

***In vitro* selection of nucleic acids and proteins: what are we learning?**

Richard W Roberts* and William W Ja

For almost a decade, *in vitro* selection experiments have been used to isolate novel nucleic acids, peptides and proteins according to their function. Selection experiments have altered our perception of molecular mimicry and catalysis, and they appear to be more facile than rational design at generating biopolymers with desired properties. New methods that have been developed improve the power of functional strategies in ways that nature has already discovered – by expanding library size and facilitating the recombination of positive mutations. Recent structural information on a number of selected and evolved molecules highlights future challenges for design via rational approaches.

Addresses

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
*e-mail: rroberts@its.caltech.edu

Current Opinion in Structural Biology 1999, **9**:521–529

<http://biomednet.com/elecref/0959440X00900521>

© Elsevier Science Ltd ISSN 0959-440X

Abbreviations

EF	elongation factor
FMN	flavin mononucleotide
HSQC	heteronuclear single quantum coherence
PCR	polymerase chain reaction
scFv	single-chain variable domain fragment
TSA	transition-state analog
VEGF	vascular endothelial growth factor

Introduction and terms of art

This review will focus on both recent results and some general conclusions from selection experiments. *In vitro* selection is a process that involves sieving a pool of molecules through repeated cycles of enrichment and amplification to recover only those molecules with a desired functional property. When performed with RNA or DNA, this process is conducted entirely *in vitro* using a combination of PCR, transcription, selection and cDNA synthesis [1]. When performed with peptides and proteins, the process may contain an *in vivo* step (as in phage display [2]) or be performed entirely *in vitro* (as in ribosome display [3] or mRNA–protein fusion systems [4,5,6]). In all cases, *in vitro* selection experiments have the functional information topologically linked to the genetic information. Indeed, for nucleic acids, this information is contained in one and the same molecule.

Generally, *in vitro* selection experiments allow all the molecules in the pool to be tested for function at the same time. By comparison, screening experiments require that each pool member be queried individually. Thus, selection experiments generally allow much larger libraries to be surveyed than screening experiments. On

the other hand, the activity of individual members in a selection experiment is not known without cloning and testing each individual.

Functional RNA and DNA

Aptamers and mimicry

A decade ago, the proposal that either an RNA or a DNA molecule could be found that mimicked the structure and/or function of a protein was heretical at best. The isolation of RNA and DNA aptamers (from *aptus*, ‘to fit’ [7]) that bind both small molecules and proteins has done much to advance the general notion that one heteropolymer (e.g. RNA or DNA) may be used to mimic the function of another (e.g. a protein).

RNA and DNA aptamers have been isolated that bind nucleic-acid-binding proteins (e.g. HIV reverse transcriptase and HIV Rev), as well as small molecules and proteins that have no natural nucleic-acid-binding activity (e.g. thrombin, vascular endothelial growth factor [VEGF] and vasopressin) [8]. These aptamers can be adapted for use *in vivo* if they are protected from degradation by the addition of phosphorothioates or by the substitution of the 2'-OH with 2'-NH₂ (e.g. [9]). Alternatively, if the mirror image of the original protein target is used with natural RNA or DNA, a mirror image, biologically stable oligonucleotide may be constructed that binds the natural target [10].

The ability of nucleic acids to mimic proteins in selection experiments is echoed in examples we now know from natural systems. One of the best examples exists in translation. There, RNA and the protein factors involved in elongation (EF-Tu•GTP•tRNA and EF-G•GTP) and termination (RF-1, RF-2 and RF-3•GTP) seem to have the same general structure [11,12]. Similarly, in eukaryotic translation initiation, a portion of the important factor eIF4G can be mimicked by an internal ribosome entry sequence (IRES) [13,14]. Hence, a great many protein motifs can be functionally mimicked by folded RNA structures. The diversity of this mimicry in both natural and non-natural structures provides experimental support for the theory that the transition from an ‘RNA world’ to a biochemistry that is dominated by proteins is possible.

Mimicry does, however, have its bounds. It follows that, if aptamers could bind any small molecule target, catalysts could be discovered by selecting library members that bind transition-state analogs (TSAs). This approach has been used to great advantage in the immune selection of catalytic antibodies [15]; however, success has been limited using RNA and DNA aptamers. The first example was a ribozyme that enhanced the isomerization of substituted

biphenyls by 88-fold over background [16]. Subsequently, RNA and DNA aptamers that bind *N*-methylmesoporphyrin (NMMP) were found to stimulate porphyrin metalation [17,18]. The scarcity of enzymes compared with catalytic antibodies may be a result of the different mechanisms used by nucleic acids to recognize small molecules compared with those used by proteins (see discussion below). On the other hand, the relative paucity of TSA-binding protein catalysts that have been isolated via phage display may imply that *in vitro* selection techniques need to be improved in order to match the ability of the immune system.

Amide- and peptide-bond-forming catalysts

One central tenet of the ‘RNA world’ hypothesis is that the mother of all enzymes, the ribosome, contains RNA in its catalytic heart. In addition to investigations that point to rRNA as being essential for ribosome activity [19–21], a number of laboratories have set out to test whether or not ribozymes could be found that are capable of catalyzing amide-bond formation. The first success involved an acyl transferase ribozyme that was capable of forming an amide bond between a 5′-NH₂-RNA and *N*-biotinyl methionine attached to the 3′ end of a hexanucleotide ($k_f \sim 0.6 \text{ min}^{-1}$) [22]. Recently, Eaton and co-workers [23] reported the isolation of amide-bond-forming ribozymes that do not require an oligonucleotide tether to the substrate, a 5′-AMP activated ester of biotin. These ribozymes do require 5′-imidazole uridine, which was incorporated uniformly in place of uridine in the RNA pool.

Although neither of these ribozymes was reported to contain sequence homology to natural RNA sequences, one of the peptidyl transferase ribozymes isolated by Zhang and Cech did [24,25•]. The clone 25 ribozyme contains an internal loop with regions of sequence identity to the central loop of domain V of 23S rRNA, a region that is central to the peptidyl transferase function of the ribosome. Surprisingly, aptamers that bind the peptidyl transferase inhibitor CcdAp–Puromycin share identity with the peptidyl transferase ribozyme and with 23S rRNA [26]. Despite the recent retraction of work arguing that domain V RNA transcripts have peptidyl transferase activity [27], the selection data represent proof that RNA has the potential to carry out a variety of amide-bond-forming reactions.

New ribozyme reactions

A number of ribozymes with new activities have been isolated recently, including molecules that are capable of catalyzing the synthesis of a nucleotide [28••], enzymes selected to mimic reactions in the spliceosome [29•] and ribozymes capable of performing a Diels–Alder cycloaddition reaction [30]. In the first example, Unrau and Bartel [28••] uncovered ribozymes that are capable of joining a tethered 4-thio-uridine nucleotide to pRpp (5′ phosphoribosyl- α -pyrophosphate), the activated ribose used in nucleotide biosynthesis ($k_{\text{cat}} = 0.13 \text{ min}^{-1}$). This reaction is almost identical to the natural reaction carried out by uracil

phosphoribosyltransferase (UPRT) to couple uracil and pRpp, making uridine 5′ phosphate.

In the second example, a selection was designed to test whether modern spliceosomal RNA U2/U6 could be used to resurrect RNA-based spliceosomal activity. Each of the four ribozyme classes found performs some kind of cleavage reaction, followed by a ligation; however, none of the selected molecules produces an activity that is similar to that of the spliceosome [29•]. Interestingly, the templates generated by the ribozymes uncovered a previously unknown activity for reverse transcriptase, nontemplated jumps in sequence.

Adding functional group diversity – ‘If I only had a histidine’

A number of laboratories have recently worked to expand the chemical diversity of RNA and DNA, either through covalent modification or through the addition of cofactors and ions to the selection mixture. Examples of the first category include ribozymes that catalyze carbon–carbon bond formation through a Diels–Alder reaction [30], as well as the previously mentioned amide-bond-forming ribozymes [23] that contain 5-pyridylmethylcarboxamid uridine and 5-imidazole uridine, respectively. Examples of the second case include the routine addition of low (micromolar to millimolar) concentrations of divalent cations, including Mg²⁺, Ca²⁺, Zn²⁺, Mn²⁺, Cu²⁺ and even Pb²⁺, or cofactors, such as histidine [31•]. The importance of these prosthetics can be seen in that their omission abrogates function.

DNA catalysts

Although DNA plays a central role in biological information storage, it is a functionally impoverished biopolymer compared with either proteins or RNA. Despite its limitations, the list of DNA catalysts has grown significantly in the past few years (see Table 1 in [32]; [33]). It is not surprising that many of the DNA catalysts isolated require a cofactor of some kind. It is surprising that the highest catalytic efficiency, $10^9 \text{ M}^{-1}\text{s}^{-1}$ (although not the highest k_{cat}), is held by the 10–23 RNA-cleaving DNA enzyme [34,35]. To date, the scope of DNA-based catalysts has significant diversity, including ligation [36], RNA phosphoester cleavage [34,37–39], DNA cleavage [40], porphyrin metalation [18], peroxidase activity [41•], fluorophore oxidation [42] and 5′-DNA phosphorylation [43•]. It is interesting to note that G-quartets seem to play a central role in the folded structure of many DNA catalysts, as well as of DNA aptamers.

Engineering allosteric enzymes and templates

One potentially powerful application of nucleic acid aptamers and catalysts is the engineering of allosterically controlled enzymes and templates. Breaker and co-workers [44,45•] have shown that, through combinatorial engineering, chimaeras between either the ATP or FMN aptamer with the hammerhead ribozyme may be allosterically regulated to function only when they bind their

Table 1

Peptides evolved from randomized sequences.

Initial pool*	Diversity [†]	Target	Result	References
MVSKGEEX ₂₇ DAQAPKA	2 × 10 ¹³	c-myc mAb	Epitope X(Q/E)XLISEXX(L/M)	(a)
Pool: X ₇ CX ₄ CX ₇ X ₇ CX ₅ CX ₆ X ₆ CX ₆ CX ₆ X ₆ CX ₇ CX ₅ X ₅ CX ₈ CX ₅ X ₅ CX ₉ CX ₄ X ₄ CX ₁₀ CX ₄	5 × 10 ⁸ (each) 3.5 × 10 ⁹ (total)	VEGF	Mimic (i) RGWVEICVADDNGMCVTEAQ (ii) GWDECDVARMWEWECFAGV	[69**]
X ₄ CX ₂ GPX ₄ CX ₄	5 × 10 ⁸	VEGF	Mimic GERWCFDGLTWVCGEES	[69**]
X ₁₂ GGGS	1.9 × 10 ⁹	Paclitaxel	Epitope HTPHP	[107]
SRX ₁₂ SR	1 × 10 ⁹	Troponin C	Epitope (V/L)(D/E)XLKXXLXXLA	[75]
SADGAX ₁₅ GAAGA	2.5 × 10 ⁸	CTLA4 mAb	Unknown (i) GFVCSGIFAVGVGRC (ii) APGVRLGCAVLGRYC	[108]
ADGAX ₆ GAAG	4 × 10 ⁷	CSF IgG	Epitope RRPFFX	[109]
X ₁₀	2 × 10 ⁹ (phage)	li peptide	Epitope WFSWGFQWW	[110]
LEX₆TS	2.8 × 10 ⁸ (clones)	PAI-1	Unknown GQFWHLTSMGSGSYFLEPFDLIST SQQRNIPLEIRDADT	[111]
LEX₁₅TS	1.4 × 10 ⁷ (clones)	PAI-1	Mimic PVSQFVFLCGHQPCFTSEHAHDVP DPAPPHHPLELITGROATPISVGMS	[111]

*Flanking sequences shown where reported. [†]Where further rounds differ, values given refer to only the initial round of selection. Values represent diversity, except where indicated (clones = number of transformants; phage = number of input phage). Sequences in bold type represent amino acids encoded by DNA sequences

corresponding to restriction sites. (a) P Burgstaller, S Hale, M Wright, RW Roberts, R Liu, JW Szostak, RW Wagner, unpublished observations. CSF, cerebrospinal fluid; CTLA4, anti-T lymphocyte co-stimulatory molecule; li, invariant chain peptide from cytosolic tail of MHC class II; PAI-1, plasminogen activator inhibitor 1.

respective small molecule. Werstuck and Green [46*] have demonstrated that allosteric mRNA templates may be generated by inserting a drug-binding aptamer into the 5'-untranslated region (5'-UTR) of an mRNA. This insertion allows both *in vitro* and *in vivo* translation to be repressed in the presence of a small molecule.

The thirst for structural information

Logically, structural analysis provides a clear path to understanding the mechanisms of recognition and catalysis performed by functional nucleic acids. Indeed, Jencks [47] argued that, for protein enzymes, the most striking generalization was that many of the expectations based on chemical studies were simply confirmed once crystal structures were in hand. In this vein, the structure of aptamers in solution solved using NMR spectroscopy has generally provided significant insight into the recognition process, indicating the presence of compact binding pockets for small molecules (for reviews, see [48–50]). Analysis of

three pockets (present in ATP, FMN and arginine aptamers) indicates that binding appears to provide somewhat less of a tight fit than protein sites and relies heavily on stacking, as opposed to hydrogen bonding [50].

The path has been somewhat bumpier when examining nucleic acid catalysts. The central difficulty is crystallizing either RNA or DNA enzymes in their active form. In the crystal structure of the 10–23 DNAzyme, the molecule has rearranged to form a dimer that is not likely to be the active conformation [51*]. The structure of the lead-dependent ribozyme crystallizes in two forms, one of which is 'pre-catalytic' and rationalizes the cleavage chemistry [52**]. Difficulties are seen in structural analyses of natural ribozymes as well. Much has been made of the conundrum that the hammerhead ribozyme structures do not rationalize the cleavage chemistry [53,54], whereas larger RNA catalysts, such as the group I intron catalytic core [55] and the hepatitis delta virus ribozyme (after cleavage) [56],

Table 2**Proteins displayed on phage.**

Protein modified	Clones*	Target(s)	Result	References
Heregulin (nine sections mutagenized separately)	$1.0\text{--}6.4 \times 10^8$	ErbB3 receptor	>50-fold higher affinity	[104**]
Knottins (CBD)	5.5×10^8	Cellulose α -Amylase Alkaline phosphatase β -Glucuronidase	Epitope Unsuccessful $K_d = 10 \mu\text{M}$ Unsuccessful	[87**]
Lipocalin (16 residues in binding site)	3.7×10^8	Fluorescein	$K_d = 35.2 \text{ nM}$ (> 10^2 -fold higher affinity)	[88*]
hGH (site 2: four sections mutagenized separately)	$1.3\text{--}7.3 \times 10^7$	hGH receptor	$K_d = 0.16 \text{ nM}$ (38-fold higher affinity)	[112]

*Values represent the number of transformants in the first round. CBD, C-terminal cellulose-binding domain; hGH, human growth hormone.

seem to provide better support for known biochemical data. Even at high resolution, several structures (before, during and after catalysis) will probably be necessary to fully understand the details of the cleavage chemistry. Overall, the structural work implies movement during the catalytic process [51*] and that this movement may be more significant in smaller ribozymes.

Peptides and proteins

Selections using phage display

The most common technique for the *in vitro* selection of peptides and proteins is currently the phage display system developed by George Smith. Sequences are expressed either in multicopy format (~1–2700 copies/phage) as N-terminal fusions to gene VIII or in 'single' copy format (0–5 copies/phage) in the N terminus of gene III on the surface of filamentous bacteriophage (for a recent review, see [2,57*]). Molecules are selected by affinity panning on immobilized substrates or receptors, and the resulting phage are mixed with bacteria to produce progeny phage for subsequent rounds. In a typical experiment, each sequence in the library is represented by many phage (>1000), with a total library size of approximately $10^7\text{--}10^9$ members. High multiplicity is essential, because the fraction of phage recovered is generally very low (typically between 0.01% and 0.0001% of the input population). This percentage can be even lower using a variation of phage display dubbed selectively infective phage or 'SIP' (reviewed in [58]). In this selection procedure, ligand–target binding is required for infectivity using bipartate gene III fusions containing the library and target.

New methods

Until recently, the library size used for the *in vitro* selection of peptides and proteins has been limited to only about one billion molecules, mainly due to the transfection limit in bacteria. In addition, phage display is not easily portable to other powerful techniques, including *in vitro* mutagenesis and recombination (see [59,60**,61*] and references

therein). By comparison, the limit with nucleic acid selections is $10^{15}\text{--}10^{16}$ sequences, because all the steps in the process are performed *in vitro*. This difference is important because the computational power of a combinatorial experiment is directly proportional to the number of molecules that can be screened. Two recently developed methods, ribosome display and mRNA–protein fusions, provide the ability to screen libraries of more than 10^{11} or 10^{13} peptides or proteins, respectively, by keeping translated proteins linked to their own mRNA (see [3,6] for reviews). A third approach, whereby information is linked via compartmentalization, has also been reported [62*].

All of these strategies utilize cell-free *in vitro* translation systems, thereby avoiding any transformation or other constraints imposed by the host cell. In ribosome display, translated proteins and mRNA remain complexed with the ribosome as a result of lowering the temperature and increasing the Mg^{2+} concentration [63]. Selections using ribosome display recently have demonstrated the affinity maturation of scFvs (single-chain variable domain fragments) from immunized libraries [64*]. One limitation of ribosome display is that libraries must be screened in the presence of the entire ribosome under conditions in which the ternary mRNA–ribosome–peptidyl-tRNA complex is stable (high Mg^{2+} and low temperature). A novel method for producing covalently linked mRNA–protein fusions has been recently introduced [4,5**]. In this approach, a mRNA library is tagged with a 3'-puromycin. After translation, the puromycin enters the peptidyl transferase site and becomes covalently linked with the protein that it encodes. This chemically robust linkage allows selections to be performed under a wide variety of conditions. Recently, this technique has been used to generate and screen a library containing approximately 2×10^{13} different peptides for binding and evolution to the anti-c-myc antibody 9E10, with enrichment factors of more than 200 per round (Table 1) (P Burgstaller, S Hale, M Wright, RW Roberts, R Liu, JW Szostak, RW Wagner, unpublished observations).

Several other technical advances are also expected to enhance the power of *in vitro* selection techniques, including improved *in vitro* recombination and mutagenesis, and interaction mapping by NMR spectroscopy. New methods for *in vitro* recombination may be technically simpler and/or more effective than the original DNA shuffling technique devised by Stemmer [65]. Among these is family shuffling [60**], whereby diversity is created by recombining two or more homologous genes. This has the advantage of creating diversity that has already been prescreened by nature to weed out mutations that are deleterious to a particular fold, generating libraries that are much richer in functional proteins. Alternatives to the original DNase-I-based shuffling protocol have also been developed, including random priming recombination (RPR) [66] and the staggered extension process [61*]. Random elongation mutagenesis (REM), the addition of a partially randomized sequence to the C terminus of the protein, is unlikely to change catalytic properties, but may be a simple approach for enhancing enzyme stability [67].

In an increasing number of cases, the conformations of selected molecules (peptides and proteins) bound to their targets have been solved by crystallography (see discussion below). NMR spectroscopy has proved to be increasingly useful for characterizing both dynamical systems and those that are recalcitrant to crystallization. Two representative examples are given. The Schultz and Wemmer laboratories [68*] have recently used transferred NOE (nuclear Overhauser enhancement) experiments to determine the geometry of two substrates in the active site of an acyl transferase antibody. Fairbrother *et al.* [69**] used two-dimensional ¹H-¹⁵N HSQC (heteronuclear single quantum coherence) experiments to define the interaction between selected peptides and the ¹H,¹⁵N-labeled target VEGF. Peptide interactions with VEGF were confirmed by a crystal structure of one of these complexes [70*]. It is important to note that HSQC experiments allowed interactions to be defined, even for one of the complexes in which the VEGF-binding peptide had limited solubility.

Peptide selections: mimicry versus linear epitopes

The results of phage peptide selections can be roughly partitioned into two classes: those sequences that conform to a linear epitope and folded peptides that mimic a discontinuous or three-dimensional binding interface (for some recent examples, see Table 1 and [2,71]). The outcome of a given selection appears to be determined by the binding target chosen. If the natural binding partner is a linear sequence, selection experiments will recover linear epitopes that match that motif. Examples of this type of outcome have been seen recently with SH3 (Src homology) domain-binding peptides [72,73], tumor-vasculature-binding peptides [74] and troponin C binders [75]. In this way, *in vitro* selection can uncover unknown binding partners or establish the biological relevance of known interactions.

An increasing number of laboratories have reported peptides or small protein ligands that can functionally mimic the three-dimensional surface that is normally bound by a full-sized protein. These examples are particularly interesting because they emphasize both the plasticity of protein surfaces and the range of structures (both in size and topology) that can functionally substitute for one another. The functional substitution of small peptides for much larger proteins runs counter to the notion that protein interactions are driven in proportion to the burial of hydrophobic surface (see [76] and references therein). One of the most compelling examples of this mimicry are the erythropoietin mimetic peptides (EMP), short (~13–20 residues) β hairpins that functionally mimic erythropoietin, a 34 kDa glycoprotein, but bear no sequence homology [77,78,79*]. Similarly, peptides or minimized proteins have been found that mimic the function of thrombopoietin [80], antagonize VEGF [69**,70*], or bind to constant regions in antibody structures [81].

The peptides isolated by both types of selection generally bind to the same regions utilized by the natural counterparts they are meant to mimic. For linear epitopes, this is perhaps not surprising. For the peptide mimics, it borders on scandalous, as, often, most of the target protein surface is available for interaction during the selection. The clear implication is that some regions of protein surface are more 'sticky' than others and, thus, act as hot spots for interaction [82].

Antibody selection: the importance of large libraries

Although complete coverage of phage antibody selection is beyond the scope of this work, a recent review highlights the importance of high-complexity starting libraries [57*]. When low-complexity libraries containing 3×10^7 clones were used, single-chain antibodies with affinities ranging from 10 to 0.1×10^{-6} μ M [83,84] were isolated. With libraries containing 10^{10} clones, affinities of 10 to 0.3 nM were found [85]. The commensurate increase in binding constant with increased library size is in line with theoretical predictions developed to estimate the binding constants from library complexity [86].

Alternative scaffolds

If peptides can be found that mimic proteins, one might expect that small, structured peptides could be used as scaffolds to display function. Rather than choose a single motif, Fairbrother *et al.* [69**] used a variety of small, disulfide-containing peptides to isolate binders to VEGF (Table 1). Similarly, knottins, small, disulfide-rich proteins have been used to generate libraries that recognize alkaline phosphatase, a protein with no wild-type knottin binding [87**].

Protein scaffolds that could be imbued with antibody diversity have also been sought, particularly because of the refractory nature of some scFvs. Skerra and co-workers [88*] have recently reported that lipocalin BBP variants

(a 174 amino acid β barrel) could be isolated that bound fluorescein with a dissociation constant (K_d) of 35 nM.

Direct selections for catalysis

There is great interest in using selection both to refine known catalysts and to isolate new ones. TSAs have been used to isolate a great diversity of antibody catalysts via *in vivo* immune selection [15]. In a few cases, *in vitro* selection has been used for affinity maturation of the catalysts using TSAs [89,90,91,92] or via a mechanism-based approach (discussed below). The potential limitations of the TSA approach are highlighted by work from the Wells' laboratory. Baca *et al.* [89] used affinity panning of an esterolytic antibody library derived from immune selection to isolate variants with two- to eightfold improvements in their binding affinity for a phosphonate TSA. Notably, a correlation between binding affinity and catalytic ability was not observed. A weaker binding variant was identified with twofold greater catalytic activity than the parent antibody, emphasizing the need for an approach that selects for catalytic activity [90].

In the direct strategy, enzymatic activity is required to either covalently attach or release the library members during selection. This approach has been successfully applied to isolate RNA and DNA catalysts that are capable of performing chemistry with rates ranging from 0.1 to more than 100 min⁻¹, significantly faster than many catalytic antibodies [33]. Lerner and co-workers [93,94] have applied this strategy to immune selection with great success. In one interesting combination of reactive and TSA approaches, phage display was used to improve catalytic antibodies designed to cleave a glycosidic bond by 700-fold ($k_{cat}/k_{uncat} = 7 \times 10^4$ versus 10^2 from immune selection) [95]. Substrate-bond cleavage generated a quinone methide that reacted with the phage to afford immobilization and selection.

Reactive approaches can also be incorporated in phage selections to afford specific release. The Schultz laboratory [96] has recently demonstrated that phage harboring active staphylococcal nuclease can be enriched 100-fold in a single round, based on cleavage of an oligonucleotide tether *in cis*. A similar strategy involving immobilization demonstrated enrichments of more than 50-fold for two model proteins using phage-displayed enzymes and the calcium-dependent binding of substrate to calmodulin [97].

Structural studies of selected and evolved peptides and proteins

There are now a significant number of selected peptides and proteins whose structures are known. We include catalytic antibodies and *in vitro* screened proteins in this discussion because the structural conclusions seem generally applicable to all functional approaches. One very striking feature is the way in which protein structure can be modulated by seemingly subtle changes in sequence.

In general, proteins that have been evolved either *in vivo* or *in vitro* contain positive mutations that are distal to the active site ([98–100]; B Spiller, A Gershenson, FH Arnold, RC Stevens, personal communication). These mutations often have the effect of remodeling the shape of the active site and may act by biasing the conformation towards lock-and-key-type recognition [99,101]. In addition, the selected mutations often make significant changes to the local structure (B Spiller, A Gershenson, FH Arnold, RC Stevens, personal communication) and can even alter topology [102].

As interesting as where mutations are is where they are not — the active site or critical residues on the interface of the proteins. For example, residues conserved in phage selections of heregulin variants (binding to the ErbB3 receptor) correlate strongly with positions that were found to be important using alanine-scanning mutagenesis [103,104]. Selection and evolution thus seem to avoid changing residues that are in direct contact with the target. Rather, beneficial mutations enhance affinity by reconfiguring essential binding residues or by forming new contacts. Indeed, these complicating secondary effects prohibit the delineation of a recognition code, even for highly modular protein contacts, such as zinc-finger–DNA interactions [105]. The importance of long-range interactions and plastic local structure are likely reasons why rational design has yet to produce the results seen with functional approaches, such as *in vitro* selection [106]. Computational approaches may significantly enhance pool design, however, allowing us to ‘stack the deck’ of initial libraries, making selections more fruitful or reducing the number of cycles required.

Conclusions

In vitro selection has proved itself to be a powerful technique. For the foreseeable future, functional approaches provide the most facile path for engineering nucleic acids and proteins alike. The new methods developed will allow larger, better libraries to be generated and more powerful ways to screen them to be elaborated — further expanding our ability to generate interesting molecules for study. A fundamental question is how the information from selection experiments can be used to further insight into biopolymer folding and function. For the present, the results of selection and screening experiments provide existence proofs and hard data on some very interesting questions, including the number of nucleotides needed to specify a catalyst, the number of amino acid changes needed to improve protein thermostability and the free energy consequences (or lack thereof!) of introducing or removing a small functional group (e.g. a methyl group) from either the macromolecule or its target. Future experiments are likely to address issues such as the generation of new function, the minimum size of protein needed to perform catalysis and the likelihood that such a protein could be found in random sequence libraries.

Acknowledgements

We thank Frances H Arnold and Raymond Stevens for providing a copy of their manuscript prior to publication.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Szostak JW, Ellington RF: **In vitro selection of functional RNA sequences.** In *The RNA World*, Edited by Gesteland RF, Atkins JF. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1993:511-533.
2. Smith GP, Petrenko VA: **Phage display.** *Chem Rev* 1997, **97**:391-410.
3. Jermutus L, Ryabova L, Plückthun A: **Recent advances in producing and selecting functional proteins by using cell-free translation.** *Curr Opin Biotechnol* 1998, **9**:534-548.
4. Roberts RW, Szostak JW: **RNA-peptide fusions for the in vitro selection of peptides and proteins.** *Proc Natl Acad Sci USA* 1997, **94**:12297-12302.
5. Liu R, Barrick J, Szostak JW, Roberts RW: **Optimized synthesis of RNA-protein fusions for in vitro protein selection.** *Methods Enzymol* 2000, **317**:in press.
This paper details significant improvements in the yield of mRNA-protein fusion synthesis. In addition, it is the first to demonstrate that a full-sized protein can be generated as a fusion and that RNA-binding peptides retain specificity as naked fusions.
6. Roberts RW: **Totally in vitro protein selection using mRNA-protein fusions and ribosome display.** *Curr Opin Chem Biol* 1999, **3**:268-273.
7. Ellington A, Szostak JW: **In vitro selection of RNA molecules that bind specific ligands.** *Nature* 1990, **346**:818-822.
8. Gold L, Polisky B, Uhlenbeck O, Yarus M: **Diversity of oligonucleotide functions.** *Annu Rev Biochem* 1995, **64**:763-797.
9. Willis MC, Collins B, Zhang T, Green LS, Sebesta DP, Bell C, Kellogg E, Gill SC, Magallanes A, Knauer S *et al.*: **Liposome-anchored vascular endothelial growth factor aptamers.** *Bioconjug Chem* 1998, **9**:573-582.
10. Williams KP, Liu X-H, Schumacher TNM, Lin HY, Ausiello DA, Kim PS, Bartel DP: **Bioactive and nuclease-resistant L-DNA ligand of vasopressin.** *Proc Natl Acad Sci USA* 1997, **94**:11285-11290.
11. Nakamura Y, Ito K, Isaksson LA: **Emerging understanding of translation termination.** *Cell* 1996, **87**:147-150.
12. Wilson KS, Noller HF: **Molecular movement inside the translational engine.** *Cell* 1998, **92**:337-349.
13. Pestova TV, Shatsky IN, Hellen CU: **Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes.** *Mol Cell Biol* 1996, **16**:6870-6878.
14. Hentze MW: **eIF4G: a multipurpose ribosome adaptor?** *Science* 1997, **275**:500-501.
15. Schultz PG, Lerner RA: **From molecular diversity to catalysis: lessons from the immune system.** *Science* 1995, **269**:1835-1842.
16. Prudent JR, Uno T, Schultz PG: **Expanding the scope of RNA catalysis.** *Science* 1994, **264**:1924-1927.
17. Conn MM, Prodent JR, Schultz PG: **Porphyrin metallation catalyzed by a small RNA molecule.** *J Am Chem Soc* 1996, **118**:7012-7013.
18. Li Y, Sen D: **A catalytic DNA for porphyrin metallation.** *Nat Struct Biol* 1996, **3**:743-747.
19. Noller HF, Hoffarth V, Zimniak L: **Unusual resistance of peptidyl transferase to protein extraction procedures.** *Science* 1992, **256**:1416-1419.
20. Samaha RR, Green R, Noller HF: **A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome.** *Nature* 1995, **377**:309-314.
21. Green R, Switzer C, Noller HF: **Ribosome-catalyzed peptide-bond formation with an A-site substrate covalently linked to the 23S ribosomal RNA.** *Science* 1998, **280**:286-289.
22. Lohse PA, Szostak JW: **Ribozyme-catalyzed amino-acid transfer reactions.** *Nature* 1996, **381**:442-444.
23. Wiegand TW, Janssen RC, Eaton BE: **Selection of RNA amide synthesis.** *Chem Biol* 1997, **4**:675-683.
24. Zhang B, Cech TR: **Peptide bond formation by in vitro selected ribozymes.** *Nature* 1997, **390**:96-100.
25. Zhang B, Cech TR: **Peptidyl-transferase ribozymes: trans reactions, structural characterization and ribosomal RNA-like features.** *Chem Biol* 1998, **5**:539-553.
The secondary structure determination of clone 25 from [24] reveals its surprising homology to three distinct sequences in the peptidyl transferase center of the ribosome.
26. Welch M, Majerfeld I, Yarus M: **23S rRNA similarity from selection for peptidyl-transferase mimicry.** *Biochemistry* 1997, **36**:6614-6623.
27. Nitta I, Kamada Y, Noda H, Ueda T, Watanabe K: **Reconstitution of peptide bond formation with Escherichia coli 23S ribosomal RNA domains.** *Science* 1998, **281**:666-669.
28. Unrau PJ, Bartel DP: **RNA-catalyzed nucleotide synthesis.** *Nature* 1998, **395**:260-263.
The further expansion of known ribozyme reactions to include nucleotide biosynthesis. A 3'-linked pRpp (the activated ribose used in nucleotide biosynthesis) is joined to 4-thiouracil ($k_{cat} = 0.13 \text{ min}^{-1}$).
29. Tuschl T, Sharp PA, Bartel DP: **Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library.** *EMBO J* 1998, **17**:2637-2650.
The attempted resurrection of RNA molecules with spliceosomal activity. Four separate catalysts were isolated, none of which corresponds to the spliceosomal mechanism. Interestingly, one has a complex three-step mechanism.
30. Tarasow TM, Tarasow SL, Eaton BE: **RNA-catalysed carbon-carbon bond formation.** *Nature* 1997, **389**:54-57.
31. Roth A, Breaker RR: **An amino acid as a cofactor for a catalytic polynucleotide.** *Proc Natl Acad Sci USA* 1998, **95**:6027-6031.
The first demonstration of free histidine as a necessary cofactor for a DNA-catalyzed (or RNA-catalyzed) reaction. The paper highlights the utility of enhancing the functional group diversity of nucleic acids.
32. Joyce GF: **Nucleic acid enzymes: playing with a fuller deck.** *Proc Natl Acad Sci USA* 1998, **95**:5845-5847.
33. Breaker RR: **In vitro selection of catalytic polynucleotides.** *Chem Rev* 1997, **97**:371-390.
34. Santoro SW, Joyce GF: **A general purpose RNA-cleaving DNA enzyme.** *Proc Natl Acad Sci USA* 1997, **94**:4262-4266.
35. Santoro SW, Joyce GF: **Mechanism and utility of an RNA-cleaving DNA enzyme.** *Biochemistry* 1998, **37**:13330-13342.
36. Cuenoud B, Szostak JW: **A DNA metalloenzyme with DNA ligase activity.** *Nature* 1995, **375**:611-614.
37. Breaker RR, Joyce GF: **A DNA enzyme that cleaves RNA.** *Chem Biol* 1995, **1**:223-229.
38. Faulhammer D, Famulok M: **Characterization and divalent metal-ion dependence of in vitro selected deoxyribozymes which cleave DNA/RNA chimeric oligonucleotides.** *J Mol Biol* 1997, **269**:188-202.
39. Geyer CR, Sen D: **Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA.** *Chem Biol* 1997, **4**:579-593.
40. Carmi N, Schultz LA, Breaker RR: **In vitro selection of self-cleaving DNAs.** *Chem Biol* 1996, **3**:1039-1046.
41. Travascio P, Li Y, Sen D: **DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex.** *Chem Biol* 1998, **5**:505-517.
The authors demonstrate that the N-methylmesoporphyrin DNA aptamer enhances the peroxidase activity of hemin by two orders of magnitude relative to free hemin.
42. Wilson C, Szostak JW: **Isolation of a fluorophore-specific DNA aptamer with weak redox activity.** *Chem Biol* 1998, **5**:609-617.
43. Li Y, Breaker RR: **Phosphorylating DNA with DNA.** *Proc Natl Acad Sci USA* 1999, **96**:2746-2751.
The authors describe the isolation of more than 50 different DNA enzymes that are capable of autophosphorylation. The enzymes in the pool demonstrate very high specificity for ATP as a substrate versus the other three nucleotides, or specificity between dNTPs versus NTPs.

44. Tang J, Breaker RR: **Mechanism for allosteric inhibition of an ATP-sensitive ribozyme.** *Nucleic Acids Res* 1998, **26**:4214-4221.
45. Soukup GA, Breaker RR: **Engineering precision RNA molecular switches.** *Proc Natl Acad Sci USA* 1999, **96**:3584-3589.
The combination of aptamers with the hammerhead ribozyme in order to create artificial allosteric ribozymes is described.
46. Werstuck G, Green MR: **Controlling gene expression in living cells through small molecule-RNA interactions.** *Science* 1998, **282**:296-298.
The use of aptamer structures in the 5'-UTR (untranslated region) of mRNAs allows the down-regulation of translation in the presence of the small-molecule target.
47. Jencks WP: **Binding energy, specificity, and enzymic catalysis: the Circe effect.** *Adv Enzymol Relat Areas Mol Biol* 1975, **43**:219-410.
48. Feigon J, Teikmann T, Smith FW: **Aptamer structures from A to Z.** *Chem Biol* 1996, **3**:611-617.
49. Patel DJ: **Structural analysis of nucleic acid aptamers.** *Curr Opin Chem Biol* 1997, **1**:32-46.
50. Marshall KA, Robertson MP, Ellington AD: **A biopolymer by any other name would bind as well: a comparison of the ligand-binding pockets of nucleic acids and proteins.** *Structure* 1997, **5**:729-734.
51. Nowakowski J, Shim PJ, Prasad GS, Stout CD, Joyce GF: **Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme.** *Nat Struct Biol* 1999, **6**:151-156.
An interesting nucleic acid structure is described that probably does not conform to the actual geometry of the 10-23 DNA enzyme.
52. Wedekind JE, McKay DB: **Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis.** *Nat Struct Biol* 1999, **6**:261-268.
The first crystal structure determination of a laboratory-isolated RNA enzyme. The structure contains two molecules, one of which is in a precatalytic geometry that can be used to rationalize the cleavage chemistry.
53. McKay DB: **Structure and function of the hammerhead ribozyme: an unfinished story.** *RNA* 1996, **2**:395-403.
54. Verma S, Vaish NK, Eckstein F: **Structure-function studies of the hammerhead ribozyme.** *Curr Opin Chem Biol* 1997, **1**:532-536.
55. Golden BL, Gooding AR, Podell ER, Cech TR: **A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme.** *Science* 1998, **282**:259-264.
56. Ferre-D'Amare AR, Zhou K, Doudna JA: **Crystal structure of a hepatitis delta virus ribozyme.** *Nature* 1998, **395**:567-574.
57. Griffiths AD, Duncan AR: **Strategies for selection of antibodies by phage display.** *Curr Opin Biotechnol* 1998, **9**:102-108.
An excellent review of antibodies displayed on phage.
58. Spada S, Plückthun A: **Selectively infective phage (SIP) technology: a novel method for *in vivo* selection of interacting protein-ligand pairs.** *Nat Med* 1997, **3**:694-696.
59. Tsang J, Joyce GF: ***In vitro* evolution of randomized ribozymes.** *Methods Enzymol* 1996, **267**:410-426.
60. Cramer A, Raillard SA, Bermudez E, Stemmer WPC: **DNA shuffling of a family of genes from diverse species accelerates directed evolution.** *Nature* 1998, **391**:288-291.
In vitro shuffling of four related cephalosporinase genes produced a recombinant with greatly enhanced activity in fewer cycles than through mutagenesis and recombination of the individual genes.
61. Zhao HM, Giver L, Shao ZX, Affholter JA, Arnold FH: **Molecular evolution by staggered extension process (STEP) *in-vitro*.** *Nat Biotechnol* 1998, **16**:258-261.
Short cycles of annealing and extension in a PCR-like reaction produce prematurely terminated products that can randomly anneal to parent strands in the next cycle, creating recombinant sequences. This novel method was used to evolve a subtilisin E variant that is equivalent to its thermophilic homolog.
62. Tawfik DS, Griffiths AD: **Man-made cell-like compartments for molecular evolution.** *Nat Biotechnol* 1998, **16**:652-656.
In vitro transcription and translation were linked via compartmentalization in reverse micelles. A selection for targeted DNA methylation using HaellI methyltransferase was demonstrated.
63. Hanes J, Pluckthun A: ***In vitro* selection and evolution of functional proteins by using ribosome display.** *Proc Natl Acad Sci USA* 1997, **94**:4937-4942.
64. Hanes J, Jermutus L, Weber-Bornhauser S, Bosshard HR, Plückthun A: **Ribosome display efficiently selects and evolves high-affinity antibodies *in vitro* from immune libraries.** *Proc Natl Acad Sci USA* 1998, **95**:14130-14135.
The first report of the construction and implementation of an immunized murine scFv library using ribosome display. The best clone appears to have 65-fold higher affinity than the sequence it was derived from.
65. Stemmer WPC: **DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution.** *Proc Natl Acad Sci USA* 1994, **91**:10747-10751.
66. Shao Z, Zhao H, Giver L, Arnold FH: **Random-priming *in vitro* recombination: an effective tool for directed evolution.** *Nucleic Acids Res* 1998, **26**:681-683.
67. Matsuura T, Miyai K, Trakulnaleamsai S, Yomo T, Shima Y, Miki S, Yamamoto K, Urabe I: **Evolutionary molecular engineering by random elongation mutagenesis.** *Nat Biotechnol* 1999, **17**:58-61.
68. Driggers EM, Liu CW, Wemmer DE, Schultz PG: **Structure of the Michaelis complex of an efficient antibody acyl transferase determined by transferred nuclear Overhauser enhancement spectroscopy.** *J Am Chem Soc* 1998, **120**:7395-7396.
Transferred NOE experiments were used to determine the structure of two bound substrates in the presence of an acyl transferase antibody.
69. Fairbrother WJ, Christinger HW, Cochran AG, Fuh G, Keenan CJ, Quan C, Shriver SK, Tom JYK, Wells JA, Cunningham BC: **Novel peptides selected to bind vascular endothelial growth factor target the receptor-binding site.** *Biochemistry* 1998, **37**:17754-17764.
Libraries of disulfide-constrained peptides displayed on phage yielded peptides that bind to the receptor-binding domain of vascular endothelial growth factor (VEGF). HSQC spectra of ¹H,¹⁵N-labeled protein indicate the sites of peptide interaction, suggesting a binding 'hot spot' that can be readily targeted.
70. Wiesmann C, Christinger HW, Cochran AG, Cunningham BC, Fairbrother WJ, Keenan CJ, Meng G, de Vos AM: **Crystal structure of the complex between VEGF and a receptor-blocking peptide.** *Biochemistry* 1998, **37**:17765-17772.
One of the selected peptides in [69**] was crystallized bound to VEGF. The peptide adopts a largely extended conformation resembling Fab fragment interaction more closely than the VEGF-receptor interaction.
71. Kay BK, Kurakin AV, Hyde-DeRuyscher R: **From peptides to drugs via phage display.** *Drug Discovery Today* 1998, **3**:370-378.
72. Sparks AB, Hoffman NG, McConnell SJ, Fowlkes DM, Kay BK: **Cloning of ligand targets: systematic isolation of SH3 domain-containing proteins.** *Nat Biotechnol* 1996, **14**:741-744.
73. Sparks AB, Rider JE, Hoffman NG, Fowlkes DM, Quillam LA, Kay BK: **Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, cactactin, p53bp2, PLC γ , Crk, and Grb2.** *Proc Natl Acad Sci USA* 1996, **93**:1540-1544.
74. Pasqualini R, Koivunen E, Ruoslahti E: **α Integrins as receptors for tumor targeting by circulating ligands.** *Nat Biotechnol* 1997, **15**:542-546.
75. Pierce HH, Schachat F, Brandt PW, Lombardo CR, Kay BK: **Identification of troponin C antagonists from a phage-displayed random peptide library.** *J Biol Chem* 1998, **273**:23448-23453.
76. Spolar RS, Ha J-H, Record MTJ: **Hydrophobic effect in protein folding and other noncovalent processes involving proteins.** *Proc Natl Acad Sci USA* 1989, **86**:8382-8385.
77. Wrighton NC, Farrell FX, Chang R, Kashyap AK, Barbone FP, Mulcahy LS, Johnson DL, Barrett RW, Jolliffe LK, Dower WJ: **Small peptides as potent mimetics of the protein hormone erythropoietin.** *Science* 1996, **273**:458-463.
78. Livnah O, Stura EA, Johnson DL, Middleton SA, Mulcahy LS, Wrighton NC, Dower WJ, Jolliffe LK, Wilson IA: **Functional mimicry of a protein hormone by a peptide agonist: the epo receptor complex at 2.8 Å.** *Science* 1996, **273**:464-471.
79. Johnson DL, Farrell FX, Barbone FP, McMahon FJ, Taulai J, Hoey K, Livnah O, Wrighton NC, Middleton SA, Loughney DA *et al.*: **Identification of a 13 amino acid peptide mimetic of erythropoietin and description of amino acids critical for the mimetic activity of EMP1.** *Biochemistry* 1998, **37**:3699-3710.
Alanine scanning and truncation were applied to the erythropoietin mimetic peptides (EMPs) isolated previously, resulting in a 13 amino acid sequence containing the minimal agonist epitope.
80. Cwirla SE, Balasubramanian P, Duffin DJ, Wagstrom CR, Gates CM, Singer SC, Davis AM, Tansik RL, Mattheakis LC, Boytos CM *et al.*:

Peptide agonist of the thrombopoietin receptor as potent as the natural cytokine. *Science* 1997, **276**:1696-1699.

81. Starovasnik MA, Braisted AC, Wells JA: **Structural mimicry of a native protein by a minimized binding domain.** *Proc Natl Acad Sci USA* 1997, **94**:10080-10085.
 82. Clackson T, Wells JA: **A hot spot of binding energy in a hormone-receptor interface.** *Science* 1995, **267**:383-386.
 83. Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G: **By-passing immunization: human antibodies from V-gene libraries displayed on phages.** *J Mol Biol* 1991, **222**:581-597.
 84. Griffiths AD, Malmqvist M, Marks JD, Bye JM, Embleton MJ, McCafferty J, Maier M, Holliger KP, Gorick BD, Hughes-Jones NC *et al.*: **Human anti-self antibodies with high specificity from phage display libraries.** *EMBO J* 1993, **12**:725-734.
 85. Vaughan TJ, Williams AJ, Pritchard K, Osbourn JK, Pope AR, Earnshaw JC, McCafferty J, Hodits RA, Wilton J, Johnson KS: **Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library.** *Nat Biotechnol* 1996, **14**:309-314.
 86. Lancet D, Sadvovsky E, Seidemann E: **Probability model for molecular recognition in biological receptor repertoires: significance to the olfactory system.** *Proc Natl Acad Sci USA* 1993, **90**:3715-3719.
 87. Smith GP, Patel SU, Windass JD, Thornton JM, Winter G, Griffiths AD: **Small binding proteins selected from a combinatorial repertoire of knottins displayed on phage.** *J Mol Biol* 1998, **277**:317-332.
- The knottins represent some of the smallest proteins known. These sequences maintain their folded conformations due to the presence of disulfide bonds (typically 2–3 out of 30 amino acids) and have potential as scaffolds to display chemical diversity. The authors screened a cellulose-binding knottin library and derived binders for alkaline phosphatase (see Table 2).
88. Beste G, Schmidt FS, Stibora T, Skerra A: **Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold.** *Proc Natl Acad Sci USA* 1999, **96**:1898-1903.
- The authors demonstrated that lipocalins may represent an alternate framework to antibodies for developing general binding proteins via *in vitro* selection.
89. Baca M, Scanlan TS, Stephenson RC, Wells JA: **Phage display of a catalytic antibody to optimize affinity for transition-state analog binding.** *Proc Natl Acad Sci USA* 1997, **94**:10063-10068.
 90. Arkin MR, Wells JA: **Probing the importance of second sphere residues in an esterolytic antibody by phage display.** *J Mol Biol* 1998, **284**:1083-1094.
- Phage display was used to improve and probe esterolytic antibody function. The report provides another example in which residues remote from the active site are important for binding and catalysis. A correlation is seen between transition state analog binding and catalytic function for single mutations.
91. Fujii I, Fukuyama S, Iwabuchi Y, Tanimura R: **Evolving catalytic antibodies in a phage-displayed combinatorial library.** *Nat Biotechnol* 1998, **16**:463-467.
 92. Gao C, Lavey BJ, Lo C-HL, Datta A, Wentworth PJ: **Direct selection for catalysis from combinatorial antibody libraries using a boronic acid probe: primary amide bond hydrolysis.** *J Am Chem Soc* 1998, **120**:2211-2217.
 93. Wirsching P, Ashley JA, Lo C-HL, Janda KD, Lerner RA: **Reactive immunization.** *Science* 1995, **270**:1775-1782.
 94. Barbas CFI, Heine A, Zhong G, Hoffmann T, Gramatikova S, Bjornestedt R, List B, Anderson J, Stura EA, Wilson IA, Lerner RA: **Immune versus natural selection: antibody aldolases with enzymatic rates but broader scope.** *Science* 1997, **278**:2085-2092.
 95. Janda KD, Lo L-C, Lo C-HL, Sim M-M, Wang R, Wong C-H, Lerner RA: **Chemical selection for catalysis in combinatorial antibody libraries.** *Science* 1997, **275**:945-948.
 96. Pedersen H, Hölder S, Sutherlin DP, Schwitter U, King DS, Schultz PG: **A method for directed evolution and functional cloning of enzymes.** *Proc Natl Acad Sci USA* 1998, **95**:10523-10528.
- A novel strategy for the direct selection of catalysts by phage display is demonstrated by attaching both the enzyme and substrate to the coat

protein of phage. Intramolecular catalysis provides the basis for selection. A model enrichment using staphylococcal nuclease is demonstrated.

97. Demartis S, Huber A, Viti F, Lozzi L, Giovannoni L, Neri P, Winter G, Neri D: **A strategy for the isolation of catalytic activities from repertoires of enzymes displayed on phage.** *J Mol Biol* 1999, **286**:617-633.
 98. Patten PA, Gray NS, Yang PL, Marks CB, Wedemayer GJ, Boniface JJ, Stevens RC, Schultz PG: **The immunological evolution of catalysis.** *Science* 1996, **271**:1086-1091.
 99. Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Steven RC: **Structural insights into the evolution of an antibody combining site.** *Science* 1997, **276**:1665-1669.
 100. Oue S, Okamoto A, Yano T, Kagamiyama H: **Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues.** *J Biol Chem* 1999, **274**:2344-2349.
 101. Romesberg FE, Spiller B, Schultz PG, Stevens RC: **Immunological origins of binding and catalysis in a Diels-Alderase antibody.** *Science* 1998, **279**:1929-1933.
 102. Cordes MHJ, Walsh NP, McKnight CJ, Sauer RT: **Evolution of a protein fold *in vitro*.** *Science* 1999, **284**:325-327.
- The substitution of two residues into the Arc repressor changes the local secondary structure from a β sheet to a short α helix, with the preservation of core packing. The results imply that modest changes in amino acid sequence may allow significant changes in local secondary structure
103. Jones JT, Ballinger MD, Pisacane PI, Lofgren JA, Fitzpatrick VD, Fairbrother WJ, Wells JA, Sliwkowski MX: **Binding interaction of the heregulin β *egf* domain with ErbB3 and ErbB4 receptors assessed by alanine scanning mutagenesis.** *J Biol Chem* 1998, **273**:11667-11674.
- A thorough study of protein specificity through exhaustive alanine scanning mutagenesis.
104. Ballinger MD, Jones JT, Lofgren JA, Fairbrother WJ, Akita RW, Sliwkowski MX, Wells JA: **Selection of heregulin variants having higher affinity for the ErbB3 receptor by monovalent phage display.** *J Biol Chem* 1998, **273**:11675-11684.
- Variants of a growth factor domain with higher affinities for its receptor revealed that beneficial mutations generally avoided critical binding residues. Evolved proteins had enhanced affinity either by reconfiguring essential binding residues or by forming new contacts.
105. Elrod-Erikson M, Benson TE, Pabo CO: **High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition.** *Structure* 1998, **6**:451-464.
- A thorough structural analysis of zinc-finger-DNA variants revealed the difficulties in determining an exhaustive recognition code defining zinc-finger DNA-binding specificity. The results cast serious doubt on developing a code for sequence-specific protein-DNA recognition.
106. Street AG, Mayo SL: **Computational protein design.** *Structure* 1999, **7**:R105-R109.
 107. Rodi DJ, Janes RW, Sanganeer HJ, Holton RA, Wallace BA, Makowski L: **Screening of a library of phage-displayed peptides identifies human Bcl-2 as a taxol-binding protein.** *J Mol Biol* 1999, **285**:197-203.
 108. Fukumoto T, Torigoe N, Kawabata S, Murakami M, Uede T, Nishi T, Ito Y, Sugimura K: **Peptide mimics of the CTLA4-binding domain stimulate T-cell proliferation.** *Nat Biotechnol* 1998, **16**:267-270.
 109. Rand KH, Houck H, Denslow ND, Heilman KM: **Molecular approach to find target(s) for oligoclonal bands in multiple sclerosis.** *J Neurol Neurosurg Psychiatry* 1998, **65**:48-55.
 110. Bremnes T, Lauvrak V, Lindqvist B, Bakke O: **Selection of phage displayed peptides from a random 10-mer library recognising a peptide target.** *Immunotechnol* 1998, **4**:21-28.
 111. Gårdsvoll H, van Zonneveld A-J, Holm A, Eldering E, van Meijer M, Danø K, Pannekoek H: **Selection of peptides that bind to plasminogen activator inhibitor 1 (PAI-1) using random peptide phage-display libraries.** *FEBS Lett* 1998, **431**:170-174.
 112. Pearce KHJ, Cunningham BC, Fuh G, Teeri T, Wells JA: **Growth hormone binding affinity for its receptor surpasses the requirements for cellular activity.** *Biochemistry* 1999, **38**:81-89.