INTERFACE

rsif.royalsocietypublishing.org

Review

Cite this article: Petrović D, Risso VA, Lynn Kamerlin SC, Sanchez Ruiz JM. 2018 Conformational dynamics and enzyme evolution. *J. R. Soc. Interface* 20180330. http://dx.doi.org/10.1098/rsif.2018.0330

Received: 10 May 2018
Accepted: 27 June 2018

Subject Category:
Reviews

Subject Areas:
evolution, biotechnology, biophysics

Keywords:
enzyme evolution, conformational ensembles, conformational shifts, multifunctional enzymes, de novo enzymes, primordial enzymes

Authors for correspondence:
Shina Caroline Lynn Kamerlin
e-mail: kamerlin@icm.uu.se
Jose M. Sanchez Ruiz
e-mail: sanchezr@ugr.es

THE ROYAL SOCIETY
PUBLISHING

Conformational dynamics and enzyme evolution

Dušan Petrović[1], Valeria A. Risso[2], Shina Caroline Lynn Kamerlin[1] and Jose M. Sanchez Ruiz[2]

[1]Department of Chemistry, BMC, Uppsala University, Box 576, 751 23 Uppsala, Sweden
[2]Departamento de Quimica Fisica, Facultad de Ciencias, University of Granada, 18071 Granada, Spain

DP, 0000-0002-1834-7358

Enzymes are dynamic entities, and their dynamic properties are clearly linked to their biological function. It follows that dynamics ought to play an essential role in enzyme evolution. Indeed, a link between conformational diversity and the emergence of new enzyme functionalities has been recognized for many years. However, it is only recently that state-of-the-art computational and experimental approaches are revealing the crucial molecular details of this link. Specifically, evolutionary trajectories leading to functional optimization for a given host environment or to the emergence of a new function typically involve enriching catalytically competent conformations and/or the freezing out of non-competent conformations of an enzyme. In some cases, these evolutionary changes are achieved through distant mutations that shift the protein ensemble towards productive conformations. Multifunctional intermediates in evolutionary trajectories are probably multi-conformational, i.e. able to switch between different overall conformations, each competent for a given function. Conformational diversity can assist the emergence of a completely new active site through a single mutation by facilitating transition-state binding. We propose that this mechanism may have played a role in the emergence of enzymes at the primordial, progenote stage, where it was plausibly promoted by high environmental temperatures and the possibility of additional phenotypic mutations.

1. Introduction

From local bond vibrations to global conformational motions, enzymes are dynamic entities, and their dynamical properties are clearly linked to their biological function [1]. The functional roles of such conformational changes include, but are not limited to, the allosteric regulation of enzyme function [2,3], motion necessary to access catalytically competent conformations [4], order–disorder transitions that can be necessary to facilitate efficient chemistry [5,6], and, in the case of catalytically promiscuous enzymes, conformational changes that allow for the catalysis of multiple reactions in the same enzyme [7,8]. The extent to which such functionally important conformational dynamics play a role in promoting enzyme *catalysis* has been the topic of vigorous debate [9–19]. Even more cryptic is the extent to which conformational diversity plays a role in allowing for enzyme *evolvability* [20,21], either through the repurposing of existing active sites, or through the emergence of completely new active sites in old enzymes. The plausible role of conformational diversity in allowing for functional plasticity was first put forward by James & Tawfik (figure 1) [21], who argued that conformational fluctuations such as, for example, side chain or loop dynamics, can lead one sequence to adopt multiple structures and multiple functions, some of which can interact with promiscuous ligands. These conformations might be rare in the conformational ensemble of the wild-type enzyme, but mutations can gradually shift the balance of populations such that any of these alternative conformations becomes the dominant one in the evolved enzymes, leading to a population shift towards favouring a new

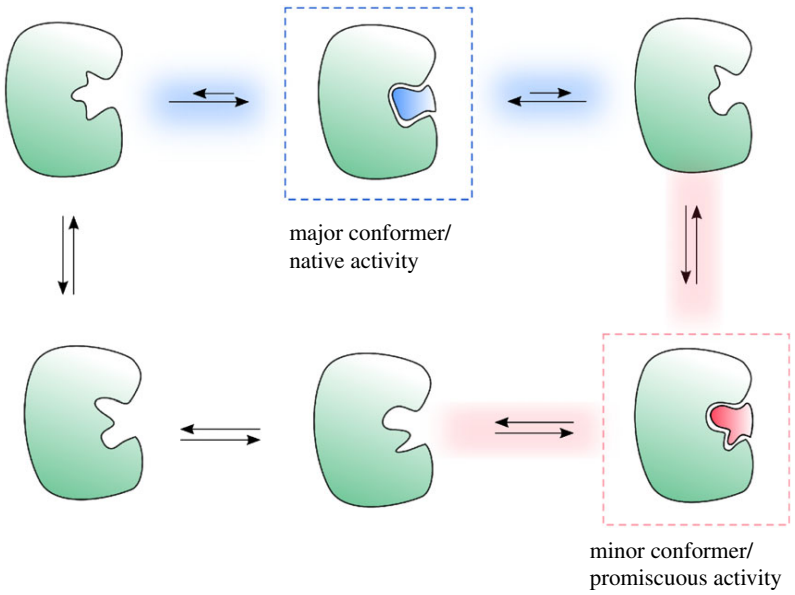


Figure 1. Illustrating the relationship between conformational dynamics and protein evolvability [20,21]. Proteins are conformationally dynamic, and can sample multiple conformations, including a major conformer (the 'native' state) that interacts with the ligand (blue), as well as minor conformers that can interact with promiscuous ligands and confer new activities to the enzyme. Conformational fluctuations, which can be both local or global in nature, lead to shifts between these different conformational substates. The alternative 'promiscuous' states may only be rarely sampled in the wild-type enzyme; however, the incorporation of mutations can gradually shift the equilibrium between these different states, until any of these alternate conformations becomes the dominant conformation in the evolved enzymes, which leads to a shift in activity. This figure, which is adapted from Tokuriki & Tawfik [20], was originally published in [19], and is reproduced with permission from Pabis *et al*. [19] and Tokuriki & Tawfik [20]. (Online version in colour.)

activity. In such a way, enzymes can vastly expand the functional diversity of a comparatively limited repertoire of sequences, thus allowing for completely new enzyme functions to emerge on pre-existing scaffolds [20,21].

From a semantic perspective, it is important to emphasize that these arguments are distinct and separate from debates about the role of enzyme dynamics in promoting the catalytic activity of a given enzyme, which focus on the feasibility of energy transfer from the conformational to the chemical coordinates, as well as the extent to which non-equilibrium effects can play a role in driving the chemistry [12,14,17]. In particular, as discussed in [19], protein flexibility and dynamics are often discussed in terms of the different time scales associated with conformational motions [22,23]. These time-scales range from fast pico- to nanosecond motions that reflect local conformational fluctuations, as well as slower micro- to millisecond (or even longer) motions that reflect global conformational changes. From an evolutionary perspective, motions in any time scale can play a role in allowing for new functions, as a native protein can be seen as an equilibrium ensemble of conformations that are to some extent related to each other, with mutations perturbing the equilibrium between these different conformations (see also [20,21]). As the conformational dynamics can change over an evolutionary trajectory, these population shifts can be 'lost' in the fully evolved enzymes. Hence, a mechanism of functional evolution based on the idea of conformational flexibility/diversity is neither consistent nor inconsistent with the idea of a 'rigid' evolved protein that populates several closely related conformations capable of efficiently catalysing the new function in an electrostatically pre-organized active site, and therefore these issues need to be addressed separately. In addition, even in a fully evolved enzyme, conformational pre-adaptation need not be complete, and as we will illustrate in this work, evolved enzymes may still retain sufficient flexibility to allow for local cooperative rearrangements in response to different substrates, which in turn would facilitate new enzyme functions.

Although still a young field, there have been several substantial advances because James and Tawfik initially presented their 'avant-garde' view of conformational diversity and enzyme evolution [21], focusing both on the role of conformational dynamics in the evolution of new enzyme functions [24–27], as well as how modulating an enzyme's dynamical properties allowed existing functions to emerge on previously non-catalytic scaffolds [28,29]. There is also increasing evidence that futile encounters and enzyme floppiness have significant impact on modulating an enzyme's reaction rate [30]. This review will focus on two issues: (1) the role of conformational dynamics in enzyme evolution, and (2) the role of conformational dynamics in allowing new functions to emerge either entirely de novo or through the repurposing of previously existing binding or enzyme functionalities. Taken together, a better understanding of conformational diversity and enzyme evolution is critical not just for advancing fundamental biochemical insights, but also for allowing for the more efficient design of novel enzymes with tailored physico-chemical properties.

2. The role of conformational dynamics/diversity in enzyme evolution

The role of conformational dynamics in enzyme catalytic cycles has been studied over many years using both computational and experimental approaches. Recent work on some selected protein systems has more specifically addressed the role of dynamics in enzyme evolution. As discussed below, these studies are based to a substantial extent on the comparison between homologous proteins from different organisms,

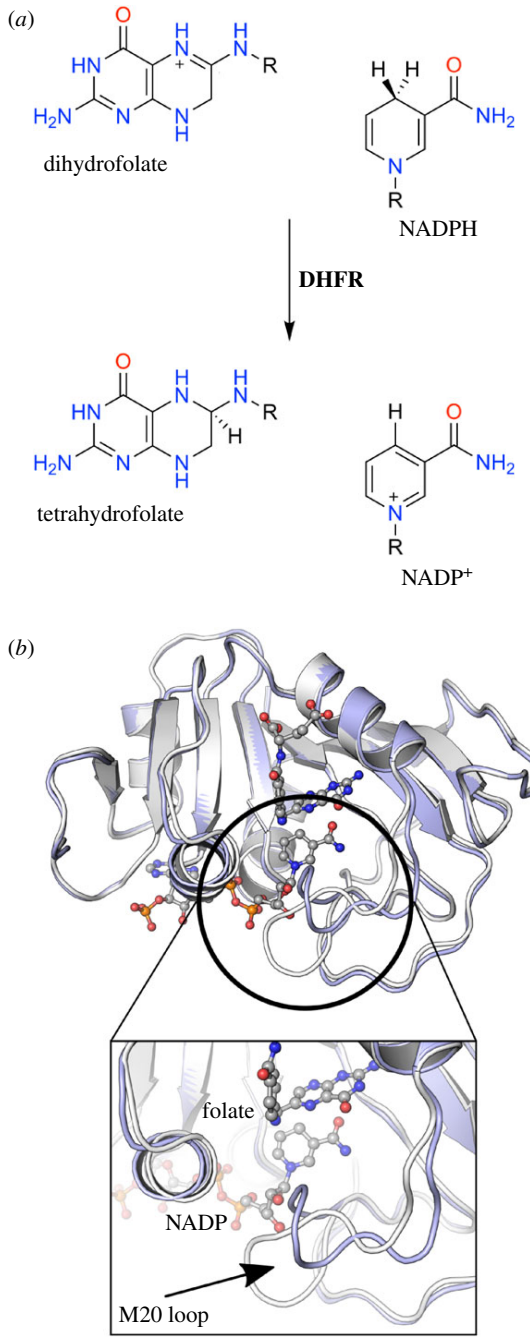


Figure 2. Dihydrofolate reductase (DHFR). (*a*) The DHFR-catalysed hydride transfer reaction. (*b*) The structure of *E. coli* DHFR in complex with folate and NADP cofactor, highlighting the mobile Met20 loop. The closed and occluded (PDB [31] IDs: 1RX2 and 1RX4, respectively [32]) conformations are shown in grey and blue, respectively.

and the analysis of evolutionary trajectories derived either from laboratory evolution experiments, or inferred from ancestral sequence reconstruction.

2.1. Dihydrofolate reductase

Dihydrofolate reductase (DHFR) catalyses the NADPH-assisted conversion of dihydrofolate (DHF) to tetrahydrofolate (THF) via the hydride transfer reaction shown in figure 2*a* [33]. In doing so, it fulfils several important functions *in vivo* [34], involving both converting nutritional folates into tetrahdyrofolates, and recycling DHF (which is in turn facilitated by the catalytic action of thymidylate synthase, the enzyme responsible for generating the nucleotide thymidine). This enzyme is typically monomeric and approximately 20 kDa in molecular weight [34], although, depending on organism, it can also take other oligomerization states [35]. It possesses a catalytically important mobile loop, the Met20 loop (figure 2), that closes over the active site during the chemical step of catalysis [34]. The comparably small system size of DHFR makes this enzyme easily tractable to both experimental and computational studies, and the unusual temperature dependence of the kinetic isotope effects for the hydride transfer reaction catalysed by this enzyme [36,37] have made it an important model system for studying tunnelling and dynamical effects in enzyme catalysis [24,38–43].

In recent years, research focus has also shifted to understanding the role of conformational dynamics in the evolution of DHFR activity in enzymes from different organisms. Specifically, while the human (*h*DHFR) and *E. coli* (*Ec*DHFR) enzymes are structurally highly similar, they have significant differences in their sequences, and the associated reaction kinetics and rate-limiting steps under physiological conditions [44–46]. Wright and co-workers studied these enzymes using an integrative approach that combined structural biology, mutagenesis, bioinformatics and cell biology techniques, in order to explore the evolutionary implications of the conformational dynamics of the different enzymes in the DHFR enzyme family [42]. In the case of the well-studied *Ec*DHFR, it is known that the enzyme undergoes multiple conformational states of the Met20 loop during the catalytic cycle, corresponding to open, closed and occluded states [32]. For comparison, the authors solved crystal structures of *h*DHFR at different stages in the catalytic cycle, and demonstrated that, in contrast with *Ec*DHFR, the corresponding loop in *h*DHFR remains locked in a closed conformation throughout the catalytic cycle [42]. In addition, the authors demonstrated that the active site cleft of *h*DHFR is more compact than that of *Ec*DHFR when bound to the same ligands, leading the authors to suggest that *h*DHFR is better pre-organized to facilitate the hydride transfer reaction catalysed by this enzyme. Finally, unlike *Ec*DHFR, *h*DHFR did not show conformational fluctuations on the millisecond time scale, suggesting that it uses either different motions or a different mechanism to facilitate ligand flux, compared to that used by *Ec*DHFR. However, ^{15}N relaxation dispersion experiments demonstrated the presence of pervasive microsecond timescale motions in many regions of the human enzyme (at rates ranging from approx. 15 000 to 30 000 s^{-1}), including regions that line one edge of the active site, and thus may play a role in ligand binding and release. In addition, *h*DHFR was demonstrated to show much larger motions in the hinges of the active-site cleft that most probably opens to accommodate substrate than the corresponding motions in *Ec*DHFR. Through sequence comparison of all (at the time) available DHFR sequences, the authors were able to link three clear regions that were linked to the flexibility. Finally, the authors' analysis showed that despite the overall structural similarity between human and *E. coli* DHFRs, the two enzymes show highly divergent dynamical properties, caused by evolutionary fine-tuning that can be plausibly linked to why *h*DHFR is unable to function effectively in the environment of an *E. coli* cell [42].

Kohen and Klinman have also performed a detailed analysis of co-evolving residues in DHFR [24], as a model system with which to probe the evolutionary aspects of enzyme dynamics. Specifically, genetic analysis of the *fol*A gene, which codes for DHFR, has allowed for a series of residues, spanning across the entire protein, and distant from each other in both sequence and physical space, to be identified as having co-evolved (i.e. they depended on each other during evolution) [47]. Separate computer simulations have suggested that these residues are also dynamically coupled to each other [48–52]. In principle, these results would indicate strong evolutionary pressure to preserve protein dynamics related to the catalytic step of the reaction. To test this hypothesis, both single and double mutants of these co-evolving residues have been generated. These were then studied using a broad range of kinetic methods, including measurements of steady-state and pre-steady-state rate constants (from both single turnover and burst experiments), as well as both kinetic isotope effects (KIE_{obs}) and intrinsic kinetic isotope effects (KIE_{int}), and the temperature dependence of these kinetic isotope effects [36,37,53,54]. These experiments indicated strong synergy in the effects of mutations at Met42, Gly121 and Phe125 on the C-H → C hydride transfer reaction catalysed by this enzyme [24]. However, in-depth analysis of the KIE suggested a more complex picture, in which there existed indeed a functional network of dynamics coupled to the chemical step catalysed by DHFR, and that coupled residues with a functional role are also co-evolving, but that not all co-evolving residues necessarily affect the chemical step (they may have co-evolved to maintain other functional properties such as folding and solubility).

However, we note that, in a contrasting view, based on extensive experimental work involving protein labelling studies, Alleman and co-workers have argued that dynamic coupling is in fact detrimental to catalysis by DHFR, and by the mesophile *Ec*DHFR [55]. By comparing DHFRs from organisms that have adapted to survive at a wide range of temperatures, the authors concluded that the dynamical coupling has been minimized during evolution and is rather a consequence of reorganizational motions that are necessary to facilitate charge transfer effects. More recently, by comparing the DHFRs from the moderate thermophile *Geobacillus stearothermophilius* (*Gs*DHFR) and the cold-adapted *Moritella profunda* (*Mp*DHFR), using both protein labelling and hybrid QM/MM studies, the authors argued instead for a need to minimize conformational dynamics by optimizing the free-energy surface of the reaction, in order to create a nearly static reaction-ready configuration of the enzyme, with optimal electrostatic properties for efficient catalysis of the hydride transfer reaction [56].

2.2. Glucose oxidase

Glucose oxidase (GOx) is an industrially important enzyme, commonly used in manufacturing glucose biosensors for diabetes control [57]. Further GOx applications are potentially for logical circuits [58] and biofuel cells [59]. Finally, GOx-based biofuel cells can in turn be engineered as small batteries to power implantable medical devices such as cardiac pacemakers, as was recently shown in several animal studies [60,61]. Owing to the many commercial applications of this enzyme, numerous directed evolution studies were conducted to optimize features such as its catalytic activity, stability or even oxygen independence [62–64].

We recently combined structural and computational work to investigate one particular directed evolution trajectory, which increased the catalytic activity of GOx by 2.6-fold [25]. Kinetic experiments and simulations of the GOx from *Aspergillus niger* (*An*GOx) have suggested the role of a key active-site residue, His516, as a catalytic base (figure 3*a*) [65,66]. However, analysis of the available crystal structures indicated that this residue is in fact conformationally mobile, with often poorly defined electron density (see e.g. PDB ID: 1GAL [67] or 1CF3 [68]). Interestingly, a catalytically favourable conformation of the active-site histidine (henceforth denoted as the 'catalytic' conformation) was observed in a related glucose dehydrogenase from *Aspergillus flavus* (*Af*GDH), which was crystallized in complex with D-glucono-1,5-lactone product [69]. However, in the structure of the GOx from *Penicillium amagasakiense* (*Pa*GOx), His516 is observed in a different conformation [68], in which its position geometrically prevents proton abstraction from the substrate (henceforth denoted as the 'non-catalytic' conformation; figure 3*a*). Using enhanced-sampling Hamiltonian replica exchange (HREX) and umbrella sampling (US) molecular dynamics (MD) simulations, we were able to track changes between the populations of the catalytic and non-catalytic conformations over the evolutionary trajectory [25] (figure 3*b*). Our simulations indicated that while in the wild-type *An*GOx both His516 conformations are almost equally populated, the directed evolution experiments significantly enriched the catalytic conformation of His516, while making the non-catalytic conformation energetically unfavourable.

Furthermore, we observed a general trend of increasing coupling between conformational motions over the evolutionary trajectory. We note that the individual residues can be either correlated (moving in the same direction), anti-correlated (moving in opposite directions) or non-correlated (motion not linked between them). Our simulations show that during the laboratory evolution of *An*GOx, there was a gradual increase in both the number of residues performing (anti)correlated motions, as well as the strength of (anti)correlation between these residues (figure 3*c* and *d*). Interestingly, most of residues showing increased anti-correlation were located on the secondary structures in the active site, and in close proximity to the mutated residues during the laboratory evolution. Such motions could represent the breathing motions of the active site, through the synchronized motion of different secondary structure features, as well as of individual residues working together on stabilizing the substrate in the active site.

2.3. 2-deoxyribose-5-phosphate aldolase

Coupled motions were also shown to be crucial for 2-deoxyribose-5-phosphate aldolase (DERA) activity [70]. DERA is a type I aldolase with main activity in the pentose phosphate pathway [71], catalysing C–C bond formation between acetaldehyde and glyceraldehyde-3-phosphate (G3P), as well as the reversed retro-aldol cleavage of 2-deoxyribose-5-phosphate (dR5P) [72,73]. *Escherichia coli* DERA (*Ec*DERA) uses a $(\beta\alpha)_8$ triosephosphate isomerase (TIM) fold, where the substrate forms a Schiff base with Lys167, and is among the most efficient aldolases [74]. Its ability to catalyse the aldol condensation reaction made DERA a popular enzyme for organic

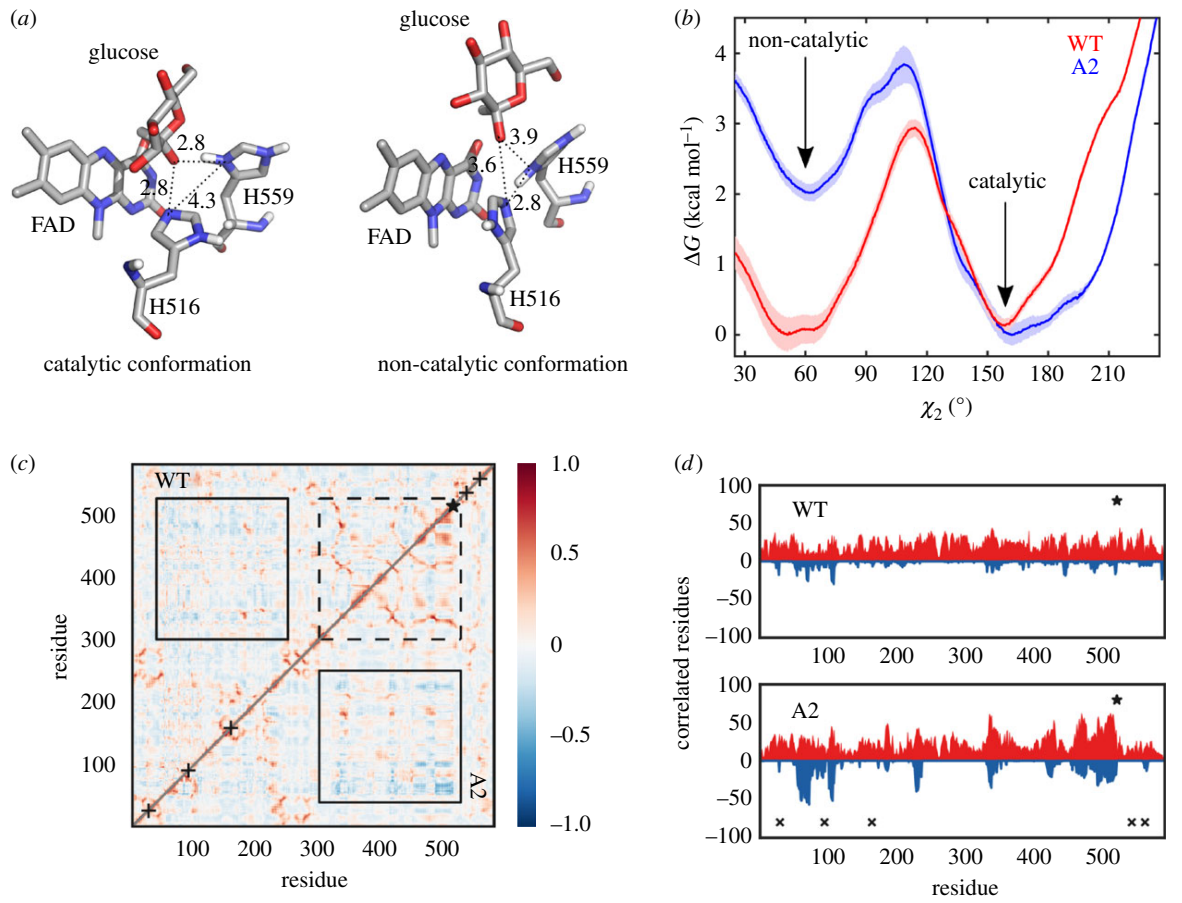


Figure 3. The role of conformational diversity in catalysis by glucose oxidase (GOx). (*a*) A comparison of the catalytic versus non-catalytic conformations of His516. (*b*) The relative free energies of the two His516 states in the wild-type enzyme and an engineered variant, A2, as a measure of the free energy for rotation around the χ_2 dihedral angle of His516. The free energy profiles were obtained from umbrella sampling calculations, as described in [25]. (*c*) A comparison of dynamic cross-correlation maps (DCCM) for wild-type GOx and the A2 variant, with the regions with the highest discrimination highlighted with rectangles. (*d*) The cumulative data from the DCCM plots shown on a per-residue basis, with the positions of key mutations highlighted with X, and that of His516 with ★. The figure is modified with permission from figures originally published in [25].

synthesis [75]. However, the absolute requirement of DERA for phosphorylated substrates limits its substrate scope, and therefore it is interesting to see if it is possible to use enzyme engineering to break the absolute dependence of this enzyme on phosphorylated substrates.

The phosphate moiety of the phosphorylated substrate binds near the N-terminus of α-helix 8, carrying a partially positive charge due to a helix dipole effect [76,77], at an atypical phosphate-binding site including the residues Ser238 and Ser239 (figure 4*a*). To investigate how important the phosphate-binding site actually *is* for the catalytic activity of the enzyme, we took two approaches: (1) replacing the two serine residues both individually and simultaneously by proline, and (2) introducing another 11 substitutions at each position using the NDT codon set. In the former case, our objective with the replacements of Ser by Pro was to both slightly adjust the position of the peptide backbone, which would in turn affect both side-chain and backbone interactions with the phosphate group of the substrate, as well as using this backbone shift in order to slightly move the N-terminal end of the associated helix, α-helix 8. Shifting this helix would, in turn, decrease the proposed additional contribution to the binding of the phosphate group by the positive dipole moment of the helix [70]. We then analysed the retro-aldolase activity of the resulting S238P, S239P and S238P/S239P mutants towards dR5P, as well as assessing the temperature dependence of the wild-type and S239P variants of this enzyme [70].

From our experimental data, we observed that, in the case of the S239P mutation, there was a negligible impact on k_{cat}, but a 100-fold increase in K_M. However, substitutions at position S238, in the case of the S238P and S238P/S239P variants, completely inactivated the enzyme (k_{cat}/K_M values of less than 0.1 $s^{-1}\,M^{-1}$). Interestingly, however, having a serine at this position appeared to not be an absolute necessity, as some activity was maintained in an S238I/S239I double mutant, albeit at 100-fold reduced catalytic efficiency (k_{cat}/K_M) [70]. To explore the origins of this effect, we performed molecular dynamics simulations, once again observing a qualitative correlation between the loss of coupled motions in the different engineered variants, and the experimentally observed changes in catalytic activity (figure 4*b*). Analysis of the overall flexibility of the enzyme in our simulations also indicated that even the single point substitution of the enzyme appeared to rigidify the overall dynamical behaviour of the enzyme, an observation supported by our analysis of the temperature dependence of the wild-type and S239P variants, which showed a strong impact on the entropic contribution to the thermodynamic parameters upon inserting a proline at position 239 [70].

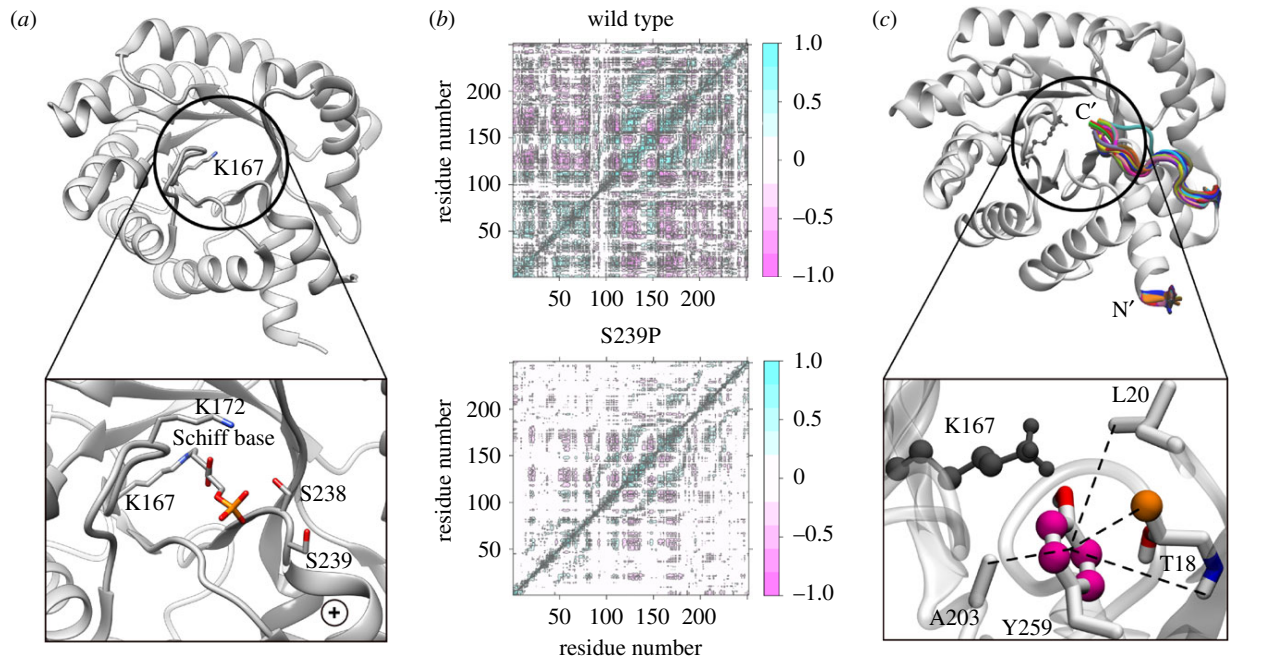


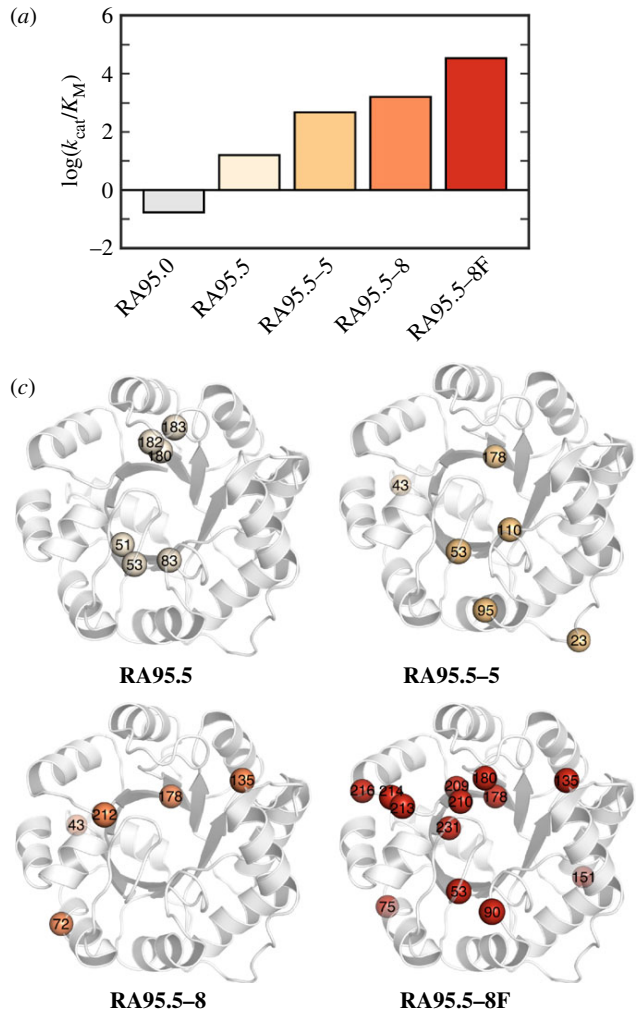
Figure 4. The role of correlated motions and conformational sampling in catalysis by 2-deoxyribose-5-phosphate aldolase (DERA). (*a*) An overview of the overall tertiary structure and the phosphate-binding site (PDB ID: 1JCJ) [74], with the position of the Schiff base K167 and the N-terminus of α-helix 8 (denoted by '+') highlighted. (*b*) Dynamic cross-correlation maps (DCCM) calculated from molecular dynamics simulations of wild-type DERA and the catalytically impaired S239P variant. (*c*) The intrinsically disordered C-terminal tail samples the closed-state conformation where Y259 protrudes into the active site, as supported by the nuclear Overhauser effect signals (dashed lines). (*a–b*) Modified from Ma *et al*. [70] with permission from the Royal Society of Chemistry. (*c*) Adapted with permission from Schulte *et al*. [78] Copyright © 2018 American Chemical Society.

In our subsequent work on DERA [78], we investigated the conformational sampling of the intrinsically disordered C-terminal tail. Specifically, further NMR experiments have shown that the same phosphate-binding motif is located on the intrinsically disordered C-terminal tail of DERA, which has not been resolved in any of the currently available crystal structures due to its high flexibility [78]. Notably, mutating the last C-terminal residue, Tyr259, into a phenylalanine, has an adverse effect on catalysis, reducing the turnover number by two orders of magnitude [74,78]. It was, therefore, assumed that the C-terminal tail could protrude into the active site (figure 4*c*), where it would have a direct impact on the chemical step. A combined NMR and MD study provided the first conclusive evidence that the C-terminal tail is, in fact, in an equilibrium between open and closed states, where the open states are quite diverse in conformation and much more populated than the closed state, where Y259 enters the active site and coordinates the phosphate moiety [78]. Furthermore, the nuclear overhauser effect (NOE) signals, used to solve the structure of the closed state, agree very well with the contacts observed in the closed state obtained from HREX-MD simulations. In such a state, Tyr259 is located close enough to the Schiff-base-forming Lys167 residue and could potentially serve in either stabilizing the transition state structure, or actively participating in the proton abstraction step.

2.4. Retro-aldolases

The importance of conformational dynamics for the evolution of enzyme activity was also showcased using MD simulations of a range of engineered retro-aldolases (RAs) [26]. Recently, several RAs were designed de novo to break the C–C bond in unnatural substrate 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone (methodol) [79,80]. In the design process, several possible catalytic mechanisms were considered, and the necessary catalytic site was designed based on the proposed mechanistic requirements and ported into several protein scaffolds. The most active designs had k_{cat}/k_{uncat} values on the orders up to 10^5 [81]. In the case of RA95, further directed evolution of the original design increased the activity by another 6 orders of magnitude over 19 rounds of directed evolution, and with the mutations at 23 different amino acid positions (figure 5) [82,83]. Osuna and co-workers studied the evolutionary trajectory of RA95, where many new mutations were located far away from the active site, yet they still had positive impact on the catalytic turnover [26].

Introducing the first six mutations into the scaffold (i.e. the so-called RA95.5 mutant) increased the catalytic efficiency more than 90 times (figure 5), and introduced a Lys residue at position 83, which replaced the *in silico* designed Lys210 as the Schiff base. After six more mutations (i.e. RA95.5–5 variant), the Schiff-base intermediate formed with Lys83 is not perfectly positioned for catalysis, which remains the case even after the insertion of five more mutations (i.e. RA95.5-8 variant, 13 rounds of directed evolution), where the catalytic efficiency of the evolved enzyme is 100 times higher than in the RA95.5 variant [82]. Finally, after yet another 13 mutations (i.e. 23 mutations in total compared to the starting point) and 6 additional rounds of directed evolution, the RA95.5-8F variant had a suitably pre-organized active site for catalysing the retro-aldol reaction, and with a k_{cat} of 10.8 s^{-1}, it reaches the catalytic proficiency of natural (retro)-aldolases such as DERA [83]. Interestingly, in parallel with the evolution of the active site in RA95, it was observed that less than half of



(b)

RA95.5	RA95.5–5	RA95.5–8	RA95.5–8F
	R23H		
	R43S	S43R	
V51Y			
E53S	S53T		T53L
		S72Y	
			R75P
T83K			
			N90D
	T95M		
	S110N		
		K135N	N135E
			S151G
	G178S	S178V	V178T
M180F			F180Y
R182M			
D183N			
			A209P
			K210L
		G212D	
			I213F
			S214F
			R216P
			L231M

Figure 5. Directed evolution of the *in silico* designed retro-aldolase RA95. (*a*) The catalytic efficiencies of various RA95 variants with (*R*)-methodol. The data in all panels in this figure were compiled from Giger *et al*. [82] and Obexer *et al*. [83]. (*b*) Mutations introduced in succeeding rounds of directed evolution of RA95. (*c*) Positions of the introduced mutations in each round of directed evolution, projected on the crystal structure of RA95 (PDB ID: 4A29) [82].

the 23 mutations introduced during the directed evolution were located at positions that were distant from the active site, indicating that their impact is rather to influence the overall evolution of conformational dynamics of the enzyme. In particular, the conformational dynamics of flexible loops on the catalytic face of the TIM-barrel fold were affected by mutations, as shown by MD simulations [26]. Over the evolutionary trajectory of RA95, these loops were observed to become substantially more rigid, while the overall enzymes themselves gained increased thermostability [26,83]. Simultaneously, while the original RA95 enzyme barely sampled the catalytically favourable conformations of the active site, the highly evolved variants were mostly pre-organized, sampling almost exclusively the catalytically favourable states of the enzyme [26].

2.5. β-lactamases

Conformational dynamics appears to play an important role in the evolution and enhancement of β-lactamase activity in both modern and ancient β-lactamases [84]. These enzymes have been the focus of substantial experimental and computational research effort due to their ability to break down β-lactam antibiotics, and the associated implications for the development of antibiotic resistance [85–95]. In particular, as global structural features appear to be conserved over large spans of evolutionary time [84], it is non-trivial to use simply structural comparisons to rationalize the origins of the large observed changes in activity. Bowman and co-workers have used Markov state models (MSMs) [96,97] in order to construct maps of the conformational ensembles adopted by variants of the modern TEM β-lactamase [98], in order to explore the extent to which the conformational diversity of this enzyme can play a role in modulating activity. These MSMs allowed for the identification of 'hidden' conformational states that appear to be functionally relevant in determining specificity but are not visible from static crystal structures. This was coupled with 'Boltzmann docking', an ensemble docking approach that can approximate the relative binding affinity of a compound by calculating the ensemble-average score for the binding affinity across a set of structural states that have each been weighted by their equilibrium probability. The simulations were used to target TEM's ability to degrade the cephalosporin antibiotic cefotaxime, and were then combined with both *in vitro* and *in vivo* experimental validation of the activities of novel TEM variants selected for their ability to stabilize different hidden 'cefotaxime states' (i.e. conformational states that are favourable for the degradation of cefotaxime). The combined experimental and computational work suggested

a correlation between the increased cefotaxime-degrading ability of the variants and the increased rigidity of the Ω-loop of TEM in the designed variants [98].

Similarly, Cortina and Kasson extracted positional mutation information from multiple microsecond time-scale molecular dynamics simulations in order to predict putative coupling between residues in another modern β-lactamase, CTX-M9, and the catalytic activity of the enzyme [99]. The simulations indicated that the motions of the bound drug and the motions of the protein as a whole are relatively decoupled. However, the use of an excess mutual information metric allowed the authors to predict 31 residues that appear to be important for the degradation of cefotaxime by this enzyme. Mutations at nine of these residues had been subjected to prior experimental testing, and a further 16 mutations (eight high-scoring mutations and eight low--scoring controls) were tested in bacteria alongside the computational work [99]. Of these, it was demonstrated that all nine previously tested mutants, and six out of the eight new prospective mutations, were able to increase the cefotaxime resistance by greater than twofold. This indicates that conformationally coupled residues can be used to modulate enzymatic activity, in this case as applied to understanding drug resistance. In a follow-up study of CTX-M9 [100], Kasson and co-workers used molecular dynamics simulations to predict allosteric mutants that increase the ability of CTX-M9 to degrade antibiotics (and thus its drug resistance), followed by experimental testing of the top computationally predicted mutants across multiple antibiotics. This was then combined with structural work, which showed that, interestingly, despite clear changes in catalytic rate and efficiency upon mutations, very little structural changes are observed in the mutant crystal structures, pointing at changes in the conformational ensemble of residues stabilizing the acyl-enzyme intermediate form during the hydrolysis of the antibiotics playing a role in mediating the observed changes in activity. The postulated changes were then explored using machine-learning analysis to identify the key changes to the conformational ensemble of the binding pocket during the hydrolysis step.

While these studies have focused primarily on modern β-lactamases, a computational study by Ozkan, Sanchez-Ruiz and co-workers focused on analysing the evolution of conformational dynamics in the transition from ancestral to modern β-lactamases [101]. Specifically, this study compared resurrected proteins corresponding to 2- to 3-Gy-old Precambrian nodes in the evolution of Class A β-lactamases, to the modern TEM-1 lactamase, which is known to be a comparatively rigid enzyme. It should be noted that the ancestral lactamases were experimentally characterized and demonstrated to be able to promiscuously catalyse the degradation of a range of antibiotics with activity levels that are similar to those of the average modern enzyme [84]. A combination of molecular dynamics simulations [101], with analysis through the 'Dynamic Flexibility Index' [102] (an approach which allows for the quantification of the contribution of each position in a protein to functionally related dynamics) showed clear changes in the conformational dynamics of the enzyme across evolutionary time. That is, while the active site of the modern lactamase was shown to be comparatively rigid, the ancestral lactamases showed much greater flexibility, in particular, in residues near the active site of the enzyme. Further principal component analysis of the conformational dynamics of these enzymes demonstrated that the ancestral β-lactamases form a cluster that is distinct from the more rigid modern TEM-1 lactamase [101]. An unrelated study by Vila and co-workers [103] employed NMR spectroscopy in order to explore the intrinsic dynamic features of different variants of a metallo-β-lactamase, metallo-β-lactamase II (BcII). The authors focus on the wild-type enzymes as well as three variants with expanded substrate scope obtained during a directed evolution trajectory and demonstrated that the enzyme has optimized the micro-to-millisecond time-scale dynamics along the evolutionary trajectory, and that the effect of individual mutations on the dynamics is epistatic. Therefore, the authors demonstrate that such conformational dynamics is an evolvable trait, and thus that proteins endowed with more dynamic active sites are also probably more evolvable [103]. Taken together, these studies show a clear role for conformational dynamics and coupled motions in modulating the evolution of existing enzyme functions, and in §3 we will also discuss how the conformational dynamics of ancestral β-lactamases can be manipulated for the generation of a de novo active site capable of proficiently catalysing a non-natural reaction [104].

3. The role of conformational dynamics/diversity in the emergence of new enzymes

Our main focus up until this section has been on studies that help unravel the role of conformational dynamics in modulating and/or enhancing already existing activities in an enzyme. A related question to this is how conformational dynamics can modulate the acquisition of new activities in existing active or binding sites, for example, through the acquisition of additional promiscuous catalytic activities, or the emergence of completely new activities on scaffolds that were previously non-catalytic. These scenarios have different biophysical and evolutionary implications and are therefore discussed below in separate sections.

3.1. The emergence of new enzymes from previously existing enzyme functionalities

It is widely accepted that most modern enzyme activities have arisen from previously existing functionalities through evolutionary processes that probably involved gene duplication. In the simplest model [105,106], one copy of the gene maintains the 'old' functionality, while the other copy evolves towards the new functionality. At the molecular level, this would involve the evolutionary adaptation of an already existing enzyme active site for the new enzyme function with the generation of a multifunctional enzyme as a necessary intermediate step. In this context, the acquisition of promiscuous activities in enzymes that catalyse phosphate hydrolysis has been an important model system for historical reasons [107–109], as well as the fact that many phosphatases are catalytic promiscuous [107]. From a computational perspective, we have extensively studied the drivers for selectivity and the emergence of promiscuous activities both among members of the alkaline phosphatase superfamily [110–112], as well as among organophosphate hydrolases. These enzymes have evolved the ability to promiscuously hydrolyse organophosphate nerve agents and pesticides, and have usually evolved from enzymes that catalyse lactone

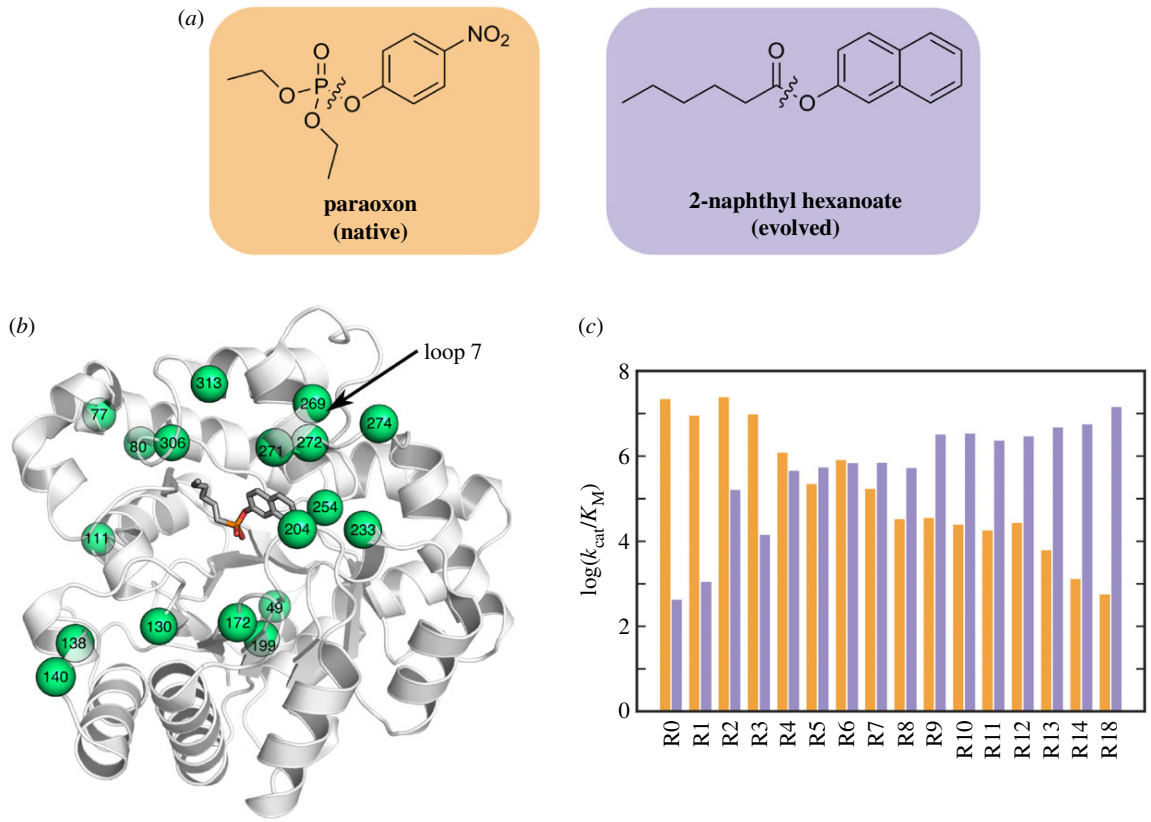


Figure 6. Directed evolution of an arylesterase. (*a*) The structures of the native (paraoxon, orange) and promiscuous (2-naphthyl hexanoate, blue) substrates. The wavy lines indicate the bond to be cleaved. (*b*) The structure of the evolved enzyme in complex with the transition state analogue (18 rounds of directed evolution, PDB ID: 4E3T [120]), with the positions of all 18 mutations denoted by the green spheres. The active-site loop 7, discussed further in the text, is also denoted. (*c*) The catalytic efficiencies of the native (orange) and promiscuous (blue) reactions over the course of the evolutionary trajectory [120].

hydrolysis or related reactions [113–115]. This work, which has been reviewed in detail in e.g. [116–119], points to the importance of both 'electrostatic flexibility' (in terms of the cooperative behaviour of the individual amino acid side chains), as well as harnessing conformational motions such as loop closure in order to facilitate the sequestration of solvent from the active site, as major drivers in allowing for these enzymes to gain new activities.

In most cases, however, the specific historical events that led to the emergence of modern enzymes are not known in detail. Here, laboratory evolution provides a valuable tool with which to obtain detailed molecular characterization of the emergence of a new enzyme function, because the intermediate protein states along the evolutionary trajectory will then be known, and can be subject to biophysical characterization. A particularly informative example of this approach is discussed below in some detail.

Tawfik and co-workers [120] have evolved an efficient arylesterase (capable of hydrolysing the C-O bond in 2-naphthyl hexanoate) from a naturally occurring phosphotriesterase (capable of hydrolysing the P–O bond in the pesticide paraoxon) that displayed a low, promiscuous level of arylesterase activity (figure 6). The first two rounds of directed evolution on this enzyme led to an increase of about 3 orders of magnitude in arylesterase activity, accompanied by a moderate decrease in the paraoxonase activity, thus leading to a multifunctional enzyme. In a hypothetical natural evolution scenario, in which both activities would contribute to organismal fitness, gene duplication at this stage would allow for the evolution of two different enzymes, each specialized in one function. In the laboratory evolution experiment, however, the selection for the new function in subsequent rounds increased the arylesterase activity in a more gradual manner and with a trade-off with the old function stronger than that observed in the first rounds, finally resulting in an efficient arylesterase with a low level of paraoxon degradation activity. This arylesterase included 18 mutations with respect to the phosphotriesterase background of which only four were located in the active site (figure 6*b*).

More recently, Tokuriki, Jackson and co-workers have performed structural characterization of several protein states along this directed evolution trajectory [28]. As a result, a clear picture of the role of dynamics in the evolution of the new function is available. In particular, this work very clearly highlights the importance of conformational dynamics in the evolution of new enzyme functions, and emphasizes that such changes in function can be achieved through enriching pre-existing conformational substates, as in the previously discussed example of GOx [25]. Specifically:

'Residues that are crucial for the emergence of a new catalytic function may initially appear as ensembles of several conformations of which many are non-productive. Substantial levels of activity may, therefore, require subsequent mutations to shift the ensemble towards productive conformations (and eliminate the non-productive ones)'. In the example of the evolution of a phosphotriesterase to an arylesterase, this is most clearly illustrated by the active-site residue at position 254. The H254R mutation appears very early in the laboratory evolution trajectory and provides a

critical interaction for the new activity, as the arginine at position 254 can form a cation-π interaction with the naphthyl leaving group. Still, X-ray crystallography showed that Arg254 can adopt two conformations, 'open' and 'bent'. Both conformations are compatible with the 'old' (phosphotriesterase) activity, but only the bent rotamer is catalytically productive for the new function, because the extended rotamer causes a steric clash with the arylester substrate. In fact, mutations that shifted the conformational ensemble of Arg254 towards the bent conformation did occur along the evolutionary trajectory. For example, a D233E mutation substantially enhanced the new function by stabilizing the bent conformation of Arg254 through the formation of a salt bridge.

'Some conformational fluctuations that are relevant for the old function may be irrelevant or even detrimental for the new function. Consequently, mutations that restrict the original fluctuations and eliminate unproductive conformations will occur along the evolutionary trajectory leading to the new function'. In the example of the evolution of a phosphotriesterase to an arylesterase, this is most clearly illustrated by the changing of the dynamics of the active-site loop, loop 7. This loop can, in principle, populate open and closed conformations. The closed conformation creates a catalytically competent cavity in which the reaction can take place. Still, the rate of original phosphotriesterase function is limited by-product release, which requires the opening of the loop. Therefore, fluctuation between the two conformations is essential (and rate-limiting) for the original activity [121]. However, it is not required for the new arylesterase activity. In fact, some of the mutations that accumulated over the evolutionary trajectory that leads to the arylesterase function (specifically L271F, L272M, F306I, I313F) generated hydrophobic contacts that stabilized the closed conformation of loop 7 [28].

'Mutations that are to varying degrees distant from the active site may play a role in the emergence of the new function'. It is clear from the two preceding paragraphs that the mere presence of suitable catalytic groups in the active site does not guarantee substantial levels of the new activity, because these groups may be sampling mostly conformations that are not catalytically competent. Additional mutations outside the active site may be required to 'freeze-out' unproductive conformations. These may be 'second-shell' mutations that introduce direct interactions with active-site residues in the 'correct' conformation, as in the examples discussed above, but they may also be mutations at more distant positions that shift the protein ensemble towards conformations that can interact favourably with the substrate and the transition state. Indeed, the fact that amino acid residues can promote catalysis through interaction with the groups directly involved in the catalytic mechanism has been known for some time [122], and more recent analyses have shown that active-site residues induce gradients of evolutionary conservation that extend far beyond the active-site region [123]. In addition, stabilizing mutations, which can occur in any part of the molecule, may be required to compensate the destabilizing effect of many functional active-site mutations, although this case would not be directly linked to the dynamics.

'Multifunctional intermediates involved in the emergence of a new function are likely to be multi-conformational'. Multifunctional ancestor enzymes were probably involved in the emergence of most modern enzyme functionalities [124], because this would avoid functional trade-offs and the concomitant detrimental effects if both the old and the new functions contribute to organismal fitness. Even when they can be accommodated in the same active site, two different functions will probably have incompatible molecular requirements in terms of active shape architecture, rotameric states of crucial residues etc. Therefore, the presence of two different enzyme functions in the same active site requires that the protein can switch between at least two different conformations, each competent for one of the functions. This is indeed observed in the bi-functional intermediates involved in the laboratory evolution of a phosphotriesterase into an arylesterase [28]. More generally, the fact that a protein molecule can switch between two different conformations should not come as a surprise given that a single mutation has been shown to completely change the overall three-dimensional structure for some designed sequences [125,126]. Such structure bistability may have played a relevant role in protein evolution [127,128].

It is clear overall, that, in many cases, conformational diversity is the key to the multifunctional characters of intermediate (ancestral) stages in the emergence of a new function. In the specific example of the phosphotriesterase to arylesterase evolution [28] discussed here, the relevant conformational diversity involved large fluctuations in side-chain and loop conformations, and was therefore apparent even at the level of the X-ray crystallography. In other cases, of course, subtler conformational effects may be involved that may not be trivial to detect without the help of extensive integrated biochemical, structural and dynamical analysis, using both computational tools as well as experimental tools such as NMR.

3.2. The emergence of new enzymes through the functionalization of binding sites

Molecular mechanisms for active-site refitting certainly exist, and are beginning to be understood owing to recent experimental and computational work. We have described these advances in the preceding sections of this review. However, a new enzyme can also emerge from the functionalization of a previously existing binding site, through an evolutionary process that does not involve a multifunctional enzyme ancestor. This mechanism is unlikely to be prevalent; that is, it is expected to be the exception, while the rule is expected to be the emergence of new enzymes from previously existing functionalities. Note that there is an enormous diversity of natural enzymes and that, at least in principle, all natural enzymes can evolve to (or be repurposed for) new functionalities. On the other hand, the number of proteins that bind small ligands and that are suitable as scaffolds for the generation of new functions (periplasmic binding proteins, for instance) appears to be more limited [129]. Still, very recent work (discussed below) on the evolutionary functionalization of binding sites provides important clues as to the role of conformational flexibility/diversity and epistasis in the emergence of new enzyme functions.

Jackson and co-workers [130] have used ancestral protein resurrection to show how the enzyme cyclohexadienyl dehydratase (CDT) evolved from a cationic amino acid-binding protein of the solute-binding protein superfamily. The authors demonstrated that the emergence and optimization

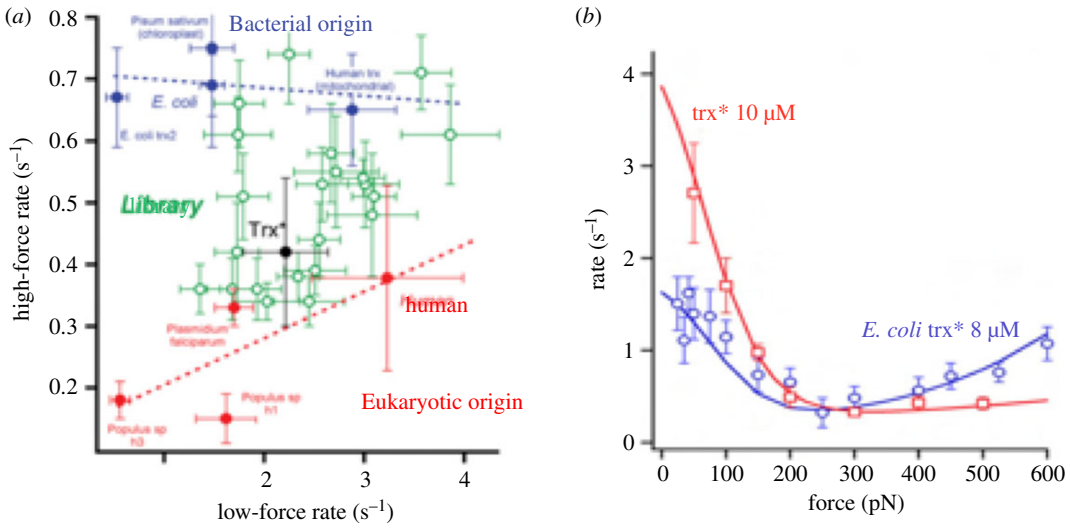


Figure 7. Atomic force microscopy mechanochemistry experiments on the reduction of disulphide groups catalysed by thioredoxin. (*a*) Rates of disulphide reduction determined at low and high applied force for several thioredoxins from bacterial (blue) and eukaryotic (red) origin. The plot also includes data (green) for 23 variants of *E. coli* thioredoxin, taken from a combinatorial library of conservative mutations derived from statistical analyses of the sequence alignment. These variants span the range from bacterial thioredoxin to eukaryotic thioredoxin behaviour, thus suggesting the existence of multiple evolutionary paths between the two regimes. (*b*) Q12 Plots of reduction rate versus applied force for *E. coli* thioredoxin and one of its variants (trx*). This variant includes 8 mutations and shows clear eukaryotic behaviour, with a depressed contribution of the chemical S_N2 mechanism prevalent in bacterial thioredoxins. This research was originally published in [133].

of the catalytic activity involved the incorporation of a desolvated general acid into the active site, which allowed for the functionally relevant chemistry to occur, as well as active-site reshaping, in order to facilitate complementarity between the enzyme and the substrate. Subsequent gains in catalytic activity were facilitated by the introduction of hydrogen-bonding networks that positioned the catalytic residues precisely in an electrostatically pre-organized environment, as well as remote substitutions that dampened the sampling of non-catalytic conformations. In fact, the modulation of conformational diversity along the evolutionary trajectory that leads to the new enzyme is essential in this case because: (i) solute-binding proteins rely on an equilibrium between closed and open conformations to regulate binding affinity and the rate of solute transport; (ii) enzyme catalysis relies on active site pre-organization and reduction of unproductive conformational sampling. Indeed, it is the closed conformation that is catalytically competent in CDT, while it is the open conformation that is favoured in the non-ligated forms of the reconstructed ancestral solute-binding proteins. Therefore, a large shift in the protein conformational ensemble is unescapably linked to the emergence of the new enzyme function.

In a related study [29], we combined ancestral protein resurrection, laboratory evolution, X-ray crystallography, NMR and MD simulations, in order to explore the evolutionary emergence of chalcone isomerase activity on a previously non-catalytic protein scaffold. This activity, which is key in plant flavonoid biosynthesis, was presumed to have evolved from a non-enzymatic ancestor of the fatty-acid binding proteins and a plant family devoid of isomerase activity. Remarkably, the binding site of the resurrected ancestors included the catalytic groups that are known to be responsible for the catalysis in the modern chalcone isomerases Q3 and yet these ancestral proteins were inactive. Emergence of activity along the evolutionary trajectories that led to the modern enzymes could be linked to subtle rearrangements of active-site residues with a concomitant substrate repositioning and to changes in the conformational ensemble of the catalytic arginine at position 34. Indeed, a consistent change in rotamer population was found to accompany the emergence of isomerase activity, as shown by NMR and MD simulations. As noted in refs. [29,130], these results have general implications for enzyme engineering as they highlight that placing catalytic residues in a cavity without taking into account dynamics is likely to lead to poor enzymes.

One of the observations we found most remarkable in [29] is that many founder mutations and subsequent trajectories can lead to the modern chalcone isomerase enzymes. This indicates a smooth evolutionary landscape and a limited impact from epistasis. This surprising result is in contrast with many previous studies that show strong epistasis in protein evolution. Still, it is not without precedent. Atomic force microscopy mechanochemistry experiments on the reduction of disulphide bridges catalysed by *E. coli* thioredoxin [131] have revealed two distinct catalytic mechanisms that are each favoured in different ranges of the applied force. They use the same catalytic residues and appear to differ mainly in the orientation of the substrate bound in the active site. Remarkably, the mechanism favoured at high force was shown to occur only in thioredoxins from bacteria [132], while it was found to be absent in eukaryotic thioredoxins, a difference that was linked to the evolution of the binding groove. Still, it was later found [133] that just a few mutations Q4 could cause *E. coli* thioredoxin to behave as a eukaryotic thioredoxin in terms of the force dependence of catalysis (i.e. low rate at high forces). Furthermore, the variants of a combinatorial library of a few conservative mutations derived from the statistical analysis of sequence alignments were found to fully span the range of thioredoxin mechanochemistry, from the bacterial behaviour to the eukaryotic behaviour (figure 7). This obviously suggests the existence of many evolutionary paths connecting the two regimes.

In the light of the studies described above, it appears that the evolution of a new function could occur with little epistasis when crucial active-site residues are already present, and that the emergence of new functions relies mostly on conformational shifts. This proposal is supported by the fact that the limited impact of epistasis can be explained in this case by a simple and, we believe, convincing mechanism. Recent work has shown that mutational effects on protein stability are evolutionarily conserved to a substantial extent [134,135]. Several plausible molecular mechanisms actually justify such conservation even among relatively distant homologues, as we have discussed in some detail [135]. It follows that the effect of a given mutation on stability may be, at least qualitatively, independent of the previous mutational background. Therefore, stability effects of mutations are likely to show limited epistasis except, of course, when the cumulative effect of destabilizing mutations violates the marginal stability threshold of the protein [136,137]. The essential point to note here is that evolutionarily shifts of protein conformational ensembles are mediated by mutational effects on the relative stabilities of the different conformations of the ensemble. It follows that, in some cases at least, limited epistasis could allow for the existence of many different mutational trajectories for these conformational transformations.

3.3. The de novo evolution of new enzyme functions

Unless we accept some form of panspermia as a plausible explanation for the origin of life on Earth, it would be clearly absurd to assume that all enzymes that have existed during the history of life in this planet descended from previously existing enzymes or binding functionalities. In fact, upon ruling out panspermia, it becomes inescapable to accept that completely new enzyme functionalities (i.e. completely new enzyme active sites) emerged de novo at some very early stage of evolution on Earth. Several estimates of the gene content in the last universal common ancestor (LUCA) support the theory that a wide diversity of complex enzymes already existed in the LUCA [138,139]. Consequently, the de novo emergence of completely new enzyme active sites must have occurred at an even earlier stage. An interesting (although certainly speculative) scenario assumes that amino acids and polypeptides originally served as cofactors of ribozymes in a primordial RNA world [140] and that, at some point, very early proteins developed, perhaps through the assembly of short polypeptides [141], and acquired the capability to catalyse 'useful' chemical reactions, eventually replacing the (less efficient) RNA enzymes as catalysts for most of the chemical reactions of life.

Regardless of the specific scenario, the presence of complex enzymes already in LUCA suggests that efficient mechanisms for the de novo emergence of protein-based enzymes must exist or, at least, that they must have existed at a primordial stage. One simple pathway towards a new enzyme functionality could be provided by the recruitment of a catalytic metal ion [142]. However, by no means are all enzymes metalloenzymes. Indeed, it is believed that 'only' about 30% of them are [143,144]. We must consider, therefore, the generation of new enzyme functionalities through mutations on non-catalytic protein scaffolds. Certainly, little is known about how these kinds of mutation-based evolutionary mechanisms occur *in vivo*, but we can perhaps derive some clues from the attempts of protein engineers to rationally design completely new active sites. Of particular relevance in this context are the several (successful) attempts at the design of de novo enzymes on the basis of the Rosetta suite of computational modelling reported by Baker's group [79,145–147]. We briefly describe some of these efforts below.

The Rosetta-based approach is best illustrated with the Kemp elimination of 5-nitrobenzioxazole to yield the corresponding *o*-cyanophenolate ion. This reaction is a simple (and activated) model of proton abstraction from carbon, a fundamental process in chemistry and biochemistry, and has been used as a benchmark for computational enzyme design. To design a Kemp-eliminase, Baker, Tawfik and colleagues [145] used (i) quantum mechanical calculations to create an idealized active site having an optimal interaction with the transition state; (ii) RosettaMatch to search for constellations of backbone positions capable of supporting the idealized active site in a set of 87 protein scaffolds; (iii) computational optimization and active site redesign. They experimentally tested 59 designs (requiring 10–20 amino acid replacements) and found detectable Kemp-eliminase activity in eight designs, with the best variant displaying a catalytic efficiency (k_{cat}/K_M) of 163 M^{-1} s^{-1}. This appears to be a rather low value when compared with the catalytic efficiency of a modern average enzyme which is around 10 000 M^{-1} s^{-1} [148].

The Rosetta design approach has been applied to other cases with similar results in terms of design efficiency. The Diels-Alder reaction is a concerted pericyclic reaction of fundamental importance in synthetic organic chemistry. To design a Diels-alderase [146], 207 scaffolds were considered as possible backgrounds for new active-site generation, 84 designs were selected for experimental characterization and only two of them (including 13 and 14 mutations) were reported to display Diels-Alderase activity. Likewise, in another work [147] 214 protein scaffolds were chosen as candidates to accommodate a new active site for ester hydrolysis, 55 designs were selected for experimental analysis and only four (including 9, 11, 13 and 20 mutations) were reported to show significant activity in the burst phase. For the design of a retro-aldolase enzyme, 72 designs with 8–20 amino acid changes in 10 different scaffolds were selected, and 32 of them showed detectable retro-aldolase activity [79].

The studies summarized above are certainly groundbreaking from a protein-engineering point of view, but raise also intriguing issues in an evolutionary context, because they seem to imply that the emergence of completely new enzyme functionalities (i.e. completely new active sites) is scarcely possible. This message is conveyed not only by the rather low design success rates reported, but also by the very large number of mutations required to achieve the new functionalities (within the 8–20 range). Natural selection can efficiently drive the amplification of an existing function through the accumulation of function-enhancing mutations, provided, of course, that this process confers a survival advantage. On the other hand, natural selection cannot act on a function before this function has emerged. Consequently, the evolutionary emergence of a new function becomes highly unlikely if it requires the simultaneous occurrence of a large number of mutations. In this case, the selective advantage would only appear when the complete set of function-related mutations is present and no survival-related factor would select the few mutational trajectories that lead to this set from the astronomically large number of all possible mutational trajectories. Clearly, therefore, the

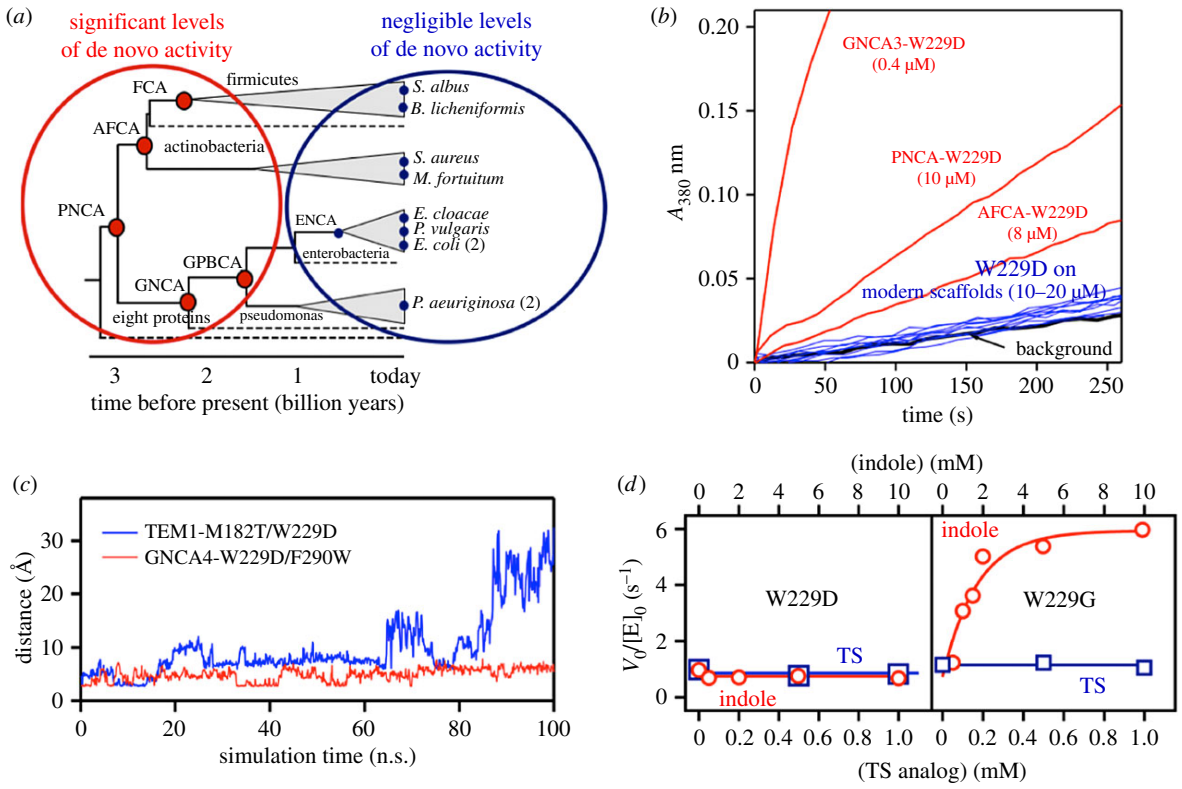


Figure 8. Using modern and resurrected ancestral β-lactamases as scaffolds for the engineering of new active sites. (*a*) Schematic of the phylogenetic tree used for ancestral β-lactamase reconstruction (see [104] for details). A minimalist design successfully generated a new active site for Kemp elimination in most of the ancestral nodes tested (red) but failed (blue) in all the modern scaffolds tested. This is clearly shown by the activity assays included in *b* (colour code is the same as in *a*; see [104] for further details). (*c*) Molecular dynamics simulations of engineered β-lactamases. The profiles of the distance between the catalytic group and a transition-state model initially placed in the active site indicate that the modern TEM-1 β-lactamase cannot stably bind the transition state over a 100 ns time scale. Indeed, experimental binding to the cavity created at the targeted 229 position is not observed in in the modern TEM-1 β-lactamase background (*d*) unless the ligand exactly matches the cavity (binding of indole to the cavity created by a W229G mutation). This indicates a link between de novo catalysis and flexibility that is further supported by NMR relaxation experiments and X-ray crystallography (see [104] for details). This figure was adapted from Risso *et al*. [104]. Reproduced here with permission from Risso *et al*. [104].

best scenario for the evolutionary emergence of a new enzyme functionality is that significant activity levels are generated through a single mutation, so that natural selection can immediately drive the enhancement of the new function. The interest in this possibility has led to several recent attempts to apply minimalist approaches to the design of de novo active sites. We describe some of these efforts below.

Korendovych, DeGrado and co-workers [149,150] have shown that the introduction of just one reactive residue in the hydrophobic cavity of the C-terminal domain of calmodulin leads to significant levels of catalysis of simple organic reactions, including Kemp elimination and ester hydrolysis. These studies took advantage of a pre-existing cavity to generate a new active site. It is possible, however, to generate both the cavity and the catalytic residue through a single mutation that replaces a partially buried hydrophobic residue with the reactive residue. This has been recently demonstrated by us [104] using β-lactamases, the enzymes responsible for the primary mechanism of resistance against lactam antibiotics, as scaffolds for engineering (figure 8). By substituting a partially buried tryptophan with an aspartic acid residue, a new active site capable of catalysing Kemp elimination and ester hydrolysis was generated. Interestingly, the new active site did not impair the antibiotic-degradation activities linked to the 'old', natural active site. Here, we actually tested a large number of different β-lactamases, including 10 different modern β-lactamases and 13 resurrected Precambrian β-lactamases. While the minimalist, single-mutation design was successful in most of the ancestral scaffolds used, it failed in all the modern β-lactamases tested. X-ray crystallography, NMR-relaxation studies, molecular dynamics simulations and experimental binding studies supported that the enhanced conformational flexibility of ancestral proteins played a crucial role in the generation of the new active site. This should not come as a surprise. Enzyme catalysis relies to some substantial extent on the stabilization of the transition state of the chemical reaction [151]. A single mutation can introduce a catalytic residue but it is unlikely to lead to a cavity tailored (i.e. pre-organized) for the transition-state shape. A flexible scaffold is therefore required to allow substrate and, most important, transition-state binding. Indeed, it is plausible that conformational flexibility was also a factor in the successful minimalist design of new activities on a calmodulin scaffold [149,150]. Calmodulin is in fact known to be conformationally flexible, as expected from its biological function, which involves its binding to a wide diversity of target enzymes [152]. We further elaborate below on the role of conformational flexibility in the emergence of new enzyme functions.

In a highly influential article, Roy Jensen proposed many years ago that primordial life probably used limited genetic

information that coded for a comparatively small number of proteins and, consequently, that primitive proteins were promiscuous generalists with broad functionalities [153]. Protein promiscuity requires the capability to bind a diversity of molecular targets (substrates and transition states) and it is, therefore, linked to the possibility of populating many different conformational states. Accordingly, conformational flexibility/diversity was probably a primordial protein trait that could have facilitated the de novo generation of enzyme activities in the first place. Hence, function generation through single mutations on flexible ancestral protein scaffolds becomes a plausible scenario for the emergence of the first enzymes. Some caveats must be noted, however. The studies summarized in the preceding paragraph targeted simple reactions for which significant rates can be obtained by introducing a single catalytic residue. Yet, substantial catalysis for many of the complex chemical reactions of life may require the cooperation of several active residues and perhaps cannot be achieved through a single mutation. This is more so when considering that many of these reactions are extremely slow in the absence of catalysts, displaying half-lives for the non-catalysed processes that can in some cases reach thousands and even millions of years [154]. It would seem that, for many chemical reactions, only enormous rate enhancements could have brought the reaction rates to primordial levels that could impact fitness and could be subject to subsequent enhancement through natural selection. This creates a kind of catch-22 situation, as it is the highly evolved enzymes that are expected to be able to produce very large increases in reaction rate. However, as we discuss below, the combination of high ancestral temperatures with the contribution of phenotypic mutations may provide an escape from the paradox.

Many plausible scenarios are consistent with ancient life being thermophilic, including, for instance, that primordial life thrived in hydrothermal vents [155] or that the ancient oceans that hosted life were hot, as supported by analyses of the isotopic composition of rocks [156]. Most chemical reactions have high activation energies and are strongly accelerated by increases in temperature. As shown by Wolfenden and co-workers, strong thermal acceleration of the uncatalysed chemical reactions of life does indeed occur [154]. Therefore, in a high-temperature primordial environment, even moderate levels of catalysis could bring reaction rates to a biologically relevant time scale. As a result, even the small rate enhancements brought about by 'poor' primordial enzymes may have been of adaptive value and may have served as starting points for subsequent evolutionary optimization through natural selection.

Errors during protein synthesis (linked to transcription or translation) occur frequently and lead to the so-called phenotypic mutations [157,158]. Phenotypic mutations are not inherited. Still, it has been proposed that they may play an evolutionary role through a kind of 'look ahead' effect [159]. That is, they may provide a crucial functional advantage that allows the organism to survive until a functionally useful mutation occurs at the genetic level. In the context of the de novo emergence of the new enzyme function, phenotypic mutations may alleviate the problems associated with the need to simultaneously have several residues to achieve catalysis. As a simple example, assume that two mutations are simultaneously required to generate enzyme catalysis for a relevant reaction. One of the mutations could occur at the genetic level, while the second could appear as a phenotypic mutation in some of the variants generated by protein synthesis errors. Accordingly, the rate of the reaction will actually increase, even if the first mutation alone does not generate catalysis. This will contribute to organismal fitness and to the spread of the first mutation through the population, thus providing a stepping stone for the second mutation to actually appear at the genetic level. This kind of evolutionary mechanism may have been particularly relevant at a very early evolutionary stage (sometimes referred to as the progenote stage [160]), when a precise relation between phenotype and genotype had not been yet established because of the low accuracy of transcription and translation.

In summary of this section, we hypothesize that the efficient emergence of new enzymes occurred at a very early stage of evolution on Earth, and was linked to the following factors: (i) Conformational flexibility of the (probably small) primordial proteins facilitated the binding of different substrates and transition states. (ii) High environmental temperatures enhanced the rates of the future biochemical reactions and brought their half-lives to a biologically relevant time scale, in such a way that even the moderate rate enhancements caused by poor primordial enzymes were of selective value. (iii) A single mutation at the genetic level could lead to a new function, either by itself or with the help of phenotypic mutations, the occurrence of which was greatly favoured at the early progenote evolutionary stage.

4. Conclusion

In this contribution, we have reviewed a large number of experimental and computational studies that bear on the evolution of enzyme function. These studies span a diversity of protein systems, and cover different evolutionary scenarios, including the adaptation of enzymes to specific host organisms, the emergence of new enzyme functions from pre-existing functionalities, and the generation of completely new active sites with enzymatic ability on previously non-catalytic scaffolds. Remarkably, a few common themes and general notions emerge from these studies, revealing the essential role of conformational dynamics in enzyme evolution. Specifically, different conformations of residues that are crucial for catalysis are often observed. Trajectories that either lead to a new function, or to the fine-tuning of an old function, typically involve the enrichment of catalytically competent conformations and/or the freezing out of non-competent conformations. Such conformational shifts are often mediated by mutations that are, to a varying degree, distant from the active site. These can be second shell mutations that introduce direct interactions with active residues, or mutations that are more distant from the active site, but that shift the protein ensemble towards catalytically productive conformations, or that are involved in evolutionary and functionally relevant coupled motions.

Following from this, the evolutionary emergence of new catalytic functions from pre-existing functions typically involves multifunctional intermediates that can switch between different conformations that are each competent for a given function. Therefore, conformational flexibility/diversity is the key to multifunctional protein ancestors.

More generally, conformational flexibility/diversity is the key to enzyme promiscuity, as it allows for the binding of

different substrates and transition states. This refers to not only the large-scale structural fluctuations that can be detected by X-ray crystallography, but also to subtler phenomena such as, for instance, the electrostatic flexibility that relaxes the perfect electrostatic pre-organization required for the highly specific catalysis of a given reaction. Such conformational flexibility is also essential to the emergence of a completely new active site through a single mutation: that is, a single mutation can generate a catalytic residue, but cannot, *by itself*, lead to an active-site cavity that is tailored for the binding of a specific chemical structure. Conformational flexibility is, therefore, required to allow substrate and, more importantly, transition-state binding. Finally, this single mutation mechanism may have played a role in the emergence of enzymes at the primordial, progenote stage, when it was plausibly assisted by high environmental temperatures and the possibility of additional phenotypic mutations.

Data accessibility. This article has no additional data. Q8

Authors' contributions. D.P. and V.A.R. performed literature searches, selected and summarized relevant studies, and wrote preliminary versions of several sections of the review dealing with specific systems. S.C.L.K. and J.M.S.R. designed the overall plan of the review, wrote several general sections of the review, and assembled and completed the final version of the manuscript.

Competing interests. We declare we have no competing interests. Q9

Funding. We received no funding for this study. Q10

Acknowledgements. This work was supported by a Wallenberg Academy Fellowship to S.C.L.K. from the Knut and Alice Wallenberg Foundation (KAW 2013.0124), by Grant RGP0041/2017 from the Human Frontier Science Program to S.C.L.K. and J.M.S.R., and by FEDER Funds and Grant BIO2015–66426-R to J.M.S.R. from the Spanish Ministry of Economy and Competitiveness.

References

1. Saleh T, Kalodimos CG. 2017 Enzymes at work are enzymes in motion. *Science* **355**, 247–248. (doi:10.1126/science.aal4632)
2. Motlagh HN, Wrabl JO, Li J, Hilser VJ. 2014 The ensemble nature of allostery. *Nature* **508**, 331–339. (doi:10.1038/nature13001)
3. Wei G, Xi W, Nussinov R, Ma B. 2016 Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. *Chem. Rev.* **116**, 6516–6551. (doi:10.1021/acs.chemrev.5b00562)
4. Osuna S, Jiménez-Osés G, Noey EL, Houk KN. 2015 Molecular dynamics explorations of active site structure in designed and evolved enzymes. *Acc. Chem. Res.* **48**, 1080–1089. (doi:10.1021/ar500452q)
5. Olsson U, Wolf-Watz M. 2010 Overlap between folding and functional energy landscapes for adenylate kinase conformational change. *Nat. Commun.* **1**, 111. (doi:10.1038/ncomms1106)
6. Schulenburg C, Hilvert D. 2013 Protein conformational disorder and enzyme catalysis. *Top. Curr. Chem.* **337**, 41–68. (doi:10.1007/128_2012_411)
7. Khersonsky O, Roodveldt C, Tawfik DS. 2006 Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* **10**, 498–508. (doi:10.1016/j.cbpa.2006.08.011)
8. Khersonsky O, Tawfik DS. 2010 Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505. (doi:10.1146/annurev-biochem-030409-143718)
9. Eisenmesser EZ, Bosco DA, Akke M, Kern D. 2002 Enzyme dynamics during catalysis. *Science* **295**, 1520–1523. (doi:10.1126/science.1066176)
10. Olsson MHM, Parson WW, Warshel A. 2006 Dynamical contributions to enzyme catalysis: critical tests of a popular hypothesis. *Chem. Rev.* **106**, 1737–1756. (doi:10.1021/cr040427e)
11. Nagel ZD, Klinman JP. 2009 A 21st century revisionist's view at a turning point in enzymology. *Nat. Chem. Biol.* **5**, 543–550. (doi:10.1038/nchembio.204)
12. Kamerlin SCL, Warshel A. 2010 At the dawn of the 21st century: is dynamics the missing link for understanding enzyme catalysis? *Proteins: Struct. Funct. Bioinform.* **78**, 1339–1375. (doi:10.1002/prot.22654)
13. Nagel ZD, Klinman JP. 2010 Update 1 of: Tunneling and dynamics in enzymatic hydride transfer. *Chem. Rev.* **110**, PR41–PR67. (doi:10.1021/cr1001035)
14. Hay S, Scrutton NS. 2012 Good vibrations in enzyme-catalysed reactions. *Nat. Chem.* **4**, 161–168. (doi:10.1038/nchem.1223)
15. Klinman JP, Kohen A. 2013 Hydrogen tunneling links protein dynamics to enzyme catalysis. *Annu. Rev. Biochem.* **82**, 471–496. (doi:10.1146/annurev-biochem-051710-133623)
16. Kohen A. 2015 Role of dynamics in enzyme catalysis: Substantial versus semantic controversies. *Acc. Chem. Res.* **48**, 466–473. (doi:10.1021/ar500322s)
17. Tuñón I, Laage D, Hynes JT. 2015 Are there dynamical effects in enzyme catalysis? Some thoughts concerning the enzymatic chemical step. *Arch. Biochem. Biophys.* **582**, 42–55. (doi:10.1016/j.abb.2015.06.004)
18. Warshel A, Bora RP. 2016 Perspective: defining and quantifying the role of dynamics in enzyme catalysis. *J. Chem. Phys.* **144**, 180901. (doi:10.1063/1.4947037)
19. Pabis A, Risso VA, Sanchez-Ruiz JM, Kamerlin SCL. 2018 Cooperativity and flexibility in enzyme evolution. *Curr. Opin. Struct. Biol.* **48**, 83–92. (doi:10.1016/j.sbi.2017.10.020)
20. Tokuriki N, Tawfik DS. 2009 Protein dynamism and evolvability. *Science* **324**, 203–207. (doi:10.1126/science.1169375)
21. James LC, Tawfik DS. 2003 Conformational diversity and protein evolution—A 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **28**, 361–368. (doi:10.1016/S0968-0004(03)00135-X)
22. Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. 2007 A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916. (doi:10.1038/nature06407)
23. Henzler-Wildman K, Kern D. 2007 Dynamic personalities of proteins. *Nature* **450**, 964–972. (doi:10.1038/nature06522)
24. Klinman JP, Kohen A. 2014 Evolutionary aspects of enzyme dynamics. *J. Biol. Chem.* **289**, 30 205–30 212. (doi:10.1074/jbc.R114.565515)
25. Petrović D, Frank D, Kamerlin SCL, Hoffmann K, Strodel B. 2017 Shuffling active site substate populations affects catalytic activity: the case of glucose oxidase. *ACS Catal.* **7**, 6188–6197. (doi:10.1021/acscatal.7b01575)
26. Romero-Rivera A, Garcia-Borràs M, Osuna S. 2017 Role of conformational dynamics in the evolution of retro-aldolase activity. *ACS Catal.* **7**, 8524–8532. (doi:10.1021/acscatal.7b02954)
27. Buller AR, van Roye P, Cahn JKB, Scheele RA, Herger M, Arnold FH. 2018 Directed evolution mimics allosteric activation by stepwise tuning of the conformational ensemble. *J. Am. Chem. Soc.* **140**, 7256–7266. (doi:10.1021/jacs.8b03490)
28. Campbell E *et al.* 2016 The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **12**, 944–950. (doi:10.1038/nchembio.2175)
29. Kaltenbach M, Burke JR, Dindo M, Pabis A, Munsberg FS, Rabin A, Kamerlin SC, Noel JP, Tawfik DS. 2018 Evolution of chalcone isomerase from a noncatalytic ancestor. *Nat. Chem. Biol.* **14**, 548–555. (doi:10.1038/s41589-018-0042-3)
30. Bar-Even A, Milo R, Noor E, Tawfik DS. 2015 The moderately efficient enzyme: futile encounters and enzyme floppiness. *Biochemistry* **54**, 4969–4977. (doi:10.1021/acs.biochem.5b00621)
31. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000 The protein data bank. *Nucleic Acids Res.* **28**, 235–242. (doi:10.1093/nar/28.1.235)
32. Sawaya MR, Kraut J. 1997 Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: Crystallographic evidence. *Biochemistry* **36**, 586–603. (doi:10.1021/bi962337c)

33. Hammes GG, Benkovic SJ, Hammes-Schiffer S. 2011 Flexibility, diversity, and cooperativity: Pillars of enzyme catalysis. *Biochemistry* **50**, 10 422–10 430. (doi:10.1021/bi201486f)
34. Schnell JR, Dyson HJ, Wright PE. 2004 Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 119–140. (doi:10.1146/annurev.biophys.33.110502.133613)
35. Bhojane PP, Duff MR, Bafna K, Agarwal P, Stanley C, Howell EE. 2017 Small angle neutron scattering studies of R67 dihydrofolate reductase, a tetrameric protein with intrinsically disordered N-termini. *Biochemistry* **56**, 5886–5899. (doi:10.1021/acs.biochem.7b00822)
36. Sikorski RS, Wang L, Markham KA, Rajagopalan PTR, Benkovic SJ, Kohen A. 2004 Tunneling and coupled motion in the *Escherichia coli* dihydrofolate reductase catalysis. *J. Am. Chem. Soc.* **126**, 4778–4779. (doi:10.1021/ja031683w)
37. Wang L, Goodey NM, Benkovic SJ, Kohen A. 2006 Coordinated effects of distal mutations on environmentally coupled tunneling in dihydrofolate reductase. *Proc. Natl Acad. Sci. USA* **103**, 15 753–15 758. (doi:10.1073/pnas.0606976103)
38. Maglia G, Allemann RK. 2003 Evidence for environmentally coupled hydrogen tunneling during dihydrofolate reductase catalysis. *J. Am. Chem. Soc.* **125**, 13 372–13 373. (doi:10.1021/ja035692g)
39. Boehr DD, McElheny D, Dyson HJ, Wright PE. 2006 The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642. (doi:10.1126/science.1130258)
40. Adamczyk AJ, Cao J, Kamerlin SCL, Warshel A. 2011 Catalysis by dihydrofolate reductase and other enzymes arises from electrostatic preorganization, not conformational motions. *Proc. Natl Acad. Sci. USA* **108**, 14 115–14 120. (doi:10.1073/pnas.1111252108)
41. Bhabha G, Lee J, Ekiert DC, Gam J, Wilson IA, Dyson HJ, Benkovic SJ, Wright PE. 2011 A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science* **332**, 234–238. (doi:10.1126/science.1198542)
42. Bhabha G *et al.* 2013 Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nat. Struct. Mol. Biol.* **20**, 1243–1249. (doi:10.1038/nsmb.2676)
43. Luk LYP *et al.* 2013 Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. *Proc. Natl Acad. Sci. USA* **110**, 16 344–16 349. (doi:10.1073/pnas.1312437110)
44. Appleman JR, Beard WA, Delcamp TJ, Prendergast NJ, Freisheim JH, Blakley RL. 1989 Atypical transient state kinetics of recombinant human dihydrofolate reductase produced by hysteretic behavior. Comparison with dihydrofolate reductases from other sources. *J. Biol. Chem.* **264**, 2625–2633.
45. Appleman JR, Beard WA, Delcamp TJ, Prendergast NJ, Freisheim JH, Blakley RL. 1990 Unusual transient- and steady-state kinetic behavior is predicted by the kinetic scheme operational for recombinant human dihydrofolate reductase. *J. Biol. Chem.* **265**, 2740–2748.
46. Beard WA, Appleman JR, Huang S, Delcamp TJ, Freisheim JH, Blakley RL. 1991 Role of the conserved active site residue tryptophan-24 of human dihydrofolate reductase as revealed by mutagenesis. *Biochemistry* **30**, 1432–1440. (doi:10.1021/bi00219a038)
47. Süel GM, Lockless SW, Wall MA, Ranganathan R. 2003 Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.* **10**, 59–69. (doi:10.1038/nsb881)
48. Radkiewicz JL, Brooks CL. 2000 Protein dynamics in enzymatic catalysis: exploration of dihydrofolate reductase. *J. Am. Chem. Soc.* **122**, 225–231. (doi:10.1021/ja9913838)
49. Agarwal PK, Billeter SR, Rajagopalan PTR, Benkovic SJ, Hammes-Schiffer S. 2002 Network of coupled promoting motions in enzyme catalysis. *Proc. Natl Acad. Sci. USA* **99**, 2794–2799. (doi:10.1073/pnas.052005999)
50. Rod TH, Radkiewicz JL, Brooks CL. 2003 Correlated motion and the effect of distal mutations in dihydrofolate reductase. *Proc. Natl Acad. Sci. USA* **100**, 6980–6985. (doi:10.1073/pnas.1230801100)
51. Wong KF, Selzer T, Benkovic SJ, Hammes-Schiffer S. 2005 Impact of distal mutations on the network of coupled motions correlated to hydride transfer in dihydrofolate reductase. *Proc. Natl Acad. Sci. USA* **102**, 6807–6812. (doi:10.1073/pnas.0408343102)
52. Hammes-Schiffer S. 2006 Hydrogen tunneling and protein motion in enzyme reactions. *Acc. Chem. Res.* **39**, 93–100. (doi:10.1021/ar040199a)
53. Cheatum CM, Kohen A. 2013 Relationship of femtosecond–picosecond dynamics to enzyme-catalyzed H-transfer. In *Dynamics in enzyme catalysis* (eds J Klinman, S Hammes-Schiffer), pp. 1–39. Berlin, Germany: Springer.
54. Singh P, Sen A, Francis K, Kohen A. 2014 Extension and limits of the network of coupled motions correlated to hydride transfer in dihydrofolate reductase. *J. Am. Chem. Soc.* **136**, 2575–2582. (doi:10.1021/ja411998h)
55. Ruiz-Pernia JJ, Luk LYP, García-Meseguer R, Martí S, Loveridge EJ, Tuñón I, Moliner V, Allemann RK. 2013 Increased dynamic effects in a catalytically compromised variant of *Escherichia coli* dihydrofolate reductase. *J. Am. Chem. Soc.* **135**, 18 689–18 696. (doi:10.1021/ja410519h)
56. Ruiz-Pernia JJ, Behiry E, Luk LYP, Loveridge EJ, Tunon I, Moliner V, Allemann RK. 2016 Minimization of dynamic effects in the evolution of dihydrofolate reductase. *Chem. Sci.* **7**, 3248–3255. (doi:10.1039/C5SC04209G)
57. Wang J. 2008 Electrochemical glucose biosensors. *Chem. Rev.* **108**, 814–825. (doi:10.1021/cr068123a)
58. Yu X, Lian W, Zhang J, Liu H. 2016 Multi-input and -output logic circuits based on bioelectrocatalysis with horseradish peroxidase and glucose oxidase immobilized in multi-responsive copolymer films on electrodes. *Biosensors Bioelectron.* **80**, 631–639. (doi:10.1016/j.bios.2016.02.010)
59. Willner I, Yan YM, Willner B, Tel-Vered R. 2009 Integrated enzyme-based biofuel cells–A review. *Fuel Cells* **9**, 7–24. (doi:10.1002/fuce.200800115)
60. Cinquin P *et al.* 2010 A glucose biofuel cell implanted in rats. *PLoS ONE* **5**, e10476. (doi:10.1371/journal.pone.0010476)
61. Zebda A *et al.* 2013 Single glucose biofuel cells implanted in rats power electronic devices. *Sci. Rep.* **3**, 1516. (doi:10.1038/srep01516)
62. Holland JT, Lau C, Brozik S, Atanassov P, Banta S. 2011 Engineering of glucose oxidase for direct electron transfer via site-specific gold nanoparticle conjugation. *J. Am. Chem. Soc.* **133**, 19 262–19 265. (doi:10.1021/ja2071237)
63. Holland JT, Harper JC, Dolan PL, Manginell MM, Arango DC, Rawlings JA, Apblett CA, Brozik SM. 2012 Rational redesign of glucose oxidase for improved catalytic function and stability. *PLoS ONE* **7**, e37924. (doi:10.1371/journal.pone.0037924)
64. Horaguchi Y, Saito S, Kojima K, Tsugawa W, Ferri S, Sode K. 2014 Engineering glucose oxidase to minimize the influence of oxygen on sensor response. *Electrochim. Acta* **126**, 158–161. (doi:10.1016/j.electacta.2013.09.018)
65. Meyer M, Wohlfahrt G, Knäblein J, Schomburg D. 1998 Aspects of the mechanism of catalysis of glucose oxidase: A docking, molecular mechanics and quantum chemical study. *J. Comput.-Aided Mol. Des.* **12**, 425–440. (doi:10.1023/A:1008020124326)
66. Leskovac V, Trivić S, Wohlfahrt G, Kandrač J, Peričin D. 2005 Glucose oxidase from *Aspergillus niger*: the mechanism of action with molecular oxygen, quinones, and one-electron acceptors. *Int. J. Biochem. Cell Biol.* **37**, 731–750. (doi:10.1016/j.biocel.2004.10.014)
67. Hecht HJ, Kalisz HM, Hendle J, Schmid RD, Schomburg D. 1993 Crystal structure of glucose oxidase from *Aspergillus niger* refined at 2.3 Å resolution. *J. Mol. Biol.* **229**, 153–172. (doi:10.1006/jmbi.1993.1015)
68. Wohlfahrt G, Witt S, Hendle J, Schomburg D, Kalisz HM, Hecht H-J. 1999 1.8 and 1.9 Å resolution structures of the *Penicillium amagasakiense* and *Aspergillus niger* glucose oxidases as a basis for modelling substrate complexes. *Acta Crystallogr. D* **55**, 969–977. (doi:10.1107/S0907444999003431)
69. Yoshida H, Sakai G, Mori K, Kojima K, Kamitori S, Sode K. 2015 Structural analysis of fungus-derived FAD glucose dehydrogenase. *Sci. Rep.* **5**, 13498. (doi:10.1038/srep13498)
70. Ma H, Szeler K, Kamerlin SCL, Widersten M. 2016 Linking coupled motions and entropic effects to the catalytic activity of 2-deoxyribose-5-phosphate aldolase (DERA). *Chem. Sci.* **7**, 1415–1421. (doi:10.1039/C5SC03666F)
71. Tozzi MG, Camici M, Mascia L, Sgarrella F, Ipata PL. 2006 Pentose phosphates in nucleoside interconversion and catabolism. *FEBS J.* **273**, 1089–1101. (doi:10.1111/j.1742-4658.2006.05155.x)
72. Hoffee PA. 1968 2-deoxyribose-5-phosphate aldolase of *Salmonella typhimurium*: purification and properties. *Arch. Biochem. Biophys.* **126**, 795–802. (doi:10.1016/0003-9861(68)90473-6)
73. Machajewski TD, Wong CH. 2000 The catalytic asymmetric aldol reaction. *Angew. Chem. Int. Ed.* **39**, 1352–1375. (doi:10.1002/(SICI)1521-3773(20000417)39:8<1352::AID-ANIE1352>3.0.CO;2-J)

74. Heine A, DeSantis G, Luz JG, Mitchell M, Wong C-H, Wilson IA. 2001 Observation of covalent intermediates in an enzyme mechanism at atomic resolution. *Science* **294**, 369–374. (doi:10.1126/science.1063601)
75. Samland AK, Sprenger GA. 2006 Microbial aldolases as C–C bonding enzymes—Unknown treasures and new developments. *Appl. Microbiol. Biotechnol.* **71**, 253. (doi:10.1007/s00253-006-0422-6)
76. Hol WGJ, van Duijnen PT, Berendsen HJC. 1978 The α-helix dipole and the properties of proteins. *Nature* **273**, 443–446. (doi:10.1038/273443a0)
77. Copley RR, Barton GJ. 1994 A structural analysis of phosphate and sulphate binding sites in proteins: Estimation of propensities for binding and conservation of phosphate binding sites. *J. Mol. Biol.* **242**, 321–329. (doi:10.1006/jmbi.1994.1583)
78. Schulte M, Petrović D, Neudecker P, Hartmann R, Pietruszka J, Willbold S, Willbold D, Panwalkar V. 2018 Conformational sampling of the intrinsically disordered C-terminal tail of DERA is important for enzyme catalysis. *ACS Catal.* **8**, 3971–3984. (doi:10.1021/acscatal.7b04408)
79. Jiang L *et al.* 2008 De novo computational design of retro-aldol enzymes. *Science* **319**, 1387–1391. (doi:10.1126/science.1152692)
80. Althoff EA *et al.* 2012 Robust design and optimization of retroaldol enzymes. *Protein Sci.* **21**, 717–726. (doi:10.1002/pro.2059)
81. Lassila JK, Baker D, Herschlag D. 2010 Origins of catalysis by computationally designed retroaldolase enzymes. *Proc. Natl Acad. Sci. USA* **107**, 4937–4942. (doi:10.1073/pnas.0913638107)
82. Giger L, Caner S, Obexer R, Kast P, Baker D, Ban N, Hilvert D. 2013 Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat. Chem. Biol.* **9**, 494–498. (doi:10.1038/nchembio.1276)
83. Obexer R, Godina A, Garrabou X, Mittl PRE, Baker D, Griffiths AD, Hilvert D. 2017 Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nat. Chem.* **9**, 50–56. (doi:10.1038/nchem.2596)
84. Risso VA, Gavira JA, Mejia-Carmona DF, Gaucher EA, Sanchez-Ruiz JM. 2013 Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. *J. Am. Chem. Soc.* **135**, 2899–2902. (doi:10.1021/ja311630a)
85. Knowles JR. 1985 Penicillin resistance: the chemistry of β-lactamase inhibition. *Acc. Chem. Res.* **18**, 97–104. (doi:10.1021/ar00112a001)
86. Strynadka NCJ, Jensen SE, Alzari PM, James MNG. 1996 A potent new mode of β-lactamase inhibition revealed by the 1.7 Å X-ray crystallographic structure of the TEM-1–BLIP complex. *Nat. Struct. Biol.* **3**, 290–297. (doi:10.1038/nsb0396-290)
87. Wladkowski BD, Chenoweth SA, Sanders JN, Krauss M, Stevens WJ. 1997 Acylation of β-lactams by class A β-lactamase: an ab initio theoretical study on the effects of the oxy-anion hole. *J. Am. Chem. Soc.* **119**, 6423–6431. (doi:10.1021/ja963678g)
88. Massova I, Mobashery S. 1997 Molecular bases for interactions between β-lactam antibiotics and β-lactamases. *Acc. Chem. Res.* **30**, 162–168. (doi:10.1021/ar960007e)
89. Lim D, Park HU, De Castro L, Kang SG, Lee HS, Jensen S, Lee KJ, Strynadka NC. 2001 Crystal structure and kinetic analysis of β-lactamase inhibitor protein-II in complex with TEM-1 β-lactamase. *Nat. Struct. Biol.* **8**, 848–852. (doi:10.1038/nsb1001-848)
90. Castillo R, Silla E, Tuñón I. 2002 Role of protein flexibility in enzymatic catalysis: quantum mechanical–molecular mechanical study of the deacylation reaction in class A β-lactamases. *J. Am. Chem. Soc.* **124**, 1809–1816. (doi:10.1021/ja017156z)
91. Wilke MS, Lovering AL, Strynadka NCJ. 2005 Β-lactam antibiotic resistance: a current structural perspective. *Curr. Opin. Microbiol.* **8**, 525–533. (doi:10.1016/j.mib.2005.08.016)
92. Hermann JC, Hensen C, Ridder L, Mulholland AJ, Höltje H-D. 2005 Mechanisms of antibiotic resistance: QM/MM modeling of the acylation reaction of a class A β-lactamase with benzylpenicillin. *J. Am. Chem. Soc.* **127**, 4454–4465. (doi:10.1021/ja044210d)
93. Xu D, Guo H, Cui Q. 2007 Antibiotic deactivation by a dizinc β-lactamase: mechanistic insights from QM/MM and DFT studies. *J. Am. Chem. Soc.* **129**, 10 814–10 822. (doi:10.1021/ja072532m)
94. Heidari-Torkabadi H, Bethel CR, Ding Z, Pusztai-Carey M, Bonnet R, Bonomo RA, Carey PR. 2015 'Mind the gap': raman evidence for rapid inactivation of CTX-M-9 β-lactamase using mechanism-based inhibitors that bridge the active site. *J. Am. Chem. Soc.* **137**, 12 760–12 763. (doi:10.1021/jacs.5b10007)
95. Pratt RF. 2016 β-lactamases: why and how. *J. Med. Chem.* **59**, 8207–8220. (doi:10.1021/acs.jmedchem.6b00448)
96. Pande VS, Beauchamp K, Bowman GR. 2010 Everything you wanted to know about Markov state models but were afraid to ask. *Methods* **52**, 99–105. (doi:10.1016/j.ymeth.2010.06.002)
97. Chodera JD, Noé F. 2014 Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **25**, 135–144. (doi:10.1016/j.sbi.2014.04.002)
98. Hart KM, Ho CMW, Dutta S, Gross ML, Bowman GR. 2016 Modelling proteins' hidden conformations to predict antibiotic resistance. *Nat. Commun.* **7**, 12965. (doi:10.1038/ncomms12965)
99. Cortina GA, Kasson PM. 2016 Excess positional mutual information predicts both local and allosteric mutations affecting beta lactamase drug resistance. *Bioinformatics* **32**, 3420–3427. (doi:10.1093/bioinformatics/btw492)
100. Latallo MJ, Cortina GA, Faham S, Nakamoto RK, Kasson PM. 2017 Predicting allosteric mutants that increase activity of a major antibiotic resistance enzyme. *Chem. Sci.* **8**, 6484–6492. (doi:10.1039/C7SC02676E)
101. Zou T, Risso VA, Gavira JA, Sanchez-Ruiz JM, Ozkan SB. 2015 Evolution of conformational dynamics determines the conversion of a promiscuous generalist into a specialist enzyme. *Mol. Biol. Evol.* **32**, 132–143. (doi:10.1093/molbev/msu281)
102. Gerek NZ, Kumar S, Ozkan SB. 2013 Structural dynamics flexibility informs function and evolution at a proteome scale. *Evol. Appl.* **6**, 423–433. (doi:10.1111/eva.12052)
103. González MM, Abriata LA, Tomatis PE, Vila AJ. 2016 Optimization of conformational dynamics in an epistatic evolutionary trajectory. *Mol. Biol. Evol.* **33**, 1768–1776. (doi:10.1093/molbev/msw052)
104. Risso VA *et al.* 2017 De novo active sites for resurrected Precambrian enzymes. *Nat. Commun.* **8**, 16113. (doi:10.1038/ncomms16113)
105. Ohno S. 1970 *Evolution by gene duplication*. Berlin, Germany: Springer.
106. Innan H, Kondrashov F. 2010 The evolution of gene duplications: Classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, 97–108. (doi:10.1038/nrg2689)
107. O'Brien PJ, Herschlag D. 1999 Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* **6**, R91–R105. (doi:10.1016/S1074-5521(99)80033-7)
108. Jonas S, Hollfelder F. 2009 Mapping catalytic promiscuity in the alkaline phosphatase superfamily. *Pure Appl. Chem.* **81**, 731–742. (doi:10.1351/PAC-CON-08-10-20)
109. Mohamed MF, Hollfelder F. 2013 Efficient, crosswise catalytic promiscuity among enzymes that catalyze phosphoryl transfer. *Biochim. Biophys. Acta* **1834**, 417–424. (doi:10.1016/j.bbapap.2012.07.015)
110. Luo J, van Loo B, Kamerlin SCL. 2011 Examining the promiscuous phosphatase activity of *Pseudomonas aeruginosa* arylsulfatase: a comparison to analogous phosphatases. *Proteins: Struct. Funct. Bioinform.* **80**, 1211–1226. (doi:10.1002/prot.24020)
111. Luo J, van Loo B, Kamerlin SCL. 2012 Catalytic promiscuity in *Pseudomonas aeruginosa* arylsulfatase as an example of chemistry-driven protein evolution. *FEBS Lett.* **586**, 1622–1630. (doi:10.1016/j.febslet.2012.04.012)
112. Barrozo A, Duarte F, Bauer P, Carvalho ATP, Kamerlin SCL. 2015 Cooperative electrostatic interactions drive functional evolution in the alkaline phosphatase superfamily. *J. Am. Chem. Soc.* **137**, 9061–9076. (doi:10.1021/jacs.5b03945)
113. Ben-David M, Sussman JL, Maxwell CI, Szeler K, Kamerlin SCL, Tawfik DS. 2015 Catalytic stimulation by restrained active-site floppiness—The case of high density lipoprotein-bound serum paraoxonase-1. *J. Mol. Biol.* **427**, 1359–1374. (doi:10.1016/j.jmb.2015.01.013)
114. Purg M, Elias M, Kamerlin SCL. 2017 Similar active sites and mechanisms do not lead to cross-promiscuity in organophosphate hydrolysis: Implications for biotherapeutic engineering. *J. Am. Chem. Soc.* **139**, 17 533–17 546. (doi:10.1021/jacs.7b09384)
115. Blaha-Nelson D, Krüger DM, Szeler K, Ben-David M, Kamerlin SCL. 2017 Active site hydrophobicity and the convergent evolution of paraoxonase activity in structurally divergent enzymes: the case of serum

paraoxonase 1. *J. Am. Chem. Soc.* **139**, 1155–1167. (doi:10.1021/jacs.6b10801)

116. Duarte F, Amrein BA, Kamerlin SCL. 2013 Modeling catalytic promiscuity in the alkaline phosphatase superfamily. *Phys. Chem. Chem. Phys.* **15**, 11 160–11 177. (doi:10.1039/C3CP51179K)
117. Pabis A, Duarte F, Kamerlin SCL. 2016 Promiscuity in the enzymatic catalysis of phosphate and sulfate transfer. *Biochemistry* **55**, 3061–3081. (doi:10.1021/acs.biochem.6b00297)
118. Pabis A, Kamerlin SCL. 2016 Promiscuity and electrostatic flexibility in the alkaline phosphatase superfamily. *Curr. Opin. Struct. Biol.* **37**, 14–21. (doi:10.1016/j.sbi.2015.11.008)
119. Petrović D, Szeler K, Kamerlin SCL. 2018 Challenges and advances in the computational modeling of biological phosphate hydrolysis. *Chem. Commun.* **54**, 3077–3089. (doi:10.1039/C7CC09504J)
120. Tokuriki N, Jackson CJ, Afriat-Jurnou L, Wyganowski KT, Tang R, Tawfik DS. 2012 Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nat. Commun.* **3**, 1257. (doi:10.1038/ncomms2246)
121. Jackson CJ, Foo J-L, Tokuriki N, Afriat L, Carr PD, Kim H-K, Schenk G, Tawfik DS, Ollis DL. 2009 Conformational sampling, catalysis, and evolution of the bacterial phosphotriesterase. *Proc. Natl Acad. Sci. USA* **106**, 21 631–21 636. (doi:10.1073/pnas.0907548106)
122. Bartlett GJ, Porter CT, Borkakoti N, Thornton JM. 2002 Analysis of catalytic residues in enzyme active sites. *J. Mol. Biol.* **324**, 105–121. (doi:10.1016/S0022-2836(02)01036-7)
123. Jack BR, Meyer AG, Echave J, Wilke CO. 2016 Functional sites induce long-range evolutionary constraints in enzymes. *PLoS Biol.* **14**, e1002452. (doi:10.1371/journal.pbio.1002452)
124. Siddiq MA, Hochberg GKA, Thornton JW. 2017 Evolution of protein specificity: insights from ancestral protein reconstruction. *Curr. Opin. Struct. Biol.* **47**, 113–122. (doi:10.1016/j.sbi.2017.07.003)
125. Alexander PA, He Y, Chen Y, Orban J, Bryan PN. 2009 A minimal sequence code for switching protein structure and function. *Proc. Natl Acad. Sci. USA* **106**, 21 149–21 154. (doi:10.1073/pnas.0906408106)
126. Sikosek T, Krobath H, Chan HS. 2016 Theoretical insights into the biophysics of protein bi-stability and evolutionary switches. *PLoS Comp. Biol.* **12**, e1004960. (doi:10.1371/journal.pcbi.1004960)
127. Grishin NV. 2001 Fold change in evolution of protein structures. *J. Struct. Biol.* **134**, 167–185. (doi:10.1006/jsbi.2001.4335)
128. Sikosek T, Bornberg-Bauer E, Chan HS. 2012 Evolutionary dynamics on protein bi-stability landscapes can potentially resolve adaptive conflicts. *PLoS Comp. Biol.* **8**, e1002659. (doi:10.1371/journal.pcbi.1002659)
129. de Wolf FA, Brett GM. 2000 Ligand-binding proteins: their potential for application in systems for controlled delivery and uptake of ligands. *Pharmacol. Rev.* **52**, 207–236.
130. Clifton BE, Kaczmarski JA, Carr PD, Gerth ML, Tokuriki N, Jackson CJ. 2018 Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein. *Nat. Chem. Biol.* **14**, 542–547. (doi:10.1101/157495)
131. Wiita AP, Perez-Jimenez R, Walther KA, Gräter F, Berne BJ, Holmgren A, Sanchez-Ruiz JM, Fernandez JM. 2007 Probing the chemistry of thioredoxin catalysis with force. *Nature* **450**, 124–127. (doi:10.1038/nature06231)
132. Perez-Jimenez R *et al.* 2009 Diversity of chemical mechanisms in thioredoxin catalysis revealed by single-molecule force spectroscopy. *Nat. Struct. Mol. Biol.* **16**, 890–896. (doi:10.1038/nsmb.1627)
133. Rodriguez-Larrea D, Perez-Jimenez R, Sanchez-Romero I, Delgado-Delgado A, Fernandez Julio M, Sanchez-Ruiz JM. 2010 Role of conservative mutations in protein multi-property adaptation. *Biochem. J.* **429**, 243–249. (doi:10.1042/BJ20100386)
134. Ashenberg O, Gong LI, Bloom JD. 2013 Mutational effects on stability are largely conserved during protein evolution. *Proc. Natl Acad. Sci. USA* **110**, 21 071–21 076. (doi:10.1073/pnas.1314781111)
135. Risso VA *et al.* 2015 Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. *Mol. Biol. Evol.* **32**, 440–455. (doi:10.1093/molbev/msu312)
136. Godoy-Ruiz R, Ariza F, Rodriguez-Larrea D, Perez-Jimenez R, Ibarra-Molero B, Sanchez-Ruiz JM. 2006 Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments. *J. Mol. Biol.* **362**, 966–978. (doi:10.1016/j.jmb.2006.07.065)
137. Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS. 2006 Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* **444**, 929–932. (doi:10.1038/nature05385)
138. Ouzounis CA, Kunin V, Darzentas N, Goldovsky L. 2006 A minimal estimate for the gene content of the last universal common ancestor—Exobiology from a terrestrial perspective. *Res. Microbiol.* **157**, 57–68. (doi:10.1016/j.resmic.2005.06.015)
139. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, Martin WF. 2016 The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 16116. (doi:10.1038/nmicrobiol.2016.116)
140. Wolf YI, Koonin EV. 2007 On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol. Direct* **2**, 14. (doi:10.1186/1745-6150-2-14)
141. Romero MLR, Rabin A, Tawfik DS. 2016 Functional proteins from short peptides: dayhoff's hypothesis turns 50. *Angew. Chem. Int. Ed.* **55**, 15 966–15 971. (doi:10.1002/anie.201609977)
142. Camprubi E, Jordan SF, Vasiliadou R, Lane N. 2017 Iron catalysis at the origin of life. *IUBMB Life* **69**, 373–381. (doi:10.1002/iub.1632)
143. Holm RH, Kennepohl P, Solomon EI. 1996 Structural and functional aspects of metal sites in biology. *Chem. Rev.* **96**, 2239–2314. (doi:10.1021/cr9500390)
144. Waldron KJ, Robinson NJ. 2009 How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **7**, 25–35. (doi:10.1038/nrmicro2057)
145. Röthlisberger D *et al.* 2008 Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195. (doi:10.1038/nature06879)
146. Siegel JB *et al.* 2010 Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313. (doi:10.1126/science.1190239)
147. Richter F *et al.* 2012 Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J. Am. Chem. Soc.* **134**, 16 197–16 206. (doi:10.1021/ja3037367)
148. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R. 2011 The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410. (doi:10.1021/bi2002289)
149. Korendovych IV, Kulp DW, Wu Y, Cheng H, Roder H, DeGrado WF. 2011 Design of a switchable eliminase. *Proc. Natl Acad. Sci. USA* **108**, 6823–6827. (doi:10.1073/pnas.1018191108)
150. Moroz YS *et al.* 2015 New tricks for old proteins: single mutations in a nonenzymatic protein give rise to various enzymatic activities. *J. Am. Chem. Soc.* **137**, 14 905–14 911. (doi:10.1021/jacs.5b07812)
151. Pauling L. 1948 Nature of forces between large molecules of biological interest. *Nature* **161**, 707–709. (doi:10.1038/161707a0)
152. Yamniuk AP, Vogel HJ. 2004 Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Mol. Biotechnol.* **27**, 33–57. (doi:10.1385/mb:27:1:33)
153. Jensen RA. 1976 Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409–425. (doi:10.1146/annurev.mi.30.100176.002205)
154. Wolfenden R, Snider MJ. 2001 The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **34**, 938–945. (doi:10.1021/ar000058i)
155. Lane N, Martin William F. 2012 The origin of membrane bioenergetics. *Cell* **151**, 1406–1416. (doi:10.1016/j.cell.2012.11.050)
156. Robert F, Chaussidon M. 2006 A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. *Nature* **443**, 969–972. (doi:10.1038/nature05239)
157. Drummond AD, Wilke CO. 2009 The evolutionary consequences of erroneous protein synthesis. *Nat. Rev. Genet.* **10**, 715–724. (doi:10.1038/nrg2662)
158. Goldsmith M, Tawfik DS. 2009 Potential role of phenotypic mutations in the evolution of protein expression and stability. *Proc. Natl Acad. Sci. USA* **106**, 6197–6202. (doi:10.1073/pnas.0809506106)
159. Whitehead DJ, Wilke CO, Vernazobres D, Bornberg-Bauer E. 2008 The look-ahead effect of phenotypic mutations. *Biol. Direct* **3**, 18. (doi:10.1186/1745-6150-3-18)
160. Woese CR, Fox GE. 1977 Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090. (doi:10.1073/pnas.74.11.5088)