1 **Distal mutations enhance catalysis in designed enzymes by facilitating**

2 **substrate binding and product release**

3

4 **Niayesh Zarifi,[1,2] Pooja Asthana,[3] Hiva Doustmohammadi,[4,5] Cindy Klaus,[1,2] Janet**

5 **Sanchez,[4,5] Serena E. Hunt,[1,2] Rojo V. Rakotoharisoa,[1,2] Sílvia Osuna,[4,5] James S. Fraser,[3]**

6 **Roberto A. Chica[1,2,*]**

7

8 [1] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Ontario,

9 Canada, K1N 6N5

10 [2] Center for Catalysis Research and Innovation, University of Ottawa, Ottawa, Ontario, Canada,

11 K1N 6N5

12 [3] Department of Bioengineering and Therapeutic Science, University of California, San

13 Francisco, San Francisco, California 94158, United States

14 [4] CompBioLab Group, Institut de Química Computacional i Catàlisi and Departament de

15 Química, Universitat de Girona, Girona, Spain

16 [5] ICREA, Catalan Institution for Research and Advanced Studies, Barcelona, Spain

17

18 [*] To whom correspondence should be addressed: Roberto A. Chica, Email: rchica@uottawa.ca

19

20 **Keywords**

21 Enzyme catalysis; enzyme design; enzyme engineering; X-ray crystallography; molecular

22 dynamics; enzyme mechanisms

23 **Abstract**

24 The role of amino-acid residues distant from an enzyme's active site in facilitating the complete

25 catalytic cycle—including substrate binding, chemical transformation, and product release—

26 remains poorly understood. Here, we investigate how distal mutations promote the catalytic cycle

27 by engineering mutants of three de novo Kemp eliminases containing either active-site or distal

28 mutations identified through directed evolution. Kinetic analyses, X-ray crystallography, and

29 molecular dynamics simulations reveal that while active-site mutations create preorganized

30 catalytic sites for efficient chemical transformation, distal mutations enhance catalysis by

31 facilitating substrate binding and product release through tuning structural dynamics to widen the

32 active-site entrance and reorganize surface loops. These distinct contributions work synergistically

33 to improve overall activity, demonstrating that a well-organized active site, though necessary, is

34 not sufficient for optimal catalysis. Our findings reveal critical roles that distal residues play in

35 shaping the catalytic cycle to enhance efficiency, yielding valuable insights for enzyme design.

36

Introduction

Enzymes are essential for life, catalyzing nearly all biochemical reactions with remarkable precision and efficiency. Understanding the molecular principles underlying enzyme catalysis has been a cornerstone of biochemistry, with decades of research providing a detailed understanding of how active-site residues orchestrate catalytic mechanisms[1, 2]. Structural, mutational, and computational studies have also demonstrated that residues located far from the active site contribute to catalysis by modulating conformational dynamics to enrich productive active-site configurations[3-12]. Despite these insights, the specific contributions of distal residues in facilitating other aspects of the catalytic cycle, such as substrate binding and product release, remain poorly understood. Elucidating the role of distal residues in these processes in natural enzymes is challenging due to the intricate allosteric networks in their structures[12] and the presence of epistatic interactions[13] shaped by millions of years of evolution. As a result, reliably predicting the functional impact of distal mutations remains a significant challenge[14], hindering our ability to fully understand and exploit enzyme function.

Over the past two decades, de novo enzyme design has emerged as an approach for engineering enzymes tailored to specific reactions[15-18]. These methods focus on constructing active sites by incorporating catalytic residues and ligand-binding pockets into protein scaffolds[19, 20], neglecting the role of distal residues. While this approach has successfully produced several de novo enzymes[15-18, 21, 22], these biocatalysts have displayed low activity, necessitating subsequent directed evolution to enhance catalytic efficiency[3, 23-27]. Throughout these evolutionary processes, numerous mutations that enhance activity were discovered, both within and far from the active site. Several studies have shed light on how evolution enhances catalysis in artificial enzymes[4-6, 26, 27]. However, they have assessed the functional impacts of distal mutations alongside active-site

60   mutations, obscuring the direct contribution to catalysis of residues far from the active site.

61   Moreover, these studies have focused on the role of mutations found by directed evolution in

62   accelerating the chemical transformation[3-6, 23, 24, 27-29], limiting insights into their role in substrate

63   binding and product release —key catalytic cycle steps that can impact efficiency. Given the well-

64   documented evolutionary paths of these de novo enzymes and the fact that mutations were

65   specifically selected by directed evolution for their ability to boost catalytic efficiency, these

66   artificial enzymes offer a unique opportunity to investigate the effects of distal mutations on the

67   catalytic cycle without complications arising from the evolutionary history of the wild-type protein

68   scaffold from which they are derived.

69   Here, we investigate the role of distal mutations in facilitating the enzyme catalytic cycle

70   by studying three de novo Kemp eliminases that have undergone directed evolution. Using mutants

71   containing either active-site or distal mutations, referred to as Core and Shell variants, respectively,

72   we analyze their effects on enzyme activity, structure, and dynamics. Our results reveal that while

73   active-site mutations establish highly organized active sites for efficient chemical transformation,

74   distal mutations enhance catalysis primarily by modulating energy barriers of substrate binding

75   and product release to facilitate the overall catalytic cycle. This is achieved through altered

76   structural dynamics that widen the active-site entrance and reorganize surface loops. Together,

77   these mutation sets work synergistically to optimize enzyme activity. Our findings suggest that

78   designing more efficient artificial enzymes will require novel strategies to balance the structural

79   rigidity essential for precise active-site alignment with the flexibility needed for efficient

80   progression through the catalytic cycle.

81

82   **Results**

4

83   *Core and Shell variants of de novo Kemp eliminases*

84   To investigate how distal mutations contribute to the catalytic cycle, we generated variants

85   of three computationally designed enzymes: HG3[18], 1A53-2[18], and KE70[15]. These enzymes,

86   generated from three distinct TIM barrel scaffolds with no sequence homology, were designed to

87   catalyze the Kemp elimination (Figure 1a), a model organic transformation frequently used to

88   benchmark enzyme design protocols. This reaction is ideal for assessing the influence of distal

89   mutations, as it occurs in a single step, simplifying mechanistic evaluations of the chemical

90   transformation. The original de novo Kemp eliminases display modest catalytic efficiencies

91   ($k_{cat}/K_M \leq 10^2\,\mathrm{M}^{-1}\,\mathrm{s}^{-1}$), requiring directed evolution to enhance activity[3, 24, 26, 27]. Directed evolution

92   improved catalytic efficiency by several orders of magnitude, mostly due to increases in k_{cat}[3, 26, 27],

93   resulting in Evolved variants containing mutations distributed throughout each enzyme's structure.

94   The original de novo enzymes, referred to as Designed variants here, lack the mutations

95   introduced by directed evolution except for two cases: HG3-Designed, which is a single point

96   mutant of HG3 (K50Q) that contains a catalytic hydrogen-bond donor not present in the original

97   de novo enzyme but that appeared during its evolution[26], and KE70-Designed, which includes two

98   insertions in active-site loops (S20a and A240a) that were manually introduced during KE70's

99   evolution[3]. We employ these Designed variants instead of the original de novo enzymes for

100   consistent comparisons within each enzyme family, ensuring identical catalytic residues and

101   sequence lengths across all related variants (Supplementary Table 1).

102   In addition to the Designed and Evolved enzymes, we created Core and Shell variants by

103   introducing non-overlapping sets of mutations derived from their respective Evolved forms (Figure

104   1b, Supplementary Table 2). Core variants contain mutations that occurred within the active site,

105   defined as residues directly interacting with the transition-state analogue in their crystal structure

106   (first shell), as well as those in direct contact with ligand-binding residues (second shell). These

107   positions are those that are typically targeted during de novo enzyme design[15, 18, 30]. By contrast,

108   Shell variants contain mutations that occurred outside the active site. This process yielded a total

109   of 12 enzymes containing two to 16 mutations across three distinct lineages (Table 1).

110

111   *Functional effects of active-site and distal mutations*

112   All enzymes could be expressed and purified in good yields (Supplementary Table 3)

113   except for 1A53-Shell, which afforded substantially lower yields and tended to precipitate when

114   stored overnight at 4 °C or when concentrated. Due to this issue, we conducted enzyme kinetics

115   immediately after purification. Core variants were 90 to 1500-fold more catalytically efficient than

116   their corresponding Designed enzymes, and only slightly less active (1.2–2-fold) than Evolved

117   variants (Figure 1c, Table 1). By contrast, Shell variants did not exhibit significant improvements

118   in k_{cat}/K_M over their Designed counterparts except for HG3-Shell, which is 4-fold more

119   catalytically efficient than HG3-Designed. These results demonstrate that active-site mutations are

120   the primary drivers of enhanced activity, while distal mutations work synergistically with active-

121   site mutations to further increase catalytic efficiency.

122   Additionally, the incorporation of active-site or distal mutations into Designed enzymes

123   did not produce consistent effects on stability. Core and Shell variants displayed either increased

124   or decreased stability relative to the Designed enzymes, with no clear trend (Table 1,

125   Supplementary Figures 1–2). Similarly, combining both sets of mutations resulted in variable

126   effects on stability, ranging from stabilization (HG3-Evolved) to destabilization (1A53-Evolved),

127   or no substantial change (KE70-Evolved) relative to the Designed enzymes. These results,

128   combined with the observation that Evolved variants are more active than Core variants, suggest

6

129   that distal mutations are primarily selected during directed evolution to enhance catalytic

130   efficiency rather than to mitigate trade-offs between activity and stability introduced by active-site

131   mutations. This conclusion challenges prior hypotheses that distal mutations are primarily

132   compensatory[31, 32], instead highlighting their critical functional contributions to enzyme catalysis.

133

134   *Structural effects*

135   To investigate how active-site and distal mutations cause the observed effects on activity,

136   we solved crystal structures of several Core and Shell variants from each lineage and compared

137   them with available structures of their Designed[5, 18], Core[4] and Evolved[3, 4, 27] counterparts. Crystals

138   of 1A53-Core, KE70-Core, and HG3-Shell, both in the presence and absence of transition-state

139   analogue 6-nitrobenzotriazole (6NBT, Figure 1d), were obtained under different conditions

140   (Supplementary Table 4). The crystals diffracted at resolutions ranging from 1.44 to 2.36 Å

141   (Supplementary Table 5). Unit cells of 1A53-Core (P31 2 1) and HG3-Shell (P21 21 21) contained

142   a single protein chain and had similar dimensions regardless of the presence of 6NBT. By contrast,

143   crystals of KE70-Core displayed variations in space group and unit cell dimensions with and

144   without 6NBT (P1 21 1 and P21 21 21, respectively), accommodating six or two protein molecules

145   in the asymmetric unit, respectively. All protein structures crystallized with 6NBT displayed clear

146   density for this ligand within the active-site pocket (Supplementary Figure 3). In the case of 1A53-

147   Core, the structure obtained in the absence of 6NBT shows clear density for a 2-(*N*-

148   morpholino)ethanesulfonic acid (MES) molecule from the crystallization buffer bound in the

149   active-site pocket, overlapping with the binding site of the transition-state analogue. In all cases,

150   there were no substantial changes to the backbone conformation caused by the introduction of

151   active-site or distal mutations (Supplementary Figure 4).

152   Comparison of bound and unbound structures reveals that the active-site configurations of

153   Core and Shell variants, except for 1A53-Core, remain mostly unchanged upon 6NBT binding,

154   with catalytic residues adopting nearly identical side-chain conformations (Figure 2a). These

155   results imply that these active sites are preorganized for catalysis in the absence of the transition-

156   state analogue. In 1A53-Core however, W110 adopts a non-productive conformation that would

157   prevent binding of 6NBT due to an approximately 120-degree rotation around χ_2 compared to its

158   conformation in the bound structure. We hypothesized that this conformational change results from

159   the presence of MES in the active site. To test our hypothesis, we performed molecular dynamics

160   simulations of this enzyme structure in the absence of the buffer molecule, which showed that

161   W110 readily adopts a conformation suitable for binding 6NBT via pi-stacking throughout the

162   simulation (Supplementary Figure 5). These results suggest that the 1A53-Core active site is also

163   preorganized for catalysis in the absence of 6NBT.

164   When comparing Core variants to their Evolved counterparts, they exhibit nearly identical

165   active sites in the presence of the transition-state analogue (Figure 2b). These findings imply that

166   active-site mutations alone are sufficient to form a binding pocket optimized for efficient chemical

167   transformation in Core variants. On the other hand, HG3-Shell's active site closely resembles the

168   suboptimal active site of HG3-Designed when 6NBT is present. These results demonstrate that

169   introducing distal mutations in either Designed or Core variants does not substantially alter the

170   active-site configuration. Yet, distal mutations enhance catalysis more effectively when introduced

171   in Core variants than in Designed variants, suggesting that they do so through subtle effects that

172   cannot be explained by analyzing average crystal structures alone.

173   To explore this possibility, we assessed the impact of distal mutations on the enzyme

174   conformational ensemble, as they have been proposed to enhance activity by enriching

8

175   catalytically productive substrates[14]. To do so, we performed ensemble refinement of the 6NBT-

176   bound structures (Supplementary Table 6) and examined the percentage of ensemble members

177   capable of forming catalytic contacts (e.g., hydrogen bonds and pi-stacking interactions) with the

178   transition-state analogue. We found that in all cases, the incorporation of active-site mutations into

179   Designed enzymes to yield Core variants greatly increases the proportion of substrates within the

180   conformational ensemble capable of establishing hydrogen bonds or pi-stacking interactions

181   between catalytic residues and 6NBT, while simultaneously making catalytic residues more

182   ordered (Figure 3, Supplementary Figure 6). Introduction of distal mutations into Designed or Core

183   variants to yield Shell or Evolved variants, respectively, also enhances the proportion of productive

184   substates and rigidifies catalytic residues but to a much lesser extent. Overall, crystallography

185   revealed that active-site mutations enhance catalysis by structuring the active site for efficient

186   chemical transformation, while distal mutations complement this effect by modestly increasing the

187   proportion of productive substates in the conformational ensemble. The minimal impact of distal

188   mutations on catalytic efficiency in Shell variants, compared to their greater effect in Evolved

189   variants, suggests that they may play a more significant role in accelerating other catalytic cycle

190   steps, such as substrate binding and product release, which can be efficiency bottlenecks.

191

192   *Mechanistic effects*

193   Next, we examined the influence of distal mutations on substrate binding and product

194   release (Figure 4a). To do so, we measured kinetic solvent viscosity effects (KSVE) on k_{cat}/K_M and

195   k_{cat}, which probe whether productive substrate binding is limited by the rate of substrate diffusion

196   and whether product release is the rate-limiting step in the catalytic cycle[33], respectively. In these

197   experiments, the rates of substrate binding, product release, and enzyme conformational changes

9

198   are expected to be sensitive to solvent viscosity. By contrast, the chemical transformation itself is

199   assumed to be independent of solvent viscosity, as it occurs within the active site, which is shielded

200   from bulk solvent effects. For these analyses, we focused on Core and Evolved variants, as the

201   activity of Designed and Shell variants was too low to yield reliable data when viscosity was

202   increased.

203   KSVE on k_{cat} could not be measured for most enzymes due to lack of saturation within the

204   substrate's solubility limit in the presence of viscogen (Supplementary Figure 7). However, KSVE

205   on k_{cat}/K_M were measurable and revealed a value of approximately 1.0 for HG3-Evolved,

206   indicating that productive substrate binding is much slower than the chemical transformation

207   (Figure 4b). By contrast, HG3-Core exhibited a KSVE of 0.59, implying that substrate diffusion

208   to form productive enzyme-substrate complexes and the chemical transformation occur at

209   comparable rates, potentially due to faster diffusion than in HG3-Evolved or slower chemical

210   transformation, or both. To explore these possibilities, we performed stopped-flow experiments to

211   monitor the binding kinetics of the transition-state analogue 6NBT and reaction product 2-

212   hydroxy-5-nitrobenzonitrile via quenching of intrinsic tryptophan fluorescence[5]. Due to strong

213   inner filter effects, binding experiments with 2-hydroxy-5-nitrobenzonitrile could not be

214   performed reliably. However, stopped-flow measurements with 6NBT (Supplementary Figure 8)

215   demonstrated that the transition-state analogue bound more rapidly to HG3-Evolved than to HG3-

216   Core, while its unbinding rate remained unchanged (Supplementary Table 7). These results

217   demonstrate that distal mutations accelerate substrate binding when incorporated into HG3-Core

218   to form HG3-Evolved.

219   To investigate the influence of distal mutations on product release, we compared relative

220   rate constants from various kinetic experiments. Given that K_M values for HG3-Core and HG3-

221　Evolved are in the millimolar range (Table 1), while dissociation constants for 6NBT are in the

222　micromolar range (Supplementary Table 7), substrate unbinding (k_2) is expected to be much faster

223　than unbinding of the transition-state analogue. Since the rate of the chemical transformation (k_3)

224　in HG3-Evolved is expected to be at least 100-fold higher than that of substrate unbinding (Figure

225　4), k_3 likely exceeds k_{cat} for this variant (Table 1). As a result, product release (k_4) becomes rate-

226　limiting, since $k_{cat} = k_3 \times k_4 / (k_3 + k_4)$. This conclusion is consistent with previous kinetic isotope

227　effect experiments on HG3-Evolved, which showed that turnover in this variant is limited by

228　product dissociation[34]. In HG3-Core, where the rate of the chemical transformation is only 1.4-

229　fold higher than that of substrate unbinding (m = $k_3 / (k_2 + k_3)$ = 0.59), product release may either

230　be rate-limiting or faster than the chemical transformation. If product release is rate-limiting, it

231　would likely be slower than in HG3-Evolved due to the lower k_{cat} value of the Core variant.

232　Alternatively, if the chemical transformation remains rate-limiting in HG3-Core, this indicates that

233　distal mutations shift the rate-limiting step from the chemical transformation in HG3-Core to

234　product release in HG3-Evolved. In either case, this analysis reveals that distal mutations in HG3-

235　Evolved not only enhance the rate of the chemical transformation, possibly through the enrichment

236　of productive substates observed in the crystallographic ensemble (Figure 3a), but also modulate

237　energy barriers related to substrate binding and product release, thereby optimizing the catalytic

238　cycle for more efficient catalysis.

239　　　　For both 1A53-Core and 1A53-Evolved, KSVE experiments on k_{cat}/K_M revealed a slope of

240　approximately 1.0, indicating that formation of productive enzyme-substrate complexes is much

241　slower than the chemical transformation in both variants (Figure 4c). Despite stopped-flow

242　experiments showing slower 6NBT binding to 1A53-Evolved (Supplementary Table 7,

243　Supplementary Figure 9), its higher k_{cat} suggests that distal mutations facilitate product release.

11

244 The higher K_M values of these enzymes (Table 1), compared to the transition-state analogue

245 dissociation constants (Supplementary Table 7), suggest that substrate unbinding is much faster

246 than 6NBT unbinding. Since the chemical transformation is at least 100-fold faster than substrate

247 unbinding for these enzymes, k_3 is expected to be much larger than k_{cat}, indicating that product

248 release is rate-limiting. These analyses show that distal mutations enhance catalysis in 1A53-

249 Evolved primarily by accelerating product release.

250   For KE70-Core, KSVE on k_{cat}/K_M revealed a near-zero slope, indicating that the diffusional

251 encounter of substrate and enzyme is much faster than the chemical transformation. In KE70-

252 Evolved, a KSVE of 0.47 indicates near equivalent rates for substrate unbinding and the chemical

253 step (Figure 4d). The nearly doubled k_{cat} in KE70-Evolved compared to KE70-Core (Table 1)

254 reflects a faster chemical transformation in the Evolved variant rather than slower substrate

255 capture, possibly caused by the enrichment of productive conformational substates observed in its

256 crystallographic ensemble (Figure 3c). While we could not perform binding assays with these

257 enzymes because they lack a tryptophan reporter in the active site, we could measure KSVE on

258 k_{cat} for KE70-Evolved, which revealed an inverse hyperbolic pattern (Figure 4e). This pattern

259 suggests a solvent-sensitive internal isomerization of the enzyme-product complex into a more

260 active conformation that is favored in a more viscous medium[33]. Since k_{cat} comprises both the

261 chemical transformation (k_3) and product release (k_4) rate constants, and no inverse hyperbolic

262 pattern was observed for KSVE on k_{cat}/K_M, we propose that product release by KE70-Evolved is

263 accelerated under viscous conditions by a diffusion-dependent conformational change of the

264 enzyme-product complex. These findings highlight the critical role of conformational dynamics in

265 accelerating product release during KE70-Evolved catalysis.

266

12

267   *Dynamical effects*

268   KSVE and stopped-flow experiments showed that distal mutations facilitate substrate

269   binding and/or product release, which could be achieved by promoting active-site opening. To test

270   this hypothesis, we conducted microsecond-timescale molecular dynamics simulations of the

271   enzymes in their unbound form (Methods). Root-mean-square fluctuations revealed that the most

272   flexible regions in all enzymes were surface loops (Supplementary Figure 10). In 1A53 and KE70

273   enzymes, the most flexible loop is located at the entrance of the active site, whereas in the HG3

274   family, it is further away. Principal component analysis (PCA) showed that both distal and active-

275   site mutations reshape the conformational landscape, allowing access to alternative conformations

276   or depopulating others (Figure 5a). For 1A53 and KE70 enzymes, one or both principal

277   components reflected movements of loops covering the active site, while in HG3, they involved

278   distal loops (Supplementary Figure 11).

279   To assess active-site opening, we analyzed the distances between surface loops covering

280   the active site (Figure 5b,c). In 1A53-Designed and 1A53-Shell, loop distance analysis revealed

281   two primary conformational states: a closed state with a median loop distance of ~10 Å,

282   corresponding to the crystallographic conformation, and an open state with a median distance of

283   ≥17 Å. Distal mutations introduced into 1A53-Designed (to form 1A53-Shell) increased the

284   openness of the open state, shifting the median loop distance by 2.5 Å. In 1A53-Core and 1A53-

285   Evolved, a third, partially open state centered at ~15 Å was observed. Incorporating distal

286   mutations into 1A53-Core further expanded the median open distance by more than 10 Å. These

287   findings suggest that distal mutations enhance active-site accessibility in 1A53 by promoting more

288   open loop conformations, which is consistent with faster product release in 1A53-Evolved.

289   For KE70-Designed and KE70-Shell, two states were observed: closed and open, with

290   median loop distances at around 7 and 12.5 Å, respectively, the latter corresponding to the

291   crystallographic conformation. Introducing distal mutations into KE70-Designed to generate

292   KE70-Shell reduced the closed conformation population from 30% to 15%, but median loop

293   distances remained unchanged. For KE70-Core and KE70-Evolved, a third conformational state,

294   which is partially open (median loop distance around 10.5 Å) is observed. While distal mutations

295   introduced into KE70-Core to produce KE70-Evolved increased the closed conformation

296   population, in contrast to their effect on KE70-Designed, the open state became more open by

297   nearly 1 Å. These results suggest that distal mutations in KE70 facilitate opening of the active site

298   by either increasing the proportion of open substates or shifting open conformations toward a more

299   open state.

300   In contrast to 1A53 and KE70 enzymes, HG3 variants exhibited minimal changes in active-

301   site loop conformations, consistent with the most flexible loop being distant from the active site

302   (Supplementary Figure 10). Since the dynamics of loops near the active site were largely

303   unaffected, we examined whether distal mutations influenced sidechain dynamics, potentially

304   altering the bottleneck radius of the active site entrance. Previous studies linked a larger bottleneck

305   radius to enhanced catalysis in HG3 variants[4]. Caver3[35] analysis of bottleneck radii in molecular

306   dynamics snapshots (Supplementary Figure 12) revealed that distal mutations increased the

307   percentage of structures in the conformational ensemble with a bottleneck radius ≥ 0.9 Å, a

308   threshold chosen because it matches the radius of small ions commonly bound in protein structures,

309   and raised the median bottleneck radius. These findings suggest that distal mutations enhance

310   active-site accessibility in HG3 enzymes by widening the entrance without significantly altering

311 loop conformations. This is consistent with the faster substrate binding observed in HG3-Evolved

312 and suggests that product release is also accelerated.

313

314 **Discussion**

315 In this study, we investigated the influence on the catalytic cycle of distal mutations

316 introduced by directed evolution across three lineages of de novo enzymes designed to catalyze

317 the same reaction. Unlike previous investigations of structural and dynamic changes along enzyme

318 evolutionary trajectories[5, 9, 36], we sought to specifically isolate the effects of distal mutations from

319 those occurring at the active site by generating Core and Shell variants. Our findings demonstrate

320 that active-site mutations are the primary drivers of enhanced catalytic efficiency, organizing the

321 active site for optimal enzyme-transition state interactions to accelerate the chemical

322 transformation. These findings are consistent with the long-standing understanding that active-site

323 residues play a dominant role in shaping enzyme activity[37]. However, distal mutations complement

324 these effects by tuning energy barriers to substrate binding and product release, optimizing the

325 catalytic cycle for efficient catalysis. This role is particularly critical when these steps are rate-

326 limiting, as observed in many natural enzymes[38-40]. In such cases, large-scale structural and

327 dynamic rearrangements across the protein scaffold become necessary to achieve further efficiency

328 gains. As a result, the benefits of distal mutations emerge only when the active site is already

329 optimized, as improved substrate binding and product release cannot compensate for slow

330 chemical transformation. Although we cannot rule out that distal mutations enhance catalysis

331 through additional mechanisms, our strategy of generating Core and Shell variants provides a

332 valuable approach for studying underexplored catalytic effects possibly mediated by distal

15

333   mutations, such as alterations to heat capacity[27], local electric fields[41], and conformational

334   entropy[42].

335   Our results highlight important challenges in de novo enzyme design. Traditional enzyme

336   design algorithms focus on carving active sites for specific reactions on preexisting protein

337   scaffolds without considering distal positions[15, 16, 18]. Even recent deep-learning-based methods for

338   enzyme design optimize distal residues for folding, stability and solubility rather than for their

339   influence on catalysis[43, 44]. While both approaches have produced functional enzymes, the resulting

340   catalytic activities remain low ($k_{cat} \leq 0.1$ s^{-1}). Meanwhile, directed evolution consistently remodels

341   designed active sites to enhance catalytic efficiency, sometimes even altering catalytic residues[25,

342   26, 29], while also fine-tuning conformational dynamics through mutations at distal positions[10, 45].

343   Our study shows that even if perfectly optimized active sites could be designed from scratch, these

344   enzymes would still fall short of their full catalytic potential without the appropriate distal

345   mutations. This finding underscores the importance of removing bottlenecks in the catalytic cycle,

346   such as substrate binding and product release, by optimizing distal positions to adjust

347   conformational dynamics and enhance activity.

348   Moving forward, we propose a framework for de novo enzyme design that optimizes the

349   active site first, potentially through ensemble-based design[30] or deep-learning techniques[43], before

350   introducing distal mutations to adjust conformational dynamics for facilitated ligand binding and

351   release. Identifying beneficial distal mutations to precisely tune structural dynamics remains

352   challenging, but computational methods capable of identifying dynamically-linked residues could

353   help[46]. Balancing the structural rigidity needed for precise active-site alignment with the flexibility

354   required for efficient substrate binding and product release will likely be essential for unlocking

355   the full catalytic potential of de novo enzymes. The Core and Shell variants developed here offer

16

356   valuable benchmarks for testing new predictive methods aimed at refining distal regions,

357   potentially guiding future developments in enzyme design.

358

359   **Methods**

360   *Protein expression and purification.* Codon-optimized (*E. coli*) and his-tagged (C-terminus) genes

361   for Kemp eliminases (Supplementary Table 8) cloned into the pET-29b(+) vector via *Nde*I and

362   *Xho*I were obtained from Twist Bioscience. Enzymes were expressed in *E. coli* BL21-Gold (DE3)

363   cells (Agilent) using lysogeny broth (LB) supplemented with 100 µg mL^{-1} kanamycin. Cultures

364   were grown at 37 °C with shaking to an optical density at 600 nm of 0.3–0.7, after which protein

365   expression was initiated with 1 mM isopropyl β-D-1-thiogalactopyranoside. Following incubation

366   for 16 h at 16 °C with shaking (250 rpm), cells were harvested by centrifugation, resuspended in

367   8 mL lysis buffer (5 mM imidazole in 100 mM potassium phosphate buffer, pH 8.0), and lysed

368   with an EmulsiFlex-B15 cell disruptor (Avestin). Proteins were purified by immobilized metal

369   affinity chromatography using Ni–NTA agarose (Qiagen) pre-equilibrated with lysis buffer in

370   individual Econo-Pac gravity-flow columns (Bio-Rad). Columns were washed twice, first with 10

371   mM imidazole in 100 mM potassium phosphate buffer (pH 8.0), and then with the same buffer

372   containing 20 mM imidazole. Bound proteins were eluted with 250 mM imidazole in 100 mM

373   potassium phosphate buffer (pH 8.0) and exchanged into 100 mM sodium phosphate buffer (pH

374   7.0) supplemented with 100 mM sodium chloride using Econo-Pac 10DG desalting pre-packed

375   gravity-flow columns (Bio-Rad). Proteins used for crystallization were further purified by gel

376   filtration in 50 mM sodium citrate buffer (pH 5.5) and 150 mM sodium chloride using an ENrich

377   SEC 650 size-exclusion chromatography column (Bio-Rad). Purified samples were concentrated

378   using Amicon Ultracel-10K centrifugal filter units (EMD Millipore) and quantified by measuring

17

379   the absorbance at 280 nm and applying Beer-Lambert's law using calculated extinction

380   coefficients obtained from the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

381

382   *Circular dichroism and thermal denaturation assays.* Circular dichroism measurements were

383   performed with a Jasco J-815 spectrometer using 650-µL aliquots of each Kemp eliminase in a 1-

384   mm path-length quartz cuvette (Jasco) at a concentration of 30–40 µM in 10 mM sodium phosphate

385   buffer (pH 7) supplemented with 100 mM sodium chloride. For structural characterization of

386   protein folds, circular dichroism spectra were acquired from 200 to 250 nm, sampled every 0.2 nm

387   at a rate of 10 nm min^{-1}. Five scans were acquired and averaged for each sample. For thermal

388   denaturation assays, samples were heated at a rate of 0.5 °C per minute, and ellipticity at 222 nm

389   was measured every 0.2 °C. T_m values were determined by fitting a 2-term sigmoid function with

390   baseline correction using nonlinear least-squares regression[47].

391

392   *Steady-state kinetics.* All assays were carried out at 27 °C in 100 mM sodium phosphate buffer

393   (pH 7.0) supplemented with 100 mM sodium chloride. Triplicate 200-µL reactions with varying

394   concentrations of freshly prepared 5-nitrobenzisoxazole (AstaTech) or 6-nitrobenzisoxazole

395   (Combi-blocks) dissolved in methanol (10 % final concentration, pH of reaction mixture adjusted

396   to 7.0 after addition of methanol-solubilized substrate) were initiated by the addition of 10 µM

397   KE70-Designed/Shell, 100 nM KE70-Core/Evolved, 25 µM 1A53-Designed/Shell, 1 µM 1A53-

398   Core/Evolved, 2 µM HG3-Designed, 5 nM HG3-Core/Evolved and 50 nM HG3-Shell. Product

399   formation was monitored spectrophotometrically in individual wells of 96-well plates (Greiner

400   Bio-One) using a SpectraMax Plus plate reader (Molecular Devices) with measurements at 380 or

401   400 nm for reactions with 5-nitrobenzisoxazole (ε = 15,800 M^{-1} cm^{-1})[18] or 6-nitrobenzisoxazole

18

402   (ε = 2,870 M^{-1} cm^{-1})[27], respectively. Path lengths for each well were calculated ratiometrically

403   using the difference in absorbance of 100 mM sodium phosphate buffer (pH 7.0) supplemented

404   with 100 mM sodium chloride and 10 % methanol at 900 and 975 nm (27 °C). Linear phases of

405   the kinetic traces were used to measure initial reaction rates. For cases where saturation was not

406   possible at the maximum substrate concentration tested (1 or 2 mM for 6-nitrobenzisoxazole or 5-

407   nitrobenzisoxazole, respectively), data were fitted to the linear portion of the Michaelis-Menten

408   model ($v_0 = (k_{cat}/K_M)[E_0][S]$) and k_{cat}/K_M was deduced from the slope. In all other cases, data were

409   fitted to the Michaelis-Menten equation to calculate individual k_{cat} and K_M parameters.

410

411   *Kinetic solvent viscosity effects (KSVE).* The effect of solvent viscosity on k_{cat}/K_M was determined

412   at 27 °C in 100 mM sodium phosphate buffer (pH 7.0) supplemented with 100 mM NaCl, using

413   sucrose as the viscogen at different concentrations (0, 20, 28, 33 %w/v). Solution viscosities were

414   approximated from published viscosity data of sucrose solutions[48]. Kinetic measurements were

415   performed as described above using 5 nM of HG3-Core and HG3-Evolved, 1 μM of 1A53-Core

416   and 1A53-Evolved, and 150 nM of KE70-Core and KE70-Evolved. Initial rates were determined

417   and fitted to the linear portion of the Michaelis-Menten model ($v_0 = (k_{cat}/K_M)[E_0][S]$) as saturation

418   was not possible for most reactions in the presence of sucrose. Reference values (k_{cat}/K_M or k_{cat}) in

419   the absence of sucrose were divided by the values obtained at different sucrose concentrations and

420   plotted against the relative buffer viscosity (η_{rel}) to give the corresponding slopes reported in

421   Figure 4. For KSVE on k_{cat}/K_M, the slope (m) corresponds to $k_3 / (k_2 + k_3)$, where k_2 is the rate

422   constant of substrate dissociation and k_3 is the rate constant for the chemical transformation (Figure

423   4a). A slope of 1 indicates that k_3 is at least 100-fold larger than k_2 (m = k_3 / k_3), while a slope of 0

424   indicates that k_3 is much smaller than k_2 (m = k_3 / k_2 = 0).

425

Stopped-flow kinetics. Stopped-flow experiments measuring intrinsic tryptophan fluorescence were used to monitor the kinetics of transition-state analogue (6NBT) binding to the Kemp eliminases. The measurements were performed using OLIS RSM 1000 rapid-scanning monochromator, equipped with a water bath to control the temperature. Changes in intrinsic Trp fluorescence upon binding and dissociation of 6NBT were monitored using an excitation wavelength of 295 nm and a long-pass 300 nm cut-off filter to detect emission. All experiments were performed at 20 °C. A solution of 10 µM HG3-Core, HG3-Evolved, 1A53-Core and 1A53-Evolved in 50 mM Sodium phosphate buffer, 100 mM NaCl pH 7, was loaded and quickly mixed varying TSA concentrations in the same buffer containing 1% (v/v) DMSO (mixing ration 1:1 resulting in final enzyme concentration of 5 µM. A significant decrease in fluorescence intensity was observed upon binding of 6NBT. Three replicate measurements were made for each 6NBT concentration and the data were fit to exponential equation $F = F_0 + Amp.\exp\left(-k_{obs}.t\right)$ using GraphPad Prism where F is fluorescence intensity, F_0 is the initial fluorescence intensity, Amp is the amplitude of the change in fluorescence and k_{obs} is the observed rate constant.

Crystallization. Enzyme variants were prepared in 50 mM sodium citrate buffer (pH 5.5) at the concentrations listed in Supplementary Table 4. For samples that were co-crystallized with the transition-state analogue, a 100 mM stock solution of 6NBT (AstaTech) was prepared in dimethyl sulfoxide (DMSO) and diluted in the enzyme solutions for a final concentration of 5 mM (5 % DMSO). For all variants except 1A53-Core, crystallization was performed using SWISS CI 3-well sitting drop plates. For 1A53-Core, the hanging drop method was used. Crystallization drops were prepared by mixing 1 µL of protein solution with 1 µL of the mother liquor and sealing the drop

448   inside a reservoir containing an additional 500 μL of the mother liquor solution. Mother liquor

449   solutions contained various precipitants, and the specific growth conditions that yielded the

450   crystals used for X-ray data collection are listed on Supplementary Table 5. For the structures

451   bound to transition-state analogue, 6NBT was added directly on top of crystal drops using Acoustic

452   dispensing with an Echo 650 liquid handler (Labcyte). The transition-state analogue was allowed

453   to diffuse in the crystal drop for approximately 10 min.

454

455   *X-ray data collection and processing.* All crystals were flash frozen in liquid nitrogen using

456   appropriate cryoprotectant and single-crystal X-ray diffraction data were collected on beamline

457   8.3.1 at the Advanced Light Source. The beamline was equipped with a Pilatus3 S 6M detector

458   and was operated at a photon energy of 11111 eV. Data was processed and scaled using XDS[49],

459   and AIMLESS[50], respectively.

460

461   *Structure determination.* We obtained initial phase information for calculation of electron density

462   maps by molecular replacement using the program Phaser[51], as implemented in v1.13.2998 of the

463   PHENIX suite[52]. Available structures of Evolved variants were used as molecular replacement

464   search models. Next, we performed iterative steps of manual model rebuilding followed by

465   refinement of atomic positions, atomic displacement parameters, and occupancies using a riding

466   hydrogen model and automatic weight optimization. For the 1A53-Core structure with and without

467   transition-state analogue, and KE70-Core with transition-state analogue, refinement also includes

468   a translation-libration-screw (TLS) model. All model building was performed using Coot

469   0.8.9.236[53] and refinement steps were performed with phenix.refine within the PHENIX suite

470   (v1.13-2998). Restraints for 6NBT were generated using phenix.elbow, starting from coordinates

471   available in the Protein Data Bank (PDB ligand ID: 6NBT). Further information regarding model

472   building and refinement, as well as PDB accession codes for the final models, are presented in

473   Supplementary Table 5.

474

475   *Ensemble refinement.* Time-averaged ensembles were generated with

476   phenix.ensemble_refinement implemented in PHENIX v.1.15.2-3472. In ensemble refinement[54],

477   local atomic fluctuations are sampled using molecular dynamics simulations accelerated and

478   restrained by electron density to produce ensemble models fitted to diffraction data. Briefly, input

479   crystal structures were edited to remove low-occupancy conformers and assign an occupancy of

480   1.0 to the remaining conformer. Following addition of riding hydrogens, parallel ensemble

481   refinement simulations were performed using various combinations of the parameters p_{TLS} (0.6,

482   0.8, 0.9, 1.0), τ_x (0.5, 1.5, 2.0) and $w_{x\text{-}ray}$ (2.5, 5.0, 10.0), where p_{TLS} describes the percentage of

483   atoms included in a translation-libration-screw (TLS) model use to remove the effects of global

484   disorder, τ_x is the simulation time-step and $w_{x\text{-}ray}$ is the coupled tbath offset, which controls the

485   extent to which the electron density contributes to the simulation force field such that the

486   simulation runs at a target temperature of 300 K. The ensemble generated from each crystal

487   structure that displayed the lowest R_{free} value was selected for analysis (Supplementary Table 6).

488

489   *Molecular dynamics simulations.* The Amber 2020 software (http://ambermd.org/) with the

490   AMBER19SB forcefield[55] was used for all simulations. Crystal structures of the enzymes were

491   used as initial structures, except for 1A53-Shell, 1A53-Evolved and KE70-Shell, which were

492   generated in silico from the 1A53-2.5 crystal structure (PDB ID: 6NW4) and KE70-Evolved

493   structure (PDB ID: 3Q2D) using the protein design software Triad[56] (Protabit LLC, Pasadena, CA,

494   USA). Amino-acid protonation states were predicted using the H++ server

495   (http://biophysics.cs.vt.edu/H++). Each enzyme structure was immersed in a pre-equilibrated

496   cubic box (10 Å edge length) filled with buffer, using the OPC water model[57], neutralized with

497   explicit counterions (Na^+, Cl^-) using the Amber 20 leap module. A two-state geometry

498   optimization approach was applied. The first stage involved the minimization of solvent molecules

499   and ion positions while imposing positional restraints on solutes using a harmonic potential (force

500   constant: 500 kcal mol^{-1} $Å^{-2}$). The second stage consisted of an unrestrained minimization of all

501   atoms in the simulation cell. The systems were gradually heated in seven 50 ps steps, each

502   increasing the temperature by 50 K, under constant-volume and periodic boundary conditions.

503   Water molecules were treated using the SHAKE algorithm, and long-range electrostatic effects

504   were modeled using the particle-mesh-Ewald method[58]. An 8 Å cut-off was applied to Lennard-

505   Jones and electrostatic interactions. Decreasing harmonic restraints were applied to the protein

506   during thermal equilibration (500, 210, 165, 125, 85, 45, 10 kcal $mol^{-1}Å^{-2}$), with temperature

507   control achieved using the Langevin scheme. A timestep of 1 fs was maintained during these

508   stages. Following the heating, equilibration of each system was conducted for 20 ns without

509   restraints at a constant pressure of 1 atm and a 2 fs timestep. After equilibration in the NPT

510   ensemble, production run MD simulations for each system in the NVT ensemble with periodic-

511   boundary conditions were performed. These production trajectories were run for 2000 ns in three

512   separate replicates.

513   Similar protocol was used to evaluate the conformation of 1A53-Core containing Models

514   of 1A53-Core containing alternate Trp110 rotamers (Supplementary Figure 5) were prepared from

515   the crystal structure by deleting the bound MES molecule and changing the rotamer in Pymol

516   (http://www.pymol.org/). Molecular dynamics simulations were performed as described above.

517   Trp110 dihedral angles throughout the simulations were measured using the python interface of

518   cpptraj package[59].

519

520   *Bottleneck radius analysis.* Caver3[35] was used to calculate substrate entry channels for

521   representative snapshots of the most populated energy minima of HG3-Designed, HG3-Core,

522   HG3-Evolved, and HG3-Shell. Parameters used for the entry channel search were 6 Å Shell depth,

523   5 Å Shell radius, clustering threshold value of 3.5 and a 0.9 Å minimum probe radius.

524

525   **Data availability**

526   Structure coordinates for all Kemp eliminases have been deposited in the RCSB Protein Data Bank

527   with the following accession codes: 1A53-Core (8FMC, 8FOQ), KE70-Core (8FMD, 8FOR) and

528   HG3-Shell (8FME, 8FOS). Source data are provided with this paper. Other relevant data are

529   available from the corresponding authors upon reasonable request.

530

531   **Acknowledgements**

549   **Author Contributions**

550   R.A.C. and N.Z. conceived the project. N.Z. and C.K. purified proteins. N.Z., C.K., R.V.R and

551   S.E.H performed enzyme kinetics experiments. N.Z., H.D. and J.S. performed molecular dynamics

552   experiments. S.O. designed molecular dynamics experiments. P.A. crystallized proteins. P.A.

553   performed X-ray diffraction experiments. P.A. and R.A.C. performed structure refinements. P.A.

554   and J.S.F. designed X-ray crystallography experiments. N.Z. and R.A.C. wrote the manuscript. All

555   authors edited the manuscript.

556

557   **Competing Interests**

558   The authors declare no competing interests.

559

560 **Table 1. Properties of various Kemp eliminases**

Enzyme	Original name[a]	# mutations	K_M (mM)[b]	k_{cat} (s^{-1})[b]	k_{cat} / K_M (M^{-1} s^{-1})[c]	k_{cat} / K_M fold increase	T_m (°C)
HG3 series							
HG3-Designed	HG3-K50Q[d]	–	N.D.	N.D.	1300 ± 90	–	51
HG3-Shell	–	9	N.D.	N.D.	4900 ± 500	4	50
HG3-Core	HG4	7	1.9 ± 0.2	230 ± 20	120000 ± 20000	90	52
HG3-Evolved	HG3.17	16	2.1 ± 0.3	320 ± 30	150000 ± 40000	120	56
1A53 series							
1A53-Designed	1A53-2	–	N.D.	N.D.	4.6 ± 0.4	–	74
1A53-Shell	–	8	N.D.	N.D.	5.0 ± 0.7	1	65
1A53-Core	–	6	2.0 ± 0.5	13 ± 2	7000 ± 3000	1500	85
1A53-Evolved	1A53-2.9	14	1.4 ± 0.2	19 ± 2	14000 ± 3000	3000	61
KE70 series							
KE70-Designed	KE70-S20a/A240a[e]	–	N.D.	N.D.	150 ± 7	–	57
KE70-Shell	–	2	1.9 ± 0.3	0.24 ± 0.02	130 ± 30	1	60
KE70-Core	–	6	0.23 ± 0.03	5.0 ± 0.2	22000 ± 4000	150	55
KE70-Evolved	R6 6/10A	8	0.35 ± 0.02	9.1 ± 0.1	26000 ± 2000	170	58

561 [a] Original names are those previously used in literature for these variants.

562 [b] Kinetic parameters for all enzymes except the 1A53 series are reported for the 5-nitrobenzisoxazole substrate. For the 1A53 series, kinetic

563 parameters are reported for 6-nitrobenzisoxazole. All reactions were performed at 27 °C in 50 mM sodium phosphate buffer (pH 7.0),

564 supplemented with 100 mM NaCl and 10 % methanol. N.D. indicates that individual parameters K_M and k_{cat} could not be determined accurately

565 because saturation was not possible at the maximum substrate concentration tested, which is the substrate's solubility limit (2 mM or 1 mM for 5-

566 nitrobenzisoxazole or 6-nitrobenzisoxazole, respectively). Data represent the average of six individual replicate measurements from two

567 independent protein batches. Errors of regression fitting, which represent the absolute measure of the typical distance that each data point falls

568 from the regression line, are provided.

569 [c] Catalytic efficiencies (k_{cat}/K_M) were calculated from individual K_M and k_{cat} parameters, or from the slope of the linear portion ([S] << K_M) of the

570 Michaelis-Menten model ($v_0 = (k_{cat}/K_M)[E_0][S]$) when saturation was not possible.

571 [d] HG3-K50Q is a point mutant of *de novo* enzyme HG3 that contains the Q50 hydrogen-bond donor identified by directed evolution. We added

572 this mutation into HG3-Designed to allow for direct comparison with Shell, Core and Evolved HG3 variants, which all contain this mutation.

573 HG3-K50Q is 9-fold more catalytically efficient than HG3 ($k_{cat}/K_M = 146$ M^{-1} s^{-1})[4].

574 [e] KE70-S20a/A240a is a double mutant of *de novo* enzyme KE70 that contains two insertions in active site loops (S20a and A240a), which were

575 manually introduced into KE70 during its directed evolution[3]. We added them into KE70-Designed to allow for direct comparison with Shell,

576 Core and Evolved KE70 variants, which all contain these insertions. These insertions increase catalytic efficiency by approximately 2-fold when

577 introduced on their own into KE70[30].

578
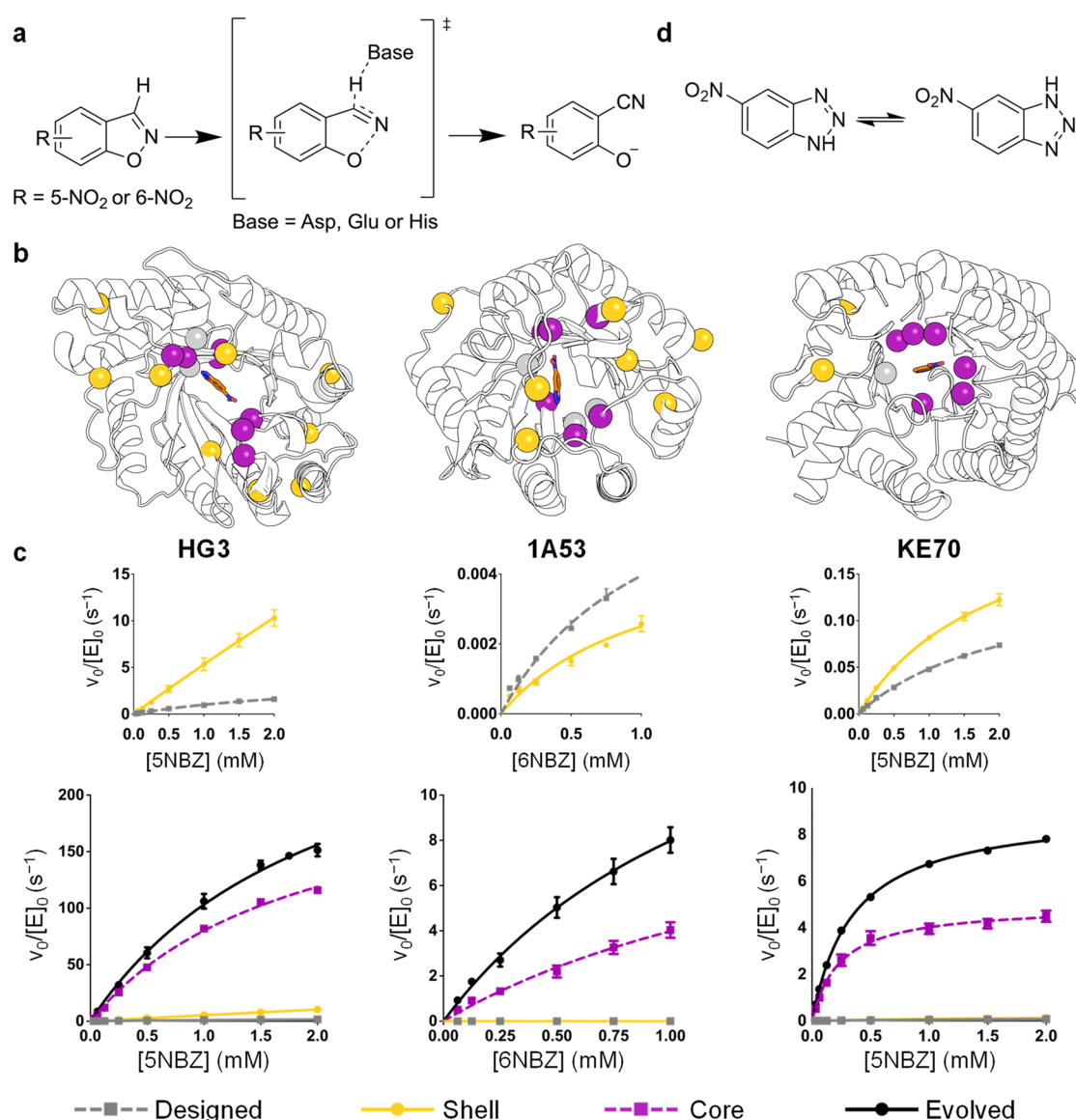
Figures

Figure 1. Core and Shell variants of Kemp eliminases. (a) Kemp eliminases catalyze the concerted deprotonation and ring opening of benzisoxazoles using an Asp, Glu or His catalytic base. (b) Core and Shell variants from the HG3, 1A53 and KE70 families contain either active-site (purple) or distal (yellow) mutations found by directed evolution, respectively. Mutations are listed on Supplementary Table 2. Catalytic residues are shown as grey spheres. (c) Michaelis–Menten plots of normalized initial rates as a function of substrate concentration are shown. Data represent the average of six or nine individual replicate measurements from two or three independent protein batches, with error bars indicating the SEM (mean ± SEM). (d) The 6-nitrobenzotriazole transition-state analogue tautomerizes in water, which allows it to bind to enzymes that react with either 5- or 6-nitrobenzisoxazole.
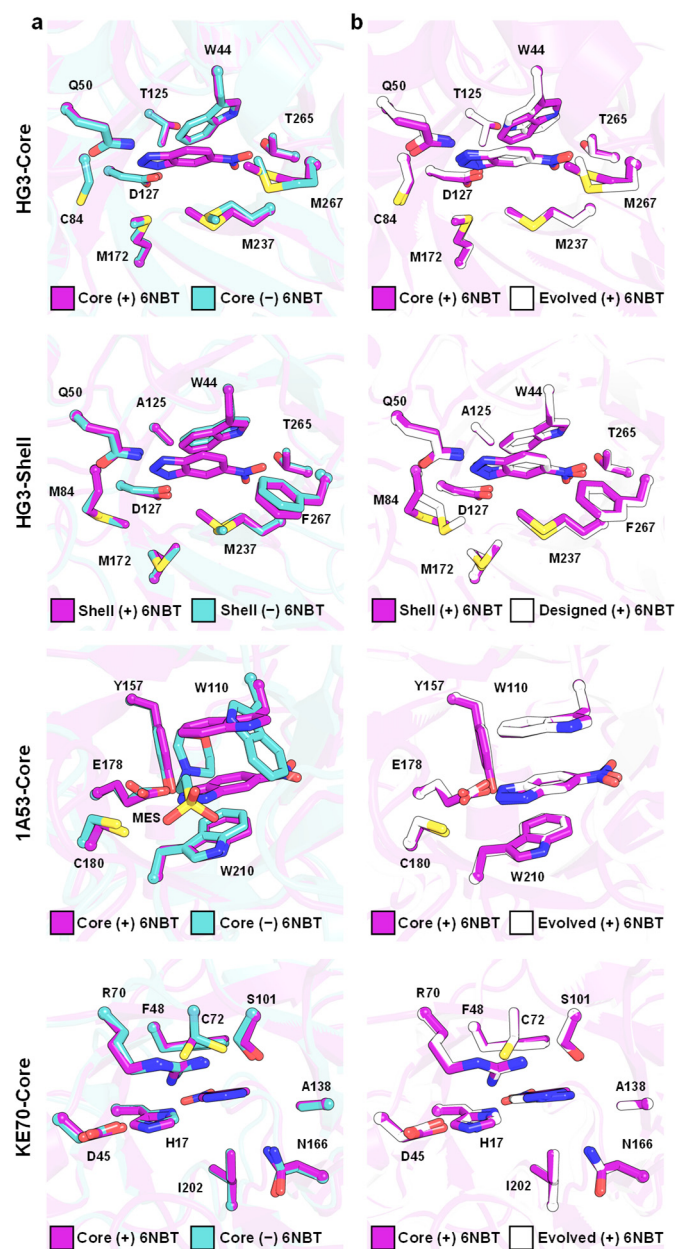
Figure 2. Crystal structures. (a) The active sites of HG3-Core, HG3-Shell and KE70-Core do not change substantially upon binding of the 6-nitrobenzotriazole (6NBT) transition-state analogue. In 1A53-Core, a 2-(*N*-morpholino)ethanesulfonic acid (MES) molecule from the crystallization buffer occupies the 6NBT binding site, inducing a conformational change in W110. (b) Core variants display active-site configurations nearly identical to those of their Evolved counterparts, while HG3-Shell displays a configuration similar to HG3-Designed. The 1A53-2.5 crystal structure (PDB ID: 6NW4)[27] was used for comparison with 1A53-Core, as 1A53-Evolved was unavailable. We re-refined the KE70-Evolved crystal structure with bound 6NBT (PDB ID: 3Q2D)[3] to model missing residues (22–25 in chain A and 3–25 in chain B) into the available density and to flip 6NBT into a productive binding pose. In all cases, only the major conformer of each first-shell active-site residue is shown, with the 6NBT transition-state analogue at the center.

28
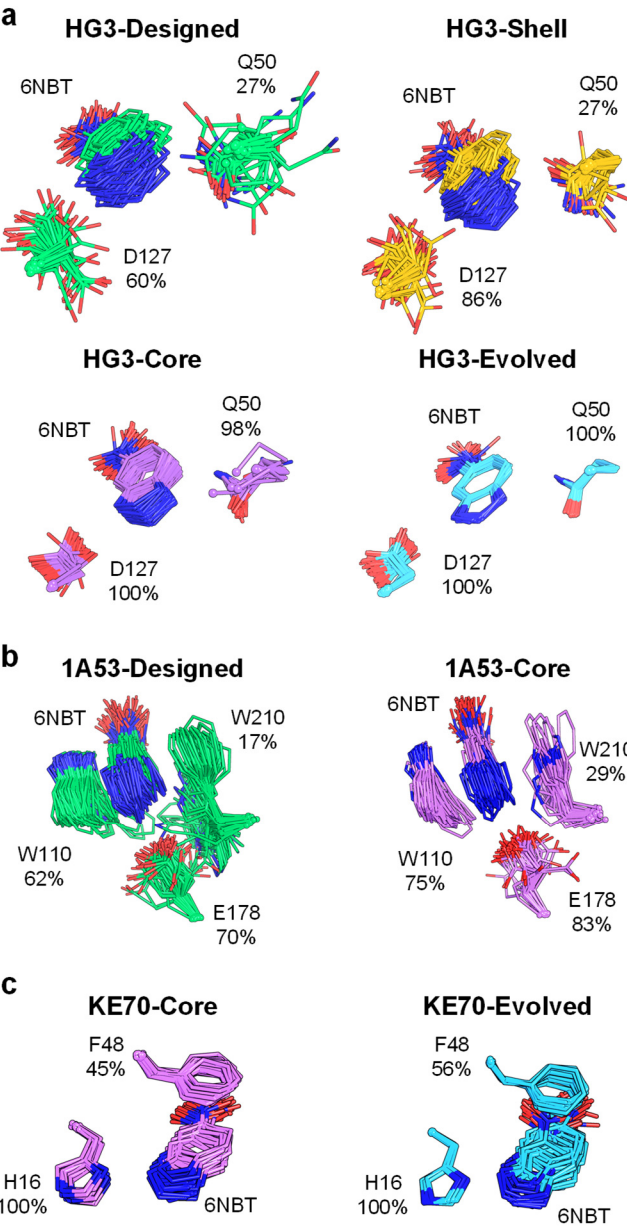
Figure 3. Ensemble refinement of Kemp eliminases. Ensemble refinements of 6NBT-bound crystal structures for (a) HG3, (b) 1A53, and (c) KE70 variants reveal that distal and active-site mutations increase the proportion of productive substates in the conformational ensemble. Percentages next to each residue indicate the fraction of substates that form hydrogen bonds or π-stacking interactions with 6NBT.
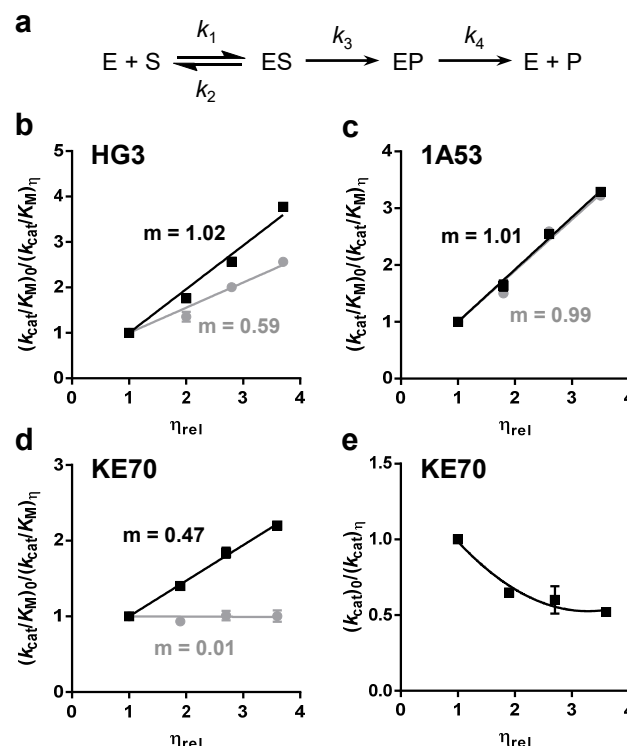
Figure 4. Kinetic solvent viscosity effects. (a) Kemp elimination reaction scheme. (b-e) Grey and black lines correspond to Core and Evolved variants, respectively. Kinetic solvent viscosity effects on k_{cat}/K_M for (a) HG3, (b) 1A53 and (c) KE70 variants show a linear dependence on relative viscosity (η_{rel}). The slope (m) corresponds to $k_3 / (k_2 + k_3)$, where k_2 is the rate constant of substrate dissociation and k_3 is the rate constant for the chemical transformation. A slope of 1 indicates that k_3 is much larger than k_2 (m = k_3 / k_3). A slope of 0 indicates that k_3 is much smaller than k_2 (m = k_3 / k_2 = 0). (e) Kinetic solvent viscosity effects on k_{cat} for KE70-Evolved show an inverse hyperbolic pattern, which is consistent with the presence of a solvent-sensitive internal isomerization of the enzyme-product complex. Data represent the average of 6 individual replicates from 2 independent protein batches, with error bars reporting the SEM.

Figure 5. Conformational landscapes of Kemp eliminases analyzed by molecular dynamics. (a) Representation of the molecular dynamics trajectories projected into the two most important principal components (PC1, PC2) based on Cα contacts for each enzyme family. Color coding represents the relative population density of the states, with blue indicating the most populated wells and red representing the least populated regions. Details on conformational changes contributing to PC1 and PC2 are shown on Supplementary Figure 11. (b) Histograms of loop distances across the entire energy landscape for KE70 and 1A53 enzymes, or from the most populated well for HG3 enzymes. Loop distances were calculated between Cα carbons of residues 21 and 78 for KE70 variants, 58 and 188 for 1A53 variants, and 90 and 275 for HG3 variants (shown as spheres in panel c). The dashed line represents the average loop distance in crystal structures of each enzyme family. The multimodal distributions were deconvoluted using a Gaussian mixture model, with the median loop distances for each conformational state indicated in Å. (c) Representative snapshots for each conformational state sampled by the enzymes. These snapshots display loop distances equivalent to the median value for each population in the distribution. The Cα carbons of residues used to calculate loop distances are shown as spheres.

References

1. Rauwerdink, A. & Kazlauskas, R.J. How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: the Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. *ACS Catalysis* **5**, 6153-6176 (2015).

2. Du, S. et al. Conformational ensembles reveal the origins of serine protease catalysis. *Science* **387**, eado5068 (2025).

3. Khersonsky, O. et al. Optimization of the In-Silico-Designed Kemp Eliminase KE70 by Computational Design and Directed Evolution. *Journal of Molecular Biology* **407**, 391-412 (2011).

4. Broom, A. et al. Ensemble-based enzyme design can recapitulate the effects of laboratory directed evolution in silico. *Nature Communications* **11**, 4808 (2020).

5. Otten, R. et al. How directed evolution reshapes the energy landscape in an enzyme to boost catalysis. *Science* **370**, 1442-1446 (2020).

6. Hong, N.-S. et al. The evolution of multiple active site configurations in a designed enzyme. *Nature Communications* **9**, 3900 (2018).

7. Fröhlich, C. et al. Epistasis arises from shifting the rate-limiting step during enzyme evolution of a β-lactamase. *Nature Catalysis* **7**, 499-509 (2024).

8. Peccati, F., Noey, E.L., Houk, K.N., Osuna, S. & Jiménez-Osés, G. Interplay Between Substrate Polarity and Protein Dynamics in Evolved Kemp Eliminases. *ChemCatChem* **16**, e202400444 (2024).

9. Campbell, E. et al. The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* **12**, 944-950 (2016).

10. Jimenez-Oses, G. et al. The role of distant mutations and allosteric regulation on LovD active site dynamics. *Nat Chem Biol* **10**, 431-436 (2014).

11. Doustmohammadi, H., Sanchez, J., Ram Mahato, D. & Osuna, S. Evolution Enhances Kemp Eliminase Activity by Optimizing Oxyanion Stabilization and Conformational Flexibility. *Chemistry – A European Journal* **n/a**, e202403747 (2024).

12. Lee, J. & Goodey, N.M. Catalytic Contributions from Remote Regions of Enzyme Structure. *Chemical Reviews* **111**, 7595-7624 (2011).

13. Miton, C.M. & Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein science : a publication of the Protein Society* **25**, 1260-1272 (2016).

14. Osuna, S. The challenge of predicting distal active site mutations in computational enzyme design. *WIREs Computational Molecular Science* **11**, e1502 (2021).

15. Rothlisberger, D. et al. Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190-195 (2008).

16. Jiang, L. et al. De novo computational design of retro-aldol enzymes. *Science* **319**, 1387-1391 (2008).

17. Siegel, J.B. et al. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309-313 (2010).

18. Privett, H.K. et al. Iterative approach to computational enzyme design. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 3790-3795 (2012).

19. Lassila, J.K., Privett, H.K., Allen, B.D. & Mayo, S.L. Combinatorial methods for small-molecule placement in computational enzyme design. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16710-16715 (2006).

20. Zanghellini, A. et al. New algorithms and an in silico benchmark for computational enzyme design. *Protein Science* **15**, 2785-2794 (2006).

21. Althoff, E.A. et al. Robust design and optimization of retroaldol enzymes. *Protein Science* **21**, 717-726 (2012).

22. Bjelic, S. et al. Computational Design of Enone-Binding Proteins with Catalytic Activity for the Morita–Baylis–Hillman Reaction. *ACS Chemical Biology* **8**, 749-757 (2013).

23. Khersonsky, O. et al. Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 10358-10363 (2012).

24. Khersonsky, O. et al. Evolutionary Optimization of Computationally Designed Enzymes: Kemp Eliminases of the KE07 Series. *Journal of Molecular Biology* **396**, 1025-1042 (2010).

25. Crawshaw, R. et al. Engineering an efficient and enantioselective enzyme for the Morita-Baylis-Hillman reaction. *Nat Chem* **14**, 313-320 (2022).

26. Blomberg, R. et al. Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* **503**, 418-421 (2013).

27. Bunzel, H.A. et al. Emergence of a Negative Activation Heat Capacity during Evolution of a Designed Enzyme. *Journal of the American Chemical Society* **141**, 11745-11748 (2019).

28. Giger, L. et al. Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nature Chemical Biology* **9**, 494-498 (2013).

29. Obexer, R. et al. Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nature Chemistry* **9**, 50-56 (2017).

30. Rakotoharisoa, R.V. et al. Design of Efficient Artificial Enzymes Using Crystallographically Enhanced Conformational Sampling. *J Am Chem Soc* **146**, 10001-10013 (2024).

31. Shoichet, B.K., Baase, W.A., Kuroki, R. & Matthews, B.W. A relationship between protein stability and protein function. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 452-456 (1995).

32. Tokuriki, N., Stricher, F., Serrano, L. & Tawfik, D.S. How Protein Stability and New Functions Trade Off. *PLOS Computational Biology* **4**, e1000002 (2008).

33. Gadda, G. & Sobrado, P. Kinetic Solvent Viscosity Effects as Probes for Studying the Mechanisms of Enzyme Action. *Biochemistry* **57**, 3445-3453 (2018).

34. Kries, H., Bloch, J.S., Bunzel, H.A., Pinkas, D.M. & Hilvert, D. Contribution of Oxyanion Stabilization to Kemp Eliminase Efficiency. *ACS Catalysis* **10**, 4460-4464 (2020).

35. Chovancova, E. et al. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLOS Computational Biology* **8**, e1002708 (2012).

36. Bunzel, H.A. et al. Evolution of dynamical networks enhances catalysis in a designer enzyme. *Nature Chemistry* **13**, 1017-1022 (2021).

37. Yang, G., Miton, C.M. & Tokuriki, N. A mechanistic view of enzyme evolution. *Protein Science* **29**, 1724-1747 (2020).

721 38. Natarajan, S.K. & Sierks, M.R. Identification of Enzyme−Substrate and Enzyme−Product
722   Complexes in the Catalytic Mechanism of Glucoamylase from Aspergillus awamori.
723   *Biochemistry* **35**, 15269-15279 (1996).
724 39. Machado, T.F.G., Gloster, T.M. & da Silva, R.G. Linear Eyring Plots Conceal a Change
725   in the Rate-Limiting Step in an Enzyme Reaction. *Biochemistry* **57**, 6757-6761 (2018).
726 40. Yao, L., Li, Y., Wu, Y., Liu, A. & Yan, H. Product Release Is Rate-Limiting in the
727   Activation of the Prodrug 5-Fluorocytosine by Yeast Cytosine Deaminase. *Biochemistry*
728   **44**, 5940-5947 (2005).
729 41. Vaissier, V., Sharma, S.C., Schaettle, K., Zhang, T. & Head-Gordon, T. Computational
730   Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminase.
731   *ACS Catalysis* **8**, 219-227 (2018).
732 42. Wankowicz S, F.J. Making sense of chaos: uncovering the mechanisms of
733   conformational entropy. *ChemRxiv* (2024).
734 43. Braun, M. et al. Computational design of highly active de novo
735   enzymes. *bioRxiv*, 2024.2008.2002.606416 (2024).
736 44. Lauko, A. et al. Computational design of serine hydrolases. *bioRxiv*,
737   2024.2008.2029.610411 (2024).
738 45. Crean, R.M., Gardner, J.M. & Kamerlin, S.C.L. Harnessing Conformational Plasticity to
739   Generate Designer Enzymes. *Journal of the American Chemical Society* **142**, 11324-
740   11342 (2020).
741 46. Duran, C., Kinateder, T., Hiefinger, C., Sterner, R. & Osuna, S. Altering Active-Site
742   Loop Dynamics Enhances Standalone Activity of the Tryptophan Synthase Alpha
743   Subunit. *ACS Catalysis*, 16986-16995 (2024).
744 47. Koepf, E.K., Petrassi, H.M., Sudol, M. & Kelly, J.W. WW: An isolated three-stranded
745   antiparallel beta-sheet domain that unfolds and refolds reversibly; evidence for a
746   structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded
747   state. *Protein science : a publication of the Protein Society* **8**, 841-853 (1999).
748 48. Swindells, J.F., Snyder, C.F., Hardy, R.C. & Golden, O.E. Viscosities of sucrose
749   solutions at various temperatures : tables of recalculated values. (Washington : For sale
750   by the Supt. of Docs., U.S. G.P.O., [1958], 1958).
751 49. Kabsch, W. XDS. *Acta Crystallographica Section D* **66**, 125-132 (2010).
752 50. Evans, P.R. & Murshudov, G.N. How good are my data and what is the resolution? *Acta*
753   *Crystallographica Section D* **69**, 1204-1214 (2013).
754 51. McCoy, A.J. et al. Phaser crystallographic software. *Journal of applied crystallography*
755   **40**, 658-674 (2007).
756 52. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and
757   electrons: recent developments in Phenix. *Acta crystallographica. Section D, Structural*
758   *biology* **75**, 861-877 (2019).
759 53. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot.
760   *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).
761 54. Burnley, B.T., Afonine, P.V., Adams, P.D. & Gros, P. Modelling dynamics in protein
762   crystal structures by ensemble refinement. *eLife* **1**, e00311 (2012).
763 55. Tian, C. et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained
764   against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory*
765   *and Computation* **16**, 528-552 (2020).

766 56. Lee, F.S., Anderson, A.G. & Olafson, B.D. Benchmarking TriadAb using targets from
767   the second antibody modeling assessment. *Protein Engineering, Design and Selection* **36**,
768   gzad013 (2023).
769 57. Izadi, S., Anandakrishnan, R. & Onufriev, A.V. Building Water Models: A Different
770   Approach. *The Journal of Physical Chemistry Letters* **5**, 3863-3871 (2014).
771 58. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for
772   Ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089-10092 (1993).
773 59. Roe, D.R. & Cheatham, T.E., III PTRAJ and CPPTRAJ: Software for Processing and
774   Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and*
775   *Computation* **9**, 3084-3095 (2013).

776