

Conformational diversity and protein evolution – a 60-year-old hypothesis revisited

Leo C. James¹ and Dan S. Tawfik²

¹Centre for Protein Engineering, Hills Road, Cambridge, UK CB2 2HQ

²Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76 100, Israel

Complex organisms have evolved from a limited number of primordial genes and proteins. However, the mechanisms by which the earliest proteins evolved and then served as the origin for the present diversity of protein function are unknown. Here, we outline a hypothesis based on the ‘new view’ of proteins whereby one sequence can adopt multiple structures and functions. We suggest that such conformational diversity could increase the functional diversity of a limited repertoire of sequences and, thereby, facilitate the evolution of new proteins and functions from old ones.

Results from many different disciplines including protein folding, NMR and fast kinetics are challenging the ‘one sequence, one structure, one function’ paradigm. Proteins – the simplified view of which assumes a single rigid fold – have been shown to exist as an ensemble of different conformations in equilibrium before encountering a substrate or hapten. This has prompted a ‘NEW VIEW’ (see Glossary) of protein structure and function in which conformational diversity provides a mechanism for controlling protein activation and permitting multi-functionality. At the same time, re-examination of enzymatic activity and the *in vitro* evolution of enzymes have highlighted the role of promiscuity in protein evolution. In this review, we discuss the ‘new view’ hypothesis and the growing evidence for CONFORMATIONAL DIVERSITY and functional PROMISCUITY. We present an evolutionary model that suggests that conformational diversity and functional promiscuity are evolvability traits that enable existing enzymes to rapidly evolve new activities. We also describe how the ‘new view’ and its most extreme manifestation – intrinsically disordered proteins – might support a scenario of co-evolution of protein fold and function, replacing the less plausible hypothesis of fold evolving before function.

The ‘new view’ of protein function

Traditionally, it is assumed that a given sequence dictates a single 3D structure upon which function is entirely dependent. This view is manifested in the ‘lock-and-key’ and ‘INDUCED-FIT’ models of action, in which the structure is entirely fixed or changes (locally, around the active site)

only after a ligand is bound (Box 1). Both models are strongly supported by the principal method used to study protein structure – X-ray crystallography. However, crystallization is, by definition, a purification process leading to the isolation of conformationally and chemically homogeneous molecules. The crystal structure of a protein might not be the only conformation adopted in solution and, owing to the influence of crystal-packing forces, might not even be the most representative. For example, the ‘lock-and-key’ model was assumed from the identical

Glossary

Numerous terms exist that describe the ability of a single protein to exhibit more than one specificity or perform more than one function. These terms are not necessarily distinct and are, in many senses, overlapping; they are often used in parallel or in a different context than described herein. Here, we attempt to define a clearer terminology and distinct uses for these terms. A detailed discussion of promiscuity, moonlighting and cross-reactivity can be found in Ref. [22].

Conformational diversity: the existence of a single protein in more than one conformation, independent of any conformational changes induced by ligand binding (pre-existing conformations or plasticity).

Cross-reactivity: refers to activities of the same protein that overlap with the original activity – for example, with substrates or ligands that resemble the primary or original target of the protein. Cross-reactivity relies on the same key binding-site or active-site interactions to bind or catalyze a process similar to the original activity of the protein. Many enzymes exhibit cross-reactivity to the extent that the ‘original’ target can hardly be defined.

Induced-fit: this model assumes that a protein exists in a single conformation that binds the ligand with low affinity. Ligand binding (i.e. formation of a low-affinity encounter complex) induces an isomerization that leads the final high-affinity complex.

Moonlighting: refers to a protein exerting a different function by using parts of the protein that are generally different than its original active site. An enzyme, for example, might moonlight as an activator by binding to a receptor using parts of the enzyme distant from its enzymatic active site. Moonlighting functions generally have an *in vivo* role.

Multi-specificity: refers to a protein exerting a similar function (such as binding) on distinctly different ligands, perhaps while using different active-site residues (for example see Ref. [16]).

‘New view’ of proteins: refers to the view of proteins as an ensemble of pre-existing structures of discrete yet similar free energies. We have borrowed this term from Dill’s ‘new view of protein folding’ [43].

Pre-equilibrium: this model assumes that a protein equilibrates between two (or more) pre-existing conformational isomers, one of which has higher affinity and therefore binds to the ligand. Thus, ligand binding shifts the equilibrium in favour of the high-affinity isomer.

Promiscuity: (or poly-reactivity) refers to a protein exerting a different function using the same active site that is responsible for its original activity [22]. Promiscuous activities typically proceed via distinctly different mechanisms than the original activity and, thereby, differ from cross-reactivity or multi-specificity (for example see Ref. [23]). Promiscuity is a property often identified by *in vitro* assays and might not necessarily bare a physiological function [22,34].

Corresponding author: Dan S. Tawfik (tawfik@weizmann.ac.il).

Box 1. The 'simplistic' versus the 'new view' of proteins

The 'simplistic view' of proteins assumes an energy diagram with a single well (a global minimum), which corresponds to the existence of a single structural conformer (Fig. 1a). The 'new view' of proteins has an energy landscape with many local minima corresponding to an ensemble of pre-existing structures with similar but discrete energy levels (plasticity; Fig. 1b). The mechanism by which conformational changes are linked to function also varies according to the two views. The induced-fit model has become part of the 'simplistic view' because it assumes that, in the absence of a ligand (L), the protein (P) adopts one

conformation only. The active conformer (P^*) is induced by ligand binding and has no existence in the absence of the ligand. The Koshland, Nemethy and Filmer (KNF) model of allostery (the sequential model) is an extension of the induced-fit model. By contrast, the 'new view', assumes that both P and P^* are pre-existing isomers (pre-equilibrium) and that ligand binding shifts the equilibrium only in favour of P^* . The Monod–Wyman–Changeux (MWC) model of allostery (the symmetry model) is based on an equilibrium between at least two pre-existing isomers.

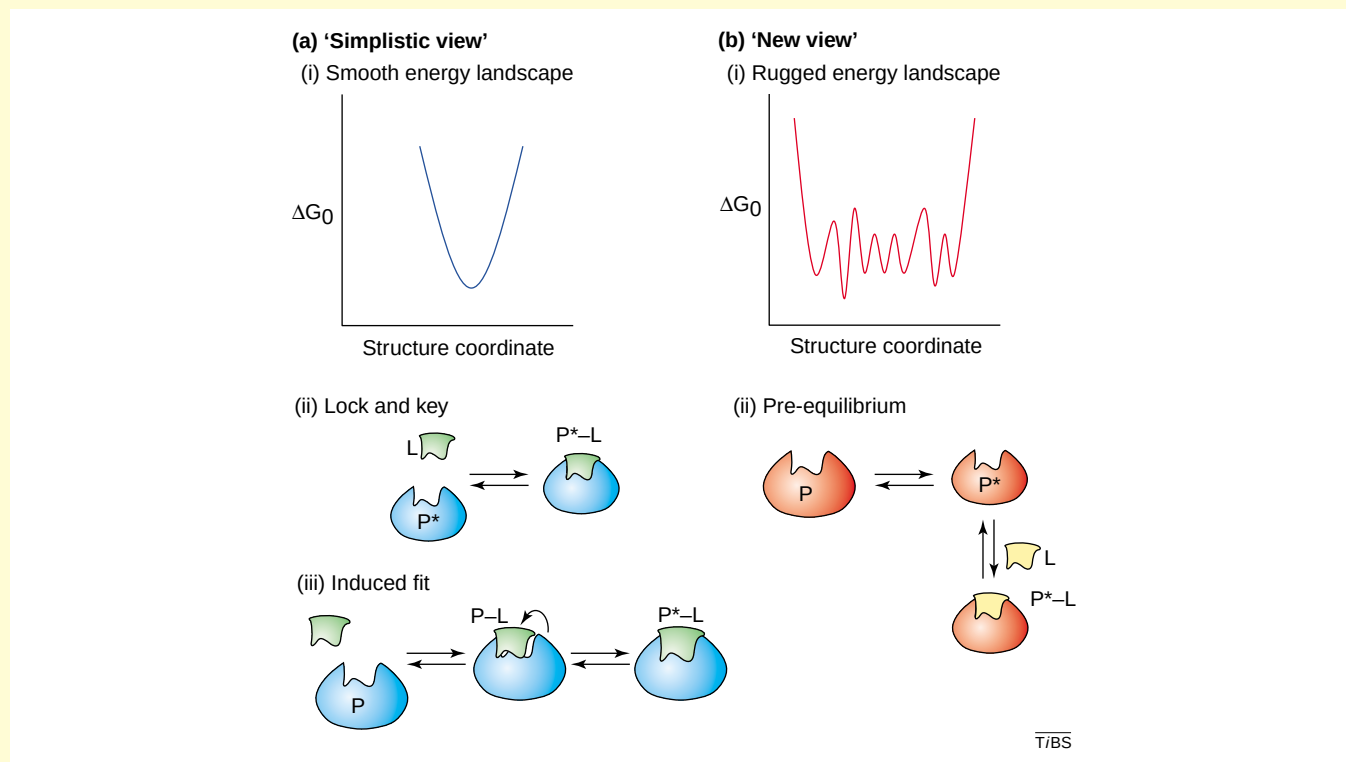


Fig. 1. Schematic energy landscapes and modes of function that represent the 'simplistic' versus the 'new view' of proteins. (a) The 'simplistic' model of proteins describes an energy landscape of a single stable conformer (i) and a function mode of either lock and key (ii) or induced fit (iii). (b) The 'new view' assumes an ensemble of conformers of similar free energy (i), and a mode of function based on an equilibrium between two (or more) pre-existing isomers, only one of which exerts function (ii).

structures of a bound and unbound antibody. However, pre-steady-state kinetics subsequently revealed a PRE-EQUILIBRIUM between two antibody isomers, only one of which binds the hapten with high affinity. Kinetics performed in the crystallization buffer indicate that it shifts the equilibrium in favour of the high-affinity isomer, explaining why the same structure was observed in the presence and absence of the hapten [1].

The magnitude of conformational diversity observed in proteins ranges from fluctuation of side chains to the movement of loops and secondary structures, and even to global tertiary structure rearrangements. NMR, in particular, has proven particularly effective in revealing the true conformational diversity of proteins. NMR analysis has demonstrated that proteins can, in fact, adopt many alternative conformations in solution in the absence of ligand. For example, the N-terminal SH3 domain of the *Drosophila* protein drk adopts ~60 different stable and distinct conformations, some of

which differ considerably [2]. Most notable are the discovery of intrinsically disordered proteins, the identification of proteins and domains that spontaneously interconvert between different secondary structures, and the identification of pre-existing active conformations in allosteric proteins.

The Monod–Wyman–Changeux (MWC) model of allosteric regulation is, to our knowledge, the first instance in which the possibility of an equilibrium between pre-existing conformations was widely recognized. This model poses a key element of the 'new view', namely, of having a single protein equilibrating freely (in the absence of a ligand) between two structural isomers (Box 1). There has always been considerable debate as to whether the different states of allosteric proteins really represent pre-equilibrium diversity, as in the MWC model, or are actually the result of induced fit, as in the Koshland–Nemethy–Filmer (KNF) model. A recent crystallographic study has confirmed the MWC model in aspartate

transcarbamylase [3]. An elegant example came from an NMR study that addressed the exchange between inactive and active conformations in a protein (NtrC) that is allosterically regulated by phosphorylation. It appears that both conformations are present in unphosphorylated NtrC, and phosphorylation merely shifts the equilibrium towards the active conformation [4].

A well-characterized example of conformational diversity at the secondary-structure level comes from work on the DNA-binding domain of the Arc repressor. The Arc-N11L mutant spontaneously interconverts between a two-stranded antiparallel β sheet and a two- 3_{10} -helix structure [5]. In the absence of ligand (DNA), these two different conformations appear to be in pre-equilibrium at almost equal proportions. Addition of DNA shifts the equilibrium towards the β -sheet form. In this case, exchange between different binding-domain conformers does not dismantle the entire Arc-N11L tertiary structure, but there are examples in which a global conformational change occurs. For example, the prion protein (PrP) interconverts between an all α -helix (PrP^c) and an all β -sheet conformation (PrP^{sc}) [6,7]. The β -sheet form (PrP^{sc}) is trapped by subsequent oligomerization, resulting in amyloid deposit and onset of disease.

Perhaps the most striking or extreme manifestation of protein-conformational diversity is in proteins that appear to be partially, or even completely, unstructured under native conditions, but form a perfectly ordered structure in the presence of ligand [8–10]. Intrinsically disordered or naturally unfolded proteins, or protein domains are thought to be common rather than rare exceptions [9]. They show that the structural differences between isomers of the same protein sequence can vary dramatically and be as global as to affect the entire structure of the protein (one conformation being fully ordered, the other disordered). From an evolutionary point of view, intrinsically disordered proteins provide a powerful demonstration of the fact that a defined 3D structure is not a prerequisite for function [8]. Thus, rather than fold having to evolve before function (an implausible scenario hampered by the fact that fold without function provides no selective advantage), function might have co-evolved with fold. This is a key feature of the evolutionary mechanisms presented in this review.

Conformational changes are intimately linked to protein function, and to enzyme function in particular [11]. This is manifested in allostery as described above and in many other cases; for instance, the two-step binding reaction of Ras with effector proteins results from an isomerization between different pre-existing conformations of the highly mobile switch I region [12]. Conformational diversity has also been observed in the form of two dimerization modes for complexes of the ribosomal protein L12; these also involve two different L12 monomer conformations and are in accordance with the dynamic role L12 plays in ribosomal translocation [13].

NMR is not the only solution-based technique to demonstrate conformational diversity. Pre-steady-state binding kinetics provided the first data indicating pre-existing isomers in equilibrium (for early examples see Refs [14,15]). These analyses reveal not only the existence of conformational diversity but also its kinetic parameters. Pre-steady-state kinetics can provide essential details regarding protein dynamism and the binding process – details that cannot be obtained by crystal-structure analysis alone [1,16]. Such analysis is, however, technically challenging and, unfortunately, unattainable in many cases. Fluorescence techniques, such as fluorescence resonance energy transfer (FRET), have also revealed the existence of multiple conformations in equilibrium [17]. Finally, single molecule techniques offer unique opportunities to monitor the activity and conformation of protein molecules. Such measurements have already shown that protein molecules exist in various states or conformations and that these exhibit different levels of activity [18,19].

Linking conformational diversity and functional diversity

An intriguing consequence of conformational diversity is that it provides a mechanism for functional diversity at the single-protein level. A protein that adopts several different conformations could, in principle, exert several different functions [20]. A significant number of proteins that exhibit multiple activities have been identified [21,22] but the mechanism behind this promiscuity or CROSS-REACTIVITY, in most cases, involves the same active-site configuration that confers the original activity. Even so, the same active-site features can be recruited for catalysis

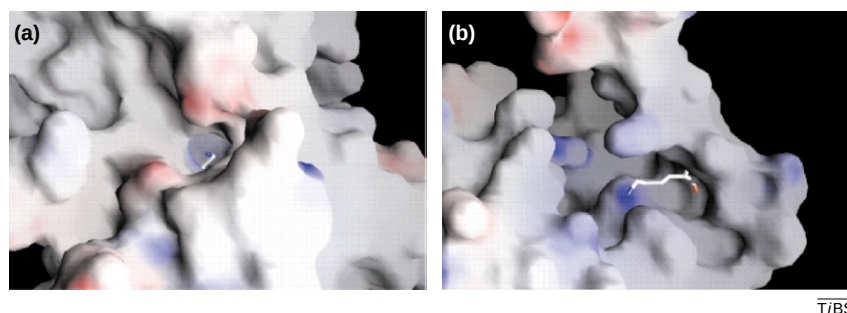


Fig. 1. Two proteins – an antibody (a) and human serum albumin (b) share a pattern of promiscuous binding and catalytic activities. Both proteins have a hydrophobic active site with a lysine side-chain that can act as a base (in catalyzing the Kemp elimination) or interact electrostatically with negatively charged ligands. The two proteins are completely unrelated in their sequence and structure (global and active site) and neither evolved towards the binding and catalytic promiscuous activities that they both exhibit [23,24]. This example indicates that active-site features might be promiscuously recruited in a mechanistic context that is completely different than the one they evolved for [22]. Reproduced, with permission, from Ref. [23].

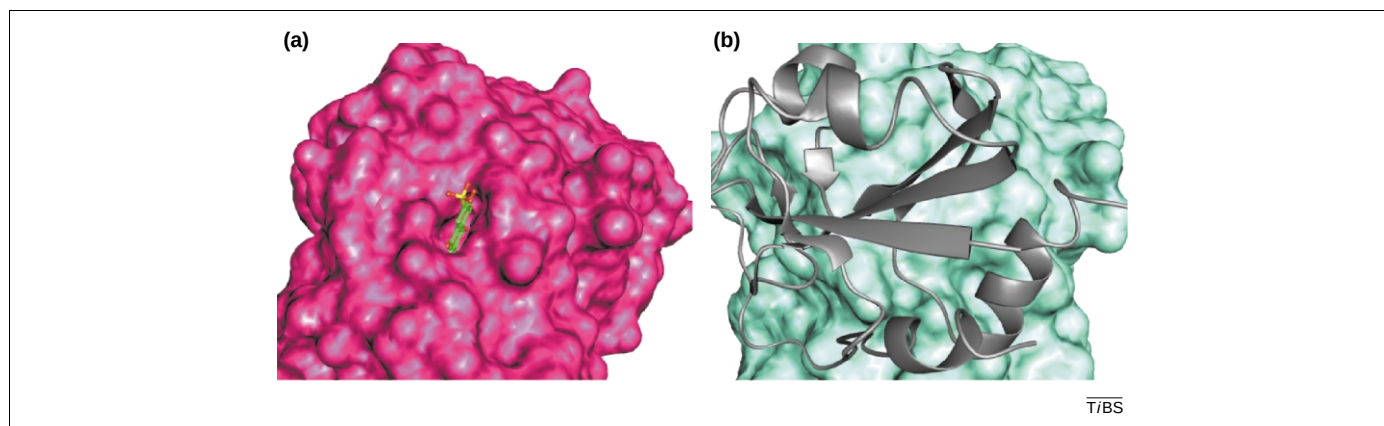


Fig. 2. Conformational diversity of antibody SPE7. Two different conformations of SPE7 allow it to bind to entirely unrelated antigens from across the structural spectrum (hapten to protein). One conformation (a) has a deep and narrow binding site that binds the hapten against which SPE7 was raised [dinitrophenol (DNP)]. The other conformation (b) has a flat binding surface that promiscuously binds a protein antigen (grey). The difference in the binding surfaces can be seen at the centre of each image; there is a deep cavity in the DNP-bound structure that is absent in the protein-bound structure. This structural analysis is supported by pre-steady-state kinetics that indicate two different pre-existing isomers. It appears that each ligand selects a different complementary isomer and, thereby, shifts the equilibrium in its direction [16].

by a very different mechanism. Therefore, the same activity can be found in completely unrelated proteins because they share similar active-site characteristics, for example, a hydrophobic active-site and a lysine side-chain that can act as a base [23,24] (Fig. 1). The identification of MULTI-SPECIFICITY or promiscuity that is mediated by conformational diversity is more difficult. The latter involves a more complex type of promiscuity whereby one protein (sequence) affords multiple structures and, thereby, multiple active-sites and functions. Nonetheless, such examples do exist. Conformational-diversity-mediated multi-specificity has been identified in an aminoglycoside kinase, in which two sub-sites are formed by the motion of a flexible active-site loop [25]. A conformational switch, involving the isomerization of a tyrosine side-chain has also been identified in the trypanosomal *trans*-sialidase; it allows this one enzyme to have two active-site configurations and thereby two activities (glycosyl hydrolase and transferase) [26,27]. Movement of side chains might also enable different binding specificities, for example, the anti-lysozyme antibody D1.3, which binds both lysozyme and an anti-idiotypic antibody through the rearrangement of several side chains [28].

The above examples are restricted to side-chain rearrangements or local adaptations, but we have recently demonstrated conformational changes that are far more significant, and involve different backbone as well as side-chain conformations. We have characterized a monoclonal antibody – dubbed SPE7 – that adopts radically different binding-site conformations and thereby binds multiple, unrelated antigens [16] (Fig. 2). The antibody was analyzed by both X-ray crystallography and pre-steady-state kinetics to reveal an equilibrium between at least two different pre-existing isomers that can both be crystallized in the absence of any ligand. One pre-existing isomer (present at ~80%) has a flat binding site, whereas the other isomer (that is 20% populated at equilibrium) possesses a narrow and deep binding site. The minor isomer binds a number of unrelated small aromatic ligands including the immunizing hapten. The dominant isomer binds a protein antigen that was identified by library

selection. The equilibrium between pre-existing isomers observed in solution (by pre-steady-state kinetics) coincides with the different crystal structures and indicates that they are not induced by crystal-packing forces.

These results have several important implications. First, they show at a detailed molecular level how conformational diversity provides a mechanism for functional diversity. Second, they illustrate that this mechanism permits a kind of promiscuity that is entirely unrelated to the primary activity. Indeed, the wide range of different antigens shown to bind a single antibody is unprecedented. Third, they demonstrate the operation of conformational diversity in the immune system in which, although sequence diversity is limited, unlimited functional diversity is required.

Conformational diversity and protein evolution

In the 1930s, Landsteiner and Pauling [44,45] proposed that proteins (antibodies in this case) can exist as an ensemble of isomers with different structures but with similar free energy. They realized that if each isomer was capable of binding to a different ligand, then functional diversity could go far beyond sequence diversity (Box 2). This idea, although considered outlandish for many decades, now proves to be remarkably prescient of the 'new view' of proteins. It also has intriguing implications with regard to protein evolution in general.

The immune response is a microcosm of protein evolution. An initial selection from a completely random (naïve) repertoire of germline antibodies of restricted diversity (10^6 – 10^8), followed by iterative rounds of mutation and selection (affinity maturation), lead to the evolution of highly specific and avid antibodies against almost any antigen presented to the immune system. Like any other evolutionary process, the immune response is totally reliant on the initial existence of at least one antibody in the germline repertoire that recognizes the antigen. Thus, the immune system is an excellent example of unlimited functional diversity evolving from limited sequence diversity (Box 2).

The limitation of sequence diversity is a fundamental issue in protein evolution [29]. The overall number of

Box 2. The Landsteiner–Pauling hypothesis

How does a limited repertoire of antibodies bind and thereby protect against an almost infinite diversity of invading antigens? The physicochemical homogeneity of serum immunoglobulins suggested that they have an identical or very similar sequence of amino acids. At the same time this was discovered, it became apparent that the immune system was capable of generating specific antibodies against almost any physiological or synthetic antigen. Karl Landsteiner (Fig. 1a) first suggested that this might be accomplished through ‘different ways of folding the same polypeptide chain’ [44]. Expanding on Landsteiner’s idea, Linus Pauling (Fig. 1b) proposed that proteins exist in an ensemble of conformations out of which functional ones are selected. Pauling supposed that parts of the immunoglobulin chains are highly flexible and ‘as a result of their amino-acid composition and order, having accessible a very great many configurations with nearly the same stability’ [45]. Put in Landsteiner’s elegant words: ‘one could surmise that changes of the amino acids make-up, in conjunction with different folding, would afford the widest scope for variation’ [44]. With the establishment of the clonal-selection theory it was realized that the combinatorial arrangement of different antibody gene segments generates a vastly diverse repertoire of antibodies from a small number of genes. But the primary germline immune response is still limited by the number of circulating B-cells, which can be as few as 10^6 . How does a limited repertoire of sequences cover such a vast area of antigen space?

In 1994, the idea of conformational diversity was revived by Foote and Milstein [46], who demonstrated that pre-existing isomers are readily observed in antibodies. Later work supported the proposal that germline antibodies complex antigen through one of many pre-existing isomers [47], and that during affinity maturation, mutations stabilizing the binding isomer are selected [48]. However, a direct demonstration that a single antibody can use different pre-existing isomers to bind different antigens has only recently been provided [16] (Fig. 2).

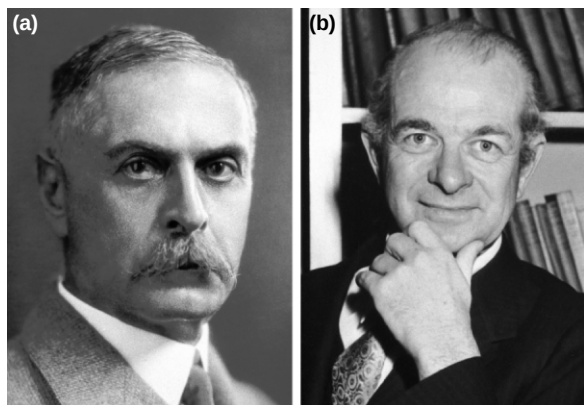


Fig. 1. (a) Karl Landsteiner, Nobel Laureate Physiology or Medicine 1930. (b) Linus Carl Pauling, Nobel laureate Chemistry. Reprinted, with permission, from The Nobel Foundation. ©The Nobel Foundation.

permutations for a 100-amino-acid protein is $\sim 20^{100}$ (or $\sim 10^{130}$), whereas the total number of different genes in the biosphere is >100 orders of magnitude smaller ($\sim 10^{13}$) [30]. Therefore, the evolutionary process samples an impossibly small fraction of sequence space but can somehow achieve great, perhaps unrestricted, functional diversity. Even if we assume that the diverse array of today’s proteins has evolved from a handful of primordial proteins, it still remains to be explained how these primordial proteins evolved in the first place. Moreover, although gene duplication explains how existing genes are freed to evolve new functions, it does not explain how these

new functions are acquired. Can mutation and drift explain their emergence? Can they account for the evolution of new proteins in nature on a decade timescale (for examples of recently evolved proteins see [22,31,32])? After all, Darwinian evolution relies on the evolving gene to confer a selective advantage from the first stage of its development. Therefore, here exists a ‘catch’: a new function can evolve only if it is already present to some extent, otherwise the organism would not survive and the evolutionary process cannot begin. Thus, it seems an attractive proposition that nature evolves new functions by recruiting existing promiscuous activities and gradually improving them. That promiscuity plays a role in protein evolution is not a new idea. Jensen [33] suggested that, under changing circumstances, promiscuous activity in an existing protein (or substrate ambiguity as he called it) might endow the organism a selective advantage and thereby enable its survival and further evolution. Gene duplication and mutation could then provide the increased genetic diversity to drive the evolution of the diverging new protein [21,33].

To establish a role for promiscuity in protein evolution it needs to be shown that promiscuous activities are common in proteins and that their magnitude can confer a selective advantage. Although several promiscuous proteins have been reported [21,22], these data are largely sporadic and accidental in origin. Moreover, promiscuity appears to be at odds with our view of the precise and specific nature of protein action – how can proteins be highly specific and proficient, and at the same time promiscuous? The precise positioning of active-site residues is thought to be the prime contributor to enzymatic catalysis. It is true, of course, that promiscuous activities can be explained by shared active-site chemistry, or by some structural or chemical resemblance between the ‘natural’ substrate (for which the protein has evolved) and the promiscuous protein [21,23] (Fig. 1). However, there are proteins with promiscuous activities that appear to be completely unrelated to the original activity. How can, for example, a single antibody bind to a series of unrelated aromatic ligands and to a protein antigen (Fig. 2)? MOONLIGHTING proteins comprise another striking example in which two completely different functions are conferred by one protein [22,34].

The ‘new view’ of proteins – whereby one sequence adopts multiple structures – allows functional and structural diversities to be interlinked to enable proteins to rapidly evolve new functions and structures. In many respects, conformational diversity and functional promiscuity can be considered to be traits of the ability of a protein to evolve [35]. When combined with the classical mechanisms of gene duplication mutation and selection, conformational diversity provides a powerful mechanism to facilitate the evolution of new functions (Box 3). Interestingly, simulations of RNA evolution also indicate that, although structural plasticity is costly (the more conformations a sequence assumes, the less time it spends in any one of them), reduction of plasticity leads to genetic canalization – namely to a dramatic decrease in variability and evolvability, and a complete halt of the

Box 3. Enzyme evolution mediated by conformational diversity and functional promiscuity: a proposed model

The proposed model suggests that conformational diversity and functional promiscuity are evolvability traits that enable existing enzymes to rapidly evolve new activities (Fig. 1). The predominant conformation of an existing enzyme catalyses the conversion of a native substrate. An alternative conformation, that is scarcely populated, has the potential to bind and promiscuously catalyze another substrate. Initially, this secondary activity provides only a limited fitness advantage because binding of the primary substrate will sequester most of the protein. This secondary activity can be improved through mutation, yet only to a limited extent, as such mutations might also decrease availability and activity for the primary substrate. However,

following a gene duplication event, one gene copy of the enzyme is free to evolve without compromising the original activity. Mutations in the duplicated gene could improve catalytic efficiency towards the new substrate by optimizing active site chemistry and also by stabilizing the promiscuous conformation. Following iterative rounds of mutation and selection, it is highly likely that the new enzyme will have completely lost the original activity and conformation. In fact, mutation and selection may lead to the new enzyme acquiring new conformations and consequently new promiscuous activities. These in turn could serve as the starting point for the evolution of yet more new enzymes.

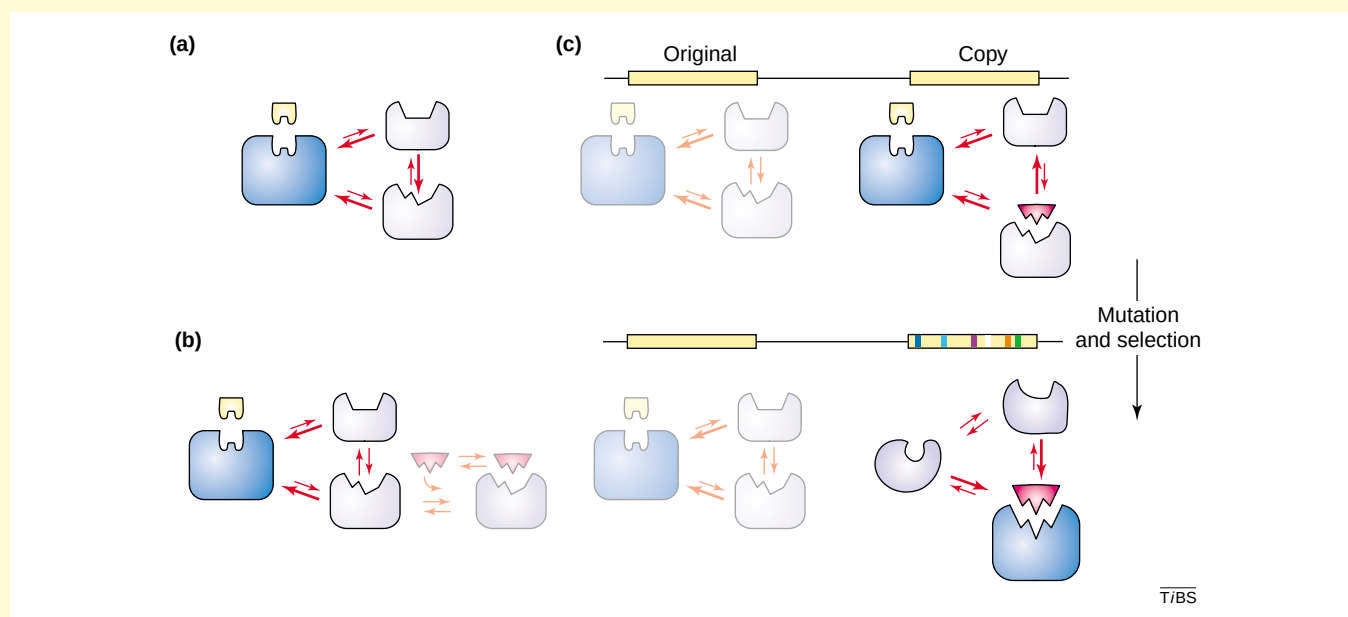


Fig. 1. Proposed model of enzyme evolution mediated by conformational diversity and functional promiscuity. (a) The enzyme is in equilibrium between different conformations. The native substrate (yellow) selects the dominant conformer (dark blue) and, thus, enzyme activity confers selective advantage. (b) An alternative conformation potentiates the binding of a second substrate (pink). The secondary activity confers a limited selective advantage under changing environmental conditions. (c) Gene duplication enables one copy to evolve improved activity with the promiscuous substrate while the original gene maintains its original function.

evolutionary process [36]. It is likely that proteins exhibit a similar pattern of behaviour.

The co-evolution of fold and function

The proposed model (Box 3) describes how existing genes could diverge to evolve new functions, but it can also be extended to explain how primordial protein precursors evolved in the first place. The emergence of the first functional proteins is a complete mystery. It is generally assumed that proteins exert function only when folded into a well-defined 3D structure. However, fold alone provides no selective advantage and, thus, cannot evolve independently. We are therefore left with an unlikely scenario in which random sampling must simultaneously produce a sequence with a discrete 3D structure and a selectable function.

The Pauling hypothesis regarding antibodies (Box 2) and the 'new view' of proteins provide a mechanism for the co-evolution of fold and function – both surmise that a defined 3D structure is not a prerequisite for function. The concept of functional non-structured proteins is best illustrated with intrinsically disordered proteins [8–10].

In such proteins, the binding and folding processes are linked [37]. During the early stages of evolution, a scenario of co-evolution of fold and function could be based on a selection for function from a repertoire of random polypeptides in perpetual order–disorder transition [38]. These loosely held structures would be based around, for example, 25–30-amino-acid loops [39] or other 'seeding' elements of structure. Such polypeptides could dynamically sample many different structural isomers of similar and relatively low free energy. The resulting conformational diversity, when multiplied by sequence diversity, would give vastly larger repertoires out of which functional polypeptides could be selected. A polypeptide, one of the transient conformers of which enables binding and catalytic conversion of a useful metabolite, would then be selected. Function would also shift the structural equilibrium, leading to higher occupancy of the active conformer. Driven by mutation and selection, these polypeptides could rapidly evolve via two parallel routes. First, the population time of the active conformer could be increased by mutations that progressively create discrete folding pathways, which would eventually lead to stable,

well-defined folds. Second, the activity of the protein could be improved directly by mutations at the active site itself. Theoretical support for this route comes from simulations of the evolutionary process that start with a fraction of random-sequence polypeptides and demonstrate that fold and function can co-evolve [38].

Concluding remarks

The 'new view' of proteins has prompted the revision of many facets of protein science. Here, we have outlined the intriguing implications that this 'new view' might have for protein evolution. The hypotheses described are supported by many properties of today's proteins, including the recent demonstration of a linkage between conformational diversity and multi-specificity in antibodies [16] (Fig. 2). Moreover, although present-day enzymes do not necessarily retain the activities of their evolutionary precursors, it appears that enzymes that retain both the original and a newly evolved activity can be readily generated in the test tube [40,41]. It is not known, however, whether these multi-specific enzymes exhibit increased conformational plasticity. Thus, a direct proof for the role of functional diversity in protein evolution is yet to be obtained. Finally, the exploitation of promiscuous activities as starting points for the directed evolution of new proteins, although scarce at present [42], could prove a fruitful approach in the future.

Acknowledgements

Financial support by the Wellcome Trust, the EU (through the END-IRPRO network) and the Israel Science Foundation is gratefully acknowledged.

References

- Lindner, A.B. *et al.* (1999) Conformational changes affect binding and catalysis by ester-hydrolysing antibodies. *J Mol Biol* 285, 421–430
- Choy, W.Y. and Forman-Kay, J.D. (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J. Mol. Biol.* 308, 1011–1032
- Macol, C.P. *et al.* (2001) Direct structural evidence for a concerted allosteric transition in *Escherichia coli* aspartate transcarbamoylase. *Nat. Struct. Biol.* 8, 423–426
- Volkman, B.F. *et al.* (2001) Two-state allosteric behavior in a single-domain signaling protein. *Science* 291, 2429–2433
- Cordes, M.H. *et al.* (2000) An evolutionary bridge to a new protein fold. *Nat. Struct. Biol.* 7, 1129–1132
- Chien, P. and Weissman, J.S. (2001) Conformational diversity in a yeast prion dictates its seeding specificity. *Nature* 410, 223–227
- Derreumaux, P. (2001) Evidence that the 127–164 region of prion proteins has two equi-energetic conformations with β or α features. *Biophys. J.* 81, 1657–1665
- Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *J. Mol. Biol.* 293, 321–331
- Uversky, V.N. (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 11, 739–756
- Tomba, P. (2002) Intrinsically unstructured proteins. *Trends Biochem. Sci.* 27, 527–533
- Hammes, G.G. (2002) Multiple conformational changes in enzyme catalysis. *Biochemistry* 41, 8221–8228
- Spoerner, M. *et al.* (2001) Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4944–4949
- Wahl, M.C. *et al.* (2000) Flexibility, conformational diversity and two dimerization modes in complexes of ribosomal protein L12. *EMBO J.* 19, 174–186
- Lancet, D. and Pecht, I. (1976) Kinetic evidence for hapten-induced conformational transition in immunoglobulin MOPC 460. *Proc. Natl. Acad. Sci. U. S. A.* 73, 3549–3553
- Fersht, A.R. and Requena, Y. (1971) Equilibrium and rate constants for the interconversion of two conformations of chymotrypsin. The existence of a catalytically inactive conformation at neutral pH. *J. Mol. Biol.* 60, 279–290
- James, L. *et al.* (2003) Antibody multi-specificity mediated by conformational diversity. *Science* 299, 1362–1367
- Chakraborty, S. *et al.* (2001) Structure and dynamics of the α -lactalbumin molten globule: fluorescence studies using proteins containing a single tryptophan residue. *Biochemistry* 40, 7228–7238
- Xie, S. (2001) Single-molecule approach to enzymology. *Single Mol.* 4, 229–236
- Edman, L. and Rigler, R. (2000) Memory landscapes of single-enzyme molecules. *Proc. Natl. Acad. Sci. U. S. A.* 97, 8266–8271
- Ma, B. *et al.* (2002) Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci.* 11, 184–197
- O'Brien, P.J. and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* 6, R91–R105
- Copley, S. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol.* 7, 1–8
- James, L.C. and Tawfik, D.S. (2001) Catalytic and binding poly-reactivities shared by two unrelated proteins: the potential role of promiscuity in enzyme evolution. *Protein Sci.* 10, 2600–2607
- Hollfelder, F. *et al.* (1996) Off-the-shelf proteins that rival tailor-made antibodies as catalysts. *Nature* 383, 60–63
- Fong, D. and Berghuis, A. (2002) Substrate promiscuity of an aminoglycoside antibiotic resistance enzyme via target mimicry. *EMBO J.* 21, 2323–2331
- Buschiazio, A. *et al.* (2002) The crystal structure and mode of action of *trans*-sialidase, a key enzyme in *Trypanosoma cruzi* pathogenesis. *Mol. Cell* 10, 757–768
- Colman, P.M. and Smith, B.J. (2002) The trypanosomal *trans*-sialidase: two catalytic functions associated with one catalytic site. *Structure* 10, 1466–1468
- Tello, D. *et al.* (1994) Structural and physicochemical analysis of the reaction between the anti-lysozyme antibody D1.3 and the anti-idiotypic antibodies E225 and E5.2. *J. Mol. Recognit.* 7, 57–62
- Griffiths, A.D. and Tawfik, D.S. (2000) Man-made enzymes – from design to *in vitro* compartmentalisation. *Curr. Opin. Biotechnol.* 11, 338–353
- Burton, S.G. *et al.* (2002) The search for the ideal biocatalyst. *Nat. Biotechnol.* 20, 37–45
- Buchbinder, J.L. *et al.* (1998) Biochemical characterization and crystallographic structure of an *Escherichia coli* protein from the phosphotriesterase gene family. *Biochemistry* 37, 5096–5106
- Seffernick, J.L. and Wackett, L.P. (2001) Rapid evolution of bacterial catabolic enzymes: a case study with atrazine chlorohydrolase. *Biochemistry* 40, 12747–12753
- Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* 30, 409–425
- Jeffery, C.J. (1999) Moonlighting proteins. *Trends Biochem. Sci.* 24, 8–11
- Joyce, G.F. (1997) Evolutionary chemistry: getting there from here. *Science* 276, 1658–1659
- Ancel, L.W. and Fontana, W. (2000) Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool.* 288, 242–283
- Dunker, A.K. and Obradovic, Z. (2001) The protein trinity – linking function and disorder. *Nat. Biotechnol.* 19, 805–806
- Yomo, T. *et al.* (1999) Gradual development of protein-like global structures through functional selection. *Nat. Struct. Biol.* 6, 743–746
- Trifonov, E.N. *et al.* (2001) Distinct stages of protein evolution as suggested by protein sequence analysis. *J. Mol. Evol.* 53, 394–401
- Matsumura, I. and Ellington, A.D. (2001) *In vitro* evolution of β -glucuronidase into a β -galactosidase proceeds through non-specific intermediates. *J. Mol. Biol.* 305, 331–339
- Jurgens, C. *et al.* (2000) Directed evolution of a (β/α) 8-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9925–9930
- Buchholz, F. and Stewart, A.F. (2001) Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nat. Biotechnol.* 19, 1047–1052

- 43 Dill, K.A. and Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4, 10–19
- 44 Landsteiner, K., (1936) *The Specificity of Serological Reactions*, reprinted 1962, Dover Publications
- 45 Pauling, L. (1940) A theory of the structure and process of formation of antibodies. *J. Am. Chem. Soc.* 62, 2643–2657
- 46 Foote, J. and Milstein, C. (1994) Conformational isomerism and the diversity of antibodies. *Proc. Natl. Acad. Sci. U. S. A.* 91, 10370–10374
- 47 Wedemayer, G.J. *et al.* (1997) Structural insights into the evolution of an antibody combining site. *Science* 276, 1665–1669
- 48 Foote, J. and Winter, G. (1992) Antibody framework residues affecting the conformation of the hypervariable loops. *J. Mol. Biol.* 224, 487–499

Articles of interest in other Current *Trends* journals

Never say never. The NIMA-related protein kinases in mitotic control

Matthew J. O'Connell, Michael J.E. Krien and Tony Hunter
Trends in Cell Biology, 2003, 13:5:221–228

14-3-3 Proteins find new partners in plant cell signaling

Michael R. Roberts
Trends in Plant Science, 2003, 8:5:218–22

Class II histone deacetylases: versatile regulators

Eric Verdin, Franck Dequiedt and Herbert G. Kasler
Trends in Genetics, 2003, 19:5:286–293

RNA polymerase holoenzyme: structure, function and biological implications

Sergei Borukhov and Evgeny Nudler
Current Opinion in Microbiology, 2003, 6:2:93–100

Definitions of enzyme function for the structural genomics era

Patricia C Babbitt
Current Opinion in Chemical Biology, 2003, 7:2:230–237

Peptide folding simulations

S Gnanakaran, Hugh Nymeyer, John Portman, Kevin Y Sanbonmatsu and Angel E García
Current Opinion in Structural Biology, 2003, 13:2:168–174

The methyl-CpG binding domain and the evolving role of DNA methylation in animals

Brian Hendrich and Susan Tweedie
Trends in Genetics, 2003, 19:5:269–277

Cell polarity: Par6, aPKC and cytoskeletal crosstalk

Sandrine Etienne-Manneville and Alan Hall
Current Opinion in Cell Biology, 2003, 15:1:67–72

Replication of human hepatitis delta virus: recent developments

John M. Taylor
Trends in Microbiology, 2003, 11:4:185–190