Chapter 1

Overview: Toward the Synthesis and Evolution of Cyclic, Unnatural mRNA Display Libraries

This chapter is adapted from the following reference (in preparation):

Takahashi, T.T., Millward, S.W., Roberts, R.W. (2007) Unnatural Display Libraries

Abstract

The development of ligands that disrupt protein-protein interactions remains a major challenge in the post-genomic era. Compounds that bind to protein surfaces with high affinity and specificity are potential tools for chemical genetics, molecular diagnostics, and therapeutic development. Traditionally, lead compounds for these applications have been designed by computational chemistry or by screening thousands of compounds from small molecule libraries based on natural products. Biological display methodologies, which can access enormous molecular diversity, are generally considered suboptimal for the design of compounds with high cellular permeability and biological stability, properties that are required for “drug-like” molecules. One of the major goals of biological display is the generation of high-diversity libraries that are biased toward evolving compounds with improved pharmacological properties. The incorporation of unnatural amino acids into peptide libraries in conjunction with macrocyclization is the first step toward the selection of highly potent compounds that reside comfortably within “drug space”. Strategies to incorporate unnatural monomers are reviewed along with their application to unnatural biological display. The advantages of cyclic peptides are discussed as well as recent efforts to generate redox-insensitive macrocyclic peptide libraries. The parallel developments of unnatural amino acid incorporation and macrocyclization may lead to wholly unnatural display libraries that can access the chemical diversity of natural products while retaining the evolvability characteristic of biological display.

Introduction

The size of the human proteome (>21000 proteins (*1*)) and the complexity of biological growth and development indicate vast networks of protein-protein interactions within the cell. The ability to selectively disrupt or attenuate these interactions is essential for understanding the global control circuits that regulate cellular behavior. One of the goals of chemical biology (chemical genomics) is the development of ligands that selectively abrogate protein-protein interactions *in vivo*, obviating the need for genomic manipulation (e.g., knockout models). In the absence of detailed structural information for novel protein targets, combinatorial methods are particularly suited to the task of rapidly designing and synthesizing high-affinity ligands.

Combinatorial design methods attempt to sample regions of chemical and conformational space using a library of diverse ligands, which can then be sieved or screened to isolate target-binding molecules. Combinatorial libraries can be crudely divided into small molecules and "biologics." Small molecule libraries are typically synthesized by solid-phase or solution-phase organic methodologies, while biologics are assembled by living organisms or from material derived from living organisms.

Small molecule libraries utilize a wide array of diverse chemical functionalities and contain low molecular weight compounds based on the scaffolds and chemical properties of known drugs and pharmacologically relevant natural products (*2-4*)**.** Small molecule libraries are particularly appropriate for the design of substrate mimics for enzyme inhibition, however their utility as modulators of protein-protein interactions (typically involving ~1,600 $Å^2$ of buried surface) remains uncertain (*5*). The structural elucidation

of “hit” compounds also remains a daunting task for high-complexity small molecule libraries.

In contrast, biologics are comprised of polymerized monomers: 20 natural amino acids for peptides and proteins, and four nucleotides for DNA and RNA. The library complexity achievable in biological display technologies (e.g., mRNA, phage, and ribosome display) is several orders of magnitude greater than synthetic small molecule libraries (compare 10^9-10^{13} for *in vitro* display libraries (*6*) versus 10^7 for small molecule libraries (*7*)), in large part due to the facile manipulation of genetic information tightly associated with each functional molecule and the great efficiency and fidelity with which the material is polymerized (Figure 1.1). Biological display can also utilize folded macromolecular scaffolds to present large surface areas for the disruption of protein-protein interactions (*8, 9*). Since the functional library member is linked to the genetic material that directs its assembly, the lead compounds obtained from biological display libraries can be readily identified and characterized. More importantly, the use of genetic material enables the evolution of the ligand through multiple rounds of selective pressure and affinity maturation, increasing the probability of finding a high affinity ligand for the target (*10-12*).

Unfortunately, these ligands often have poor pharmacokinetics owing to their low cell permeability and poor *in vivo* stability. To convert these initial leads into useful biological probes and/or pharmacological agents, medicinal chemistry must be employed to improve stability, solubility, and permeability without sacrificing the molecule’s original function. While medicinal chemistry has had a number of successes (*13*), compound optimization is non-trivial and the stepwise nature of chemical substitution

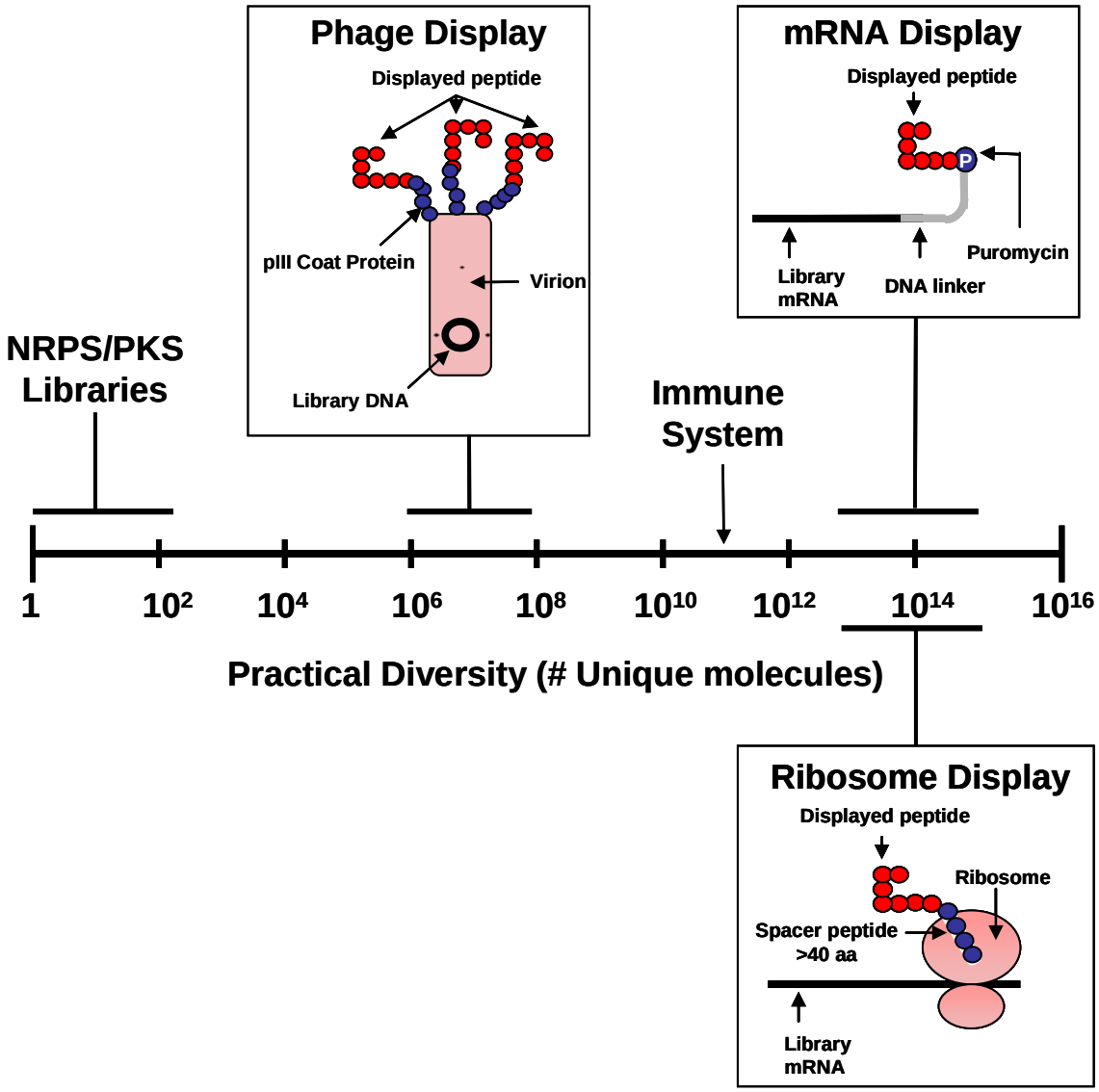

Figure 1.1: Biological display Methodologies. The scale shows the maximum practical initial diversity for the most common biological display methods.

often precludes access to co-variation effects between two or more positions. Indeed, one of the most powerful aspects of biological display is the ability to vary the chemical composition of multiple synergistic positions simultaneously, effectively identifying positions that are energetically coupled (*14-16*).

The Goals of Unnatural Biological Display

Instead of optimizing the physical characteristics of a lead molecule after it has been isolated, a more facile approach is to construct libraries from a monomer set that accurately reflects the chemical characteristics of "drug space": in effect, medicinal chemistry at the front end. While some medicinal chemistry will likely be employed on unnatural biologics after isolation, judicious choice of the monomer alphabet used to construct the library coupled with selection for improved physicochemical properties (such as protease resistance) may result in "pre-optimized" lead compounds. An ideal technology platform for ligand selection would combine the superior library complexities and evolvability in biological display systems with the more drug-like properties of unnatural amino acids (Uaas), including non-natural backbone linkages (N-methyl amino acids, α-hydroxy acids), α-carbon stereoisomers (D-amino acids), and non-natural side chains. We believe that this platform will generate high-affinity, bioactive ligands with superior pharmacokinetic properties, as compared to natural display methodologies.

This chapter describes the current progress and steps required for the creation of unnatural display libraries. We begin by addressing the chemical functionalities present in known pharmacaphores and potential routes for incorporating these functionalities into proteins as unnatural amino acids. Next, we review efforts to integrate these properties into biological display libraries using co-translational unnatural amino acid incorporation

and post-translational chemical modification. Finally, we discuss macrocyclization as a means to limit the conformational flexibility of peptides and its emergence as a tool for biological display.

In Vitro Selection as a Tool for Drug Design

Combinatorial methods for ligand isolation are currently the best approach for obtaining functional molecules. *In vitro* selection harnesses the power of biological synthesis and genetic amplification to enable the creation of libraries in excess of trillions of unique molecules (Figure 1.1). Typically, a library of biologics is sieved against a target and the functional molecules amplified, eventually resulting in subset of molecules that bind the target. SELEX enables the selection of RNA and DNA (*17*), while methods such as phage display (*18*), DNA display (*19*), ribosome display (*20*), and mRNA display (*21*) allow the selection of peptides and proteins. In general, much larger libraries can be generated using totally *in vitro* methods, since *in vivo* techniques, such as phage display, require transformation of the library into cells. Previous reviews have highlighted the many successes of these methods in generating ligands to a diverse set of targets (*6, 22*).

In order to make these biological libraries more drug-like, Uaas must either be incorporated during the synthesis of the library, or chemically derivatized after library synthesis. While unnatural amino acids can be incorporated into phage display libraries (*23*), it is difficult to incorporate more than a few Uaas because each Uaa must have its own orthogonal synthetase and tRNA. Likewise, techniques such as phage display, ribosome display, and CIS require associated proteins, which make the libraries difficult to modify post-translationally. mRNA display thus represents the most flexible method for synthesis of totally unnatural libraries since Uaas can be introduced either co-

translationally or through post-translational modification. Additionally, since the linkage between mRNA and peptide in mRNA display is mediated by a small molecule, the conversion of natural mRNA display to unnatural mRNA display only requires that the translation system be adapted to incorporate unnatural monomers.

Choosing an Unnatural Alphabet

One of the challenges in constructing an unnatural display library lies in choosing an alphabet of natural and unnatural monomers to localize the library within "drug space". Unfortunately, this begets the more complex problem of defining drug space within the nearly infinite boundaries of chemical space. Statistical surveys of known pharmacaphores have identified recurrent molecular frameworks and chemical moieties that provide an empirical sketch of drug space (*24-26*)**.** Alternatively, pharmacaphores may be described by a set of physico-chemical parameters correlated with favorable solubility and permeability (e.g., Lipinski's Rule of 5) (*27*)**.** While these principles may suggest which Uaa side chains may be useful, their bias toward small molecule drug discovery somewhat mitigates their applicability in peptide-based drug design.

Given the bias in drug discovery toward small molecules with molecular weights less than 500 Da, it may prove useful to use known bioactive peptides as models for designing unnatural peptide display libraries. Nonribosomal peptides fall outside many of the parameters used to define small molecule drug space, yet remain potent bioactive compounds. (Figure 1.2A) For example, the cyclic undecapeptide cyclosporin violates two of Lipinski's rules, yet functions as an orally bioavailable immunosuppressant (*28*). Nonribosomal peptides are synthesized by large multi-enzyme complexes and contain a wide variety of nonproteogenic side chains and connectivities (reviewed in (*29, 30*)).

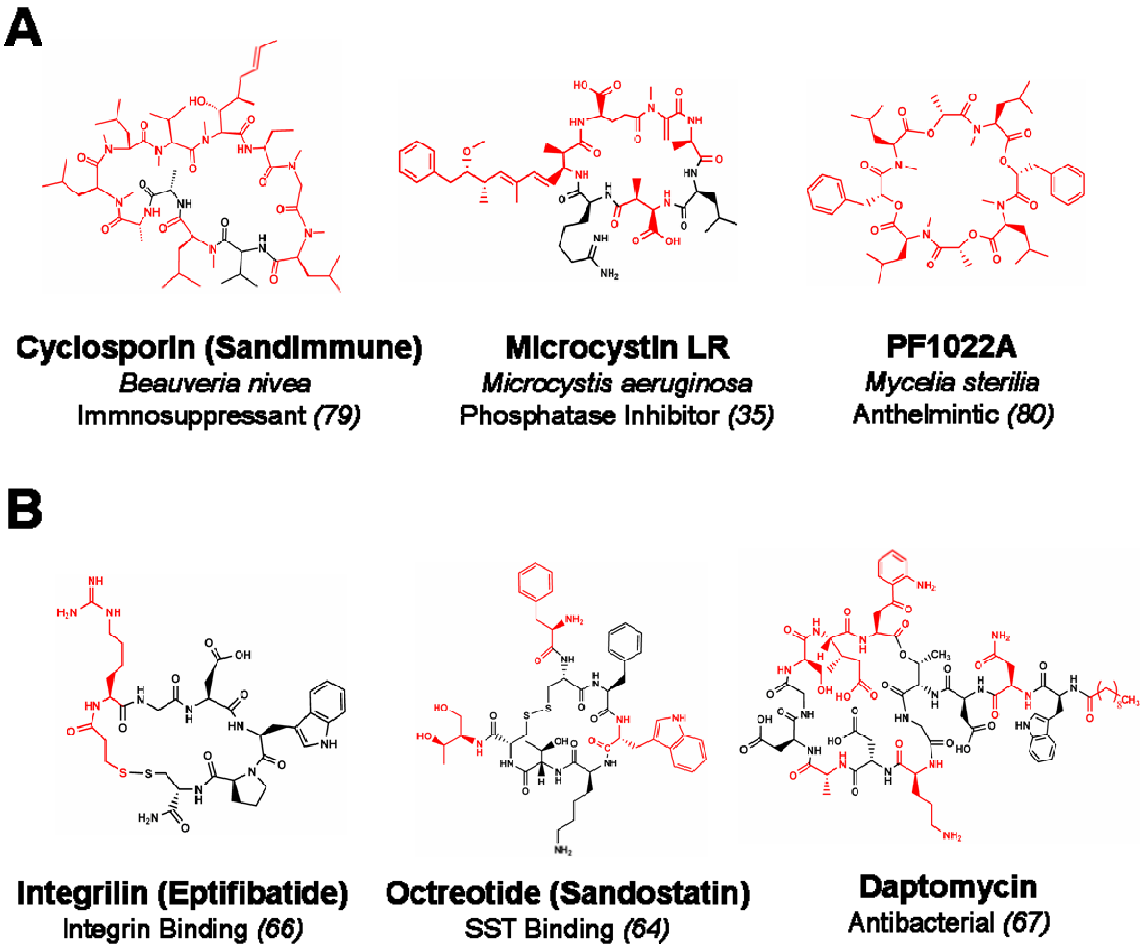

Figure 1.2: (A) Bioactive nonribosomal peptides and (B) approved therapeutic peptides. Unnatural (nonproteogenic) amino acids are shown in red. (*35*, *64*, *66*, *67*, *79*, *80*)

Enzymatic macrocyclization (*31*) is a characteristic of many nonribosomal peptides and has previously been shown to impart higher affinity and/or biological activity to linear peptide sequences (Reviewed in (*32*)**).** D-amino acids, and to a lesser extent, N-methyl amino acids, are frequently found in nonribosomal peptides and have been shown to greatly enhance the stability of linear peptide sequences to proteolytic digestion (*33, 34*). Unnatural side chains, such as the electrophilic N-methyldehydroalanine residue in microcystin, also contribute to the bioactivity of nonribosomal peptides (*35*) (Figure 1.2A). Combinatorial biosynthesis of NRPS gene clusters has been proposed as means to generate NRP libraries *(36),* although technical challenges significantly limit their diversity *(37).*

Rewriting the Genetic Code

Currently, most unnatural amino acids are introduced co-translationally into proteins and peptides through nonsense suppression. In nonsense suppression, a Uaa is first charged onto a suppressor tRNA that recognizes a stop codon (UAG, UAA, or UGA) then the charged tRNA is translated in the presence of an mRNA containing a stop codon at the desired site of Uaa incorporation. In essence, the Uaa "fills in a blank" in the genetic code, which exists because most organisms lack nonsense suppressor tRNAs. Extension of this method would allow the incorporation of up to three additional Uaas at each of the three stop codons.

However, a different strategy must be employed to synthesize wholly unnatural peptides or proteins, since there are no other natural blanks in the genetic code. First, a synthetic blank in the genetic code must be introduced: i.e., a sense codon that does not code for a natural amino acid must be generated. This newly blank sense codon is then

reassigned and suppressed by a Uaa. Using sense suppression, several or possibly even all the codons could be recoded for Uaas.

In order to generate a codon that is a synthetic blank in the genetic code, the endogenous aminoacyl-tRNA charged with the natural amino acid must be prevented from entering the ribosome. Several strategies are possible: (1) removal of the tRNA and/or natural amino acid, (2) competition with an exogenously added Uaa or Uaa-charged tRNA, (3) Inhibition or removal of the aminoacyl synthetase, or (4) a combination of two or more of the above strategies.

Removal of the endogenous tRNA can be achieved through the use of a column that binds to tRNA selectively. Jackson and coworkers describe the use of an ethanolamine-sepharose column that is effective in depleting >90% of the tRNA from the rabbit reticulocyte lysate (*38*). The cationic resin is specific for negatively-charged tRNA and does not bind ribosomes. Addition of tRNA restores the activity of the depleted lysate. However, aminoacyl tRNA-synthetases (AARS) are not removed by this procedure and any exogenously added aminoacyl-tRNA may be edited by the synthetases.

Suppression of natural amino acid incorporation may be accomplished either by addition of Uaas or Uaa-charged tRNAs. By adding high concentrations of the Uaa, the natural amino acid can be totally replaced. Most of the work detailing competition experiments is performed by addition of Uaas to minimal media used to grow *E. Coli*, resulting in the nearly complete replacement of the natural amino acid (*39*). While the replacement of a natural amino acid depends partly on the concentration of Uaa added, it is also dependent on the codon bias present in cells. In nature, not all codons exist in

equal concentrations; some codons are used more (or less) frequently than the average. In theory, the codons that are used less frequently could be reassigned with minimal incorporation of the natural amino acid.

Frankel and Roberts performed a selection in rabbit reticulocyte lysate to identify codons that would be the most amenable for substitution (*40*). mRNA display was used to select GNN codons that would allow higher levels of incorporation of a tRNA charged with biocytin. The selection identified the Val GUA codon as the most amenable to sense suppression. Further experiments recoded the GUA and GCU codons to N-methyl Phenylalanine, resulting in peptides with marked protease resistance (*41*) (Chapter 3).

While competition can be used to reassign several sense codons, it is impractical for rewriting the complete genetic code with a fidelity approaching that of natural translation. Suppression is often incomplete (leaky) and the natural amino acid can be incorporated, resulting in the synthesis of heterogeneous products. Also, large amounts of aminoacyl tRNA or Uaa must be added to the translation mixture in order to drive Uaa incorporation.

Alternately, elimination of endogenous aminoacyl-tRNA synthetases precludes the formation of natural aminoacyl tRNA while preventing deacylation of unnatural aminoacyl-tRNA. Prokaryotic (*42, 43*) and eukaryotic (*44*) reconstituted translation systems have been developed where the essential translation components (initiation factors, elongation factors, and ribosomes) are purified away from amino acids, tRNAs, and AARSs. These minimal translation systems allow any combination of amino acids (both natural and unnatural), aminoacylated tRNA, and AARSs to be added back to the system. Initial experiments using an *E. Coli* derived translation system focused on

reconstitution of the translation system with natural amino acids. The PURE system was reconstituted with all twenty AARSs and demonstrated that microgram-quantities of proteins could be synthesized (*43*). Omission of release factor 1, which recognizes the UAG stop codon, increased the incorporation of valine via a UAG-suppressor tRNA. Independently, Forster and Blacklow demonstrated that supplementing the core translation system with charged tRNAs allowed the synthesis of short peptides (*42*).

More recently, several sense codons have been reassigned for unnatural amino acid incorporation. Addition of three Uaa-charged tRNAs to a reconstituted extract enabled the synthesis of unnatural tetrapeptides, which were analyzed by HPLC (*45*). Several backbone analogs, such as N-methyl (methyl on the amide nitrogen), α-methyl (methyl on the Cα carbon), and α-hydroxy (OH- instead of NH_2-), have been incorporated (*46*).

The Szostak lab utilized the PURE system in the context of mRNA display to rewrite twelve sense codons with unnatural amino acids, generating a wholly unnatural mRNA-peptide fusion whose identity was confirmed by MALDI-MS (*47*). These experiments demonstrated the feasibility of constructing wholly unnatural mRNA display libraries using sense codon suppression in a reconstituted translation extract.

Generating the Aminoacyl-tRNAs

Once a synthetic blank is generated in the genetic code, this blank codon can then be reassigned to another amino acid. To do this, a cognate tRNA charged with the amino acid must be added to the translation reaction. This aminoacyl tRNA can either be synthesized chemically or generated through catalysis with aminoacyl tRNA synthetases or ribozymes.

Most UAA-charged tRNAs are generated by chemical methods (Figure 1.3). First, the UAA is attached to the dinucleotide pdCpA and then ligated via T4 RNA Ligase to the 3' end of a tRNA lacking the last two nucleotides (-rCrA). This method has been described in detail and allows the greatest flexibility, as virtually any amino acid can be charged onto a tRNA (*48-51*). However, the method is time consuming and costly, thereby limiting the number of UAA-acylated tRNAs that can be synthesized.

Another strategy for obtaining Uaa-charged tRNA is to modify a natural amino acid after it is charged on the tRNA, but before it is used in translation. N-methyl amino acyl-tRNAs have been synthesized by methylation of bulk aminoacylated *E. coli* tRNA (*52*). Unfortunately, since the transformation is carried out indiscriminately on the bulk tRNA, this method forces the user to choose between an entirely natural amino acid library and an entirely N-methylated amino acid library

In nature, aminoacyl tRNAs are synthesized by the aminoacyl synthetases (AARSs), which enforce the genetic code by ensuring that a specific amino acid is charged to its cognate tRNA. AARSs can be evolved to charge an orthogonal tRNA with a specific Uaa (*53*), or can be engineered to broaden their substrate specificity, often by either expanding the amino acid binding pocket or by mutating editing domains (*54, 55*). Many non-intuitive UAAs are charged by the synthetases, as discovered by a MALDI-TOF MS-based screening of AARS charging activities (*56*).

One advantage of this strategy is that the aminoacyl tRNAs can be generated *in situ* merely by adding the UAAs and synthetases to the translation reaction, and is therefore technically easy to implement. This strategy has been used to synthesize peptides containing twelve UAAs (*47*). However, this method is limited to UAAs that

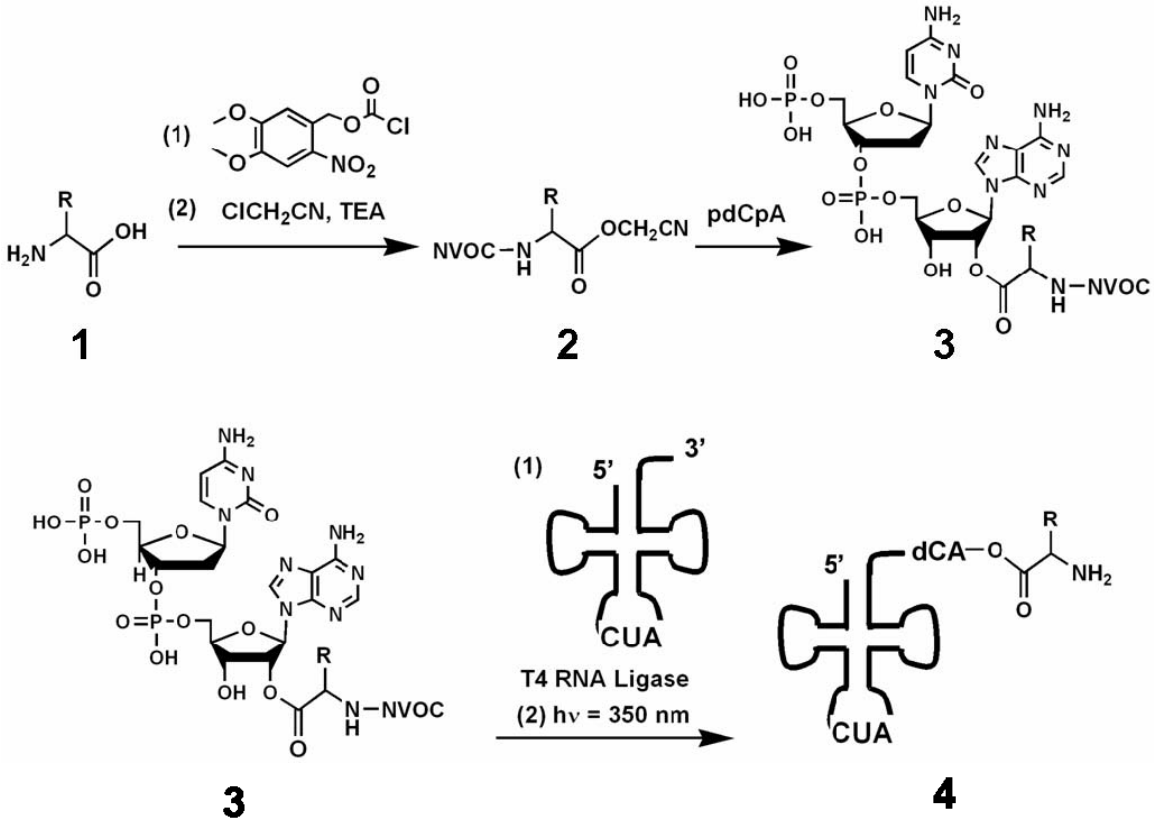

Figure 1.3: Chemical aminoacylation of nonsense suppressor tRNA. The α-amino group on **1** is protected with a photolabile NVOC group and the carboxylate is activated as a cyanomethyl ester to form compound **2**. Coupling to the pdCpA dinucleotide generates compound **3** (note the 2' adduct is shown although reaction may occur at either the 2' or 3' position). The NVOC-aminoacyl-pdCpA is enzymatically ligated to a suppressor tRNA lacking the terminal CA dinucleotide. The NVOC-aminoacyl-tRNA is deprotected by irradiation at 350 nm. to form the aminoacyl-tRNA **4**. Adapted from (*48-51*)

are recognized by the synthetases, and therefore Uaas similar to the natural amino acids. Also, it is not possible to introduce multiple analogs of the same amino acid type, at the same time (for example, two methionine analogs).

Joining Unnatural Amino Acids and Selection

Once the ribosome can be coaxed into synthesizing libraries of unnatural peptides, these peptides must be then selected for function. The amalgam of these two powerful methods was first demonstrated using nonsense suppression and mRNA display (*57*). In this proof of concept experiment, Roberts and coworkers charged a UAG-suppressor tRNA with biocytin (biotinyl-lysine) and selected for binding for streptavidin. After two rounds of selection, the library was primarily composed of sequences containing the UAG stop codon, demonstrating that the unnatural amino acid was incorporated in the peptide and was functional (Chapter 2).

Recently, a selection was performed using a 24-amino acid alphabet. Muranaka et al. used three tRNAs containing four-base codons and one UAG-suppressor tRNA to perform a selection involving four UAAs (*58*). Selection for binding to streptavidin resulted in one peptide containing p-benzoylphenylalanine. However, when a 10-mer peptide derived from the winning sequence was synthesized chemically, it lacked binding, suggesting that the whole peptide is necessary for activity.

Selections using mRNA display with wholly unnatural alphabets are imminent. Szostak and coworkers were able to rewrite twelve sense codons by introducing amino acid analogs that were charged by the AARSs (see above) (*47*). Peptides containing up to twelve unnatural amino acids were synthesized as mRNA-peptide fusions, and the identities of the displayed peptides confirmed by MALDI-TOF MS. Although some side

products (most likely resulting from incorporation of natural amino acids) are seen, these experiments demonstrate the feasibility of selecting a wholly unnatural peptide from a biological display library.

Cyclization and Drug Design

Macrocyclization of a linear peptide sequence often confers desirable biochemical and pharmacological properties. The conformational constraint imparted by cyclization can reduce the conformational entropy lost upon binding (*59*). This property has been exploited to increase the affinity/potency of natural linear peptide sequences for their biological targets (*60*, *61*). Conformational constraint, in conjunction with unnatural amino acid incorporation, has been shown to increase the stability of linear peptide sequences to proteolytic degradation (Reviewed in (*62*)). Macrocyclization has also been found to increase the cell permeability of peptides and peptidomimetics. Recent work by Lokey and co-workers suggests that the conformational restraint in hexapeptide macrocycles facilitates the formation of intramolecular hydrogen bonds leading to an increase in membrane permeability (*63*).

Many peptide-based drugs employ macrocyclization as a means to reduce conformational constraint and improve bioavailability (Figure 1.2B). Octreotide (Sandostatin), an unnatural analog of somatostatin, is considerably more selective for a subset of SST receptors than its parent compound (*64*). Integrelin (Eptifibatide) is a disulfide-cyclized peptide derived from barbourin, a large (73 mer) peptide found in rattlesnake venom (*65*). The conformational constraint imparted by the disulfide cross-link is essential for this compound's high affinity and specificity for GPIIb-IIIa integrin (*66*)**.** Daptomycin, a potent cyclic lipopeptide antibiotic from *Streptomyces roseosporus*,

has been approved for the treatment of both MRSA and VRSA infections (Reviewed in (*67*)).

Cyclization and Biological Display

Disulfide-mediated cyclization has been routinely employed in phage display selections, and at present remains the only practical method for constraining phage display libraries (*68-70*). While disulfide-linked macrocycles can be employed as pharmacaphores (e.g., Octreotide, Eptifibatide, Oxytocin), their utility in the selection of intracellular ligands is limited by the oxidative lability of the disulfide bond (*71*). Alternate methods of cyclization, akin to those found in natural products, could generate redox-insensitive macrocycles with expanded chemical and structural properties.

It has previously been observed that peptide-nucleic acid conjugates can be subjected to a wide range of chemical transformations without modification of the nucleic acid component (*72*, *73*). Li and Roberts exploited this property to chemically append a small molecule drug to a linear mRNA display library (*74*). While the yield for this reaction was modest, the experiment indicated that a judicious choice of reagents could enable the specific chemical modification of bulk mRNA display libraries. This work also suggested that the submicromolar concentrations of peptide library typically encountered in biological display are compatible with organic reactions normally carried out at millimolar concentrations.

Millward and co-workers extended this methodology by employing a homobifunctional cross-linker to cyclize large mRNA display libraries via the N-terminal amino group of methionine and the ε-amino group of a lysine side chain (*75*). In the absence of a functional assay, the extent of cyclization in model peptide fusions was

determined by incorporating a α-hydroxy acid between the cross-linking positions and using NH_4OH to cleave non-cyclized molecules. Formation of the desired product was confirmed by MALDI-TOF MS. Finally, a series of libraries were constructed and the number of residues between the cross-linking positions varied to determine the effect of cycle size on cyclization efficiency. In this way, it was determined that libraries with up to ten random positions could be efficiently cyclized using the optimized reaction conditions (Chapter 4). This methodology has since been employed in the selection of cyclic peptides with high affinity for G_α subunits (Chapter 5).

Seebeck and Szostak (*76*) have recently developed a second cyclization strategy for biological display based on the previously reported cross-linking of reactive cysteines by α, α'-dibromo-*m*-xylene (*77*). The peptide used in this study was synthesized using a reconstituted translation system and the resulting macrocyclic product was verified by MALDI-TOF MS. It is reasonable to expect that this methodology is adaptable to mRNA display and provides another alternative to the redox-labile disulfide bond.

Current strategies for the post-translational cyclization of biological display libraries are limited to chemistries that are compatible with the reactive groups of the 20 natural amino acids. However, the facile incorporation of multiple unnatural amino acids in reconstituted translation extracts will greatly expand the options for generating peptide macrocycles. For instance, the [3+2] Huisgen cycloaddition between azides and terminal alkynes ("click" chemistry) represents a highly specific, thermodynamically favorable chemistry for the macrocyclization of peptide libraries. Both azides (*23*) and terminal alkynes (*47*) have been incorporated into peptides within a biological display context.

Incorporation of both functionalities within a peptide sequence offers the possibility of macrocyclization via a 5-membered triazole ring.

Future Outlook

Recent experiments have demonstrated that the elements for unnatural display–*in vitro* selection and unnatural peptide synthesis by the ribosome–have been established. Once an unnatural selection is completed, we should be able to address the question as to how much of an improvement an unnatural alphabet confers upon a library and its resultant compounds. The choice of monomer alphabet will likely be important and it is possible that multiple selections using different alphabets will be needed to address this issue. Intellectually, unnatural selection can experimentally address why nature chose the natural 20 amino acids and if the natural alphabet is optimized for the procurement of functional protein-interacting molecules, or if it is indeed a "frozen accident" as suggested by Crick (*78*). In any case, the compounds resulting from these unnatural libraries are likely to be superior lead compounds for chemical genetics and therapeutic applications.

This thesis is divided into chapters where a different aspect of unnatural display is described. A brief summary of each chapter is described below:

Chapter 2

Chapter 2 is adapted from the first published account of unnatural amino acid incorporation in a biological display system. The incorporation of biocytin into mRNA-peptide fusions and the subsequent selection for streptavidin binding confirm that nonsense suppression is compatible with mRNA display technology.

Chapter 3

This chapter describes previously published work where sense suppression is employed to direct the synthesis of N-methyl phenylalanine polymers in the context of an mRNA-peptide fusion. These "encodamers" display considerable resistance to proteolytic degradation and provide evidence that multiple unnatural amino acids may be incorporated within a single mRNA-peptide fusion.

Chapter 4

This chapter includes previously published work describing a post-translational route for cyclizing mRNA-peptide fusions. The cyclization chemistry was confirmed by MALDI-TOF MS and represents one of the first published applications of this technique to biological display systems. The incorporation of an α-hydroxy acid by nonsense suppression enabled a quantitative evaluation of the cyclization chemistry in both model mRNA-peptide fusions and trillion-member mRNA display libraries.

Chapter 5

This chapter presents experiments which unify unnatural amino acid incorporation, post-translational cyclization, and mRNA display. Selection against Gαi1 with a trillion-member macrocyclic library resulted in an extremely high-affinity cyclic peptide (K_i = 1 nM) with marked protease resistance. These experiments suggest that mRNA display, coupled with unnatural amino acid incorporation and post-translational macrocyclization, have the potential to generate potent, bioactive lead compounds for chemical biology and drug development.

References

(1) Hopkins, A. L., and Groom, C. R. (2002) The druggable genome. *Nat Rev Drug Discov 1*, 727-30.

(2) Boldi, A. M. (2004) Libraries from natural product-like scaffolds. *Curr Opin Chem Biol 8*, 281-6.

(3) Goodnow, R. A., Jr., Guba, W., and Haap, W. (2003) Library design practices for success in lead generation with small molecule libraries. *Comb Chem High Throughput Screen 6*, 649-60.

(4) Shang, S., and Tan, D. S. (2005) Advancing chemistry and biology through diversity-oriented synthesis of natural product-like libraries. *Curr Opin Chem Biol 9*, 248-58.

(5) Yin, H., and Hamilton, A. D. (2005) Strategies for targeting protein-protein interactions with synthetic agents. *Angew Chem Int Ed Engl 44*, 4130-63.

(6) Takahashi, T. T., Austin, R. J., and Roberts, R. W. (2003) mRNA display: ligand discovery, interaction analysis and beyond. *Trends Biochem Sci 28*, 159-65.

(7) Lam, K. S., Lebl, M., and Krchnak, V. (1997) The "One-Bead-One-Compound" Combinatorial Library Method. *Chem Rev 97*, 411-448.

(8) Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., Grutter, M. G., and Pluckthun, A. (2004) High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol 22*, 575-82.

(9) Xu, L., Aha, P., Gu, K., Kuimelis, R. G., Kurz, M., Lam, T., Lim, A. C., Liu, H., Lohse, P. A., Sun, L., Weng, S., Wagner, R. W., and Lipovsek, D. (2002)

Directed evolution of high-affinity antibody mimics using mRNA display. *Chem Biol 9*, 933-42.

(10) Daugherty, P. S., Chen, G., Iverson, B. L., and Georgiou, G. (2000) Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. *Proc Natl Acad Sci U S A 97*, 2029-34.

(11) Hanes, J., Schaffitzel, C., Knappik, A., and Pluckthun, A. (2000) Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat Biotechnol 18*, 1287-92.

(12) Keefe, A. D., and Szostak, J. W. (2001) Functional proteins from a random-sequence library. *Nature 410*, 715-8.

(13) Silverman, R. B. (1992) The Organic Chemistry of Drug Design and Drug Action.

(14) Austin, R. J., Xia, T., Ren, J., Takahashi, T. T., and Roberts, R. W. (2003) Differential modes of recognition in N peptide-boxB complexes. *Biochemistry 42*, 14957-67.

(15) Carter, P. J., Winter, G., Wilkinson, A. J., and Fersht, A. R. (1984) The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (Bacillus stearothermophilus). *Cell 38*, 835-40.

(16) Schreiber, G., and Fersht, A. R. (1995) Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J Mol Biol 248*, 478-86.

(17) Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature 346*, 818-22.

(18) Scott, J. K., and Smith, G. P. (1990) Searching for peptide ligands with an epitope library. *Science 249*, 386-90.

(19) Halpin, D. R., and Harbury, P. B. (2004) DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol 2*, E174.

(20) Mattheakis, L. C., Bhatt, R. R., and Dower, W. J. (1994) An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proc Natl Acad Sci U S A 91*, 9022-6.

(21) Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A 94*, 12297-302.

(22) Levin, A. M., and Weiss, G. A. (2006) Optimizing the affinity and specificity of proteins with molecular display. *Mol Biosyst 2*, 49-57.

(23) Tian, F., Tsao, M. L., and Schultz, P. G. (2004) A phage display system with unnatural amino acids. *J Am Chem Soc 126*, 15962-3.

(24) Bemis, G. W., and Murcko, M. A. (1996) The properties of known drugs. 1. Molecular frameworks. *J Med Chem 39*, 2887-93.

(25) Bemis, G. W., and Murcko, M. A. (1999) Properties of known drugs. 2. Side chains. *J Med Chem 42*, 5095-9.

(26) Hajduk, P. J., Bures, M., Praestgaard, J., and Fesik, S. W. (2000) Privileged molecules for protein binding identified from NMR-based screening. *J Med Chem 43*, 3443-7.

(27) Lipinski, C. L., F, Dominy, BW; Feeney, PJ. (1997) Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Advanced Drug Delivery Reviews 23*, 3-25.

(28) Owens, J. (2003) Chris Lipinski discusses life and chemistry after the Rule of Five. *Drug Discov Today 8*, 12-6.

(29) Finking, R., and Marahiel, M. A. (2004) Biosynthesis of nonribosomal peptides1. *Annu Rev Microbiol 58*, 453-88.

(30) Walsh, C. T., Chen, H., Keating, T. A., Hubbard, B. K., Losey, H. C., Luo, L., Marshall, C. G., Miller, D. A., and Patel, H. M. (2001) Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol 5*, 525-34.

(31) Kohli, R. M., and Walsh, C. T. (2003) Enzymology of acyl chain macrocyclization in natural product biosynthesis. *Chem Commun (Camb)*, 297-307.

(32) Rizo, J., and Gierasch, L. M. (1992) Constrained peptides: models of bioactive peptides and protein substructures. *Annu Rev Biochem 61*, 387-418.

(33) Adessi, C., Frossard, M. J., Boissard, C., Fraga, S., Bieler, S., Ruckle, T., Vilbois, F., Robinson, S. M., Mutter, M., Banks, W. A., and Soto, C. (2003) Pharmacological profiles of peptide drug candidates for the treatment of Alzheimer's disease. *J Biol Chem 278*, 13905-11.

(34) Hamamoto, K., Kida, Y., Zhang, Y., Shimizu, T., and Kuwano, K. (2002) Antimicrobial activity and stability to proteolysis of small linear cationic peptides with D-amino acid substitutions. *Microbiol Immunol 46*, 741-9.

(35) MacKintosh, R. W., Dalby, K. N., Campbell, D. G., Cohen, P. T., Cohen, P., and MacKintosh, C. (1995) The cyanobacterial toxin microcystin binds covalently to cysteine-273 on protein phosphatase 1. *FEBS Lett 371*, 236-40.

(36) Keller, U., and Schauwecker, F. (2003) Combinatorial biosynthesis of non-ribosomal peptides. *Comb Chem High Throughput Screen 6*, 527-40.

(37) Clugston, S. L., Sieber, S. A., Marahiel, M. A., and Walsh, C. T. (2003) Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. *Biochemistry 42*, 12095-104.

(38) Jackson, R. J., Napthine, S., and Brierley, I. (2001) Development of a tRNA-dependent in vitro translation system. *RNA 7*, 765-73.

(39) Link, A. J., and Tirrell, D. A. (2005) Reassignment of sense codons in vivo. *Methods 36*, 291-8.

(40) Frankel, A., and Roberts, R. W. (2003) In vitro selection for sense codon suppression. *RNA 9*, 780-6.

(41) Frankel, A., Millward, S. W., and Roberts, R. W. (2003) Encodamers: unnatural peptide oligomers encoded in RNA. *Chem Biol 10*, 1043-50.

(42) Forster, A. C., Weissbach, H., and Blacklow, S. C. (2001) A simplified reconstitution of mRNA-directed peptide synthesis: activity of the epsilon enhancer and an unnatural amino acid. *Anal Biochem 297*, 60-70.

(43) Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol 19*, 751-5.

(44) Pestova, T. V., and Hellen, C. U. (2005) Reconstitution of eukaryotic translation elongation in vitro following initiation by internal ribosomal entry. *Methods 36*, 261-9.

(45) Forster, A. C., Tan, Z., Nalam, M. N., Lin, H., Qu, H., Cornish, V. W., and Blacklow, S. C. (2003) Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc Natl Acad Sci U S A 100*, 6353-7.

(46) Tan, Z., Forster, A. C., Blacklow, S. C., and Cornish, V. W. (2004) Amino acid backbone specificity of the Escherichia coli translation machinery. *J Am Chem Soc 126*, 12752-3.

(47) Josephson, K., Hartman, M. C., and Szostak, J. W. (2005) Ribosomal synthesis of unnatural peptides. *J Am Chem Soc 127*, 11727-35.

(48) Bain, J. D., Diala, E. S., Glabe, C. G., Dix, T. A., and Chamberlin, A. R. (1989) Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *J Am Chem Soc 111*, 8013-8014.

(49) Gallivan, J. P., Lester, H. A., and Dougherty, D. A. (1997) Site-specific incorporation of biotinylated amino acids to identify surface-exposed residues in integral membrane proteins. *Chem Biol 4*, 739-49.

(50) Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins. *Science 244*, 182-8.

(51) Robertson, S. A., Ellman, J. A., and Schultz, P. G. (1991) A general and efficient route for chemical aminoacylation of transfer RNAs. *J Am Chem Soc 113*, 2722-2729.

(52) Merryman, C., and Green, R. (2004) Transformation of aminoacyl tRNAs for the in vitro selection of "drug-like" molecules. *Chem Biol 11*, 575-82.

(53) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli. *Science 292*, 498-500.

(54) Datta, D., Wang, P., Carrico, I. S., Mayo, S. L., and Tirrell, D. A. (2002) A Designed Phenylalanyl-tRNA Synthetase Variant Allows Efficient in Vivo Incorporation of Aryl Ketone Functionality into Proteins. *J Am Chem Soc 124*, 5652-5653.

(55) Tang, Y., and Tirrell, D. A. (2002) Attenuation of the Editing Activity of the Escherichia coli Leucyl-tRNA Synthetase Allows Incorporation of Novel Amino Acids into Proteins in Vivo. *Biochemistry 41*, 10635-10645.

(56) Hartman, M. C., Josephson, K., and Szostak, J. W. (2006) Enzymatic aminoacylation of tRNA with unnatural amino acids. *Proc Natl Acad Sci U S A 103*, 4356-61.

(57) Li, S., Millward, S., and Roberts, R. (2002) In vitro selection of mRNA display libraries containing an unnatural amino acid. *J Am Chem Soc 124*, 9972-3.

(58) Muranaka, N., Hohsaka, T., and Sisido, M. (2006) Four-base codon mediated mRNA display to construct peptide libraries that contain multiple nonnatural amino acids. *Nucleic Acids Res 34*, e7.

(59) Khan, A. R., Parrish, J. C., Fraser, M. E., Smith, W. W., Bartlett, P. A., and James, M. N. (1998) Lowering the entropic barrier for binding conformationally flexible inhibitors to enzymes. *Biochemistry 37*, 16839-45.

(60) Eckert, D. M., Malashkevich, V. N., Hong, L. H., Carr, P. A., and Kim, P. S. (1999) Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. *Cell 99*, 103-15.

(61) Kohli, R. M., Takagi, J., and Walsh, C. T. (2002) The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides. *Proc Natl Acad Sci U S A 99*, 1247-52.

(62) Veber, D. F., and Freidinger, R. M. (1985) The design of metabolically-stable peptide analogs. *Trends in Neurosciences 8*, 392.

(63) Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P., and Lokey, R. S. (2006) Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J Am Chem Soc 128*, 2510-1.

(64) Pawlikowski, M., and Melen-Mucha, G. (2004) Somatostatin analogs–from new molecules to new applications. *Curr Opin Pharmacol 4*, 608-13.

(65) Scarborough, R. M., Rose, J. W., Hsu, M. A., Phillips, D. R., Fried, V. A., Campbell, A. M., Nannizzi, L., and Charo, I. F. (1991) Barbourin. A GPIIb-IIIa-specific integrin antagonist from the venom of Sistrurus m. barbouri. *J Biol Chem 266*, 9359-62.

(66) Scarborough, R. M., Naughton, M. A., Teng, W., Rose, J. W., Phillips, D. R., Nannizzi, L., Arfsten, A., Campbell, A. M., and Charo, I. F. (1993) Design of potent and specific integrin antagonists. Peptide antagonists with high specificity for glycoprotein IIb-IIIa. *J Biol Chem 268*, 1066-73.

(67) Baltz, R. H., Miao, V., and Wrigley, S. K. (2005) Natural products to drugs: daptomycin and related lipopeptide antibiotics. *Nat Prod Rep 22*, 717-41.

(68) Dennis, M. S., Eigenbrot, C., Skelton, N. J., Ultsch, M. H., Santell, L., Dwyer, M. A., O'Connell, M. P., and Lazarus, R. A. (2000) Peptide exosite inhibitors of factor VIIa as anticoagulants. *Nature 404*, 465-70.

(69) Dennis, M. S., Herzka, A., and Lazarus, R. A. (1995) Potent and selective Kunitz domain inhibitors of plasma kallikrein designed by phage display. *J Biol Chem 270*, 25411-7.

(70) Giebel, L. B., Cass, R. T., Milligan, D. L., Young, D. C., Arze, R., and Johnson, C. R. (1995) Screening of cyclic peptide phage libraries identifies ligands that bind streptavidin with high affinities. *Biochemistry 34*, 15430-5.

(71) Gilbert, H. (1995) *Methods in Enzymology 251*, 8-28.

(72) Halpin, D. R., Lee, J. A., Wrenn, S. J., and Harbury, P. B. (2004) DNA display III. Solid-phase organic synthesis on unprotected DNA. *PLoS Biol 2*, E175.

(73) Sakurai, K., Snyder, T. M., and Liu, D. R. (2005) DNA-templated functional group transformations enable sequence-programmed synthesis using small-molecule reagents. *J Am Chem Soc 127*, 1660-1.

(74) Li, S., and Roberts, R. W. (2003) A novel strategy for in vitro selection of peptide-drug conjugates. *Chem Biol 10*, 233-9.

(75) Millward, S. W., Takahashi, T. T., and Roberts, R. W. (2005) A general route for post-translational cyclization of mRNA display libraries. *J Am Chem Soc 127*, 14142-3.

(76) Seebeck, F. P., and Szostak, J. W. (2006) Ribosomal synthesis of dehydroalanine-containing peptides. *J Am Chem Soc 128*, 7150-1.

(77) Timmerman, P., Beld, J., Puijk, W. C., and Meloen, R. H. (2005) Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *Chembiochem 6*, 821-4.

(78) Crick, F. H. (1968) The origin of the genetic code. *J Mol Biol 38*, 367-79.

(79) Fliri, H., Baumann, G., Enz, A., Kallen, J., Luyten, M., Mikol, V., Movva, R., Quesniaux, V., Schreier, M., Walkinshaw, M., et al. (1993) Cyclosporins. Structure-activity relationships. *Ann N Y Acad Sci 696*, 47-53.

(80) Scherkenbeck, J., Jeschke, P., and Harder, A. (2002) PF1022A and related cyclodepsipeptides–a novel class of anthelmintics. *Curr Top Med Chem 2*, 759-77.