Chapter 3
Biorecognition Molecules: Types and Molecular Basis and Development of Specificity

Robert E. Collins and Aitziber L. Cortajarena

3.1 Introduction

As introduced in Chap. 1 many biomolecules are involved in molecular recognition processes. These molecules include proteins, peptides, and nucleic acids. In the current chapter we introduce in detail the different recognition molecules found in nature. The chapter analyzes the recognition processes that they mediate and the key aspects of such recognition, including affinity and specificity. Finally, we hint about the tools that can be used in order to modify and expand the natural biorecognition diversity. The chapter aims to provide an overview of the biomolecular complexity and the array of biorecognition functions available in nature or by design, focusing mostly into the two major recognition moieties: proteins and peptides and nucleic acids. We cover some of the biorecognition pairs that are used in the applications described in the following book chapters.

Most importantly, we aim to provide the basis for a thoughtful selection of appropriate biorecognition moieties to add functionality to different polymeric platforms considering the final application.

A. L. Cortajarena (✉)
Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA-Nanociencia),
CNB-CSIC-IMDEA Nanociencia Associated Unit Unidad de Nanobiotecnología, Cantoblanco,
28049 Madrid, Madrid, Spain
e-mail: aitziber.lopezcortajarena@imdea.org

R. E. Collins
Physical Sciences Department, Esatern Connecticut State University, 83 Windham St,
Willimantic, CT 06226, USA
e-mail: collins.r.e@gmail.com

3.2   Proteins

Proteins, also called polypeptides, are genetically encoded linear polymers built from 20 standard amino acids. An amino acid contains an invariant backbone, which consists of a carbon flanked by an amine and a carboxylic acid group. This carbon attaches to the variable amino acid side chains, which include charged, polar and nonpolar groups. Following dehydration during amino acid synthesis, a planar peptide bond is formed. The resulting amide grants each amino acid residue in the polymer a hydrogen bond donating and accepting group. Although the lowest energy state of some proteins is an extended unstructured form, most either assemble colinearly with other polypeptides (e.g., collagen), or fold into a compact state. Assembly and folding is typically driven by the so-called "hydrophobic effect," the sequestration of hydrophobic amino acid side chains away from water, in order to minimize the entropically disfavored ordering of water around groups with which it cannot interact. Folding maximizes the hydrogen bonding of main chain and side chain polar groups. Although the peptide bond is planar, other bonds in the protein can rotate in all conformations that do not result in clashes with other atoms. Predicting protein structure de novo solely from sequence remains difficult and has unpredictable results.

Proteins are critical to life, serving as catalysts, structural supports, carriers, sensors, and scaffolds. In this chapter, we focus on proteins that have been utilized as binding platforms. Excluding covalent modification of proteins, protein–protein, or protein–ligand binding is driven by the same fundamental forces that drive protein folding: hydrogen bond formation, ionic bonding, the hydrophobic effect, and van der Walls interactions. Ionic bonds are generally (mis)taken to mean the interactions of the monovalent positively and negatively charged side chains. Aromatic hydrophobic side chains have recently been recognized to have quadrupolar ionic nature, as the pi orbitals above and below aromatic rings are negatively charged, leaving the ring edges positively charged. This can provide a quite unique hydrophobic-yet-charged interaction interface seen, for example, in nicotine binding to acetylcholine receptor and proteins that bind methylated lysines [1, 2]. The presence of ions at interfaces, and metal coordination chemistry is another consideration, as polyvalent cations are observed in protein–protein and protein–ligand interfaces, and have been included in design strategies of polymeric assemblies and scaffolds [3–8]. Although the forces are the same, the frequencies of amino acid pairings differ in protein cores and protein–protein interfaces. Ionic bonds and hydrophobic interactions, in particular, tryptophan-proline pairs, tend to be overrepresented in the latter [9].

At any given protein recognition site, any or all of these bond types may be present. One might predict respective bond energies to sum evenly. In this case, interface size would directly correlate with binding strength. The previously exposed surface area that is buried when proteins interact does, in fact, correlate with affinity up to a certain size; however, at any given size of buried area, there is an enormous range of actual binding affinities, from very weak to quite strong interactions [10]. Clearly, not all interactions are equal. Considerations of the difficulty of accurately

modeling the affinities of interactions have been extensively reviewed elsewhere [11]. The four primary noncovalent bond types above are listed roughly in descending order of strength; however, local environments can alter bond strength. For example, an ionic bond in a desolvated protein–protein interface is shielded from water, and therefore binds with much greater strength than a solvent-exposed one. This may, in part, explain why interfaces with the same buried surface area and similar amino acid content can have affinities ranging across many orders of magnitude. Binding free energy "hot-spots" that have unexpectedly large contributions to binding affinities have been identified in natural complexes, and are a desirable but elusive target in designed interfaces. Although different residue types (primarily charged or aromatic amino acids) have been found in hot-spots, protection from solvent by other residues is the feature that best appears to typify hot-spots [12–14].

Affinity, Avidity, and Specificity Affinity is a measure of strength of binding of a molecule to its ligand. The affinity of a protein interaction with its ligand is quantified as a dissociation constant (Kd), in molar units (M). Quantitatively, Kd values can be determined as the ratio of the off-rate constant (how quickly the protein dissociates from bound ligand) and the on-rate constant (how quickly the protein binds ligand). Alternatively, at equilibrium, Kd is the concentration of ligand that results in equal concentrations of free protein and complex. As the inverse of an association constant, it describes the susceptibility of a complex to dissociation. Lower values indicate tighter binding. As described in Chap. 1, physiological binding constants can range from the femtomolar to the millimolar. Complexes with femtomolar affinities are long lived, for example, biotin-streptavidin complexes exhibit a half-life of 35 h under physiological conditions [15]. Assuming diffusion-limited on-rates, complexes with millimolar affinities will have millisecond half-lives at best.

While affinity properly refers to the binding of a single ligand to a single protein, avidity is considered with multivalent binding. Avidity is the cumulative effects of multiple affinities working together. In order for multivalent complexes to dissociate, each subunit must simultaneously reach the "off" state. Avidity is sometimes referred to as functional affinity, in that the association and dissociation of multiple binding units and their ligands can be measured. Avidity has been noted in the generation of engineered antibodies, where single-chain fragments (scFV) are selected for, and then are later reassembled into IgG class antibodies, which have two binding sites, or are simply dimerized into "diabodies." In these cases, the Kd values for the affinity of the scFV (single unit) are hundreds to thousands times weaker than the avidity of the diabody or antibody) [16]. With high starting affinities, bivalent or higher order binding soon becomes practically irreversible. The degree to which functional avidity assays can, therefore, obscure the true underlying affinity has led to criticisms against these techniques (especially with indirect assays) [17]. Nevertheless, it is clear that conversion of monovalent binders to polyvalent binders may boost effectiveness when engineering binding scaffolds for a target.

Specificity, as opposed to affinity and avidity, lacks a formal quantifiable unit. It is simply the binding of one ligand to the exclusion of others. Specificity can be experimentally assessed, for example by pull down, where hopefully one ligand and

not many is isolated from the cellular milieu. In protein design, engineering, and selection, specificity often seems to come about through the selection of functional high affinity binders. The perfection of a binding site for one ligand should tend to physically and chemically gate out others. Moreover, truly nonspecific binders may be lost in purification or fail to behave in selection assays. In the following sections, we highlight the development of high affinity binding platforms, and a few cases where specificity was assessed and enhanced.

In the case studies we present in this chapter, as high-affinity binders are progressively selected and refined, structural features that promote binding emerge. More surface area is buried, and more fruitful bonds emerge, particularly in the satisfaction of constrained hydrogen bond distances and geometries. Side chains become better positioned, requiring less rearrangement (and entropic penalty) to bind. Finally, solvent excluding seals emerge around hot-spot residues, guaranteeing their full, uninterrupted binding energies.

3.2.1 Common Natural Protein Recognition

As efforts to use proteins as polymeric platforms for recognition, and as therapeutic agents began, the first source of tools were naturally occurring proteins. As detailed in Chap. 1, the affinities of the protein ligand interactions found in nature range from the transient to the near permanent. This is according to the physiological role of the protein itself. Interactions of cellular signaling processes must be short lived, so that when a signal is terminated, those complexes transmitting that signal dissociate and return to a resting state. Antibody–antigen complexes should remain intact until immune processes are recruited and complete their job. In this section, we highlight some of the most commonly used naturally occurring protein–ligand interactions.

3.2.1.1 Streptavidin–Biotin

Streptavidin is a protein isolated from bacteria that binds the small molecule biotin with one of the strongest noncovalent interactions known ($K_d = 10^{-14}$ M). This remarkable affinity is achieved through an extensive network of hydrogen bonds with the polar moieties of biotin and hydrophobic and van der Waals interactions mediated by conserved tryptophans packing against biotin (Fig. 3.1b). These interactions are reinforced by the closing of a loop over the pocket. For streptavidin to release biotin, concerted loop movement, solvation of apolar surfaces, and the breaking of the hydrogen bond network must occur [23, 24].

Because of the high affinity and specificity of this interaction, the biotin–streptavidin interaction is often used in scaffold engineering and protein isolation. Biotin can be synthetically attached to DNA, RNA, or proteins in vitro, and in vivo, proteins can be tagged with biotin acceptor peptides which are posttranslationally

modified with biotin in the presence of biotin ligases [25–27]. A small peptide sequence has also been selected for streptavidin affinity in the absence of biotin [28, 29]. Additionally, engineered streptavidins of varying affinities and valence (monomers, tetramers, monovalent tetramers) have been developed, allowing control of scaffold design [30, 31]. These features have been used to generate biosensors, facilitate nanotube assembly in controlled orientation, and the controlled assembly of extended protein lattices, for example [3, 32, 33].

3.2.1.2   Antibody–Antigen

Antibodies are the protein products of the adaptive immune system that exist as either secreted or cell-bound forms. Antibody diversity is far greater than the number of genes in a genome. This diversity comes about by combination of several genetic cassettes and somatic mutation of the sequence in those cassettes. Antibodies (Fig. 3.1a) consist of conserved structural regions and variable antigen-binding regions. Typically, antibodies used in research are IgG, which contain two binding sites per antibody complex (Fab) (Fig. 3.1a), but other antibody classes contain more. Upon introduction of an antigen, the basal diversity of the immune system allows weak recognition of an antigen. Cells displaying antibodies that recognize the antigen are stimulated to proliferate, and their antibodies further diversify. The expansion and diversification of successful binders is termed affinity maturation, and this term has been adopted to describe synthetic selection processes that proceed by genetic diversification and selection of superior binders [34]. Figure 3.2 depicts methods for the affinity maturation process. In short, loose contacts are replaced by more fruitful, direct interactions that bury additional hydrophobic surface area, and protect inner binding contacts that may constitute a solvent-shielded "hot-spot."

Antibodies are typically generated in one of two ways. Polycolonal antibodies are the product of all cell lineages that recognize an antigen. An animal is injected with antigen (often in one primary and several booster injections). Following development of high-affinity antibodies, serum is harvested, and antibodies purified. Polyclonal antibodies may recognize different sites on a large antigen, and will range in affinities. Conversely, monoclonal antibodies are produced by a single cell lineage. An animal (typically mouse) is injected with antigen, and immortal hybridoma cell lines are generated from the animal's spleen cells. Cloning of antibody producing cells allows for isolation of a single antibody against the antigen.

Protocols for the chemical modification of antibodies are abundant and have nearly a 60 year history. Additionally, in the design of antibody-based binding scaffolds, antibody-binding modules have been borrowed from bacteria that use them as countermeasures to the adaptive immune system. Protein A is derived from *Staphylococcus aureus* and has high affinity for the conserved (Fc) region of most antibodies. Protein G is produced by Streptococcal species and has Fc and Fab affinities. Physiologically, these block conserved regions of antibodies from their normal immune functions, but antigen binding should be unimpaired. Both are used in antibody purification and immobilization [35, 36].
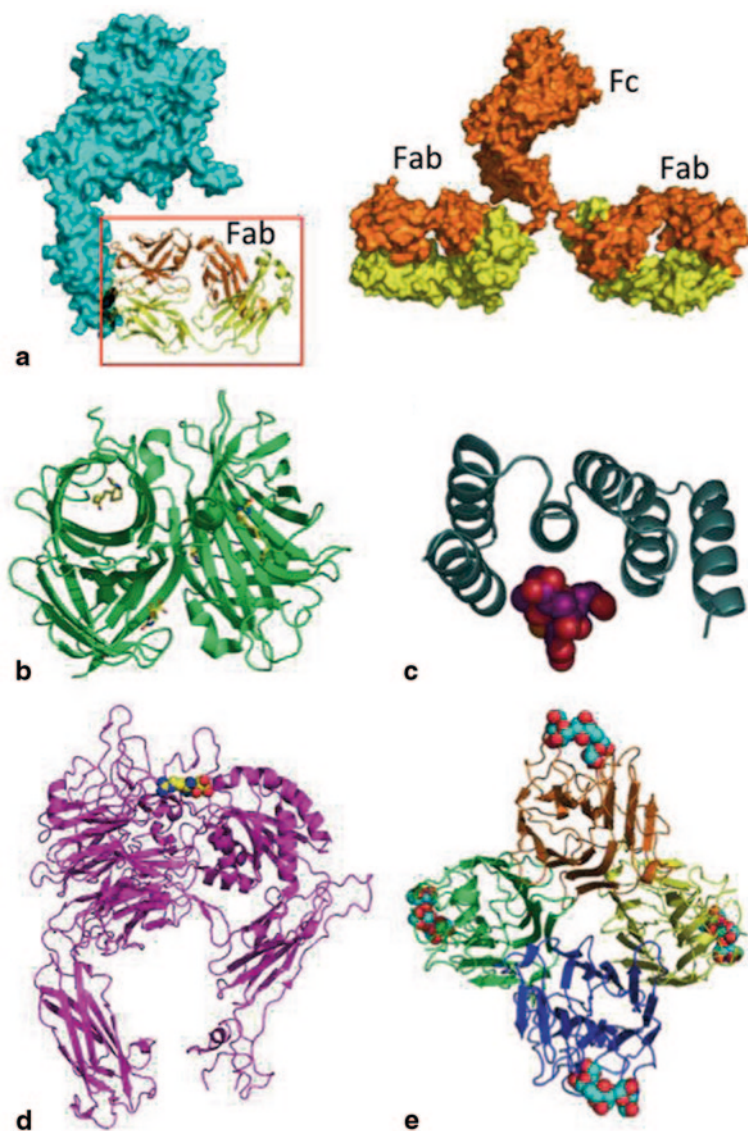
Fig. 3.1 Collection of structures that represent different types of biorecognition interactions mediated by proteins. **a** Crystal structure of Herceptin monoclonal antibody antigen-binding fragment (Fab) represented as *ribbons* in *yellow* and *orange* for the light and heavy chains respectively. The structure shows the Herceptin antibody bound to the extracellular region of Her2 shown as cyan surface (PDB ID: 1N8Z) [18]. The *right panel* shows the structure of a complete antibody with the light and heavy chains displayed as *yellow* and *orange* surfaces, respectively. The two antigen recognition regions Fab, and the constant region (Fc) are shown. The Herceptin anti-Her2 antibody example represents a high affinity interaction between two proteins with a dissociation constant of 0.1 nM. Heceptin is commonly used in breast cancer treatment. **b** Crystal structure of

Human antibodies against therapeutic targets are desirable as drugs (called biologics), as they should avoid immune responses to nonhuman immunoglobins (Fig. 3.1a). As conventional development of human or humanized antibodies would be impossible, in vitro approaches were developed. Libraries of the isolates variable fragments of an antibody (scFv) were expressed on the surface of phage (viruses that infect bacteria) or bacteria. Physical isolation of successful binders was followed by rounds of diversification and reisolation. The result produced antibodies with affinities comparable to antibodies produced by affinity maturation in vivo. When reconstituted into antibodies, even higher apparent affinity (avidity) was observed. The highly successful biologic Humira (human monoclonal antibody in rheumatoid arthritis) was developed by this approach [37–40].

3.2.1.3 Integrin–RGD

Proteins that contain the Arg-Gly-Asp peptide sequence (RGD) are recognized by the integrins that serve as receptors for them and constitute a major recognition system for cell adhesion [41]. The RGD sequence is the cell attachment site of a large number of adhesive extracellular matrix, blood, and cell surface proteins, and nearly half of the over 20 known integrins recognize this sequence in their adhesion protein ligands. The integrin-binding activity of adhesion proteins can be reproduced by short synthetic peptides containing the RGD sequence (Fig. 3.1d). The binding affinity of this interaction is relatively low with dissociation constant values of $10^{-5}-10^{-3}$ M. The structural basis of the interaction provides insights on the promiscuity of RGD-binding integrins [42]. Many ligands are shared by this subset of integrins, but the ligand affinity varies, presumably reflecting the fit of the ligand RGD conformation with the specific α-β integrin binding pockets. RDG is a ubiquitous adhesion sequence, therefore numerous new biomaterials and surfaces coated with RGD peptides have been used to control the cell adhesion in vitro and in vivo.

---

the streptavidin-biotin complex. The streptavidin protein is a tetramer shown in *green ribbons*. Each barrel binds one biotin molecule the small biotin molecules (one per streptavidin barrel) are shown as *yellow sticks*. (PDB ID: 1MK5) [19]. This interaction represents a high affinity interaction with a dissociation constant of $K_d = 10^{-14}$ M between a protein and a small molecule. **c** Crystal structure of a small designed protein binding module in complex with its target peptide. The designed tetratricopeptide repeat (*TPR*) module is shown in *green* as *ribbon* representation. The target C-terminal Hsp90 peptide is shown as *spheres* in *purple* (PDB ID: 3KD7) [20]. This is an example of a low affinity interaction with micromolar dissociation constant and therefore a short half-live, these interactions mediate transient and dynamic interactions. **d** Crystal structure of the integrin-RGD peptide biorecognition complex. The Integrin alpha 5 chain head piece recognition fragment is shown as *magenta ribbon* and the RGD peptide shown in *yellow spheres* (PBD ID: 3VI4) [21]. **e** Crystal structure of the lectin Concanavalin A in complex with sugar ligands. The four subunits of the Concanavalin A tetramer are shown in different colors: *green*, *blue*, *orange*, and *yellow*. Each subunit binds through specific hydrogen bonds and van der Waals interactions a trimannoside molecule shown as *cyan spheres*. PDB ID: 1CVN) [22]
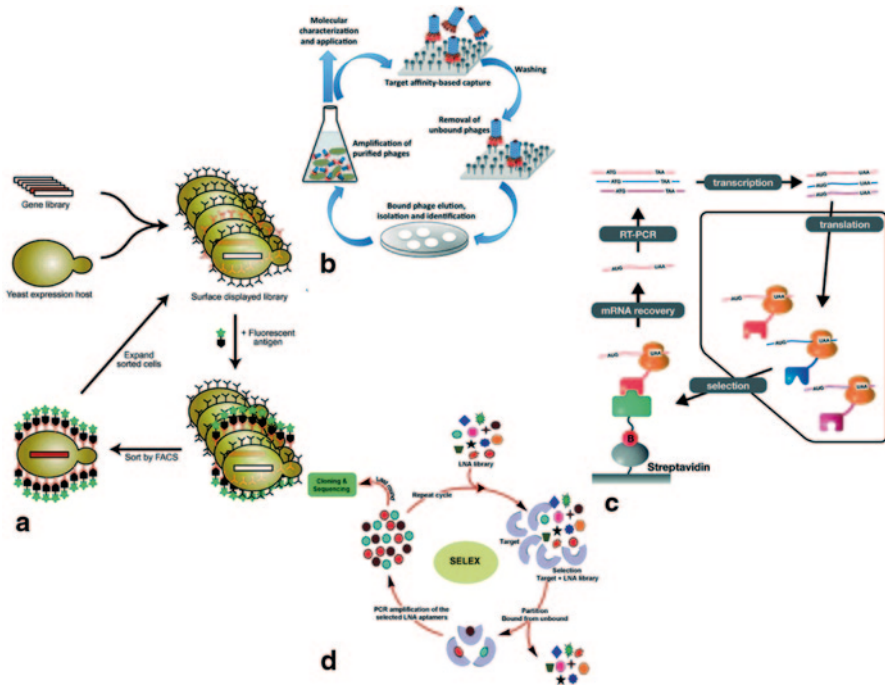
Fig. 3.2 Examples of the most commonly used methods for in vitro evolution methods for selection of high-affinity binders. **a** Schematic of yeast display technology, in which a protein of interest is expressed on the surface of yeast. From a library of potential ligand molecules high-affinity protein ligands can be isolated [52]. Iterative rounds of cell sorting by fluorescence activated cell sorting (*FACS*) and expansion by cell culture are applied to yield enrichment of binders from a combinatorial library. The cells recovered are expanded and subjected to multiple rounds of FACS/expansion. Figure reproduced from [53]. **b** Schematic of phage display technique for affinity-based selection of protein–protein, protein–peptide, and protein–DNA interactions. A protein or library of variants of a protein of interest is displayed on the surface of phage. The phages displaying the proteins are then screened against the target molecule of interest. After several screening and amplification rounds new or high-affinity ligands can be developed. Figure reproduced from [54]. **c** Schematic representation of ribosome-display selection methodology. A DNA of the library encoding the protein variants of interest is transcribed in vitro and the resulting mRNA is used for in vitro translation to synthesize the encoded proteins. The absence of the stop codon at the end of the protein results in the protein connected to the tRNA. The mRNA-ribosome-protein ternary complexes are used for affinity selection on an immobilized target. The genetic information of binders is rescued by RT-PCR yielding a PCR product. After several cycles highly specific and pure binders are selected [55]. Figure reproduced from [56]. **d** Schematic of systematic evolution of ligands by exponential enrichment (*SELEX*). SELEX is used for the selection of nucleic acids including single-stranded DNA, RNA, or aptamers that specifically bind to a target ligand. Synthetic libraries of nucleic acids are screened to select for nucleic acids that bind the target molecule. After several cycles of selection and PCR-amplification specific high affinity binders are identified. Figure reproduced from [57]

3.2.1.4 Lectins

Lectins are proteins that bind carbohydrates and mediate cell to cell contacts. These contacts are strong by the combination of many individual lectins–sugar weak and specific interactions. Similarly, these interactions can open many doors in their use in biotechnology and for the generation of specific biorecognition surfaces [43–45]. For example, concanavalin A (ConA) is a commercially available lectin that binds specifically to defined glycosyl moieties found in various sugars, glycoproteins, and glycolipids (Fig. 3.1e). ConA is widely used as a tool in sugar biorecognition, detection, and sensing applications [46-49]. Recent works on the carbohydrate–lectin recognition mechanisms open the door to the efficient design of new lectin-based sugar recognition modules [50, 51].

3.2.2 Engineering to Expand the Binding Repertoire and Affinity of Natural Scaffolds

Protein engineers soon exhausted the natural repertoire of protein–ligand in their designs, and sought to engineer new ones. In this section, we describe approaches that take a protein scaffold from nature, and functionalize it with new amino acids that allow binding to a new target. Typically, this is done by repeatedly generating large libraries of variants and selecting for increasingly high-affinity binders. Figure 3.2 summarizes methods used to accomplish this goal. Successful binding proteins displayed on the surfaces of phage, bacteria, or yeast, or attached to ribosomes, can be physically isolated by binding to immobilized ligand. Proteins displayed on cells can be mixed with fluorescent ligand and sorted in a fluorescent activated cell sorter (FACS). As the gene that produced the protein is coisolated in all cases, DNA sequences are obtained, allowing for determination of successful substitutions across many clones, and serves as a foundation for the next round of variation and selection.

Scaffold choice itself is important. Scaffolds with sufficient surface area to bury and, in particular, ones with loops that project into their target, seem to frequent studies that successfully isolate high-affinity binders. Stable scaffolds that express well, that are stable, and that are structurally predictably are desirable. To this end, consensus sequence design has been used. Many sequences of the same domain are aligned, and the consensus construct is generated. Consistent with the evolutionary hypothesis that the most conserved residues are responsible for structural cohesion of the domain, these consensus designs are often extremely stable, even more so than their natural counterparts [58–60]. The least-conserved hypervariable positions are what define the different binding functions of each member of the protein family, and are, therefore, varied in the libraries in order to obtain novel functions (Fig. 3.1c).

Another deliberate design feature that improves binding is the incorporation of charges complimentary to the ligand being bound. Long-range electrostatic interac-

tions apparently attract and position the ligand before the specific interactions of the binding site lock it in. On rates, and therefore affinities, are dramatically increased in electrostatically optimized systems of very different folds [61–63].

The success of these approaches cannot be understated. Several scaffolds have emerged into mature technologies with commercial and even some clinical success. They demonstrate antibody-like affinities in the nanomolar to picomolar range. Their specificities are precise. Further, they can be engineered with modules coupled together for avidity effects. Modules with different affinities can be coupled for bi- or multi-specific binding.

The basis of higher and lower affinity binders are not fully understood, but the structural biology work on many biorecognition complexes has provided valuable information to decipher the key elements in the biorecognition interactions. Figure 3.3 provides detailed structural comparison between high- and low-affinity binders.

Affibodies are based on a Protein A scaffold. This is a helical bundle which originally had positions on two alpha helices randomized. Selection primarily has been performed by rounds of phage display (Fig. 3.2b). Affibodies have been commercialized, resulting in over 200 publications with applications ranging from protein capture and purification to enzyme inhibition to a clinical trial for an anti-Her2neu imaging reagent [66–70].

DARPins (designer ankyrin repeat proteins) are based on a consensus ankyrin repeat scaffold. A single ankyrin repeat is typically 33 amino acids and consists of a β-turn, followed by two antiparallel α-helices and a loop leading to the next repeat. Hydrophobic and polar interactions within a repeat, and with neighboring repeats, form a stable, extended hydrophobic core mediated by the stacking of the α-helices, while the more flexible loops project out [58, 71, 72]. Like other consensus designs, DARPins are very stable [58]. Their ability to accept any residue type in loop positions has allowed the development of DARPin libraries, from which high-affinity binders of therapeutic and diagnostic targets have been selected by ribosome display (Fig. 3.2c; [73–75]).

Adnectins are based on a fibronectin iii fold, resulting in antibody-like loops projecting out of a core comprised of sheets. They have been selected by a variety of methods including yeast display (Fig. 3.2a) and phage display (Fig. 3.2b). Adnectins are commercialized and have entered clinical trials [76–78]. Many other scaffolds have been designed and selected for novel binding affinities. The goal of this review is to highlight a few cases and their development. Other scaffolds, design principles, and structural considerations [79] have been extensively reviewed elsewhere [80].

3.2.2.1 Computationally Designed Proteins

Rather than randomizing the amino acids on a scaffold, the Baker lab has selected scaffolds to complement a computer-designed binding site [65]. DIG (digoxigenin), a drug and easily incorporated DNA/RNA modification commonly used by mo-
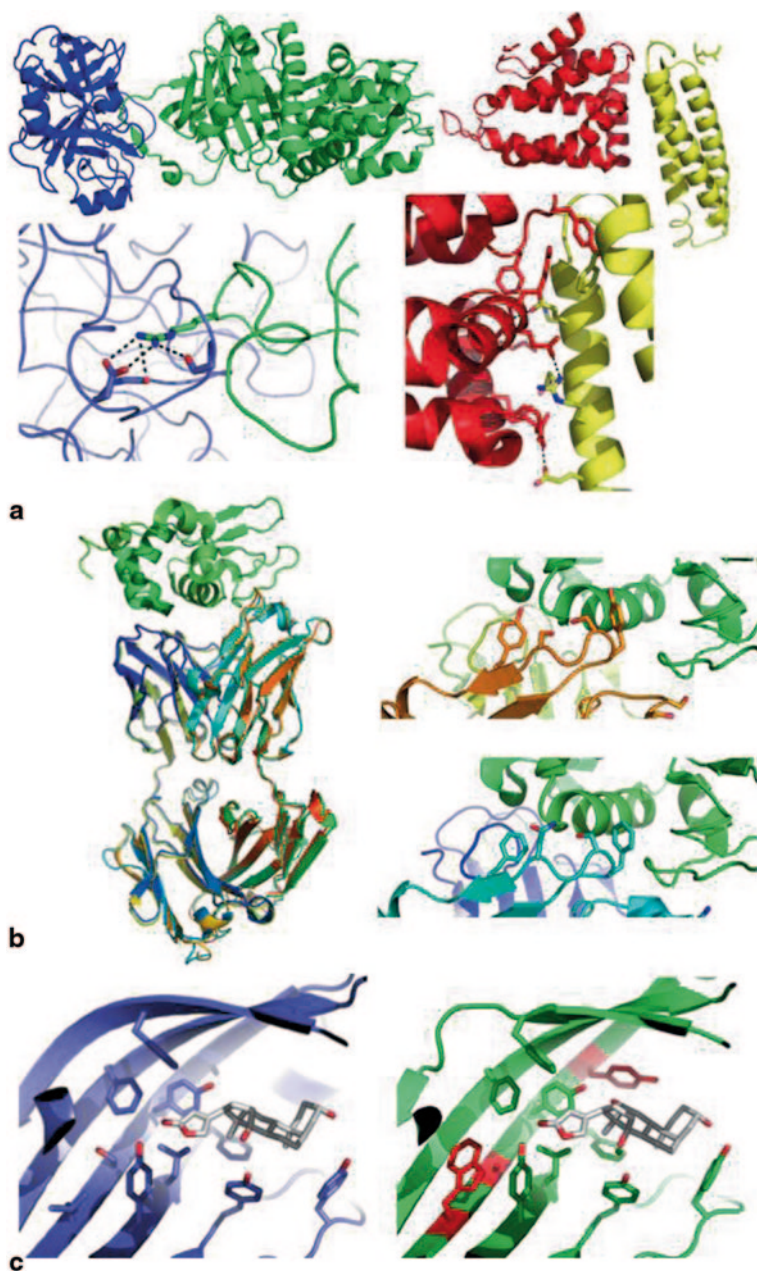
Fig. 3.3 Structural comparisons of higher and lower affinity binders. **a** A potential natural hot-spot versus surface binding is apparent when comparing the natural protein–protein complexes in proteinase–proteinase inhibitors (PDB ID: 1OPH) and a SNARE-adaptor complex (PDB ID: 2V8S). Both are protein complexes that bury almost the same surface area (1359 and 1333 Å^2), respectively. The two complexes have similar hydrophobic content at their interfaces. However, 1OPH

lecular biologists was selected as a target. "Disembodied" amino acids were aligned with DIG and optimized for binding using the ROSETTA program. Potential scaffolds from the protein database were computationally screened for their ability to harbor the desired amino acids in favored orientations. Designer proteins comprised of the selected amino acids mated to the scaffold were generated, and selected using surface display on yeast and fluorescent-activated cell sorting. From an initial affinity of ~10 μM, additional rounds of design, randomization, and selection resulted in pM to low nM affinities (Fig. 3.3c). Specificity for DIG versus other steroids of similar shapes and chemistry was developed by positive designs that inserted residues incompatible with binding undesirable ligands. Although the Baker lab sampled known protein backbone conformations in this study, it seems totally de novo protein designs are imminent [81].

3.2.3 Peptides

Peptides are distinct from proteins in their small size and general lack of folding or even secondary structure. Specific high-affinity interactions do not necessarily require protein-sized molecules. For example, antibody epitopes that mediate nanomolar binding with precise specificity can be as short as 6 amino acids. The biotin acceptor peptide (described in 2.1.1) has high affinity and specificity for streptavidin. The short lengths of peptides allow comprehensive sequence libraries to be generated. Peptide binding can be selected for by panning phage-display libraries. Libraries of random or near random sequence can also be queried as fusions to proteins in yeast-two hybrids or functional screens. Alternatively, libraries can be plated on glass, generating peptide microarrays. In these microarrays, target binding to immobilized peptides of known identity can be directly detected [82]. Peptides selected for binding are often referred to as "aptamers", which can cause some

---

(*left*) features binding 4000 times tighter than seen in 2V85 (*right*). A potential hot-spot (*inset*) is formed by a salt bridge. This interaction is clearly isolated from bulk solvent by surrounding polar and hydrophobic interactions. The SNARE-adaptor complex features as many interactions along long helices, but side chain and main chain interactions are solvent adjacent. Values from [10]. **b** During antibody affinity maturation, antibody affinity for lysozyme (*green left*) improved 36 times, resulting in nanomolar affinities. Only a few of the amino acid substitutions that occur during affinity maturation are at the antibody–antigen interface (shown as *sticks*). The immature (*right top*, PDB ID: 1NDM) and fully matured (*right bottom*, PDB ID: 1NDG) complexes bury a similar solvent-accessible surface area (overlay, *left*). The major change observed is due to the mutation of a noncontact residue that allows a key loop to move closer to the antigen. This allows replacement of tyrosines that participate in loose water-mediated hydrogen bond networks with phenylalanine that directly bind the antigen, burying additional nonpolar surface area and protecting a neighboring hot-spot. Other substitutions in the same region provide some additional contacts [64]. **c** Selection of a picomolar affinity small molecule (DIG) binding (*right*, PDB ID: 4J9A proceeded through lower affinity units (*left*, PDB ID: 4J8T). Substitutions are colored *red*. A 16–28X improvement in affinity cannot be accounted for by the single substitution that contacts DIG. Noncontact mutations, in particular a leu to trp substitution help position binding residues in optimal geometries, and lock them into binding conformations, decreasing the entropic and kinetic penalty on binding affinity if residues need to reorient prior to binding [65]

confusion as DNA and RNA modules are also called aptamers in some contexts. Peptide aptamers have been generated to inhibit cellular processes and to act as biosensors and diagnostics ([83], reviewed in [84]). Interestingly, peptides have been developed with specificity and fairly high affinity for different surfaces, including polystyrene, carbon nanotubes, and glass [85].

3.3   Nucleic Acids

Nucleic acids in addition to their fundamental role to encode the genetic information in living systems have some intrinsic features that make them useful biorecognition molecules. The Watson–Crick base pairing between G and C and A and T bases by hydrogen bonds permits the formation of stable double-stranded nucleotide chains (Fig. 3.4a). These interactions are very specific and follow very simple rules with only 4 blocks A, T, G, and C and two possible interactions G–C and A–T, therefore can be easily used to engineer and program interactions based on the DNA base pairing.

In order to generate different binding specificities, nucleic acids present less chemical functionality than proteins but large libraries of oligonucleotides with different sequences can be generated relatively easily. In addition to the simple base pair complementarity, nucleic acids can fold into complex three-dimensional structures (Fig. 3.4b, 3.4c). Those structures have been shown to be able to bind specific target molecules (Fig. 3.4c; [86, 87]). Large libraries of nucleotide sequences will potentially encode binding molecules to almost any potential target. Advanced selection tools including Systematic Evolution of Ligands by EXponential Enrichment (SELEX) are used to select from complex libraries DNA molecules that bind specifically the different targets of interest (Fig. 3.2d; [88]). Aptamers are small single-stranded nucleic acids that fold in defined structure (Fig. 3.4d). High-affinity aptamers have been developed to bind a wide variety of molecules and even complex systems such as live cells [82, 89]. The affinity and specificity of those molecules can be evolved and selected to match the desired properties. The expansion of the natural repertoire of nucleotides by the introduction of unnatural nucleotides can increase the functionality and affinity of aptamers by providing additional chemical and structural diversity beyond that available with modified natural nucleotides [90].

In addition, to use nucleic acids as stapling and binding molecules, there are new advanced technologies for the generation of complex structures based on the assembly of DNA molecules exploiting the same simple base paring complementarity [91]. Nowadays using these methodologies, that include the DNA origami, scientist can create almost any type of DNA-based shapes [92]. The DNA structures could be functionalized primarily through selective attachment of functional groups, such as biotin [93–95]. These novel technologies will allow not only to generate complex biorecognition platforms but also to pattern chemical reactions at the surfaces.
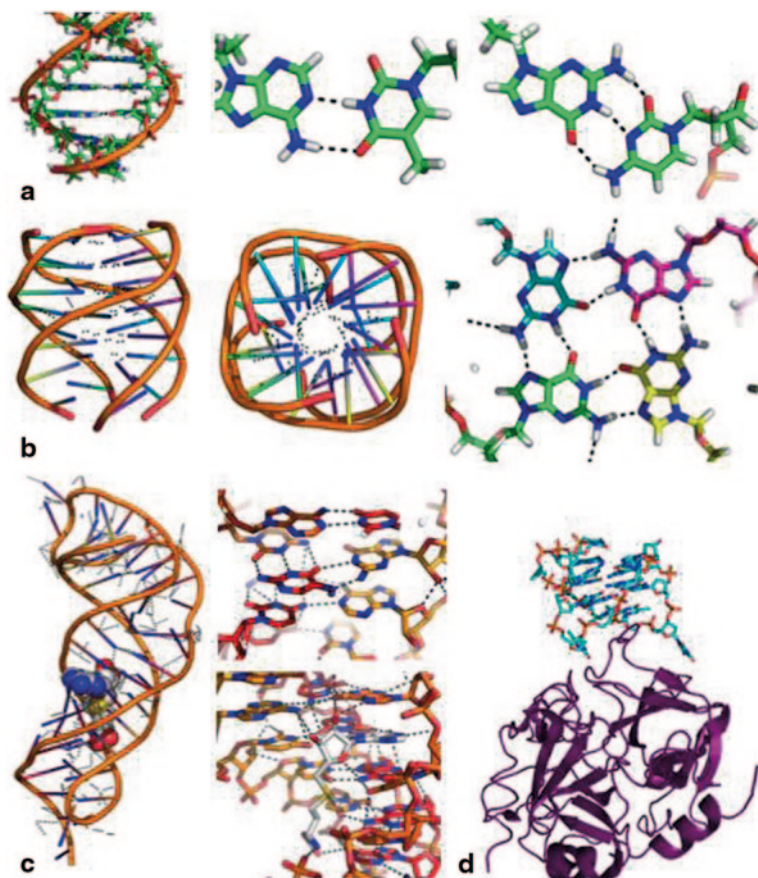
Fig. 3.4 Nucleic acids-based interactions. **a** A typical B-DNA helix and Watson–Crick A:T and G:C base pairs. (PDB ID: 2VAH). Hydrogen bonds are drawn as *black dashed lines*. Bases stack in a regular orientation, and traditional geometries are observed. **b** Quadruplex DNA from side on and top orientations. (PDB ID: 139D). Base pairing in quadruplex DNA is extensive, and involves additional hydrogen bonds not seen in double stranded DNA helices. **c** The S-adenosylmethionine binding domain of an RNA riboswitch (PDB ID: 2QWY). S-adenosylmethionine is rendered with *grey* carbons. The structure is an RNA helix formed by one RNA that loops back on itself. Non-Watson–Crick base pairs stabilize the binding site (*upper right*). S-adenosylmethionine binding includes extensive hydrogen bond with the nucleotides, and base stacking of the adenosine ring (*lower right*). **d** Crystal structure of an aptamer, a structure oligonucleotide, recognizing its molecular target. The figure shows the overall structure of the best-known aptamer, the thrombin-binding aptamer (*TBA*) in complex with thrombin. Thrombin molecule is representedin *purple*. TBA molecule is represented on *top* as *cyan sticks* (PDB ID: 3QLP) [96]

3.4   Summary and Conclusions

In summary, the natural variety of molecular recognition modules provides a wide tool-set for encoding and grafting specific biorecognition activities into polymeric surfaces. The versatility of the modules permits the construction of biorecognition platforms with unique binding properties in terms of affinity and specificity. The key aspects of biorecognition including binding affinity, avidity, and specificity, need to be comprehended and considered for the selection of the optimal molecular pairs for each application. The fast development of biomolecular engineering techniques is expanding the natural diversity even more and allows exploring a new landscape of molecules. The near future will deliver a complete new collection of tailored interactions and biorecognition molecules with fine-tuned properties that will allow scientist to build biorecognition systems "à *la carte*".

References

1. Dougherty, D.A. The cation-pi interaction. Acc Chem Res, 2013, 46, 885–893.
2. Hughes, R.M.; Wiggins, K.R.; Khorasanizadeh, S.; Waters, M.L. Recognition of trimethyl-lysine by a chromodomain is not driven by the hydrophobic effect. Proc Natl Acad Sci U S A 2007, 104, 11184–11188.
3. Sinclair, J.C.; Davies, K.M.; Venien-Bryan, C.; Noble, M.E. Generation of protein lattices by fusing proteins with matching rotational symmetry. Nat Nanotechnol 2011, 6, 558–562.
4. Salgado, E.N.; Radford, R.J.; Tezcan, F.A. Metal-directed protein self-assembly. 2010, 43, 661–672.
5. Mejías, S.H.; Sot, B.; Guantes, R.; Cortajarena, A.L. Controlled nanometric fibers of self-assembled designed protein scaffolds. Nanoscale 2014, 6, 10982–10988.
6. Sinclair, J.C. Constructing arrays of proteins. Curr Opin Chem Biol 2013, 17, 946–951.
7. Bozic, S.; Doles, T.; Gradisar, H.; Jerala, R. New designed protein assemblies. Curr Opin Chem Biol 2013, 17, 940–945.
8. Gradisar, H.; Bozic, S.; Doles, T.; Vengust, D.; Hafner-Bratkovic, I.; Mertelj, A.; Webb, B.; Sali, A.; Klavzar, S.; Jerala, R. Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. Nature chemical biology 2013, 9, 362–366.
9. Yan, C.; Wu, F.; Jernigan, R.L.; Dobbs, D.; Honavar, V. Characterization of protein-protein interfaces. Proteins 2007, 27, 59–70.
10. Chen, J.; Sawyer, N.; Regan, L. Protein-protein interactions: General trends in the relationship between binding affinity and interfacial buried surface area. *Protein Sci* **2013**, *22*, 510–515.
11. Kastritis, P.L.; Bonvin, A.M. On the binding affinity of macromolecular interactions: Daring to ask why proteins interact. *J. R. Soc. Interface*. **2013**, *10*, 20120835
12. Bogan, A.A.; Thorn, K.S. Anatomy of hot spots in protein interfaces. *J Mol Biol* **1998**, *280*, 1–9.
13. Desrosiers, D.C; Peng, Z.Y. A binding free energy hot spot in the ankyrin repeat protein gabpbeta mediated protein-protein interaction. *J Mol Biol* **2005**, *354*, 375–384.
14. Keskin, O.; Ma, B.; Nussinov, R. Hot regions in protein–protein interactions: The organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol.* **2005**, *345*, 1281–1294.

15. Chilkoti, A.; Tan, P.H.; Stayton, P.S. Site-directed mutagenesis studies of the high-affinity streptavidin-biotin complex: Contributions of tryptophan residues. *Proc Natl Acad Sci U S A* **1995**, *92*, 1754–1758

16. Rudnick, S.I.; Adams, G.P. Affinity and avidity in antibody-based tumor targeting. *Cancer Biother Radiopharm*, **2009**, *24*, 155–161.

17. Mattes, M.J. Binding parameters of antibodies: Pseudo-affinity and other misconceptions. *Cancer Immunol Immunother* **2005**, *54*, 513–516.

18. Cho HS; Mason K; Ramyar KX; Stanley AM; Gabelli SB; Denney DW Jr; DJ., L. Structure of the extracellular region of her2 alone and in complex with the herceptin fab. *Nature* **2003**, *421*, 756–760.

19. Hyre DE; Le Trong I; Merritt EA; Eccleston JF; Green NM;Stenkamp RE; PS., S. Cooperative hydrogen bond interactions in the streptavidin-biotin system. *Protein Sci* **2006**, *15*, 459–467.

20. Cortajarena, A.L.; Wang, J.; Regan, L. Crystal structure of a designed tpr module in complex with its peptide-ligand. *Febs J* **2010**, *277*, 1058–1066.

21. Nagae, M.; Re, S.; Mihara, E.; Nogi, T.; Sugita, Y.; Takagi, J. Crystal structure of α5β1 integrin ectodomain: Atomic details of the fibronectin receptor. *J Cell Biol* **2012**, *197*, 131–140.

22. Naismith, J.H.; Field, R.A. Structural basis of trimannoside recognition by concanavalin a. *J Biol Chem* **1996**, *271*, 972–976.

23. Weber, P.C.; Ohlendorf, D.H.; Wendoloski, J.J.; Salemme, F.R. Structural origins of high-affinity biotin binding to streptavidin. *Science* **1989**, *243*, 85–88.

24. DeChancie, J.; Houk, K.N. The origins of femtomolar protein-ligand binding: Hydrogen-bond cooperativity and desolvation energetics in the biotin-(strept)avidin binding site. *Journal of the American Chemical Society* **2007**, *129*, 5419–5429.

25. Cronan, J.E., Jr.; Reed, K.E. Biotinylation of proteins in vivo: A useful posttranslational modification for protein analysis. *Methods in enzymology* **2000**, *326*, 440–458.

26. Chapman-Smith, A.; Cronan, J.E. Jr. In vivo enzymatic protein biotinylation. *Biomolecular engineering* **1999**, *16*, 119–125.

27. de Boer, E.; Rodriguez, P.; Bonte, E.; Krijgsveld, J.; Katsantoni, E.; Heck, A.; Grosveld, F.; Strouboulis, J. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 7480–7485.

28. Schmidt, T.G.; Skerra, A. One-step affinity purification of bacterially produced proteins by means of the "strep tag" and immobilized recombinant core streptavidin. *Journal of chromatography. A* **1994**, *676*, 337–345.

29. Schmidt, T.G.; Koepke, J.; Frank, R.; Skerra, A. Molecular interaction between the strep-tag affinity peptide and its cognate target, streptavidin. *Journal of molecular biology* **1996**, *255*, 753–766.

30. Howarth, M.; Chinnapen, D.J.; Gerrow, K.; Dorrestein, P.C.; Grandy, M.R.; Kelleher, N.L.; El-Husseini, A; Ting, A.Y. A monovalent streptavidin with a single femtomolar biotin binding site. *Nature methods* **2006**, *3*, 267–273.

31. Wu, S.C.; Wong, S.L. Engineering soluble monomeric streptavidin with reversible biotin binding capability. *J Biol Chem* **2005**, *280*, 23225–23231.

32. Lyonnais, S.; Goux-Capes, L.; Escude, C.; Cote, D.; Filoramo, A.; Bourgoin, J.P. DNA-carbon nanotube conjugates prepared by a versatile method using streptavidin-biotin recognition. *Small* **2008**, *4*, 442–446.

33. Chen, H.M.; Lin, C.W. Hydrogel-coated streptavidin piezoelectric biosensors and applications to selective detection of strep-tag displaying cells. *Biotechnology progress* **2007**, *23*, 741–748.

34. Sundberg, E.J. Structural basis of antibody-antigen interactions. *Methods in molecular biology* **2009**, *524*, 23–36.

35. Graille, M.; Stura, E.A.; Corper, A.L.; Sutton, B.J.; Taussig, M.J.; Charbonnier, J.B.; Silverman, G.J. Crystal structure of a staphylococcus aureus protein a domain complexed with the fab fragment of a human igm antibody: Structural basis for recognition of b-cell receptors and

superantigen activity. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 5399–5404.

36. Sjobring, U.; Bjorck, L.; Kastern, W. Streptococcal protein g. Gene structure and protein binding properties. *J Biol Chem* **1991**, *266*, 399–405.

37. Marks, J.D.; Hoogenboom, H.R.; Bonnert, T.P.; McCafferty, J.; Griffiths, A.D.; Winter, G. By-passing immunization. Human antibodies from v-gene libraries displayed on phage. *Journal of molecular biology* **1991**, *222*, 581–597.

38. Hawkins, R.E.; Russell, S.J.; Winter, G. Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *Journal of molecular biology* **1992**, *226*, 889–896.

39. Vaughan, T.J.; Williams, A.J.; Pritchard, K.; Osbourn, J.K.; Pope, A.R.; Earnshaw, J.C.; McCafferty, J.; Hodits, R.A.; Wilton, J.; Johnson, K.S. Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat Biotechnol* **1996**, *14*, 309–314.

40. Vaughan, T.J.; Osbourn, J.K.; Tempest, P.R. Human antibodies by design. *Nat Biotechnol* **1998**, *16*, 535–539.

41. Rouslahti, E. Rgd and other recognition sequences for integrins. *Annu Rev Cell Dev Biol.* **1996**, *12*, 697–715.

42. Takagi, J. Structural basis for ligand recognition by rgd (arg-gly-asp)-dependent integrins. *Biochemical Society transactions* **2004**, *32*, 403–406.

43. Yu K; Creagh AL; Haynes CA; JN., K. Lectin interactions on surface-grafted glycostructures: Influence of the spatial distribution of carbohydrates on the binding kinetics and rupture forces. *Anal. Chem.* **2013**, *85*, 7786–7793.

44. Gabius HJ; Siebert HC; André S; Jiménez-Barbero J; H., R. Chemical biology of the sugar code. *Chembiochem.* **2004**, *7:5*, 740–764.

45. Wittmann V; Pieters RJ. Bridging lectin binding sites by multivalent carbohydrates. *Chem. Soc. Rev.* **2013**, *42*, 4492–4503.

46. García R; Rodríguez R; Montesino R; Besada V; González J; JA., C. Concanavalin a- and wheat germ agglutinin-conjugated lectins as a tool for the identification of multiple n-glycosylation sites in heterologous protein expressed in yeast. *Anal. Biochem* **1995**, *231*, 342–348.

47. Cella LN; Chen W; Myung NV; A., M. Single-walled carbon nanotube-based chemiresistive affinity biosensors for small molecules: Ultrasensitive glucose detection. *J. Am. Chem. Soc.* **2010**, 14.

48. Duke J; Janer L; M., C. Sem visualization of glycosylated surface molecules using lectin-coated microspheres. *J Electron Microsc Tech* **1985**, *2*, 173–174.

49. Schiefer HG;Krauss H;Brunner H;U, G. Ultrastructural visualization of surface carbohydrate structures on mycoplasma membranes by concanavalin a. *J Bacteriol* **1975**, *124*, 1598–1600.

50. Ponader D; Maffre P; Aretz J; Pussak D; Ninnemann NM; Schmidt S; Seeberger PH; Rademacher C; Nienhaus GU; L., H. Carbohydrate-lectin recognition of sequence-defined hetero-multivalent glycooligomers. *J. Am. Chem. Soc.* **2014**, *136*, 2008–2016.

51. Sato Y; Yoshioka K; Murakami T; Yoshimoto S; O., N. Design of biomolecular interface for detecting carbohydrate and lectin weak interactions. *Langmuir* **2012**, *28*, 1846–1851.

52. Boder, E.; Wittrup, D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology* **1997**, *15*, 553–557.

53. Boder ET; Raeeszadeh-Sarmazdeh M; JV., P. Engineering antibodies by yeast display. *Arch Biochem Biophys.* **2012**, *526*, 99–106.

54. Singh A; Poshtiban S; S., E. Recent advances in bacteriophage based biosensors for food-borne pathogen detection. *Sensors* **2013**, *13*, 1763–1786.

55. Hanes, J.; Plückthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **1997**, *94*, 4937–4942.

56. Kanamori, T.; Fujino, Y.; Ueda, T. Pure ribosome display and its application in antibody technology. *Biochim Biophys Acta* **2014**, *1844*, 1925–1932.

57. Veedua, R.N.; Wengel, J. Locked nucleic acid nucleoside triphosphates and polymerases: On the way towards evolution of lnaaptamers. *Mol. BioSyst.* **2009**, *5*, 797–792.

58. Kohl, A.; Binz, H.K.; Forrer, P.; Stumpp, M.T.; Pluckthun, A.; Grutter, M.G. Designed to be stable: Crystal structure of a consensus ankyrin repeat protein. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 1700–1705.

59. Jarymowycz, V.A.; Cortajarena, A.L.; Regan, L.; Stone, M.J. Comparison of the backbone dynamics of a natural and a consensus designed 3-tpr domain. *Journal of biomolecular NMR* **2008**, *41*, 169–178.

60. Wezner-Ptasinska, M.; Krowarsch, D.; Otlewski, J. Design and characteristics of a stable protein scaffold for specific binding based on variable lymphocyte receptor sequences. *Biochimica et biophysica acta* **2011**, *1814*, 1140–1145.

61. Kajander, T.; Sachs, J.N.; Goldman, A.; Regan, L. Electrostatic interactions of hsp-organizing protein tetratricopeptide domains with hsp70 and hsp90: Computational analysis and protein engineering. *J Biol Chem* **2009**, *284*, 25364–25374.

62. Selzer, T.; Albeck, S.; Schreiber, G. Rational design of faster associating and tighter binding protein complexes. *Nature structural biology* **2000**, *7*, 537–541.

63. Schreiber, G.; Fersht, A.R. Rapid, electrostatically assisted association of proteins. *Nature structural biology* **1996**, *3*, 427–431.

64. Li Y; Li H; Yang F; Smith-Gill SJ; RA., M. X-ray snapshots of the maturation of an antibody response to a protein antigen. *Nat Struct Biol* **2003**, *10*, 482–488.

65. Tinberg, C.E.; Khare, S.D.; Dou, J.; Doyle, L.; Nelson, J.W.; Schena, A.; Jankowski, W.; Kalodimos, C.G.; Johnsson, K.; Stoddard, B.L., *et al*. Computational design of ligand binding proteins with high affinity and selectivity. *Nature* **2013**, *501*, 212–216.

66. Orlova, A.; Magnusson, M.; Eriksson, T.L.; Nilsson, M.; Larsson, B.; Hoiden-Guthenberg, I.; Widstrom, C.; Carlsson, J.; Tolmachev, V.; Stahl, S., *et al*. Tumor imaging using a picomolar affinity her2 binding affibody molecule. *Cancer research* **2006**, *66*, 4339–4348.

67. Ahlgren, S.; Orlova, A.; Wallberg, H.; Hansson, M.; Sandstrom, M.; Lewsley, R.; Wennborg, A.; Abrahmsen, L.; Tolmachev, V.; Feldwisch, J. Targeting of her2-expressing tumors using 111in-aby-025, a second-generation affibody molecule with a fundamentally reengineered scaffold. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine* **2010**, *51*, 1131–1138.

68. Wahlberg, E.; Lendel, C.; Helgstrand, M.; Allard, P.; Dincbas-Renqvist, V.; Hedqvist, A.; Berglund, H.; Nygren, P.A.; Hard, T. An affibody in complex with a target protein: Structure and coupled folding. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 3185–3190.

69. Nord, K.; Gunneriusson, E.; Ringdahl, J.; Stahl, S.; Uhlen, M.; Nygren, P.A. Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat Biotechnol* **1997**, *15*, 772–777.

70. Nord, K.; Nord, O.; Uhlen, M.; Kelley, B.; Ljungqvist, C.; Nygren, P.A. Recombinant human factor viii-specific affinity ligands selected from phage-displayed combinatorial libraries of protein a. *European journal of biochemistry/ FEBS* **2001**, *268*, 4269–4277.

71. Mosavi, L.K.; Cammett, T.J.; Desrosiers, D.C.; Peng, Z.Y. The ankyrin repeat as molecular architecture for protein recognition. *Protein science: a publication of the Protein Society* **2004**, *13*, 1435–1448.

72. Mosavi, L.K.; Minor, D.L., Jr.; Peng, Z.Y. Consensus-derived structural determinants of the ankyrin repeat motif. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 16029–16034.

73. Binz, H.K.; Amstutz, P.; Kohl, A.; Stumpp, M.T.; Briand, C.; Forrer, P.; Grutter, M.G.; Pluckthun, A. High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol* **2004**, *22*, 575–582.

74. Zahnd, C.; Pecorari, F.; Straumann, N.; Wyler, E.; Pluckthun, A. Selection and characterization of her2 binding-designed ankyrin repeat proteins. *The Journal of biological chemistry* **2006**, *281*, 35167–35175.

75. Boersma, Y.L.; Chao, G.; Steiner, D.; Wittrup, K.D.; Pluckthun, A. Bispecific designed ankyrin repeat proteins (darpins) targeting epidermal growth factor receptor inhibit a431 cell proliferation and receptor recycling. *The Journal of biological chemistry* **2011**, *286*, 41273–41285.

76. Lipovsek, D. Adnectins: Engineered target-binding protein therapeutics. *Protein engineering, design & selection: PEDS* **2011**, *24*, 3–9.

77. Khan, J.A.; Camac, D.M.; Low, S.; Tebben, A.J.; Wensel, D.L.; Wright, M.C.; Su, J.; Jenny, V.; Gupta, R.D.; Ruzanov, M., *et al*. Developing adnectins that target src co-activator binding to pxr: A structural approach toward understanding promiscuity of pxr. *Journal of molecular biology* **2015**.

78. Ackermann, M.; Morse, B.A.; Delventhal, V.; Carvajal, I.M.; Konerding, M.A. Anti-vegfr2 and anti-igf-1r-adnectins inhibit ewing's sarcoma a673-xenograft growth and normalize tumor vascular architecture. *Angiogenesis* **2012**, *15*, 685–695.

79. Gilbreth, R.N.; Koide, S. Structural insights for engineering binding proteins based on non-antibody scaffolds. *Current opinion in structural biology* **2012**, *22*, 413–420.

80. Hosse, R.J.; Rothe, A; Power, B.E. A new generation of protein display scaffolds for molecular recognition. *Protein science: a publication of the Protein Society* **2006**, *15*, 14–27.

81. MacDonald, J.T.; Maksimiak, K.; Sadowski, M.I.; Taylor, W.R. De novo backbone scaffolds for protein design. *Proteins* **2010**, *78*, 1311–1325.

82. Tombelli, S.; Mascini, M. Aptamers as molecular tools for bioanalytical methods. *Curr Opin Mol Ther*. **2009**, *11*, 179–188.

83. Reverdatto, S.; Rai, V.; Xue, J.; Burz, D.S.; Schmidt, A.M.; Shekhtman, A. Combinatorial library of improved peptide aptamers, clips to inhibit rage signal transduction in mammalian cells. *PloS one* **2013**, *8*, e65180.

84. Colas, P. The eleven-year switch of peptide aptamers. *Journal of biology* **2008**, *7*, 2.

85. Aubin-Tam, M.E.; Appleyard, D.C.; Ferrari, E.; Garbin, V.; Fadiran, O.O.; Kunkel, J.; Lang, M.J. Adhesion through single peptide aptamers. *The journal of physical chemistry. A* **2011**, *115*, 3657–3664.

86. Ellington, A.D.; Szostak, J.W. In vitro selection of rna molecules that bind specific ligands. *Nature* **1900**, *346*, 818–822.

87. Kimoto, M.; Yamashige, R.; Matsunaga, K.; Yokoyama, S.; Hirao, I. Generation of high-affinity DNA aptamers using an expanded genetic alphabe. *Nature biotechnology* **2013**, *31*, 453–457.

88. Tuerk C; L., G. Systematic evolution of ligands by exponential enrichment: Rna ligands to bacteriophage t4 DNA polymerase. *Science* **1990**, *249*, 505–510.

89. Bunka DH; PGl., S. Aptamers come of age – at last. *Nat Rev Microbiol*. **2006**, *4*, 588–596.

90. Kimoto, M.; Yamashige, R.; Matsunaga, K.; Yokoyama, S.; Hirao, I. Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat Biotechnol* **2013**, *31*, 453–457.

91. Seeman, N.C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **1991**, *350*, 631–633.

92. Rothemund, P.W.K. Folding DNA to create nanoscale shapes and patterns *Nature* **2006**, *440*, 297–302

93. Busuttil K;Rotaru A;Dong M;Besenbacher F; KV., G. Transfer of a protein pattern from self-assembled DNA origami to a functionalized substrate. *Chem Comm* **2013**, *49*, 1927–1929.

94. Voigt NV; Tørring T; Rotaru A; Jacobsen MF; Ravnsbaek JB; Subramani R; Mamdouh W; Kjems J; Mokhir A; Besenbacher F, *et al*. Single-molecule chemical reactions on DNA origami. *Nat Nanotechnol* **2010**, *5*, 200–203.

95. Wang R; Nuckolls C; SJ., W. Assembly of heterogeneous functional nanomaterials on DNA origami scaffolds. *Angew Chem* **2012**, *51*, 11325–11327.

96. Krauss, R.; Merlino, A.; Giancola, C.; Randazzo, A.; Mazzarella, L.; Sica, F. Thrombin–aptamer recognition: A revealed ambiguity. *Nucleic Acids Res* **2011**, *39*, 7858–7867.