as demonstrated by transmission electron microscopy (TEM), circular dichroism (CD) and thioflavin T (ThT) fluorescence data (Fig. 1, Fig. S2-S6, Supplementary Information). Moreover, when the peptide self-assembles in the presence of Zn^{2+} , the resulting fibrils promote ester hydrolysis (Table 1, Fig. S3a, Supplementary Information). Encouraged by this finding we tested the assemblies formed by **1** in the presence of Zn^{2+} for carbon dioxide hydration. The dependence of the initial rate of reaction on carbon dioxide concentration follows the Michaelis-Menten model (Fig. 1c) with the catalytic efficiency ($k_{\text{cat}}/K_{\text{M}}$) of $3.3 \pm 1.2 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$. This level of activity is quite remarkable – it is to our knowledge the highest reported to date in any CA mimic.^{8,16-18} In fact, considering the low molecular weight of the active unit, **1**- Zn^{2+} has a specific activity that is right in the middle of the range exhibited by naturally occurring enzymes (Table 1). The observed catalytic activity is solely due to the self-assembled species: after centrifugation at 100,000 g the supernatant showed no catalytic activity (Fig. S7, Supplementary Information).

Next, we have undertaken a structure-morphology-activity study to better understand the different factors that influence catalysis. Preserving the His/His/Gln motif while changing the

hydrophobic core to phenylalanine only (peptides **2** and **3**) resulted in increased CO₂ hydration activity, especially when the Gln was positioned on the C-terminal side of the metal binding motif. Previously we and others observed that K_M values for hydrolytic reactions promoted by related self-assembling peptides can be lowered by incorporation of tyrosine residues.^{19,20} Moreover, tyrosine plays a prominent role for proton transfer in native carbonic anhydrases.²¹ Substituting glutamine for tyrosine had a sequence dependent impact on catalytic parameters – in position two (peptide **4**) it lowered K_M, yet in position six (peptide **5**) it led to significant drop in CO₂ hydration activity, albeit with drastically increased ability to promote ester hydrolysis. Next, we explored how these limited sequence activity relationships can be prospectively applied to improve the activity further by extending the sequence of the HFHF zinc binding core by adding FY at the N-terminus and QF at the C-terminus to simultaneously increase the k_{cat} and lower the K_M. The resulting 9-residue peptide **6** (Ac-FYFHFHFQF-NH₂) showed catalytic efficiency in CO₂ hydration of $1.29 \pm 0.39 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$, that is nearly quadruple that of **1**-Zn²⁺, exceeds the reported value of HCAIII and is on par with HCAII in terms of specific activity. Interestingly, the K_M value for CO₂ hydration shown by the most active assemblies is very similar to the one found in natural enzymes. Like all previously reported functional model of carbonic anhydrases, the effective pK_a of the catalytic nucleophile is fairly high- ~ 8.7 as determined from the pH profile of its hydrolytic activity (Fig. 2H).^{17,18} Despite high sequence similarity, the fibrils formed by different peptides have distinct morphologies as shown by TEM and CD. Supramolecular assembly is key to catalysis: introduction of methyl groups into the backbone amides to preclude b-sheet formation produces hydrolytically inactive peptide Ac-FYF_{NMe}HF_{NMe}HFQF-NH₂ (Fig. S11, Supplementary Information). The subtle backbone and side chain modification are clearly important for activity as closely structurally related assemblies formed by Ac-IHIHQI-NH₂ (as well as all others tested)

show no CA activity, despite having a remarkably similar metal coordination sphere to the one seen in carbonic anhydrase that served as an inspiration for its design (Fig 1a).^{11,22,23}

Table 1 | Kinetic parameters for pNPA hydrolysis (pH 8) and CO₂ hydration (pH 9.5). Conditions: 50 mM CHES, pH 9.5, 25 μ M thymol blue.

	pNPA hydrolysis			CO ₂ hydration				Ref.
	$k_{\text{cat}} \times 10^{-2}$ (s ⁻¹)	K_M (mM)	k_{cat}/K_M (M ⁻¹ s ⁻¹)	k_{cat} (s ⁻¹)	K_M (mM)	$k_{\text{cat}}/K_M \times 10^5$ (M ⁻¹ s ⁻¹)	Specific activity (kat g ⁻¹) ^a	
Ac-QFHFHWG-NH ₂ (1)	0.3 $\pm$ 0.025	0.3 $\pm$ 0.07	9.2 $\pm$ 2.6	3,170 $\pm$ 960	9.5 $\pm$ 4.5	3.3 $\pm$ 1.2	1.6	This work
Ac-FQFHFHF-NH ₂ (2)	0.4 $\pm$ 0.03	0.5 $\pm$ 0.07	8.4 $\pm$ 1.3	8,550 $\pm$ 3110	23 $\pm$ 8.9	3.7 $\pm$ 0.7	3.7	This work
Ac-FHFHFQF-NH ₂ (3)	1.1 $\pm$ 0.06	1.1 $\pm$ 0.1	9.5 $\pm$ 1.0	5,230 $\pm$ 1190	8.2 $\pm$ 3.1	6.4 $\pm$ 2.0	3.1	This work
Ac-FYFHFHF-NH ₂ (4)	0.3 $\pm$ 0.03	0.15 $\pm$ 0.06	18.7 $\pm$ 7.8	3190 $\pm$ 1390	7.8 $\pm$ 6	4.1 $\pm$ 2.5	1.7	This work
Ac-FHFHFYF-NH ₂ (5)	0.9 $\pm$ 0.16	0.3 $\pm$ 0.14	29.6 $\pm$ 14.8	1270 $\pm$ 400	7.2 $\pm$ 4	1.8 $\pm$ 0.77	1.8	This work
Ac-FYFHFHFQF-NH ₂ (6)	1.2 $\pm$ 0.12	0.13 $\pm$ 0.06	92 $\pm$ 43	11,760 $\pm$ 2876	9.1 $\pm$ 3.6	12.9 $\pm$ 3.9	4.3	This work
Phe-Zn-Phe ^b	0.2	0.17	77	7.8	8.1	0.01	0.013	⁸
Hg/Zn TriL9CL23H	n.r.	n.r.	1.4	1,800 $\pm$ 430	10.0 $\pm$ 2.4	1.8 $\pm$ 0.3	0.095	¹⁷
Zn ^{II} α_3 DH ₃	n.r.	n.d.	n.r.	134 $\pm$ 8	3.5 $\pm$ 0.6	0.4 $\pm$ 0.05	0.011	¹⁸
HCA I ^c	22 $\pm$ 1	3 $\pm$ 0.01	753 $\pm$ 31	200,000	4.0	500	5.6	^{16,24}
HCA II ^d	81 $\pm$ 6.1	30.5 $\pm$ 2.1	2607 $\pm$ 85	1,400,000	9.3	1500	31	^{16,24}
HCA III	n.r.	n.r.	n.r.	10,000	33.0	3	.11	¹⁶

^a Specific activity values calculated from the kinetic parameters at 16 mM of CO₂. ^b 20 mM Tris, pH 8, 50 μ M phenol red ^c pH 7.5 ^d pH 7.4 n.r. – not reported

Cryo-EM analysis of **1** produced a 2.2 Å resolution map of the supramolecular assemblies formed by **1**-Zn²⁺ (Fig. 1i-l, Table S2, Fig. S8, Supplementary Information). It shows a well-defined double helical structure formed by a tetrameric peptidic unit. The strands are arranged in a very rare parallel/parallel form (Class 3) with the N-terminal glutamine residues tightly packing against each other forming a ternary fold. The electron density is best modelled by averaging two different conformations of metal coordinating histidines (Fig. 2). The metal binding site could not be

unambiguously modelled from cryo-EM data alone but is consistent with a tetrahedral zinc binding site with a nucleophilic water molecule bound to the zinc ion (Fig. 1) supported by ^{111}mCd PAC spectroscopy (Table S3, Fig. S9, Supplementary Information). The PAC signal is described by a single low frequency nuclear quadrupole interaction (NQI) with the parameters $\omega_0 = 0.069$ rad/ns and $\eta = 0.30$, that are in excellent agreement with the predicted values for the model ($\omega_0 = 0.075$ rad/ns and $\eta = 0.47$). The apparent inequivalence of the His side chains in the two types of peptide strands pronounced in the cryo-EM structure is a necessary result of the requirement for the tetrahedral metal ions site that require symmetry breaking when using a two-His monomeric unit. This is commonly observed in homodimeric enzymes in nature, where the two monomers have nearly identical structures, except near the active site, where symmetry is necessarily lifted to create a catalytic site (e.g., in HIV protease).

To gain insight into molecular determinants of the enzymatic level catalysis showed by the self-assembling peptides we performed structural and computational studies on the most catalytically active fibrils formed by **6**. The chemical shifts of the backbone atoms obtained from solid state NMR (ssNMR) experiments show perfect agreement with a typical cross β -sheet structure and provided us with backbone dihedral angles for subsequent structural modeling. We have also observed two different sets of histidine ^{15}N chemical shifts, one indicative of a bridging histidine residue, consistent with the two types of zinc sites observed previously in a ssNMR structure of a related catalytic amyloid (Ac-IHVHLQI-NH₂).²² The ssNMR-derived constraints were then used to build a model of **6**-Zn²⁺ using molecular dynamics (MD) simulations (Fig. 3d). The 12-peptide unit assembly was modeled as a bilayer and equilibrated using 500 ns all-atom MD simulations (Fig. 2b). The resulting structure adopted a twisted (twist angle of $\sim 6.5^\circ$) cross - β fibril supported by hydrogen bonds between the adjacent monomers (Fig. 2b) stabilized by interstrand hydrogen

bonds and a tight hydrophobic core formed by intercalating, zipper-like phenylalanine residues. The Zn^{2+} ions coordinate to three His residues, just like in CA, however unlike the enzyme, the zinc sites are connected by bridging, deprotonated histidine, that lowers the overall charge and contributes to peptide assembly stability (Fig. 2e). The vacant coordination sites are occupied by water molecules. The non-bridging His residue forms a hydrogen bond with the side chain of Gln.

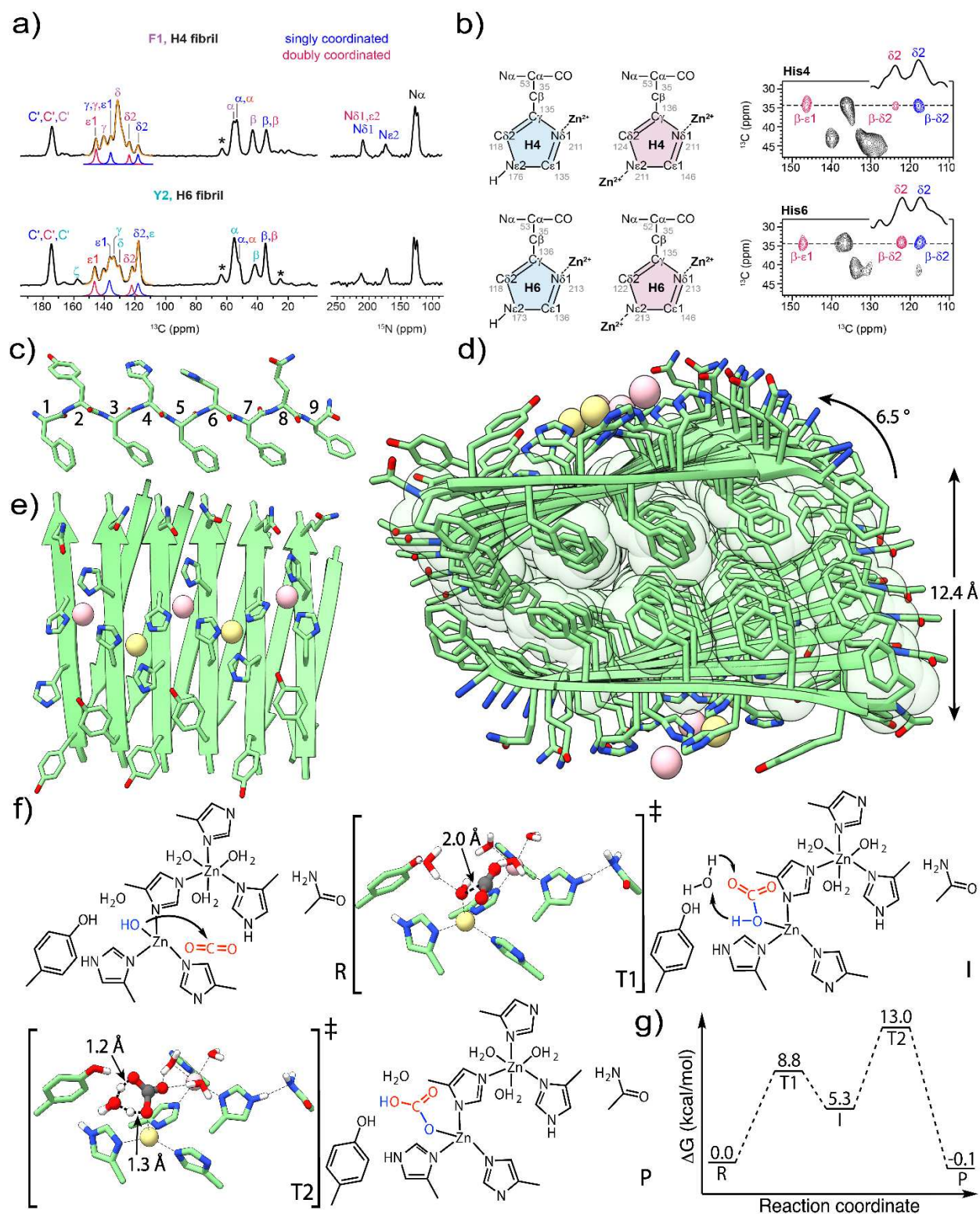


Fig. 2 | Zinc coordination by His4 and His6 residues in Ac-FYFHFHFQF-NH₂ fibrils. a) 1D ¹³C and ¹⁵N CP spectra of two catalytic fibrils at pH8 with different labeling schemes: Phe1 and His4 (top) as well as Tyr2 and His6

(bottom). Spinning sidebands are indicated with asterisks. Spectral deconvolution was conducted for of ^{13}C CP spectra to better resolve the signals of singly (blue) and doubly (magenta) coordinated His residues. The simulated spectra (yellow) closely matched the experimental data (black). Carbon sites from other two residues, Phe1 and Tyr2 are marked in purple and turquoise, respectively. b) Summary of zinc-coordination and $^{13}\text{C}/^{15}\text{N}$ chemical shifts of His4 and His6. c) Selected regions of 2D ^{13}C - ^{13}C correlation spectra highlighting the $\text{C}\beta$ - $\text{C}\delta 2$ (β - $\delta 2$) cross peaks for singly (blue) and doubly (magenta) coordinated His4 (top) and His6 (bottom) residues. Dashed lines show the positions where 1D cross sections were extracted, showing almost equal peak area for the $\text{C}\delta 2$ sites in singly and doubly coordination. d,e) MD equilibrated structure of 6-Zn^{2+} model. f) QM/MM optimized structures of all maxima and minima for the reaction. g) Potential Energy Surface diagram of the reaction including energies in kcal/mol.

While structural studies unequivocally point to two types of zinc ions in the fibrils with 1:1 metal to peptide stoichiometry, the dependence of activity on the zinc concentration is consistent with 1:2 metal to peptide stoichiometry (Fig. S10, Supplementary Information). This discrepancy is consistent with different functional properties of the two metal sites: one site being catalytic (C-site) and the other one playing a purely structural role (S-site). Analysis of the MD model is very much in agreement with this hypothesis. In the C-site, a Tyr residue is ideally positioned to polarize the Zn^{2+} -bound water molecule and generate the hydroxyl nucleophile (**R** state, Fig. 2f). The metal binding and interaction with the Tyr residue substantially lowers the pK_a value of the water and facilitates spontaneous formation of the nucleophile, a general strategy employed by many hydrolytic enzymes.²⁵ To further probe mechanistic aspects of catalysis promoted by the peptide assembly we performed quantum mechanics/molecular mechanics (QM/MM) calculations of CO_2 hydration. In the first step, the hydroxyl nucleophile attacks a CO_2 molecule to generate a Zn^{2+} -bound bicarbonate species containing intermediate (**I**) with a barrier of 8.8 kcal/mol (Fig. 2g). In **I**, the hydroxyl nucleophile doesn't leave the metal ion, but the Zn-O distance is significantly elongated by 0.09 Å. Additionally, it is stabilized by hydrogen bonding with the surrounding water molecules that mimics the role of the second coordination shell Thr199 in carbonic anhydrase. This step is endergonic by 5.3 kcal/mol due to the unfavorable positioning of the hydroxyl proton in **I**. In the next step, this proton can be transferred through either Lipscomb (via bridging water)

or Lindskog (through the rotation of the bicarbonate group) mechanisms.²⁶⁻²⁸ The barrier for the Lipscomb mechanism is computed to be 5.0 kcal/mol lower than the one for the Lindskog mechanism. From **I**, the proton transfer via transition state (**T2**) occurs with a barrier of 7.7 kcal/mol *i.e.* 13.0 kcal/mol from **R**. It is the rate-determining step of the entire mechanism as proposed for the mechanism of CA.²⁹ Additionally, the computed barrier is only slightly higher than the barrier computed for the enzyme.²⁹ The creation of the bicarbonate ion in the product (**P**) is thermoneutral. In **P**, the bicarbonate is coordinated to the Zn²⁺ and stabilized by water molecules from the adjacent non-catalytic sites followed by its replacement by water to initiate the next cycle.

To establish whether highly active catalytic amyloids can be derived from sequences of other natural proteins, we looked at another ancient, essential and highly conserved family of enzymes – superoxide dismutase (Fig. S14, Supplementary Information). These enzymes incorporate a redox active metal ion (usually copper) in their active site and are amongst the fastest enzymes known, with typical catalytic efficiencies above 10⁹ M⁻¹ s⁻¹. Copper ions promote this reaction efficiently but sequestering it within an enzyme lowers cellular toxicity and tunes the properties of the metal site to achieve maximum efficiency. Analysis of the sequences of archaeal superoxide dismutases reveals a highly concerted HGIHIH motif (Fig. 3).

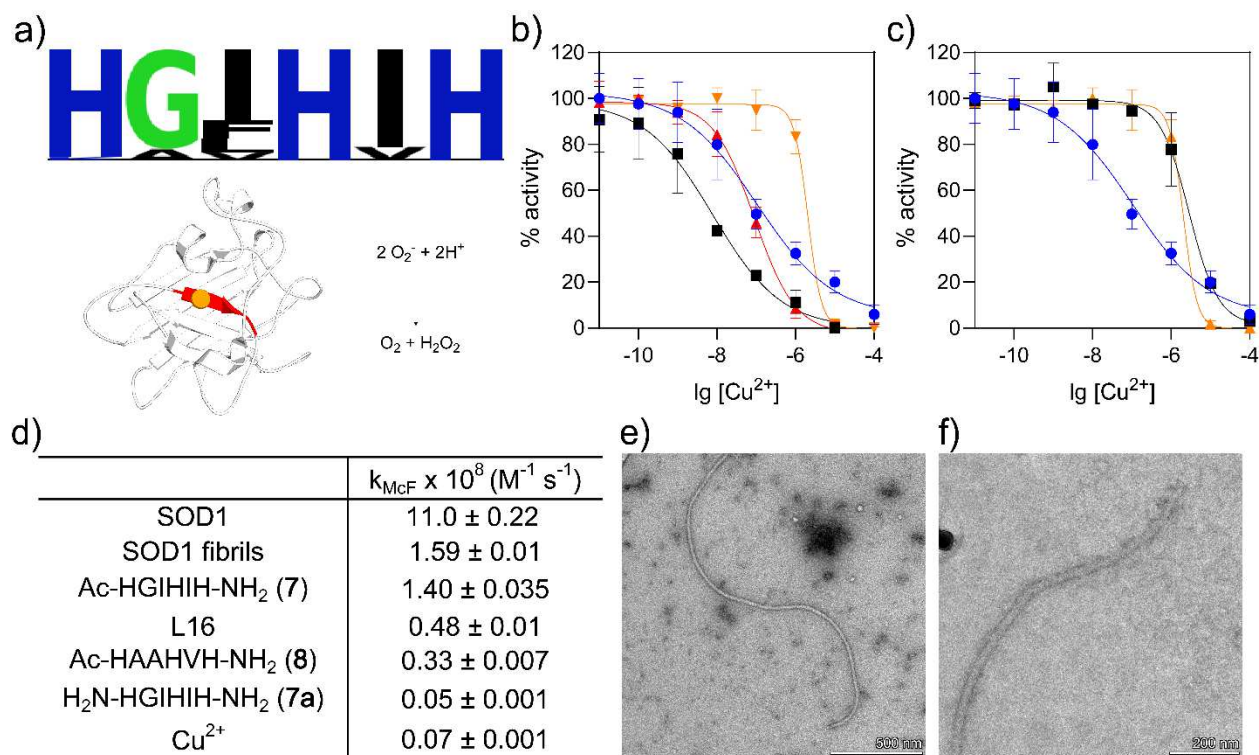


Fig. 3 | Superoxide dismutase activity of amyloid species. A- structure of SOD1 showing a conserved β -sheet at the metal binding site with a corresponding sequence alignment. B- McCord-Fridovich superoxide dismutase assay for bovine SOD1 (black), bovine SOD1 amyloids (red), Cu/Ac-HGIHIH-NH₂ (7) (blue) and free Cu²⁺ (orange). Conditions: 50 mM HEPES, pH 7.8, 25 μ M xanthine, 0.1 μ M xanthine oxidase, 250 μ M NBT. C- McCord-Fridovich superoxide dismutase assay for 7 (blue) and 7a (black), and free copper in buffer (orange). Conditions: 50 mM HEPES, pH 7.8, 25 μ M xanthine, 0.1 μ M xanthine oxidase, 250 μ M NBT. D- Experimentally determined k_{McF} values. TEM micrographs of E- Fibrillated bovine superoxide dismutase (30 μ M protein, 50 mM HEPES, pH 7.8) and F- Ac-HGIHIH-NH₂ (100 μ M fibrils, 100 μ M Cu²⁺, 50 mM HEPES, pH 7.8).

Synthetically prepared peptide Ac-HGIHIH-NH₂ (7) self-assembles in the presence of copper as indicated by CD, ThT and electron microscopy data (Fig 3, Fig S15, Supplementary Information). The assemblies formed by 7 in the presence of copper are highly reactive in promoting superoxide dismutation, approaching the diffusion limit, with the k_{McF} of $1.4 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$. SOD activity in various model systems has been extensively benchmarked and this level of activity is the highest among all SOD peptide mimics under the same conditions (Table S4). The observed superoxide

dismutase activity of **7** is due solely to the self-assembled species, as the supernatant obtained after centrifugation at 100,000 g shows no activity above the blank (Fig. S16, Supplementary Information). Non-self-assembling peptides H₂N-HGIHIIH-NH₂ (**7a**) and Ac-HAAHVH-NH₂ (**8**) with the same pattern of metal coordinating residues show much diminished activity (Figure 3D, Figures S17, S18 Supplementary Information) further supporting a correlation between activity and self-assembly.

The high catalytic activity of amyloids formed by peptides derived from natural proteins begs the question whether a possible evolutionary connection between short amyloids and modern-day proteins could manifest itself in amyloids formed by natural proteins. Amyloids formed by SOD1 have been previously characterized structurally, so we tested them for superoxide dismutase activity. The reported cryoEM structure of full-length SOD1 fibrils shows the conserved active site present in a disordered region of the amyloid fibril, accessible to the substrate.³⁰ Upon fibrillization, fibrils formed by SOD1 promote superoxide dismutation with k_{McF} comparable to that shown by **7** possibly indicating similarity of the structures observed in the fibrils and the peptide assemblies (Fig 3B). While enzymes aggregated through domain swapping and the introduction of tags have been shown to retain their catalytic activity, demonstrating this with a native enzyme in an amyloid state has not been to our knowledge reported before.³¹ This observation contributes to the ongoing discussion of the role of amyloids in the Anfinsen's thermodynamic picture.³²⁻³⁵

Despite the explosive growth of structural and functional knowledge on proteins, mere replication of enzymatic activity in simplified systems remains elusive and the solution to the problem of creating efficient enzymes *a la carte* is still not available.³⁶⁻³⁹ Thus, so far *all* major successes in the enzyme design field have been using unnatural, and in the vast majority of cases

activated, substrates.⁴⁰⁻⁴² Here we demonstrate that conserved catalytic motifs derived from natural enzymes retain sufficient information to reconstruct highly active supramolecular catalysts outside the globular protein fold. Self-assembly is key as mere replication of the sequence is not sufficient to achieving high activity. Amyloid-like peptide assemblies serve as a functional type of protein fold that supports productive arrangements of functional groups, including water molecules, akin to what is found in natural enzymes. Even in such simple systems, minor modifications of peptide sequences allow for accessing various assembly morphologies with various levels of activities. Despite differences in the metal coordination between catalytic amyloids and natural enzymes, the overall mechanism appears to be quite similar as indicated by QM/MM studies, supporting the notion that prebiotic peptide assemblies could effectively promote chemical transformations. These findings establish conserved enzyme motifs as a rich source of functional catalytic assemblies and provide a general strategy for designing efficient supramolecular enzymes. The ease with which we discovered highly efficient supramolecular catalysts of important chemical transformations opens possibilities for rapid discovery of an array of practically applicable peptide catalysts. More broadly, they demonstrate that complex catalytic function can emerge from remarkably simple self-assembling peptide architectures, providing an experimentally tractable framework for investigating the fundamental principles that underlie enzymatic catalysis.

References

- 1 Riddle, D. S. *et al.* Functional rapidly folding proteins from simplified amino acid sequences. *Nat. Struct. Mol. Biol.* **4**, 805-809 (1997).
- 2 Keefe, A. D. & Szostak, J. W. Functional proteins from a random-sequence library. *Nature* **410**, 715-718 (2001).
- 3 Greenwald, J. & Riek, R. On the Possible Amyloid Origin of Protein Folds. *J. Mol. Biol.* **421**, 417-426 (2012).
- 4 Maury, C. P. J. Amyloid and the origin of life: self-replicating catalytic amyloids as prebiotic informational and protometabolic entities. *Cell. Mol. Life Sci.* **75**, 1499-1507 (2018).

- 5 Greenwald, J., Kwiatkowski, W. & Riek, R. Peptide Amyloids in the Origin of Life. *J. Mol. Biol.* **430**, 3735-3750 (2018).
- 6 Rout, S. K., Rhyner, D., Riek, R. & Greenwald, J. Prebiotically Plausible Autocatalytic Peptide Amyloids. *Chem. Eur. J.* **28**, e202103841 (2022).
- 7 Makam, P. & Gazit, E. Minimalistic peptide supramolecular co-assembly: expanding the conformational space for nanotechnology. *Chem. Soc. Rev.* **47**, 3406-3420 (2018).
- 8 Makam, P. *et al.* Non-proteinaceous hydrolase comprised of a phenylalanine metallo-supramolecular amyloid-like structure. *Nat. Catal.* **2**, 977-985 (2019).
- 9 Wang, Y. *et al.* Fmoc-diphenylalanine gelating nanoarchitectonics: A simplistic peptide self-assembly to meet complex applications. *J. Colloid Interface Sci.* **636**, 113-133 (2023).
- 10 Su, J. *et al.* Single-site iron-anchored amyloid hydrogels as catalytic platforms for alcohol detoxification. *Nat. Nanotechnol.* **19**, 1168-1177 (2024).
- 11 Rufo, C. M. *et al.* Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* **6**, 303-309 (2014).
- 12 Makam, P. *et al.* Single amino acid bionanozyme for environmental remediation. *Nat. Commun.* **13**, 1505 (2022).
- 13 Zozulia, O. & Korendovych, I. V. Semi-Rationally Designed Short Peptides Self-Assemble and Bind Hemin to Promote Cyclopropanation. *Angew. Chem. Int. Ed.* **59**, 8108-8112 (2020).
- 14 Zozulia, O., Marshall, L. R., Kim, I., Kohn, E. M. & Korendovych, I. V. Self-Assembling Catalytic Peptide Nanomaterials Capable of Highly Efficient Peroxidase Activity. *Chem. Eur. J.* **27**, 5388-5392 (2021).
- 15 Smith, K. S., Jakubzick, C., Whittam, T. S. & Ferry, J. G. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 15184-15189 (1999).
- 16 Krishnamurthy, V. M. *et al.* Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein-Ligand Binding. *Chem. Rev.* **108**, 946-1051 (2008).
- 17 Zastrow, M. L., Peacock, A. F. A., Stuckey, J. A. & Pecoraro, V. L. Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.* **4**, 118-123 (2012).
- 18 Cangelosi, V. M., Deb, A., Penner-Hahn, J. E. & Pecoraro, V. L. A De Novo Designed Metalloenzyme for the Hydration of CO₂. *Angew. Chem. Int. Ed.* **53**, 7900-7903 (2014).
- 19 Lengyel, Z., Rufo, C. M., Moroz, Y. S., Makhlynets, O. V. & Korendovych, I. V. Copper-Containing Catalytic Amyloids Promote Phosphoester Hydrolysis and Tandem Reactions. *ACS Catal.* **8**, 59-62 (2017).
- 20 Al-Garawi, Z. S. *et al.* The amyloid architecture provides a scaffold for enzyme-like catalysts. *Nanoscale* **9**, 10773-10783 (2017).
- 21 Ferraroni, M., Angeli, A., De Luca, V., Capasso, C. & Supuran, C. T. Kinetic and structural studies of gamma-carbonic anhydrase from the oral pathogen *Porphyromonas gingivalis*. *J. Struct. Biol.* **217**, 108154 (2025).
- 22 Lee, M. *et al.* Zinc-binding structure of a catalytic amyloid from solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6191-6196 (2017).
- 23 Marshall, L. R. *et al.* Synergistic Interactions Are Prevalent in Catalytic Amyloids. *ChemBioChem* **21**, 2611-2614 (2020).
- 24 Innocenti, A. *et al.* Investigations of the esterase, phosphatase, and sulfatase activities of the cytosolic mammalian carbonic anhydrase isoforms I, II, and XIII with 4-nitrophenyl esters as substrates. *Bioorg. Med. Chem. Lett.* **18**, 2267-2271 (2008).

- 25 Vallee, B. L. & Auld, D. S. Zinc: biological functions and coordination motifs. *Acc. Chem. Res.* **26**, 543-551 (2002).
- 26 Lipscomb, W. N. Structure and catalysis of enzymes. *Annu. Rev. Biochem* **52**, 17-34 (1983).
- 27 Lindskog, S. & Coleman, J. E. The catalytic mechanism of carbonic anhydrase. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 2505-2508 (1973).
- 28 Chegwidden, W. R., Carter, N. D. & Edwards, Y. H. *The Carbonic Anhydrases : New Horizons* (Birkhäuser, 2000).
- 29 Bottoni, A., Lanza, C. Z., Miscione, G. P. & Spinelli, D. New Model for a Theoretical Density Functional Theory Investigation of the Mechanism of the Carbonic Anhydrase: How Does the Internal Bicarbonate Rearrangement Occur? *J. Am. Chem. Soc.* **126**, 1542-1550 (2004).
- 30 Wang, L.-Q. *et al.* Cryo-EM structure of an amyloid fibril formed by full-length human SOD1 reveals its conformational conversion. *Nat. Commun.* **13**, 3491 (2022).
- 31 Sambashivan, S., Liu, Y., Sawaya, M. R., Gingery, M. & Eisenberg, D. Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure. *Nature* **437**, 266-269 (2005).
- 32 Jiang, Y. X. & Eisenberg, D. S. To what extent is Anfinsen's Thermodynamic Hypothesis consistent with the formation and polymorphism of amyloid fibrils? *J. Mol. Biol.*, 169364 (2025).
- 33 Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci* **24**, 329-332 (1999).
- 34 Dobson, C. M. & Karplus, M. The fundamentals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.* **9**, 92-101 (1999).
- 35 Ke, P. C. *et al.* Half a century of amyloids: past, present and future. *Chem. Soc. Rev.* **49**, 5473-5509 (2020).
- 36 Marshall, L. R., Zozulia, O., Lengyel-Zhand, Z. & Korendovych, I. V. Minimalist de Novo Design of Protein Catalysts. *ACS Catal.* **9**, 9265-9275 (2019).
- 37 Marshall, L. R. & Korendovych, I. V. Catalytic amyloids: Is misfolding folding? *Curr. Opin. Chem. Biol.* **64**, 145-153 (2021).
- 38 Korendovych, I. V. & DeGrado, W. F. Catalytic efficiency of designed catalytic proteins. *Curr. Opin. Struct. Biol.* **27**, 113-121 (2014).
- 39 Listov, D. & Fleishman, S. J. De novo enzyme design: Controlling structure to design function. *Curr. Opin. Struct. Biol.* **98**, 103252 (2026).
- 40 Bhattacharya, S. *et al.* NMR-guided directed evolution. *Nature* **610**, 389-393 (2022).
- 41 Lauko, A. *et al.* Computational design of serine hydrolases. *Science* **388**, eadu2454 (2025).
- 42 Studer, S. *et al.* Evolution of a highly active and enantiospecific metalloenzyme from short peptides. *Science* **362**, 1285-1288 (2018).

Methods

Chemicals and reagents

All reagents and buffers were purchased from Ambeed, Fisher Scientific, VWR, Sigma Aldrich, Chem-Impex, Biobasic, Inc. and Santa Cruz Biotechnology, Inc. and used without additional purification. Buffers were made using MilliQ water (Millipore Elix 3 instrument). Superoxide dismutase and xanthine oxidase (from bovines) were obtained from Sigma Aldrich. Labeled ^{15}N and ^{13}C amino acids were purchased from CortecNet.

Multiple Sequence Alignment (MSA) analysis

We used carbonic anhydrase from *T. ammonificans* (PDB 4COQ), copper superoxide dismutase (CuSOD) from *C. albicans* (PDB 4N3U) or copper-zinc superoxide dismutase (Cu/Zn SOD) from *H. sapiens* (PDB 4B3E) as starting points for structural searches. Protein homologs were identified using the NCBI BLASTP webserver with standard search parameters, except increasing maximum results to 5000. (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) Results were partitioned according to taxonomic kingdom. Protist sequences were specifically analyzed using the Chinese National Center for Bioinformation P10K BLAST. (<https://ngdc.cncb.ac.cn/blast/p10k/>). We removed duplicate sequences, and aligned the data using MAFFT (v7, <https://mafft.cbrc.jp/alignment/server/large.html>).⁴³ Sequence conservation profiles were rendered using a custom Python visualization tool. Protein sequence alignments generated from BLAST-derived datasets were imported into the software in FASTA format and processed to remove duplicate sequences and non-informative alignment positions. The alignment tool parsed each column of the multiple sequence alignment (MSA) and calculated amino acid frequencies at each position. Sequence logos were generated from the resulting residue frequency matrices using an information-content-based representation, in which the height of each amino acid symbol reflects

its relative conservation within the alignment. For each alignment position, amino acid frequencies were converted to information content values using Shannon entropy calculations:

$$R_i = \log_2(K) - H_i$$

where (R_i) represents the total information content at position i , and (H_i) is the Shannon entropy of the amino acid distribution at that position. Individual residue heights were calculated by multiplying the information content by the fractional abundance of each amino acid at that position. Thus, highly conserved residues produced larger symbols, whereas variable positions produced smaller or absent symbols. The resulting sequence logos were rendered directly within the custom alignment visualization tool using the calculated residue matrices. Logos were formatted with consistent dimensions, residue coloring, and axis parameters to enable comparison between independent sequence datasets. Alignment regions were exported as high-resolution vector graphics for subsequent figure preparation.

Peptide synthesis and characterization.

Peptides were synthesized using fluorenylmethyloxycarbonyl (Fmoc) solid-phase synthesis, as previously described for the synthesis of hydrophobic peptides.⁴⁴ Rink amide resin (0.65 mmol/g) was swollen for 20 min before synthesis. Each synthetic step consisted of: 1) Double Fmoc deprotection using 20 % 4-methyl piperidine in DMF 2) coupling of target amino acid using a mixture of DIPEA and HCTU for 20 min 3) washing twice in DMF 4) N-terminal acetylation for 5 min with a capping solution (DIPEA: acetic anhydride: DMF; 2:1:100) to react with any unreacted N-termini. After synthesis, the resin was washed with dichloromethane and dried under a stream of nitrogen. Resin cleavage and sidechain deprotection was achieved simultaneously by treatment with a mixture of trifluoroacetic acid (TFA), triisopropylsilane (TIPS) and water (95:2.5:2.5 v/v) for 2 h at room temperature. The crude peptides were precipitated and washed

with ice cold diethyl ether and purified using a reverse phase high-performance liquid chromatography (HPLC) system (Varian ProStar 210) with a C4 preparative column (Vydac) using a linear gradient of elution solvent A (0.1 % TFA, 99.9% Millipore H₂O) and solvent B (90% acetonitrile, 9.9% H₂O, 0.1 % TFA). Identities were confirmed using matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry on a Bruker Autoflex III Smartbeam MALDI-TOF spectrometer. Purity was evaluated with an Agilent 1260 instrument with an analytical Zorbax Eclipse XDB-C18 column (2.6 mm x 150 mm).

Preparation of peptide stocks and solutions.

Purified lyophilized peptides were dissolved in 10 mM hydrochloric acid to make a 1 mM stock solution, with concentrations determined using the absorbance at 214 nm (and 280 nm for peptides containing Trp) on an Agilent Cary 60 UV-Vis spectrophotometer. Extinction coefficients for the peptides at these wavelengths were calculated using literature values.⁴⁵ The buffered working stocks of the peptides were prepared by dilution of the stock peptide solution into the appropriate buffer. Working solutions were prepared immediately before use in the case of 0 h experiments, or prepared and then aged on the bench at room temperature.

SOD amyloid formation and purification.

Full-length bovine SOD1 was fibrillated using a method adapted from Wang *et. al.*³⁰(27) In short, full length bovine Cu-Zn SOD (30 μ M) (Sigma Aldrich) was incubated in 20 mM Tris-HCl buffer (pH 7.4) with 5 mM TCEP and shaken at 37 °C for 48 hours. These fibril solutions were then diluted to 10 mL and dialyzed in cassettes (100 kDa MWCO) against 4L of fresh buffer (50 mM HEPES, pH 7.8) for 24 hours (3 x buffer changes) to exclude any remaining folded enzyme. The samples were then concentrated to ~2 mM in a centrifugal filter (VWR). The concentration of the

fibrils was verified using the extinction coefficient at 280 nm ($1490 \text{ M}^{-1} \text{ cm}^{-1}$) using an Agilent Cary 60 UV-Vis spectrophotometer before use in assays.

pNPA hydrolysis.

Kinetic measurements were performed using a BioTek Eon microplate spectrophotometer by monitoring the increase in absorbance from the product (*p*-nitrophenol) at 405 nm at 25 °C in 96-well plates. *p*-nitrophenol acetate was used from a 0.1 M stock in acetonitrile, with final acetonitrile concentration 2% in all reaction mixtures. 50 μL of the working peptide solution was pipetted into wells, with 150 μL of freshly prepared substrate solution (0.195-0.75 mM) added to initiate the reaction. The reported results were obtained by averaging three independent measurements of two peptide stocks. Extinction coefficients of 1,700, 4,300, 9,100, 12,700 and $16,600 \text{ M}^{-1}\text{cm}^{-1}$ were used to determine the product concentration at pH values of 6, 6.5, 7, 7.5 and 8, respectively, at 405 nm.¹¹(II) An extinction coefficient of $18,700 \text{ M}^{-1}\text{cm}^{-1}$ was used at pH 8.5 and above at 405 nm. Kinetic parameters (k_{cat} and K_{M}) were obtained by fitting the data to the Michaelis–Menten equation ($v_0 = k_{\text{cat}}[E]_0[S]_0/(K_{\text{M}} + [S]_0)$). pH profile studies were performed using the following buffers at a concentration of 25 mM: MES (pH 6–6.5), HEPES (pH 7–7.5), Tris (pH 8–8.5), TAPS (pH 9–9.5) and CAPS (pH 10–10.5)) Product formation was monitored for 5 min every 30 s. The pH profile data were fitted the following equation:

$$\frac{k_{\text{cat}}}{K_{\text{M}}} = \frac{(k_{\text{cat}}/K_{\text{M}})_{\text{max}} * 10^{-pK_a}}{10^{-pH} + 10^{-pK_a}}$$

where $k_{\text{cat}}/K_{\text{M}}$ is the observed enzymatic efficiency at a particular pH value, pK_a is the pK_a value for the protonation/deprotonation step and $(k_{\text{cat}}/K_{\text{M}})_{\text{max}}$ is the corresponding maximum activity.

CO₂ hydration kinetics.

CO₂ hydration was performed following the matched indicator procedure first described by Khalifah.^{17,46} Buffer-indicator pairs used were as follows: pH 9.5 CHES/thymol blue (590 nm). Each sample was prepared in 0.1 M buffer, 50 μM indicator, and 5 μM peptide-Zn complex, to yield in cell concentrations of 50 mM buffer, 25 μM indicator, and 2.5 μM Zn-peptide complex. We followed the buffer factor determination described by Pecoraro and co-workers and found our buffer factor in agreement with their findings (0.09).^{17,18} CO₂ solutions were prepared by bubbling through MilliQ water for ten minutes with the use of a fritted glass gas bubbler. For kinetic characterization, a stock of saturated CO₂ (33 mM) was prepared freshly each time and then diluted to the required concentration. Dilutions were carried out by mixing MilliQ water with the saturated CO₂ in gas tight syringes. Experiments were performed by mixing peptide/metal/buffer/indicator or buffer/peptide/indicator (in the case of blanks) solutions in equal proportions with aqueous CO₂ solutions. Upon mixing, 10 seconds of absorbance data were recorded, but only the initial linear portion of the curve was used to calculate initial rates. After calculating the slope of the initial portion, these were multiplied by our calculated buffer factor to yield the initial rate. At least 5 repeats for each concentration were recorded and averaged, and blank shots for each concentration are subtracted from the peptide shots in the calculation of the Michaelis-Menten parameters.

Superoxide dismutase activity

The superoxide dismutase (SOD)-like function of the tested peptides was evaluated by their ability to inhibit the reduction of nitro blue tetrazolium (NBT) by superoxide radicals, which were enzymatically generated in situ using the xanthine/xanthine oxidase system.⁴⁷ In this assay, superoxide radicals produced from the oxidation of xanthine by xanthine oxidase reduce NBT to a blue-colored formazan product. This reaction is monitored by measuring the increase in absorbance at 560 nm (Agilent BioTek Synergy H1 plate-reader). All experiments were performed

in 50 mM HEPES buffer at pH 7.8, containing a final concentration of 250 μ M NBT and 37.5 μ M xanthine. The reaction was initiated by adding 50 μ L of xanthine oxidase (XO) solution—comprised of 4 925 μ L of 50 mM HEPES buffer (pH 8) and 75 μ L of XO—to 50 μ L of Solution 2, which contained 5 mL of 50 mM HEPES (pH 8), 5 mL of 2 mM NBT (in the same buffer), and 75 μ L of xanthine from a 200 μ M stock. Under these control conditions (without enzyme or copper), this mixture consistently produced a reduction in NBT at a rate of 4×10^{-4} OD₅₆₀ per second. To assess superoxide scavenging activity, peptide samples complexed with Cu(II) were tested across a concentration series from 100 μ M down to 0.01 nM. Kinetic measurements of absorbance at 560 nm were recorded continuously over a three-minute interval, and the initial rate of change in OD₅₆₀ was used to quantify the extent to which each peptide–copper complex inhibited NBT reduction relative to the control.

SOD like activity was quantified by determining the IC₅₀ value, which represents the concentration required to achieve 50% suppression of NBT reduction. IC₅₀ values were determined by plotting percentage of inhibition against the concentration of peptide-Cu(II) complex analyzing the resulting curve by dose dependent Hill plot. The data were fit to the following equation using GraphPad:

$$Y = A_{\min} + (A_{\max} - A_{\min}) / (1 + 10^{((\log IC_{50} - X) * \text{Hill Slope})})$$

where, Y: Observed response, X: Logarithm (base 10) of the peptide-Cu(II) complex concentration, A_{max}: Maximum response plateau, A_{min}: Minimum response plateau, LogIC₅₀: Logarithm of the inhibitor concentration that produces a response halfway between Top and Bottom, Hill Slope: Slope factor indicating the steepness of the curve.

Ultracentrifugation.

Samples were prepared as described above. Samples were centrifuged at 100,000 g for 1 hour (Beckman Optima XPN1000). After ultracentrifugation, the supernatant was decanted from the centrifuge tubes and used for subsequent experiments.

Circular Dichroism.

CD spectra were collected on a Jasco J-715 CD spectrometer using continuous scan mode and a quartz cuvette with a 0.1 cm path length. Each spectrum is an average of four scans. Samples contained peptide (25 μ M) and ZnCl₂ (25 μ M) in buffer (5 mM HEPES, pH 8). Care was taken so the sample absorbance never exceeded 1.5 at all wavelengths to produce reliable ellipticity values.

Thioflavin T fluorescence.

Fluorescence spectra were obtained on an Agilent Cary Eclipse fluorescence spectrophotometer operating in the steady-state mode at 25°C with emission and excitation band width set to 5 nm. The measurements were taken using a quartz cuvette with 5 mm excitation and 5 mm emission path lengths. Samples were prepared by mixing peptide (1 mM stock at pH 2, 20 μ L) with buffer (25 mM Tris, pH 8.0, 0.1 mM Zn) to a final concentration of 100 μ M, and ThT (2 mM) was then added to a final concentration of 25 μ M, and the samples were vortexed before measurement.

Transmission Electron Microscopy.

Peptide samples were prepared according to the same procedures described in the kinetic assays. Sample aliquots of 10 μ l were adsorbed onto glow-discharged carbon-coated, 200-mesh copper grids (Ted Pella) for 2-5 minutes. Grids were then stained with two 5 μ l drops of 2% (w/v) uranyl acetate. After drying, samples were viewed with a JEOL-JEM 2100F Field Emission electron microscope at an acceleration voltage of 200 kV. Electron micrographs were recorded on a Gatan OneView 4K CCD camera.

Solid-state NMR

Solid-state NMR experiments were conducted on a Bruker Avance 400 MHz (9.4 T) spectrometer equipped with a 4 mm MAS HCN probe. Spectra were acquired under magic-angle spinning (MAS) conditions at 11.2 kHz for ^{13}C and 10 kHz for ^{15}N at 297 K. ^{13}C chemical shifts were externally referenced to the TMS scale by calibrating the CH_2 signal of adamantane to 38.48 ppm. ^{15}N chemical shifts were referenced to the liquid ammonia scale by calibrating the methionine amide resonance of N-formyl-Met-Leu-Phe-OH to 127.88 ppm. Typical radiofrequency field strengths were 80-100 kHz for ^1H decoupling, 62.5 kHz for ^1H hard pulses, 50-62.5 kHz for ^{13}C , and 41 kHz for ^{15}N . One-dimensional (1D) spectra were recorded using cross-polarization (CP) with contact times of 1 ms for ^{13}C and 1.5 ms for ^{15}N . Four different peptides labeled with U- ^{13}C , ^{15}N amino acids in positions a) 1 and 4 (denoted F1H4); b) 2 and 6, respectively (denoted Y2H6); c) 3 and 8 (denoted F3Q8); and d) 5 (denoted F5) were used for chemical shift assignments. Spectral deconvolution of ^{15}N CP spectra for Phe1/His4 and Tyr2/His6 residues was performed using DMfit software (version dmfit#20191211) to resolve signals from the minor conformers.⁴⁸ Two-dimensional (2D) ^{13}C - ^{13}C PDSD spectra with a 50 ms mixing time were conducted for resonance assignments. Backbone torsion angles (ϕ and ψ) of the fibrils were predicted using TALOS-N.⁴⁹

Computational Procedure

The solid-state NMR structure of the IHVHLQI amyloid (PDB ID: 5UGK) was used as the starting point to build a model for the FYFHFHFQF amyloid.²² The IHVHLQI peptide was first mutated to generate the FYFHFHFQF sequence and then replicated into a bilayer with 6 strands followed by addition of terminal caps. In the first step, the bilayer in the metal-free form was equilibrated for 200 ns through all-atom MD simulations in an aqueous solution treated using the TIP4PEW

water model.⁵⁰ These simulations were performed utilizing both CHARMM36m and AMBER14sbFF force fields using the GROMACS 2022 program.⁵¹⁻⁵³ In the next step, the Zn^{2+} ions were added to the MD equilibrated structure to build assembly of Zn^{2+} -HIS₃ metal centers using the neutral (HID) and charged (HIX) forms of HIS residues on the solvent exposed side of the bilayer (Fig. S12). These metal centers were connected by the charged HIX residues. HIX was parametrized separately for both CHARMM36m and AMBER14sbFF forcefields. The structure of HIX was optimized at the MP2/6-31+G** level of theory in gas-phase using the GAUSSIAN 16 program.^{54,55} The optimized structure was then used to compute Merz-Kollman RESP charges.⁵⁶ For this residue, the CHARMM36 forcefield parameters were computed using the CHARMM General Force Field (CGENFF) algorithm while Antechamber module and Generalized Amber Force Field (GAFF) for AMBER parameters.^{57,58}

In the models of 6- Zn^{2+} amyloid, ψ and ϕ dihedral constraints were taken from ¹³C and ¹⁵N NMR experiments and they were again equilibrated for 500 ns under the aforementioned conditions. In MD simulations, the coulombic and Van Der Waals interactions were calculated using the smooth Particle Mesh Ewald method or SPME with a cutoff of 10.0 Å each.^{59,60} Appropriate isotropic Pressure Coupling and Temperature Coupling re-scalable models (Velocity Rescale thermostat and C-Rescale barostats) were utilized for equilibration.^{61,62} All MD simulations were run with a step size of 2 fs/step in a Periodic Boundary Condition (PBC) corrected cubic box of dimension 9 Å. Structural constraints were calculated with the LINCS algorithm.⁶³ The computed root-mean-square-deviations (RMSD) confirmed the equilibration of trajectories within the simulation time frames. Important parameters like hydrogen bonding patterns, solvent accessible surface area and backbone dihedrals were used to analyze the results. In these structures, the hydrophobic sidechains of top and bottom sheets intercalate into a zipper like conformation to

minimize solvent accessible space. They form on average 95-97 hydrogen bonds to stabilize the overall structure. The Define Secondary Structure of Proteins (DSSP) analysis showed a predominantly β -sheet rich structure ($> 90\%$). The amyloid provides approximately 69.8 nm^2 of solvent exposed surface area with high contributions from TYR, HID, HIX and GLN residues. The most equilibrated structures provided by these simulations were subsequently used for QM/MM calculations using two-layer hybrid ONIOM method.

All hybrid two-layer QM/MM (ONIOM) calculations were performed using the Gaussian 16 software package.^{55,64-66} The QM region includes both catalytic and non-catalytic metal sites formed by two Zn^{2+} ions, five histidines, one tyrosine and one glutamine residues. Based on their locations in MD simulations, five water molecules surrounding the metal centers are also incorporated in the quantum region. The ONIOM optimization procedure uses macro/micro interactions, and the electrostatic interactions between the QM and the MM parts were treated using electronic embedding.⁶⁷ The QM region was optimized without any geometric constraint using the hybrid B3LYP exchange-correlation functional.⁶⁸ The Lanl2Dz basis set that includes the Hay-Wadt effective core potential (ECP) was used for the Zn atoms and the 6-31G basis for all other atoms.^{69,70} Grimme's function with the Becke-Johnson damping effect (GD3BJ) was utilized to incorporate the dispersion effects.⁷¹ The remaining atoms were included in the MM region and modeled using the AMBER forcefield as implemented in Gaussian 16 software. Hessians were calculated at the same level of theory as the optimizations to confirm the nature of the stationary points along the reaction coordinate. The transition states were confirmed to have only one imaginary frequency corresponding to the reaction coordinates. Zero-point vibrational, thermal, and entropy corrections were added to the final energies (at 298.15 K and 1 atm).

Cryo-EM sample preparation

Ac-QFHFHWG-NH₂ (1.2 mM peptide, 1.2 mM Zn, 25 mM Tris, pH 8) was prepared as described prior. The protein sample was then applied to Quantifoil 300 Copper mesh R1.2/1.3 grids (SPI) pre-treated using a PELCO easiGlow instrument (Ted Pella). The grid was flash frozen into liquid ethane using a Leica EM GP2 plunge freezer (Leica), with sensor back blotting for 5 seconds, under an environment of relative humidity of 95% and 6 °C.

Cryo-EM data acquisition.

Grid screening was performed on a Talos Glacios microscope (Thermo Fisher Scientific). The best grid was used for a 24-hour data collection on a Titan Krios microscope (Thermo Fisher Scientific) with the post-column BioContinuum energy filter (Gatan) and a K3 direct electron detector (Gatan). A total of 6,817 movies were collected using SerialEM at a nominal magnification of 105,000x and pixel size 0.827 Å, with a 10 eV imaging filter slit.⁷² The accumulated total dose was ~50 e⁻/Å² for each movie stack and fractionated into 50 frames. The defocus range of the images was set to be -0.8 to -2.2 μm.

Cryo-EM data processing.

Cryo-EM data were processed in Relion.⁷³⁻⁷⁵ Raw movies were motion-corrected and dose-weighted, then averaged into single micrographs. The Contrast Transfer Function (CTF) parameters were calculated using CTFFIND4.⁷⁶ 3,348,297 particles were auto-picked by Topaz, using a model trained with the particles manually picked from 20 micrographs.⁷⁷ After two rounds of 2D classification, 990,497 particles remained. The initial 3D model was generated using selected 2D classes, assuming it is a left-handed helix with a crossover of 330 nm. After multiple rounds of 3D classification, particle polishing, and CTF refinement, 3D auto-refine with a stack of 37,444 particles and post-processing generated a reconstruction with an overall resolution of 2.3 Å, according to the gold-standard Fourier shell correlation (FSC) using the 0.143 criterion.⁷⁸

The fibril structure adopts a C2 symmetry, with a refined helical twist of -2.97 degrees and a helical rise of 4.76 Å. The final map used for model building and refinement was sharpened with DeepEMhancer.⁷⁹ More details are listed in the cryo-EM Table S2.

Preparation of sample for ^{111m}Cd PAC spectroscopy

^{111m}Cd was produced by irradiation of ¹⁰⁸Pd with α -particles at the University Hospital cyclotron in Copenhagen and separated from the target as described previously.⁸⁰ This led to a solution of ca. 100 MBq ^{111m}Cd in ca. 50 μ L MilliQ water at the start of the chemical sample preparation.

The sample for PAC spectroscopy was prepared by mixing 20 μ L of the aqueous solution of ^{111m}Cd with 5 μ L 5 mM CdCl₂ and 20 μ L 1 M HEPES, pH 8. Next, MilliQ water was added to get a sample volume of 100 μ L, and 100 μ L 4 mM Ac-QFHFHWG-NH₂ fibrils in suspension in 10 mM HCl solution, to achieve final concentrations of 0.1 mM CdCl₂, 2 mM QFHFHWG (fibrils), 100 mM Hepes in a total volume of 200 μ L. pH was measured after the PAC experiment to be 7.7. The sample was left to incubate for 15 min before the PAC measurement, and the sample was at room temperature during the measurement.

PAC instruments and data analysis

The general theory of PAC spectroscopy may be found in the literature.^{81,82} An analogue PAC setup with six BaF₂ detectors was used.⁸³ The time resolution was 0.9 ns and the time-per-channel was 0.562 ns. The distance from the sample center was 26.9 mm to the detector heads. Data analysis was carried out with the Winfit and Prelude programs (provided by Prof. T. Butz) using 550 data points (excluding the first 5 points due to systematic errors in these). A Lorentzian line shape accounted for line broadening, given by the parameter δ , due to a static distribution of EFGs. Fourier transformation of the data and fits were carried out using 550 points and a Keiser-Bessel

function with the window parameter equal to 4. The angular overlap model adapted to analysis of PAC data was applied to predict the PAC parameters given the experimentally determined structure of the Ac-QFHFHWG-NH₂ fibrils.⁸⁴

Data availability

The cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession code xxx (submission in progress). The coordinates generated in this study are deposited in the Protein Data Bank (PDB) under accession code xxxx (submission in progress). Biological materials are available on request.

All data supporting the findings of this study are available from the Zenodo repository at 10.5281/zenodo.21986196. The repository contains the datasets underlying the figures and conclusions of this study.

Code availability

Code used for sequence analysis and data processing is available at GitHub <https://github.com/liamrmarshall7-pixel/SequenceConservation.git>.

Methods references

- 43 Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160-1166 (2017).
- 44 Korendovych, I. V. *et al.* De Novo Design and Molecular Assembly of a Transmembrane Diporphyrin-Binding Protein Complex. *J. Am. Chem. Soc.* **132**, 15516-15518 (2010).
- 45 Kuipers, B. J. & Gruppen, H. Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *J. Agric. Food Chem.* **55**, 5445-5451 (2007).
- 46 Khalifah, R. G. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *J. Biol. Chem.* **246**, 2561-2573 (1971).
- 47 Mathieu, E. *et al.* Rational De Novo Design of a Cu Metalloenzyme for Superoxide Dismutation. *Chem. –Eur. J.* **26**, 249-258 (2020).
- 48 Massiot, D. *et al.* Modelling one and two-dimensional solid-state NMR spectra. *Magn. Reson. Chem.* **40**, 70-76 (2002).
- 49 Shen, Y. & Bax, A. Protein structural information derived from NMR chemical shift with the neural network program TALOS-N *Methods Mol. Biol.* **1260**, 17-32 (2015).
- 50 Horn, H. W. *et al.* Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **120**, 9665-9678 (2004).
- 51 Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71-73 (2017).
- 52 Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696-3713 (2015).
- 53 Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25 (2015).
- 54 Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **124** (2006).
- 55 Gaussian 16 Rev. C.01 (Wallingford, CT, 2016).
- 56 Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269-10280 (2002).
- 57 Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671-690 (2010).
- 58 Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **3**, 198-210 (2013).
- 59 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092 (1993).
- 60 Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593 (1995).

- 61 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126** (2007).
- 62 Bussi, G., Zykova-Timan, T. & Parrinello, M. Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *J. Chem. Phys.* **130** (2009).
- 63 Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472 (1997).
- 64 Maseras, F. & Morokuma, K. IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.* **16**, 1170-1179 (1995).
- 65 Svensson, M. *et al.* ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and Pt(P(t-Bu)₃)₂ + H₂ Oxidative Addition. *J. Phys. Chem.* **100**, 19357-19363 (1996).
- 66 Chung, L. W. *et al.* The ONIOM Method and Its Applications. *Chem. Rev.* **115**, 5678-5796 (2015).
- 67 Mayhall, N. J., Raghavachari, K. & Hratchian, H. P. ONIOM-based QM:QM electronic embedding method using Löwdin atomic charges: Energies and analytic gradients. *J. Chem. Phys.* **132** (2010).
- 68 Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648-5652 (1993).
- 69 Francel, M. M. *et al.* Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **77**, 3654-3665 (1982).
- 70 Hay, P. J. & Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* **82**, 299-310 (1985).
- 71 Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
- 72 Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36-51 (2005).
- 73 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519-530 (2012).
- 74 Fernandez-Leiro, R. & Scheres, S. H. W. A pipeline approach to single-particle processing in RELION. *Acta Crystallogr. D Struct. Biol.* **73**, 496-502 (2017).
- 75 Scheres, S. H. W. Amyloid structure determination in RELION-3.1. *Acta Crystallogr. D Struct. Biol.* **76**, 94-101 (2020).
- 76 Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216-221 (2015).
- 77 Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153-1160 (2019).
- 78 Henderson, R. *et al.* Outcome of the first electron microscopy validation task force meeting. *Structure* **20**, 205-214 (2012).
- 79 Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).

- 80 Hemmingsen, L. *et al.* Cd-substituted horse liver alcohol dehydrogenase: catalytic site metal coordination geometry and protein conformation. *Biochemistry* **34**, 7145-7153 (2002).
- 81 H. Fauenfelder, R. M. S. *Alpha-, Beta and Gamma-Ray Spectroscopy*. 997 (1965).
- 82 Hemmingsen, L., Sas, K. N. & Danielsen, E. Biological Applications of Perturbed Angular Correlations of γ -Ray Spectroscopy. *Chem. Rev.* **104**, 4027-4062 (2004).
- 83 Butz, T., Saibene, S., Fraenzke, T. & Weber, M. A “TDPAC-camera”. *Nucl. Instrum. Methods Phys. Res. A* **284**, 417-421 (1989).
- 84 Bauer, R., Jensen, S. J. & Schmidt-Nielsen, B. The angular overlap model applied to the calculation of nuclear quadrupole interactions. *Hyperfine Interact.* **39**, 203-234 (1988).

Acknowledgments

This work was supported by the National Institutes of Health (GM119634 (I.V.K.), R01AI173270 (T.W.)) and the Welch Foundation (AA-2198-20240404 (IVK)). The authors would like to thank the Center for Microscopy and Imaging at Baylor University for technical support during microscopy and image analysis. We thank the Structural Biology Laboratory and the Cryo-Electron Microscopy Facility at UT Southwestern Medical Center, which are partially supported by grant RP220582 from the Cancer Prevention & Research Institute of Texas (CPRIT) for cryo-EM studies. A portion of this research was supported by NIH grant R24GM154185 performed at the Pacific Northwest Center for Cryo-EM (PNCC) by Omar Davulcu.

Author contributions L.R.M., T.W., R.P., L.H. and I.V.K. designed the experiments. L.R.M. performed multiple sequence analysis and wrote software to visualize the data. L.R.M., S.B., P.D. and S.P.M. performed peptide synthesis and characterization as well as kinetic studies. L.F.S., P.R. and R.P. designed and performed computational simulations and analyzed the results. M.C.D.W. and T.W. performed MAS NMR experiments and NMR data analysis. Y.L. and H.Y.J.F. conducted cryoEM experiments, as well as data analysis and model creation. M.L.J. and L.H. performed ^{111}mCd PAC spectroscopy and analyzed the data. L.R.M. and I.V.K. wrote the manuscript with input of all the authors.

Competing interests The authors declare that they have no competing interests.

Supplementary information The online version contains supplementary material available at <https://doi.org/>

Correspondence and requests for materials should be addressed to Ivan V. Korendovych

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ConservedcatalyticmotifsencodeenzymescombinedSlv3.pdf](#)