High Throughput Screening and mRNA-Display Selection

for the Identification of Biologically Functional Molecules

By
Satoko Yamazaki

Dissertation

Submitted to the Faculty of Mathematics and Natural Sciences of
the Rheinische Friedrich-Wilhelms University of Bonn
in partial fulfillment of the requirements for
the degree of
DOCTOR OF NATURAL SCIENCE (Dr. rer. nat.)

2006
Bonn, Germany

Prepared with the permission of the Faculty of Mathematics and Natural Science of the Rheinische Friedrich-Wilhelms University of Bonn

1. Referee: Prof. Dr. Michael Famulok
2. Referee: Priv.-Doz. Dr. Thomas Kolter
3. Referee: Prof. Dr. Andreas Gansäuer
4. Referee: Prof. Dr. Michael Hoch

Day of Graduation: 27 March, 2006

Index

1. Summary

Biologically functional molecules with desired properties have proven to be invaluable tools for investigating biological systems. Moreover, the identification of those functional molecules is a crucial task in modern drug discovery. These goals can be achieved by two major approaches: "screening" and "selection".

Chapter 2 introduces one attractive method for "screening" to identify small molecule inhibitors of a given protein target from a small organic chemical library.

The reverse transcriptase (RT) plays a crucial role in the early steps of the life cycle of HIV. Therefore HIV-1 RT serves as one of the major drug targets in anti-HIV therapy. Although current therapy for HIV-infected patients involving a combination of three or more of the following drugs: RT inhibitors, protease inhibitors, and/or viral entry inhibitors, has been highly successful, its long-term efficacy is severely limited by the emergence of drug-resistant variants of HIV. One way to address this problem is to identify new type of drugs with different inhibitory mechanisms to combat HIV. Thus, HIV-1 RT served a valuable target in the herein depicted screening approach.

Protein-dependent alosteric ribozymes (or reporter ribozymes) are powerful tools which are able to sense molecular interactions, e.g. protein-protein, protein-nucleic acid, or protein-metabolite interactions in real-time. The expedience of this assay format makes it applicable to and suitable for high throughput screening in order to search for new molecules which disrupt such interactions. Reporter ribozyme FK-1 is a rationally engineered hammerhead ribozyme in a fusion construct with anti-RT aptamer, which can detect the presence of HIV-1 RT selectively in a domain specific manner. Thus, FK-1 enables to identify compounds via FRET-based fluorescence readout, which compete for the binding to HIV-1 RT with aptamer sequences inserted in the ribozyme construct.

Through screening of 2500 small molecules using the reporter ribozyme (FK-1)-based assay, three potential inhibitors (**28F6, 3E4 and 2E10**) of HIV-1 RT were identified with *in vitro* inhibitory concentration that gave half-maximal activity (IC_{50}) of 2-5 μM on DNA-dependent DNA polymerase activity. The three compounds were re-synthesized and further investigated to evaluate their selectivity and to elucidate the inhibitory mechanism. The selectivity was determined by testing inhibitory effect of the compound on other RTs (AMV and MMLV) and DNA polymerases from both prokaryotic (Klenow Fragment) and eukaryotic (human DNA polymerase β) sources.

Among the three compounds, **3E4** showed not only the lowest IC_{50} value for HIV-1 RT (2.1 µM) but the highest discrimination from all the other RT and DNA polymerases (10~50 fold) in terms of DNA-dependent DNA polymerase activity. The selectivity of **3E4** with the cognate RT from HIV-2 was a factor of approximately 4.5. In addition, the interaction between **3E4** and HIV-1 RT was observed by a surface plasmon resonance-based biosensor.

Moreover, a cell-based assay for phenotypic analysis of sensitivity of HIV-1 to the three inhibitors using a self-inactivating virus vector system demonstrated that **3E4** inhibited marker gene expression in concentration dependent manner with a IC_{50} value of 5.3 µM. The effect of **3E4** was further confirmed with the similar reduction of HIV-1 replication and infectivity *in vivo* using a HIV wild type strain.

In summary, the data strongly indicates that the small molecule **3E4** identified by reporter ribozyme-based screening is a highly specific and cellular active HIV-1 RT inhibitor, which has significant potentiality as novel type of anti-HIV drug. Furthermore, the approach using a reporter ribozyme assay could be broadly applicable as screening format.

In chapter 3, the challenge of a "selection" approach has been performed with the aim to evolve functional proteins from a random sequence protein library.

Prior to the selection, the protein library consisting of a random 88 amino acid sequence was constructed by short cassettes ligations to avoid stop codons and frameshift mutations. The design of secondary structural patterns like α-helix and β-strand was introduced into the random regions.

A display technology called mRNA-display was exploited to prepare protein library, where each protein sequence was covalently linked through its carboxy terminus to the 3' end of its encoding mRNA. Starting with a library that contained 1×10^{11} - 5×10^{12} different proteins, three *in vitro* selections were attempted to identify protein binders to small molecules such as cibacron blue 3GA and coenzymes (NADP and CoA). The selection against cibacron blue 3GA yielded enriched fractions retaining on the column matrix after the ninth cycle. These sequences could not be eluted by NAD, the structural analogue of cibacron blue 3GA, even at higher concentration. The small fractions eluted with NAD from this column have been cloned and sequenced, resulting in unexpectedly high frequency of frameshifts and internal stop codons. The alignment of full length intact sequences exhibited no conserved motif but the most appearance of αβαβ type of cassette. The remaining column fractions were able to be eluted with free cibacron blue 3GA, implying those sequences might be specific cibacron blue binders. Other

selections against both NADP and CoA did not enrich protein binders.

The observed problems through the experiment concerning poor solubility and fold instability of mRNA-displayed proteins should be improved for further investigations.

2. Reporter Ribozyme toward Screening of Small Molecule Inhibitors of Human Immunodeficiency Virus Type-1 Reverse Transcriptase with anti-Viral Activity

2.1. Introduction

2.1.1. High Throughput Screening towards Drug Discovery

Efficient identification of biologically active molecules, ranging from small molecules to macromolecules such as nucleic acids and proteins, can greatly enhance the efficiency to explore biological systems in genomics and proteomics as well as the development of desired therapeutic agents in drug discovery. To pursue this basic goal, a number of innovative approaches have been evolved. Additionally, these technologies are fuelled by an expanded number of targets derived by molecular biology and an increased number of compounds generated by combinatorial chemistry. Since Szostak[1] and Gold[2] have developed a method to chemically synthesize a large pool of random nucleotide sequences with a complexity of approximately 10^{15} to isolate molecules with highly specific ligand-binding properties, the concepts of combinatorial chemistry have been extended to biological macromolecules including nucleic acids and proteins (see chapter 3). All the current advances in chemistry and biology principally increase the capability for researchers to obtain any types of molecules which fulfill desired properties through proper designing of the selection strategy or assay manipulations.

One dominant approach to identify interesting molecules within large combinatorial libraries is "screening", pioneered by the idea of systematically searching for drugs advocated by Paul Ehrlich.[3] The underlying concept of this technique is described in Figure 1A. In general, this strategy requires individual assessment of all members from the library in a biological assay and functional detection of hit compounds against a large background. Pharmaceutical industries exhaustively exploit the great advantage of high throughput screening (HTS) to identify reasonable starting points of small molecules (lead compounds) for further elaboration towards effective drugs. A series of improvements in the areas of automation and miniaturization have largely accomplished to screen compound libraries with the number of 10^4-10^5 entities per day.[4] Not only to screen small molecule libraries, HTS technologies can streamline the rapid identification of interesting protein variants.[5]

4

In contrast, another fascinating approach called "selection" has been developed by mimicking the culling process associated with true Darwinian evolution (see chapter 3 for a detailed discussion). Initially this approach has been applied for isolation of nucleic acids with high affinity and specificity, alternatively referred as SELEX (systematic evolution of ligands by exponential enrichment).[2, 6-9] The strategy is nowadays routinely utilized for combinatorial libraries of nucleic acids and proteins or peptides. As described in Figure 1B, selection strategies exploit conditions favoring the exclusive survival of desired variants, eliminating negative samples, thereby lead to isolate only the characteristic product. Darwinian evolution is able to be achieved via consecutive rounds of selection and amplification. The major advantage of selection is that more library members can be analyzed simultaneously and homogeneously. Usually protein libraries with the diversity up to 10^{10-13} and nucleic acids libraries up to 10^{15} are accessible within a single experiment.[10]
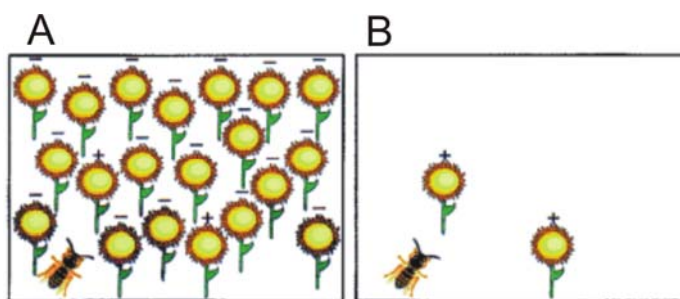
Figure 1. Different search modes for finding nectar-producing flowers (marked with +). Random screening (**A**): The bumblebee must check each flower individually to differentiate producers from the large number of non-producers (marked with -). Selection (**B**): The bee only encounters the desired flowers because nectar production was a prerequisite for plant growth in this case. The figure is reproduced from Taylor et al..[10]

As mentioned above, any drug discovery procedure requires that in an individual assay the difference between a "positive" and "negative" is recorded in an output signal in a reliable fashion (Figure 1). To identify appropriate ligands for a protein target, the desired protein-small molecule interaction has to be detected. A number of biochemical assays feasible for HTS to measure molecular interactions have been established and are listed in Table 1.

The choice of an assay for screening needs to be guided by the goal of the screening endeavor and the practical constraints for the target of interest. These constrains include amount of proteins available, substrate requirements for the enzyme and capacity for

follow-up chemistry and biology after the screen is completed.

Drug screening assays can be divided into two types: those that use labeled proteins and/or ligands and those that are label-free (a label is a reporter moiety e.g. a fluorescent or radioactive group).[11] Although labels make protein–ligand interactions easier to observe, they can also be difficult to introduce, which increases the time and expense associated with the preparation and/or measurement. Additionally, introductions of labels can perturb native properties of the substrates which hamper the detection of meaningful interactions. Label-free detection methods are therefore preferred because they obviate the disadvantages described above. However, such measurements can be more difficult to perform: without a label, a larger amount of both protein and compound have to be produced, and often the analysis in label-free measurements is slower.

Although a variety of methods to detect a molecular-binding event are available for HTS formats (Table 1), a generic method which can meet all desired conditions is not yet available. Each has advantages and disadvantages. The following points should be considered for preferable screening assays: 1) homogeneous formats, 2) no labeling on small molecules and proteins, 3) *in situ* production of the proteins to skip purifications in large parallel quantities, and 4) the possibility for automation without modifications to each system.[11]

Method	Label	Throughput	Phase	Protein amount	Description
Fluorescence polarization	Yes	High	Solution	High	Measures the change in tumbling rate for a compound when it is bound to a protein using loss-of-polarization of incident light.
Fluorescence perturbation	No	Low	Solution	High	Measures the change in fluorescence (usually of tryptophan residues) on a protein caused by proximity of a bound compound.
Fluorescence correlation Spectroscopy	Yes	High	Solution	Medium	Measures the correlated changes in fluorescence properties that occur when a labeled compound is bound to a labeled protein.
Radioligand binding assay	Yes	High	Solid	High	Measures retention of a compound on a surface holding a protein. Detection is by means of a radioactive isotope that is incorporated into the compound.
Nuclear magnetic resonance	No	Low	Solution	High	Measures changes in the nuclear spin of protons on a protein that occur when a compound is bound nearby.
Mass spectrometry	No	Medium	Solution	Medium	Measures the change in molecular weight observable in a mass spectrometer when a compound is bound to a protein.
Surface plasmon resonance	No	Low/Medium	Solid	High	Measures the change in refractivity of a metal surface when a compound binds to a protein that is immobilized on that surface.
Isothermal titration calorimetry	No	Low	Solution	High	Measures the amount of heat produced when a compound binds to a protein.
Differential scanning calorimetry	No	Low	Solution	High	An alternative method of measuring the heat released when a compound binds to a protein.
Small-molecule microarray	Yes	High	Solid	Medium	Compounds are immobilized on a surface. A protein is incubated with the surface and localizes to where the protein-interacting compounds are bound.
Protein microarray	Yes	High	Solid	Low/none	Proteins are immobilized on a surface and a labeled compound is incubated with the surface. The compound localizes to where the compound-binding proteins are bound.
Three-hybrid system	Yes	High	In cells	None	A target is covalently linked to an adaptor moiety that interacts with a DNA-binding protein. If the test compound can bind to a test protein that is fused to a DNA activation domain, transcription of a reporter gene is initiated.

Table 1. Methods for measuring the affinity of small molecules to proteins. The table is reproduced from Stockwell et al..[11] All the references for every method are listed there.

2.1.2. Allosteric Ribozyme Sensing Molecular Interactions

RNA is a highly sophisticated biomolecule that can serve for sequence specific recognition, fold into three dimensionally structures, catalyze chemical reactions, and control gene expression.[12-15] This multifunctionality, additionally driven by the simple production via both enzymatic and chemical synthesis, has been extensively applied to design and engineer a diverse range of novel biological RNA tools.[16, 17]

Ribozymes are naturally occurring RNAs that catalyze phosphordiester cleaving/forming or peptide-bond-forming reactions.[14, 18, 19] So far, nine classes of natural ribozymes have been discovered. Among them, the hammerhead ribozyme has been well characterized for its functional properties[20-23] as well as for its three-dimensional structure.[24-26] The secondary structure of a natural hammerhead

ribozyme is shown in Figure 2.
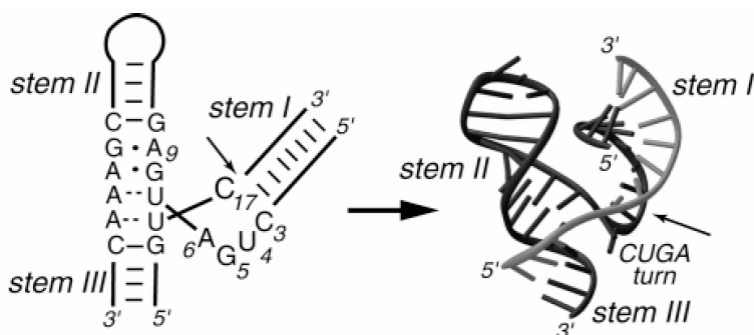
Figure 2. Secondary and tertiary structure of a hammerhead ribozyme. Universally conserved bases are shown as letters. *Dashed lines and dots* represent non–Watson-Crick interactions between bases; *solid lines* represent Watson-Crick base pairs. The site of cleavage is marked with an *arrow*. The figure is reproduced from Doherty et al..[18]

Such detailed structural information provides a solid conceptual foundation for manipulating the hammerhead ribozyme and an invaluable resource for engineering ribozymes that function by various mechanisms of allosteric regulation. Allosteric ribozymes usually carry an effector-binding site (binding domain) that is separate from its active site (reporter domain) (Figure 3). Ligand binding to this allosteric site induces a conformational change in the adjoining enzyme domain that enhances or inhibits catalytic reaction. Because of its multifaceted functions, allosteric ribozymes can adopt versatile ligands as effectors to modulate catalytic event. Allosteric ribozymes dependent on metal ion,[27] nucleic acids,[28] small molecules,[29-31] and proteins[32-35] have been generated. Concerning the effector-binding site, aptamer sequences, ssRNA owing ligand specificity and affinity, can be generally engineered to incorporate into the ribozyme context. Aptamers commonly demonstrate adaptive binding in which a conformational change in the aptamer accompanies ligand interaction,[36] transmitting dynamic conformational change to catalytically active site. Currently two strategies are applied to develop such allosteric ribozymes: rational design or combinatorial techniques (selection).

Intriguingly, after considerable endeavors for engineering allosteric ribozymes have been spent, the concept of allosteric ribozymes was recently discovered in natural systems, so-called "riboswitches".[37] Bacteria exploit non-coding regions of mRNA to control gene expression. Riboswitches can capture the target metabolite and undergo allosteric change in the structure. In the case of GlcN6P riboswitch, the *glm*S ribozyme can cleave off a portion of the 5' UTR of the *glm*S gene, which encodes for

glutamine-fructose-6-phosphate amidotransferase, upon binding of glucosamine-6-phosphate (GlcN6P), the product of the amidotransferase.[38] The feedback mechanism leads to reduced expression of the GlmS protein, serving as a novel metabolite-responsive ribozyme.

Recently, new aspects of allosteric ribozymes are emerging as novel biologically functional tools for various applications. Two years before the discovery of the GlcN6P riboswitch Thompson et al. have created an artificial gene regulatory switch in which a Group I ribozyme dependent upon theophyline was conjugated to thymidylate synthase gene.[39] Exogeneously added theophyline was able to control *in vivo* splicing in a target specific manner. The most straightforward application of allosteric ribozymes is their engineering to biosensors due to the dynamic conversion of binding event into catalytic reaction, which is easily detectable. In an approach to prepare a prototype biosensor array, seven different allosteric ribozymes, each engineered from hammerhead ribozymes to be triggered by the corresponding ligand (such as metal ions, cofactors), were immobilized onto gold surface, that report the presence and concentrations of targets even in complex chemical or biological mixtures.[40] Another example using an allostericly controlled Class I ligase for diagnostic purpose have been reported.[41] The authors constructed a "half-ribozyme" for ultra-sensitive detection of virus RNA from hepatitis C virus (HCV) in zeptomole concentration. This system has been investigated for its compatibility with an immunodiagnostic platform with equivalent sensitivity.[42] Allosteric ribozymes can even monitor post-translational modifications of proteins in cell lysate, through engineering of the ribozymes to be activated specifically upon binding to either the unphosphorylated form of the protein or the phosphorylated form.[43] More lately, highly sophisticated design of allosteric ribozymes lead to the construction of computational systems, including AND, OR, YES and NOT Boolen logic functions.[44]

As discussed above, the concept of allosteric ribozymes is highly compatible to high throughput screening (HTS) assays. Recent studies report the elaborate design of assay setups that make it possible to identify protein inhibitors via allosteric ribozyme-based signaling. Srinivasan used a ADP responsive allosteric ribozyme ("RiboReporter") to monitor the enzymatic reaction of a protein kinase.[45] Thereby, new protein kinase inhibitors or compounds that modulate the activity of any enzyme which is involved in ADP metabolism can in principle be identified. The authors demonstrated that RiboReporter was able to rediscover a known protein kinase inhibitor through screening. Another approach of HTS of protein inhibitors is to convert the inherent properties of an aptamer, such as specific binding, domain recognition and functional inactivation, into

small molecules using competitive assay formats.[46]

Figure 3 describes the concept of competitive screening assays via allosteric ribozymes (reporter ribozymes). Hartig et al. applied a human immunodeficiency virus type-1 (HIV-1) Rev dependent hammerhead reporter ribozyme to screen an 96-membered antibiotics library for the molecules that can disrupt the interaction between Rev and its cognate aptamer.[46] They also prepared a conversely regulated reporter ribozyme through incorporating a Rev binding element (RBE) into an effector binding site. Both reporter ribozymes were able to identify the same inhibitor Coumemycine A1, which efficiently suppressed HIV-1 replication *in vivo* as well.

Besides the topics discussed above, reporter ribozyme-based assays provide several additional advantages compared to conventional screening approaches: real-time signal detection, no requirements of labeling of proteins and their possible interaction partners, homogeneous performance, and target-independent assay development.[47-49] As discussed in §2.1.1, screening often requires to setup independent technology platforms for every distinct purpose. Based on the diverse acceptability for effector molecules and the facility of engineering both rationally and combinatorially, allosteric ribozymes hold the potential to establish general HTS assay formats with fewer obstacles and force to obtain highly specific inhibitors. Lately, the rational design of hairpin allosteric ribozymes to generate sensors for detecting microRNA[50] and metabolite[51] has been reported. These concepts would also increase the flexibility of designing for the ribozyme-based assay formats.
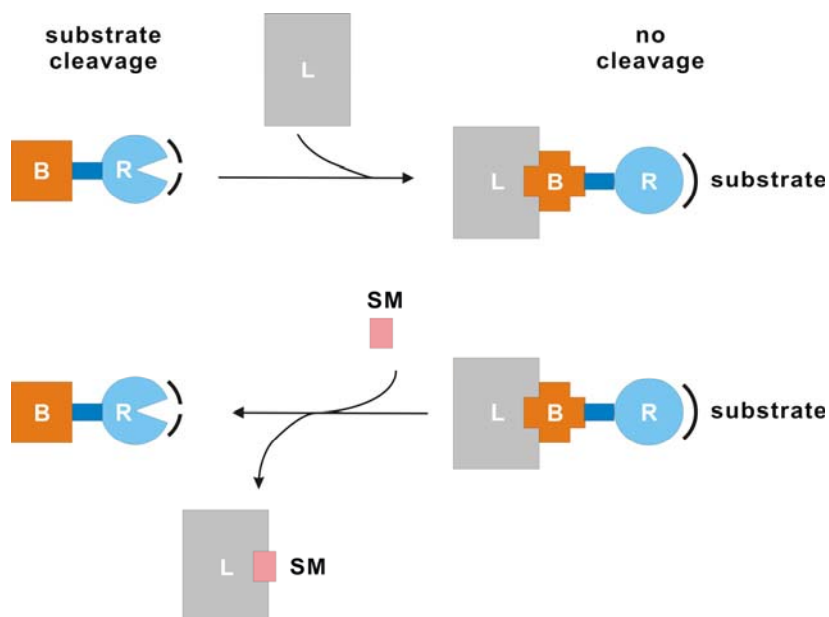
Figure 3 Schematic of an allosteric ribozyme and reporter ribozyme. The allostric ribozyme is illustrated by the combination of effector binding domain (**B**; orange) and reporter domain (**R**; blue). In the upper equation, the ligand (**L**; gray) addition induces a conformational change of the reporter domain, resulting in the inhibition of substrate (black line) cleavage. The lower equation represents a reporter ribozyme used in a competitive assay. The small molecule (**SM**; pink), which interferes with the interaction of ligand-ribozyme, restores the active conformation of ribozyme. Then, the substrate cleavage can be observed.

2.1.3. HIV and AIDS

Since the isolation of human immunodeficiency virus type-1 (HIV-1),[52, 53] efforts and studies of biology, biochemistry, and structural biology have been focused to control the AIDS (acquired immunodeficiency syndrome) epidemic. Modern drug discovery and development have markedly reduced the morbidity and mortality due to infection with HIV-1 in the past few years. HIV infection is characterized by a prolonged asymptomatic period of years to decades, which is followed by the fatal illness of AIDS including wasting, neurological impairment, and opportunistic infections and malignancies.

HIV is a member of the lentivirus genus and classified into retroviruses that possess complex genomes and exhibit cone-shaped capsid core particles. HIV's genome is encoded by RNA, which is reverse-transcribed to viral DNA by the viral reverse

transcriptase (RT) upon entering a new host cell. The virus consists of 15 proteins and 9.7 kb of RNA.[54, 55] The general features of the HIV replication cycle is shown in Figure 4. The early phase begins with the recognition of the target cell CD4 receptor by the mature virion and involves all processes leading to the inclusion and integration of the proviral DNA into the chromosome of the host cell. The late phase includes all the events from transcription of the integrated proviral DNA to virus budding and maturation.

Although all steps in the HIV life cycle are potential targets for anti-retroviral agents, current anti-retroviral therapy mainly uses three classes of drugs targeting RT, protease (PR), and viral fusion.[56] The most promising approach is termed "highly active antiretroviral therapy (HAART)" and contains a combination of RT and PR inhibitors. The administration of this kind of drug cocktail can reduce viral loads to undetectable levels for longer periods like two years. However, HAART cannot completely eradicate HIV from the body. This can result in long-term toxicity and eventually leads to the emergence of new drug-resistant strains.[57, 58] Thus, it is of outmost importance to define better drugs by new approaches.[59, 60] For example, one promising virally encoded target is the integrase.[61] Another attractive approach is to block the entry of HIV-1 through chemokine receptors.[62, 63] Other HIV-1 regulatory and accessory proteins are also gaining new interest, such as Vif.[64-66] Even cellular factors involved in HIV replication can serve as new targets to inhibit the viral replication, as demonstrated with an inhibitor against ataxia-talengiectasia-mutated (ATM) kinase.[67] Furthermore, the recent advantages in gene therapy provide RNA decoys, aptamers, and RNA interference technologies towards potential anti-HIV-1 agents.[68, 69]
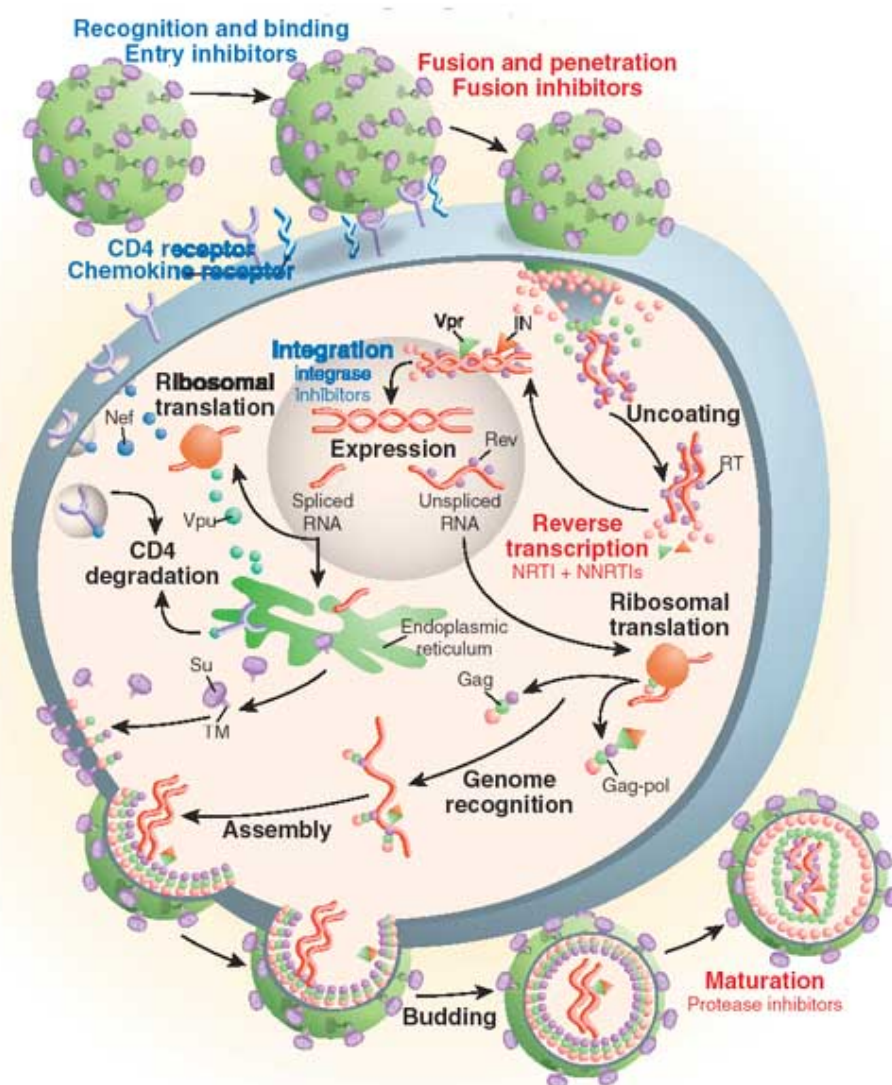
Figure 4. HIV-1 life cycle. The figure is reproduced from Pomerantz et al..[56]

2.1.4. RT as a Major Drug Target

As discussed above, RT is one of the major drug targets for anti-HIV-1 therapy. This enzyme catalyzes the synthesis of proviral DNA using the viral RNA as a template in the early phase. HIV-1 RT[55, 70] catalyzes three enzymatic reactions: RNA-dependent DNA polymerase activity (RDDP), DNA-dependent DNA polymerase activity (DDDP), and degradation of the RNA part of a DNA/RNA hybrid (RNase H activity). These

13

catalytic activities are confined only to p66/p51 heterodimeric forms (66kDa and 51 kDa respectively). Crystal structures[71, 72] of HIV-1 RT indicate that both subunits share four subdomains in common, which are termed "finger", "palm", thumb", and "connection", while an extra C-terminal domain required for RNase H activity exists only in p66 subunit (Figure 5). Highly conserved regions were found in the finger and palm subdomains of p66, providing the positions for primer/template binding as a helix clamp.[71] The most important feature of the RT is the lack of a proofreading function, resulting in 100-fold lower fidelity than the cellular DNA polymerases. The severe menace of drug resistance of HIV-1 is caused by this intrinsic property.
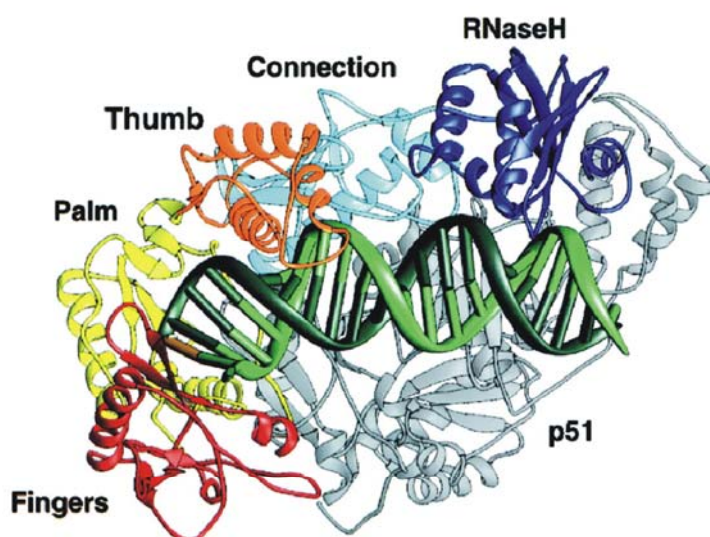
Figure 5. The crystal structure of the HIV-1 RT catalytic complex. The domains of p66 are in color: fingers (red), palm (yellow), thumb (orange), connection (cyan), and RNase H (blue); p51 is in gray. The DNA template strand (light freen) contains 25 nucleotides, and the primer strand (dark green) contains 21 nucleotides. The dNTP is in gold. The figure is reproduced form Huang et al..[72]

As already indicated by the first approved anti-HIV-1 drug AZT (*zidovudine*), RT serves as an excellent target for anti-viral strategies because RT is virus-specific, essential for virus replication, and acting in the early phase. The present RT inhibitors[57, 70] include nucleoside analogue RT inhibitors (NRTIs) and non-nucleoside RT inhibitors (NNRTIs). NRTIs act, after intracellular phosphorylation by cellular kinases, as competitive inhibitors with respect to the dNTP substrates. On the other hands, NNRTIs bind allosterically to a hydrophobic pocket located 10-15 Å from the active site.[73] The examples of NRTIs and NNRTIs are listed in Figure 6. Although a number of RT inhibitors with higher anti-viral effects have been identified, the currently used anti-viral

treatment results in the emergence of drug resistant HIV-1 strains. Therefore, new classes of inhibitors acting with different mechanisms are still needed to be developed to combat HIV.
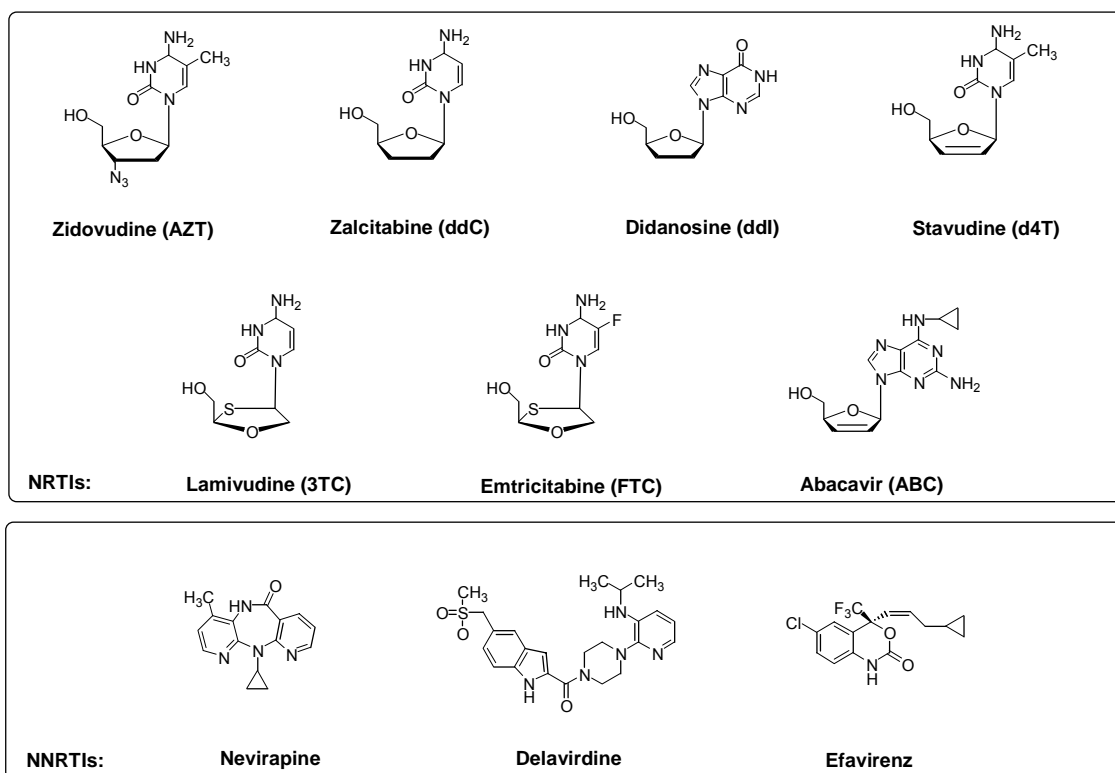
Figure 6. The structures of approved NRTIs and NNRTIs.

2.2. Aim of Study

Screening of low molecular weight compounds in massive parallel fashion is the state-of-the-art methodology for the identification of biologically functional molecules and thus, spurs endeavors in drug discovery. However, for every screening project a suitable assay format has to be developed.

Recently, effector-dependent reporter ribozymes have proven to be powerful tools which can report molecular interactions, e.g. protein-protein, protein-aptamer, or protein-metabolite interactions, in real-time. The assays are broadly applicable to and suitable for high throughput screening (HTS) to search for new molecules which disrupt such interactions. Reporter ribozyme FK-1 is a rationally engineered allosteric ribozyme, which contains an aptamer sequence that binds to an enzyme namely HIV-1 RT selectively in domain specific manner. The molecular interaction between HIV-1 RT and an aptamer sequence inserted into the ribozyme construct are detectable through the generation of a fluorescent signal.

The aim of this study is the development of HTS based on the reporter ribozyme FK-1 as an assay format to demonstrate its general utility in screening for new drug like molecules. First, the assay has to be applicable to automation of liquid handling with sufficient reliability and sensitivity, which is a prerequisite of HTS. Of particular interest is to identify novel types of HIV-1 RT inhibitors, inheriting the inhibitory properties of the aptamer, by using the ribozyme-based screening assay. Subsequent studies of the identified active hit compounds would include hit validation and characterization. For this undertaking, the re-synthesis of the hit compounds is essential for the chemical identification of hit compounds. Further characterization for selectivity and inhibitory elucidation should be performed with those synthesized compounds.

2.3. Results

2.3.1. Principle of Reporter Ribozyme-Based Screening Assay

The reporter ribozyme construct FK-1 (see Figure 7), which is regulated by HIV-1 RT, was previously published by Hartig et al..[32] FK-1 is a rationally designed allosteric ribozyme, which is able to detect the presence of HIV-1 RT selectively in a domain specific manner. Allosteric regulation of ribozyme cleavage activity results from the insertion of an aptamer sequence into stem II of the hammerhead ribozyme. The RNA aptamer sequence[74] was isolated by the SELEX procedure (Systematic Evolution of Ligands by Exponential Enrichment) and subsequent analysis of the aptamer revealed highly specific binding to HIV-1 RT with an extraordinarily low *Kd* value of 25 pM and the characteristic folding into a pseudoknot structure.[75] The allosteric construct FK-1 can fold into two structures depending of HIV-1 RT; the ribozyme-based structure and the pseudoknot aptamer-based structure (see Figure 7). The ribozyme-based structure of FK-1 is preferred in the absence of HIV-1 RT by the stable hairpin loop conformation of the inserted aptamer sequence in the ribozyme fusion construct. While the pseudoknot structure is induced by the addition of HIV-1 RT, and subsequent disruption of stem II leads to inhibition of cleavage activity of the hammerhead ribozyme.

Additionally, highly selective detection of HIV-1 RT by FK-1 was demonstrated. Only HIV-1 RT showed significant inhibition of ribozyme activity, compared with several noncognate proteins. Noteworthy is that no inhibition was observed with the homologous RT of HIV-2 in the same assay. FK-1 functions in protein selective as well as in a domain specific manner. The detection of interaction between FK-1 and HIV-1 RT was competitively influenced with primer/template complex due to the fact that free RNA aptamer sequence locates in the primer/template binding site of HIV-1 RT as indicated by the crystal structure.[76] This effect was not observed with nevirapine, one of the NNRTIs, which binds to the site near to, but distinct from, the polymerase active site.[73] All the features of FK-1 evidenced that the functions derived from the free aptamer were maintained in the fusion ribozyme context.

Those intriguing findings inspired us to use FK-1 as a high throughput screening (HTS) format for identification of inhibitors of HIV-1 RT, which can bind to HIV-1 RT in a domain specific fashion. Figure 7 describes the principle of reporter ribozyme (FK-1)-based screening assay.

17
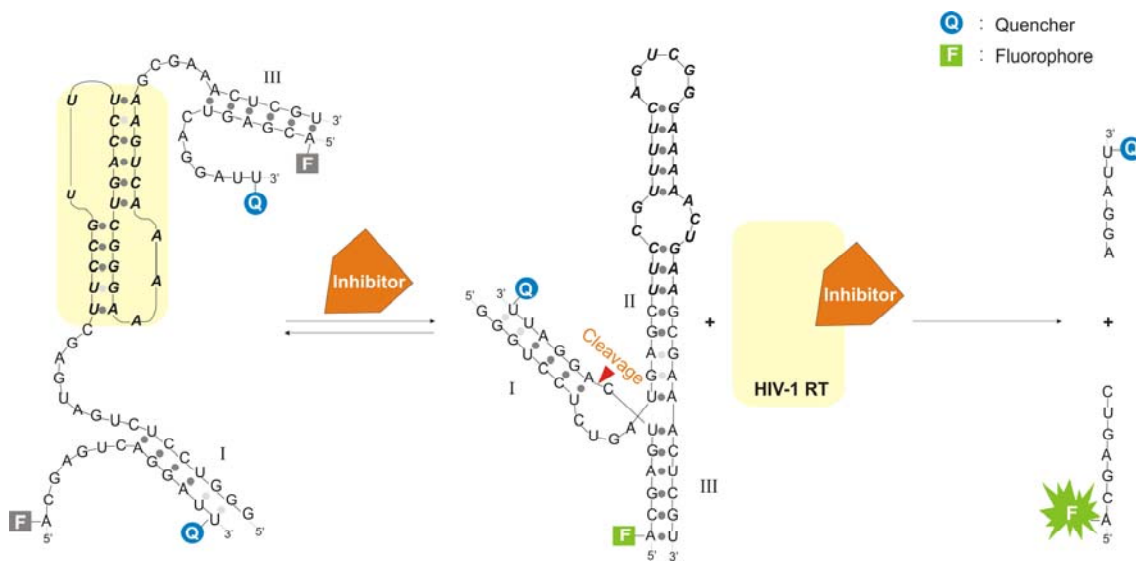
Figure 7. Structural interchange and cleavage activity of reporter ribozyme (FK-1).[32] Aptamer sequences inserted into the hammerhead ribozyme are highlighted in boldface italics. FRET-substrates carry 5'-fluorophore (F) and 3'-quencher (Q). The arrow indicates the FRET-substrate cleavage site. In the presence of HIV-1 RT (yellow), the aptamer sequence adopts a pseudoknot, disrupting the formation of stem II (left). Addition of a potential HIV-1 RT inhibitor (orange), which interferes with the interaction between HIV-1 RT and the aptamer by binding to HIV-1 RT, induces the active conformation of the hammerhead ribozyme (middle). Subsequent cleavage of the FRET-labeled substrate results in the generation of a fluorescent signal increase that can be detected in real time (right).

The inactive state of FK-1 in the presence of HIV-1 RT is restored by the addition of a potential HIV-1 RT inhibitor, which is able to bind to HIV-1 RT by competing with the binding site of the aptamer inserted into FK-1. The structural conversion of FK-1 to the active form results in the hammerhead ribozyme cleavage activity.

In the reporter ribozyme assay, the catalytic cleavage reaction of ribozyme can be detected by using the substrate oligonucleotides labeled with two fluorescent dyes at both ends; one end with fluorophore and the other with quencher.[77] The spatial proximity of the two dyes causes fluorescence quenching of the donor fluorophore by fluorescence resonance energy transfer (FRET). The cleavage of the FRET-labeled substrates can then be monitored by time-dependent increase of fluorescence in real-time.

Moreover, the reporter ribozyme-based assay has no requirement for labeling of compounds or proteins and allows simple establishment of automatization.[46, 78] In particular, FK-1 serves enhanced sensitivity which arose from the low picomolar range

affinity of the aptamer-HIV-1 RT complex in addition to the signal amplification by catalytic cleavage reaction via the ribozyme. These features should render FK-1 to be a highly suitable format for employment of HTS assays, which had to be investigated within this thesis.

2.3.2. Demonstration of FK-1 Ribozyme-Based Assay for Drug Screening

To test the suitability to employ FK-1 in screening assays, competition experiments with the known HIV-1 RT inhibitor (-)-epigallocatechin gallate were carried out (Figure 8). This compound is a natural product, and its inhibitory effect on HIV-1 RT have been reported by Nakane et al.[79] Inhibition of HIV-1 RT by (-)-epigallocatechin gallate was observed by the authors at a concentration required for 50 % inhibition (IC_{50}) of 0.045 µM. As illustrated in Figure 8, the inhibitory effects of (-)-epigallocatechin gallate were detected in a proof of principle experiment by exploiting FK-1. However, as reported by Nakane et al. other RTs and DNA polymerases such as eukaryotic DNA polymerase α, β, and γ were also inhibited but at more than five-fold higher values of IC_{50}. The detailed mechanism of HIV-1 RT inhibition by the compound is unclear. However, having observed the competitive mode of inhibition with respect to the primer/template complex by (-)-epigallocatechin gallate and the slight inhibition on DNA polymerase γ, the authors suggested catechin binding to HIV-1 RT through the recognition of the primer/template complex binding site.
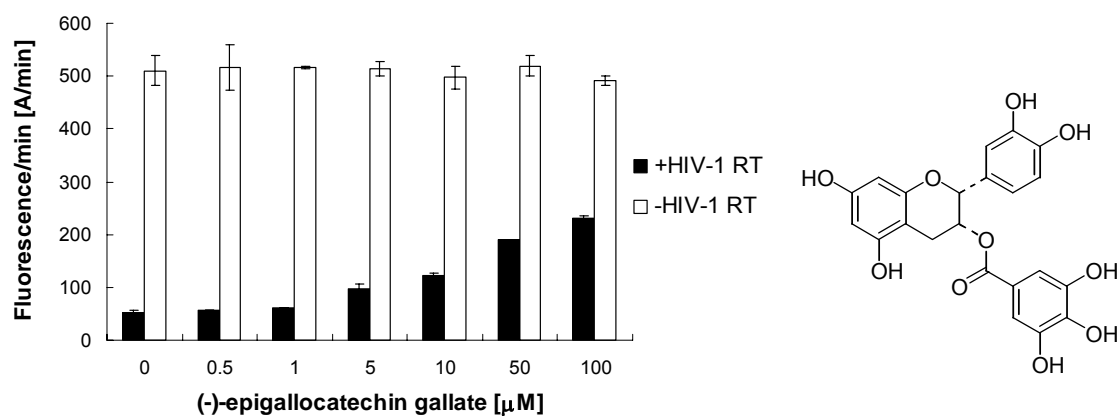
Figure 8. Competition of reporter ribozyme FK-1 from HIV-1 RT by increasing concentrations of the inhibitor (-)-epigallocatechin gallate. Black columns: reactivation of the FK-1-catalyzed cleavage activity by increasing amounts of (-)-epigallocatechin gallate in the presence of HIV-1 RT. White columns: control reactions for full cleavage activity of FK-1 alone and in the presence of (-)-epigallocatechin gallate at the indicated concentrations. The structure is shown in the right side. The mean values and the error bars are from duplicate measurements.

As shown in Figure 8, the sensitivity of FK-1 depending on HIV-1 RT was evidently supported in the absence of the catechin compound. The reactivation of FK-1 cleavage activity was observed with the addition of catechin compound in a concentration dependent manner. With 100 μM of the compound, half of the ribozyme activity was restored by interfering with the interaction between HIV-1 RT and the aptamer. These results strongly indicate that the compound is most likely to bind to HIV-1 RT, and demonstrate that the assay can be suitably used to identify inhibitors of HIV-1 RT.

Furthermore, this model system implies how to evaluate potential hit compounds in the screening assay. The relative activity (A_{rel}) of FK-1 at a certain concentration of a compound can be expressed by the proportion of ribozyme activity in the presence of HIV-1 RT versus full ribozyme activity in the absence of HIV-1 RT. Thus, by comparison of the value of A_{rel} with compound to A_{rel} without compound (usually the solvent DMSO was added in the assay for 0 μM of compound), the inhibitory effects of a given compound can be accessed. For example, the A_{rel} for 100 μM of (-)-epigallocatechin gallate was found to be 0.5, while A_{rel} for DMSO addition was 0.1. Theoretically, the value of A_{rel} should vary from 0 to 1. Hit compounds can be identified by values of A_{rel} that are higher than A_{rel} of DMSO and possibly near to 1.

2.3.3. Optimization of Robot Manipulation

Prior to perform HTS, the FK-1 assay was applied to automated liquid handling system using a robotic manipulation. The reaction mixtures and small molecule compounds from the library were mixed by a robot automatic liquid handling system before initiation of the ribozyme reaction. Primary test operations resulted in considerably varying data output under the same assay conditions. It turned out that several plastic materials (e.g. reservoir, tubes) available for robot manipulations seemed to have affinity either to RNA or to proteins. All these problems were successfully circumvented through employment of equipment that contains ceramic multiple tips for the robot autosampler and the preparation of reaction mixtures in glass vials.

2.3.4. High Throughput Screening of Small Molecule Library

A small molecule library derived from 2500 compounds (Comgenex I), each at 100 µM concentration, were screened using the FK-1 ribozyme-based assay with optimized robot manipulations. The relatively high concentration was used because it permits observation of even weak effects of compounds in an initial evaluation of the library. To identify hit compounds, the ribozyme relative activity (A_{rel}; see §2.3.1 and §4.6.2) was evaluated by comparison to A_{rel} obtained with DMSO.

Figure 9 describes an example of screening results. Most of the library members yielded averaged A_{rel} values within in the range of A_{rel} from DMSO (ca. 0.2 ± 0.1). In the primary screening, the compound showing a value of $A_{rel} > 0.4$ were identified as potential hit compounds, which were marked in black in Figure 9. Applying this rule to the primary screening, a frequent occurrence of hit compounds (ca. 8 % of the total number of the library compounds) was observed. Subsequently, in a secondary screening all selected compounds identified in the primary screening with a $A_{rel} > 0.4$ were re-screened at 50 µM concentration. This compels to identify more active compound and less fault-positive output. Finally, 67 molecules were identified as active compounds by employing the FK-1 reporter ribozyme-based assay. This marks 2.7 % of hit occurrence.
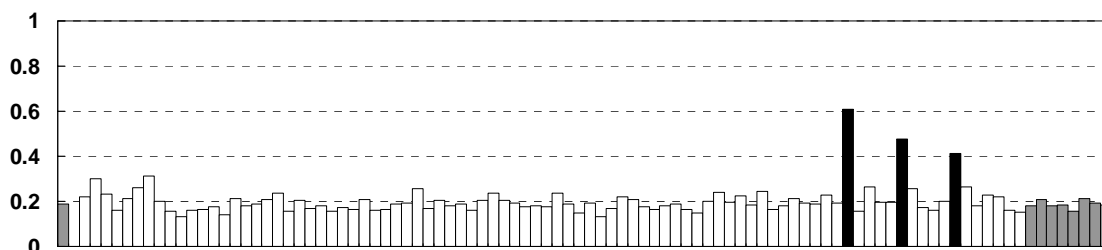
Figure 9. A_{rel} values of library compounds. Each bar represents the mean value from duplicate measurement for one compound from the library. The assay was performed in the presence of 100 µM of compound (white bars). The controls with the addition of DMSO were shown in gray bars. The bars with relative values (A_{rel}) higher than 0.4 were marked in black.

Several of the identified compounds were examined in the FK-1 assay to see whether they recover the ribozyme cleavage activity in a concentration dependent manner. For instance, two hit compounds (**2E10** and **28F6**), which were identified after the secondary screening, recovered fluorescence generation of FK-1 depending on the concentration increment (Figure 10). On the other hand, the compound **16A3**, which is one of the inactive compounds selected randomly from the library members, did not influence the ribozyme cleavage activity. These results indicate that the FK-1 based assay format is suitable to monitor inhibition in a concentration dependent manner.
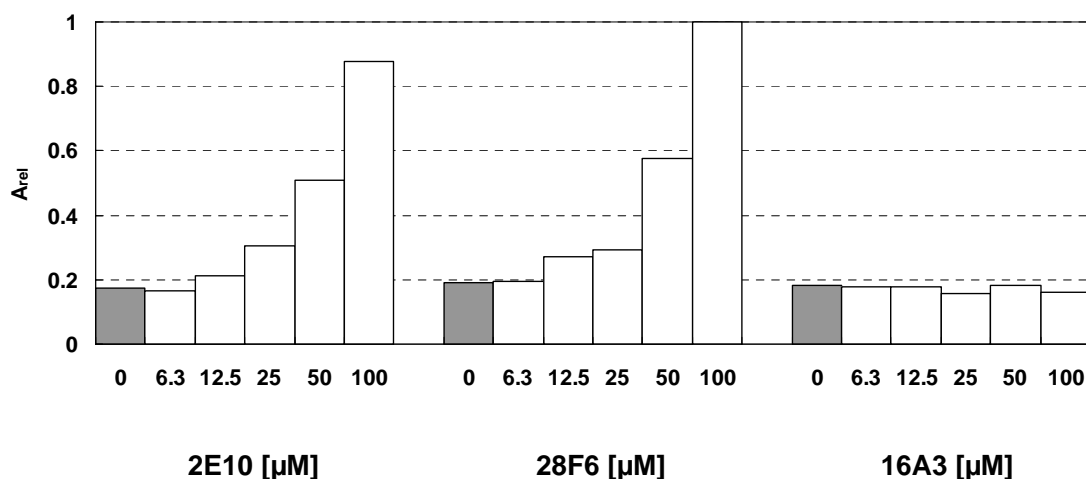
Figure 10. Concentration dependent reactivation of FK-1 ribozyme activity by the hit compounds identified in the screening. **2E10** and **28F6** were identified as hit compounds, and **16A3** was randomly chosen from inactive compounds by hit evaluation in the same library.

2.3.5. *in vitro* HIV-1 RT Inhibition by Hit Compounds

In general, the affinity of the ligand to its target correlates with inhibition of enzyme activity. Thus, an activity-based assay to evaluate the final 67 hit molecules was designed to examine their ability to inhibit the DNA polymerase activity of HIV-1 RT. According to the crystal structure of HIV-1 RT in the complex with pseudoknot RNA aptamer sequence, it was shown that the position of the pseudoknot RNA overlapped the binding surface of duplex DNA substrate.[76]

Taking into account this information, the 67 hit compounds were initially evaluated for the *in vitro* inhibitory activity towards the DNA-dependent DNA polymerase reaction of HIV-1 RT using a 5'-[^{32}P]-DNA primer/DNA template complex. The effect of the compound was analyzed in the range of the concentration of 5 nM-50 µM, as shown in Figure 11. Seven molecules out of 67 resulted in strong inhibition of DNA-dependent DNA polymerase activity.
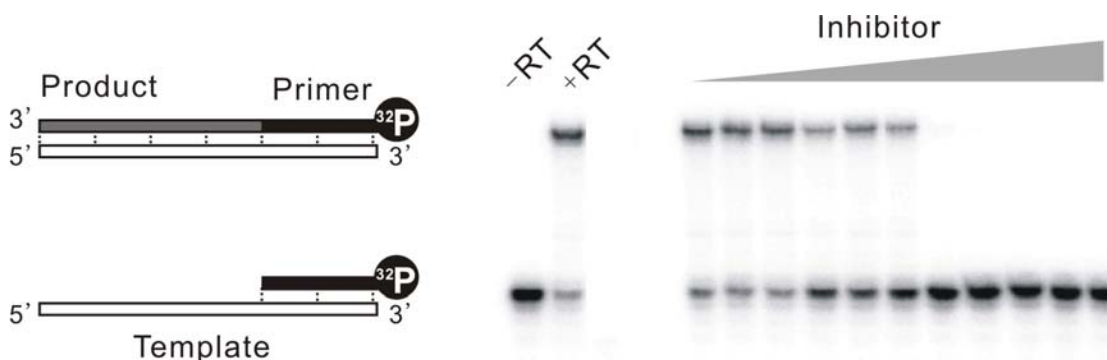
Figure 11. Inhibition of DNA polymerase activity of HIV-1 RT by small molecule inhibitors. In the presence of HIV-1 RT, 5'-end labeled primer is extended employing either DNA or RNA template (second lane from the left; +RT). The control lane in the absence of HIV-1 RT shows the primer position (left lane; -RT). Polymerization reactions were inhibited by the addition of increasing concentrations of inhibitors (right lanes), which were able to be detected by radioactive intensity of elongated products. In this example, the lanes including inhibitor shows the inhibition by **28F6** at the concentration of 0.5-50 µM.

The idioms and structures of the seven hit compounds are shown in Figure 12. Interestingly, partial structural similarity was observed with some compounds. The compounds had more than 70 % purity by HPLC analysis except **1B11**, which yielded 2 major peaks.
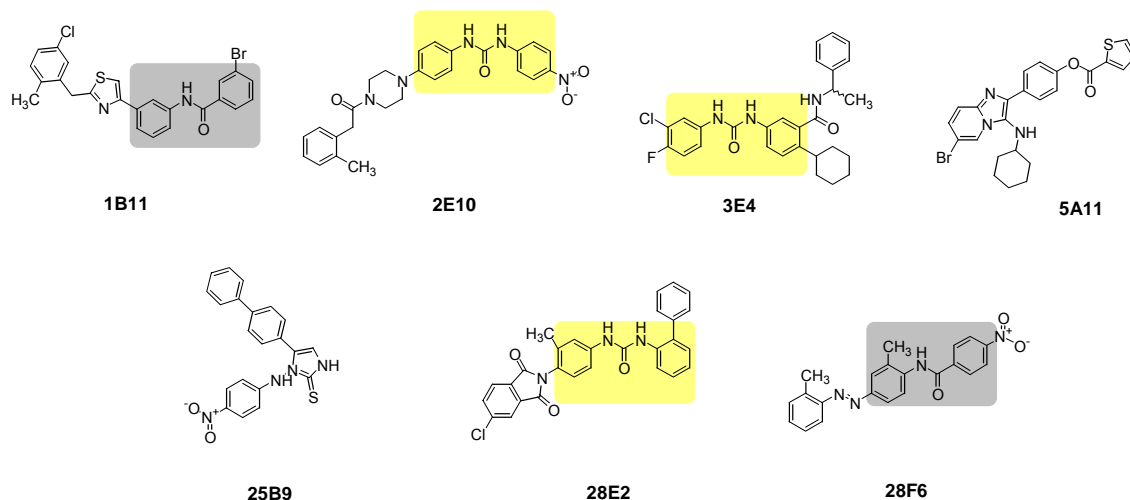
Figure 12. The structures of seven compounds, which showed evident inhibition on DNA-dependent DNA polymerase activity of HIV-1 RT *in vitro*. The colored groups (yellow and gray) represent structural similarity among several compounds.

To compare the effects of the compounds on HIV-1 RT action, the inhibitory concentrations that gave half-maximal activity (IC_{50}) for the above seven compounds were determined from the dose-dependent inhibition curves by using the method shown in Figure 11 (Table 2). Specific inhibitory effects with those molecules were confirmed by control experiment, where **18B11**, which was selected arbitrarily from the library, exhibited no inhibition on RT under the same assay condition. Three potential inhibitors (**3E4**, **2E10**, and **28F6**) which gave relatively low IC_{50} in the range of 3-4 µM were chosen for further analysis and characterization.

compound	IC$_{50}$ [μM]
3E4	3.11
2E10	3.49
28F6	3.99
1B11*	5.12
28E2	5.97
5A11	5.98
25B9	9.33
1B11*	10.3

Table 2. Inhibition of DNA-dependent DNA polymerase activity of HIV-1 RT by seven inhibitors. *: two major compounds derived from **1B11** by HPLC separation.

2.3.6. Synthesis of 28F6, 3E4, and 2E10

To pursue the process of stage-by-stage drug development, the activity of the hit compounds has to be followed up with an identity and purity evaluation. Re-synthesis of the authentic samples permits to confirm the activity and validity of the corresponding hits. The compounds ensured through this objective assessment are able to progress into the subsequent investigations.

For those reasons, the three most promising compounds (**28F6, 3E4, and 2E10**) were re-synthesized to yield sufficient quantities in high purity.

First, the coupling of amine **1** and benzoylchloride **2** in the presence of triethylamine gave **28F6** in one step (Scheme 1).
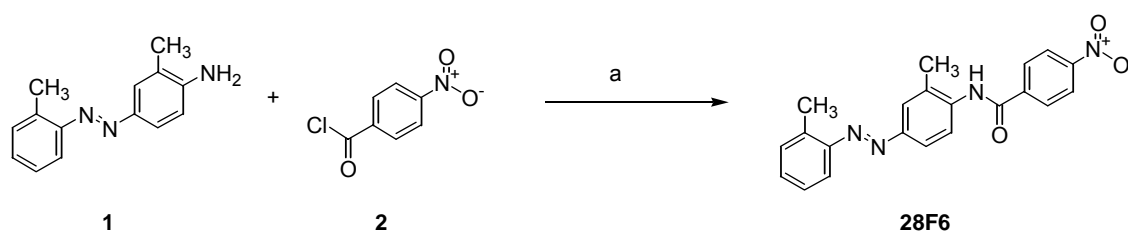
Scheme 1. Synthesis of **28F6**. Conditions: (a) Et$_3$N, CH$_2$Cl$_2$; 18%.

The synthetic strategy of **3E4** is outlined in Scheme 2. The amide **5** was prepared by the

coupling of benzoic acid **3** and the benzylamine **4** in the presence of EDAC. Subsequent treatment of **5** with piperidine yielded **6**. Catalytic hydrogenation of the nitro group in **6** was required to generate amine substituted product **7**, which was then condensed with 3-chloro-4-fluorophenylisocyanate. The reaction proceeded smoothly in the polar aprotic solvent THF at room temperature, giving the final product **3E4**.
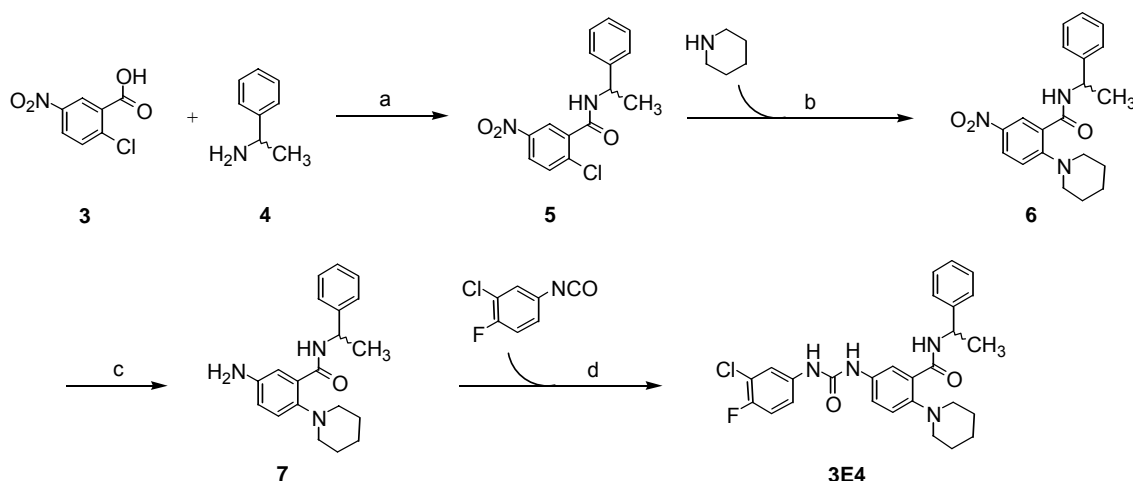
Scheme 2. Synthesis of **3E4**. Conditions: (a) EDAC, DMF; 35%. (b) pyridine, reflux; 88%. (c) Pd/C, H_2, EtOH; 89%. (d) THF, rt; 65%.

Target **2E10** was able to be synthesized using a similar procedure as employed for **3E4** (Scheme 3). *O*-tolylacetic acid (**8**) was coupled with piperazine **9** in the presence of *N*-hydroxybenzotriazole and EDAC to yield amide **10**. The amino group in **11** was introduced by catalytic hydrogenation of the nitro group in **10** using Pd/C. Final condensation of **11** and 4-nitrophenyl isocyanate gave assymmetric diphenyl urea **2E10**.
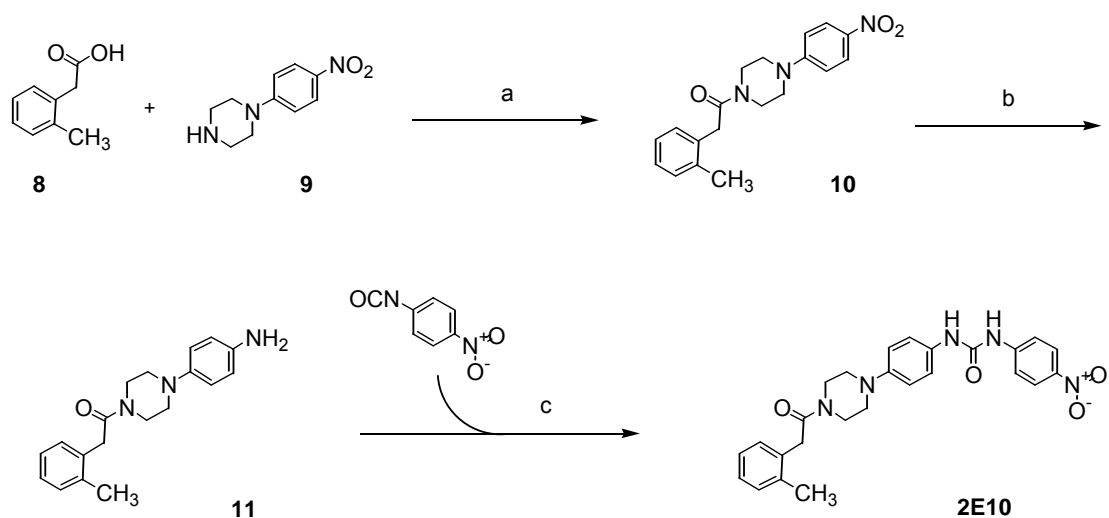
Scheme 3. Synthesis of **2E10**. Conditions: (a) EDAC, N-hydroxybenzotriazole, DMF, CH_2Cl_2; 82%.
(b) Pd/C, H_2, EtOH; 86% (c) THF, DMSO, rt; 82%.

Having all the synthetic compounds in hand, their inhibitory effects on DNA-dependent
DNA polymerase activity of HIV-1 RT were re-examined and compared the effects of
synthesized compounds to those of the library samples. IC_{50} values for both library and
synthetic compounds are listed in Table 3. The obtained data indicates that the
synthesized **28F6**, **3E4**, and **2E10** exhibit approximately equivalent activities as the
library samples.
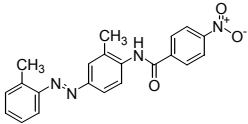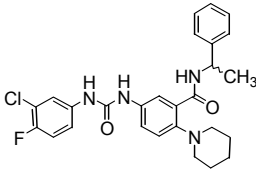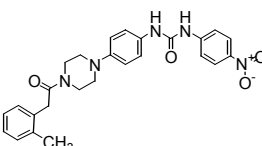
IC_{50} [µM]	28F6	3E4	2E10
Comgenex I library	3.99 ± 0.27	3.11 ± 0.55	3.49 ± 0.15
synthesized compound	5.01 ± 0.19	2.10 ± 0.58	3.72 ± 0.07

Table 3. Inhibition of DNA-dependent DNA polymerase activity of HIV-1 RT by both library and
re-synthesized compounds. IC_{50} values are the average of duplicate analyses. The structures of **28F6**,
3E4, and **2E10** are also shown.

2.3.7. Selectivity of 28F6, 3E4, and 2E10 *in vitro*

The selectivity of the small molecule inhibitors was evaluated by testing their inhibitory effects on various reverse transcriptases (HIV-2, avian myeloblastosis virus (AMV), and moloney murine leukemia virus (MMLV)) and DNA polymerases from both prokaryotic (Klenow Fragment (KF)) and eukaryotic (human DNA polymerase β (pol β)) sources. It was of interest to know if the small molecules identified via the FK-1 assay maintain the selective inhibitory property as the anti-HIV-1 RT aptamer, which was incorporated into a part of reporter ribozyme sequence. It is reported that the free anti-HIV-1 RT aptamer is capable of discriminating HIV-1 RT from other RTs such as HIV-2, AMV, and MMLV RTs.[74, 75] HIV-1 and HIV-2 are classified as lentivirus, while AMV and MMLV are in C-type retroviruses, depending upon the particle shapes. The selectivity of the pseudoknot aptamer towards HIV-1 is remarkable, because HIV-2 has 60 % sequence homology to HIV-1. Noteworthy, HIV-2 RT shows resistance to NNRTIs due to the structural difference in the NNRTI binding pocket.[80]

DNA polymerases were also investigated because their inhibition by the compound is prone to lead to the cell toxicity. Additionally, from amino acid sequence comparisons as well as crystal structure analyses, E. coli DNA polymerase I (Pol I), whose large fragment is KF, pol β, and RTs belong to different families. However, all DNA polymerases and RTs utilize an identical two metal ion catalyzed mechanism.[81-83] The results might provide a clue to understand the inhibitory mechanism of the identified compounds.

Figure 13 describes the inhibition of the three compounds on DNA-dependent DNA polymerase activity (DDDP) of RTs (HIV-1, HIV-2, AMV, MMLV) and DNA polymerases (KF, pol β) as well as the inhibition of three compounds on RNA-dependent DNA polymerase activity (RDDP) of RTs. IC_{50} values obtained from the dose-response curves in Figure 13 are summarized in Table 4.

Figure 13 (see p.29). Effects of the indicated compounds on DDDP and RDDP activities of various DNA polymerases and reverse transcriptases. Dose-response curves for DDDP inhibition by **28F6** (**A**), **3E4** (**C**), and **2E10** (**E**) as well as dose-response curves for RDDP inhibition by **28F6** (**B**), **3E4** (**D**), and **2E10** (**F**) are represented. The DDDP activity was assayed by using [α-^{32}P]-labeled DNA primer/DNA template, while RDDP activity was assayed by using [α-^{32}P]-labeled DNA primer/RNA template. The symbols used in the graphs are as follows; HIV-1 RT (■), HIV-2 RT (●), AMV RT (▲), MMLV RT (▼), Klenow fragment 3'-5' exo⁻ (◄), and human DNA polymerase β (►). The mean values and the error bars are the results from two independent experiments.
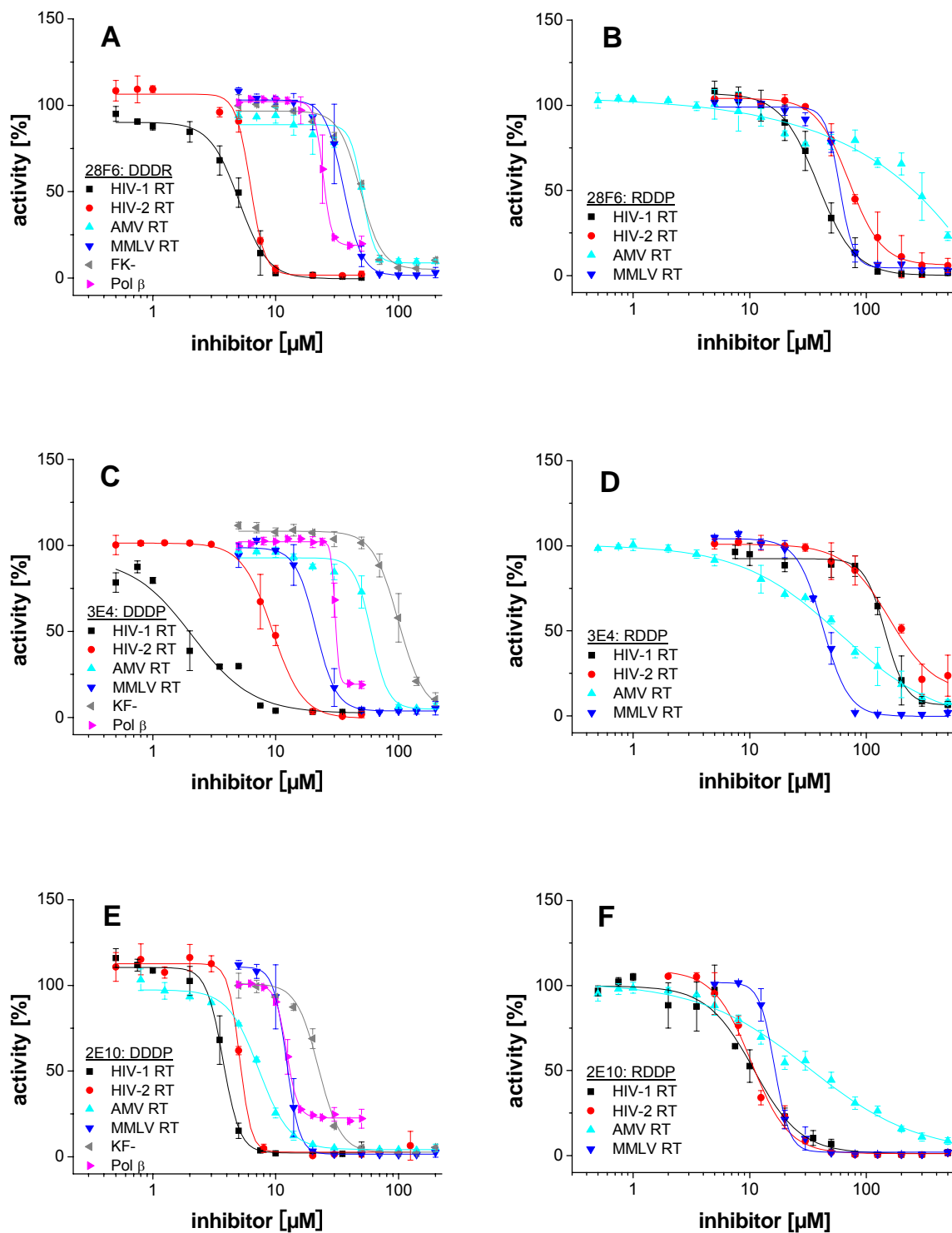
As a result, **28F6** showed moderate selectivity (4-10 fold) to HIV-1 (IC_{50} = 4.99 µM) and HIV-2 RT (6.21 µM) over other RTs and DNA polymerases to inhibit DDDP (Figure 13A, Table 4). While higher concentration of **28F6** (IC_{50} > 30 µM) was required to inhibit RDDP activities even with no selectivity (Figure 13B, Table 4).

Considerably enhanced selectivity was observed with **3E4** to inhibit DDDP (Figure 13C, Table 4). Compared to the IC_{50} value for HIV-1 RT (IC_{50} = 2.1 µM), AMV RT, MMLV RT, KF, and pol β were inhibited with 10-50 fold higher IC_{50} values. Moreover, **3E4** exhibited selectivity to HIV-1 RT over HIV-2 RT in the respect to DDDP inhibition. The difference between HIV-1 and HIV-2 in the observed IC_{50} values is a factor of about 4.5. IC_{50} values for RDDP inhibition by **3E4** were in the range of 40-160 µM with slightly stronger inhibition for AMV and MMLV compared to HIV-1 and 2 (Figure 13D, Table 4).

2E10 inhibited DDDP of all the tested polymerase and RTs with less selectivity and with IC_{50} values of 3.7-22.8 µM (Figure 13E, Table 4), and thus, turned out to be the inhibitor with the lowest selectivity among the tested three compounds. In contrast to **28F6** and **3E4**, RDDP activities of all RTs were inhibited with lower IC_{50} (10-30 µM) by **2E10** (Figure 13F, Table 4) with slightly less effect on DDDP for each RT tested.

The control compound **20H3** showed considerably higher IC_{50} values or no inhibition on DDDP and RDDP of all the enzymes used in this assay (data not shown).

Overall, the data presented in Figure 13 and Table 4 indicates that all three compounds identified as hits exhibited effects to inhibit DDDP stronger than RDDP activities. With respect to inhibition of DDDP, **3E4** showed not only the lowest IC_{50} value for HIV-1 RT but the highest discrimination from all the other RT and DNA polymerases which were investigated.
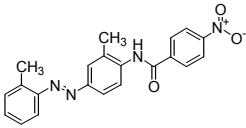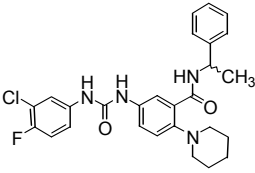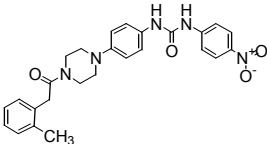
IC_{50} [µM]		28F6	3E4	2E10
HIV-1	DD	5.01 ± 0.19	2.10 ± 0.58	3.72 ± 0.07
RT	RD	38.4 ± 0.94	145 ± 5.06	10.8 ± 0.89
HIV-2	DD	6.21 ± 0.18	9.41 ± 0.21	5.09 ± 0.07
RT	RD	71.2 ± 1.48	160 ± 19.1	10.3 ± 0.50
AMV	DD	50.8 ± 1.70	64.1 ± 2.64	7.29 ± 0.18
RT	RD	> 100	63.7 ± 7.79	32.0 ± 4.79
MMLV	DD	35.6 ± 1.06	21.2 ± 0.93	12.2 ± 0.06
RT	RD	59.2 ± 1.78	42.5 ± 1.23	16.4 ± 0.39
KF⁻	DD	48.9 ± 2.60	99.7 ± 3.25	22.8 ± 0.53
Pol β	DD	24.2 ± 0.23	30.4 ± 0.77	12.3 ± 0.08

Table 4. IC_{50} values of **28F6**, **3E4**, and **2E10** for different DNA polymerases and reverse transcriptases. The structures of **28F6**, **3E4**, and **2E10** are also represented. DD; DNA-dependent DNA polymerase activity, RD; RNA-dependent DNA polymerase activity.

2.3.8. Binding Analysis of 28F6, 3E4, and 2E10 via SPR

Although DDDP activity of HIV-1 RT was nicely inhibited by **28F6**, **3E4**, and **2E10** with acceptable selectivity, all the data obtained so far are not able to support direct interaction between the small molecules and HIV-1 RT. The possible inhibition mode could be the intercalation of the compounds into the primer/template DNA double helix or binding to the major or minor groove of the DNA helix. In order to eliminate this hypothesis, inhibitor binding to the immobilized primer/template complex was investigated via surface plasmon resonance (SPR) (Figure 14). As a control, the known DNA double helix binder SYBR green was used to observe a binding response with the primer/template complex (Figure 14D). The concentration dependent binding signals were indeed obtained through SYBR green injection to the chip. In contrast, Figure 14A-C clearly show that none of the three inhibitors bind to the primer/template

31

complex immobilized on the chip surface in the range of 50 nM – 5 µM. Theses results strongly support a mechanistic model that inhibitions results from binding of the compounds to HIV-1 RT or the HIV-1 RT-primer/template ternary complex.
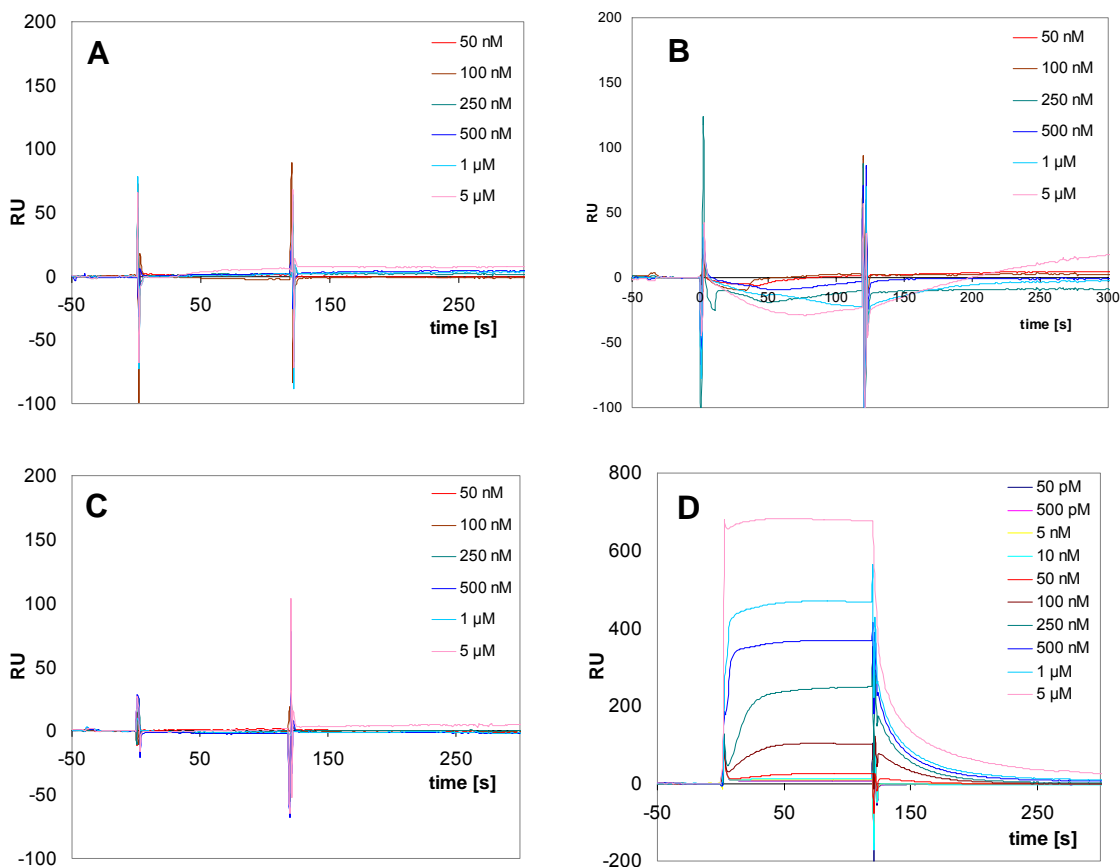
Figure 14. BIAcore sensorgrams showing no binding or binding of small molecules to the primer/template complex immobilized on the chip surface. **28F6** (**A**), 3E4 (**B**), and **2E10** (**C**) were injected onto the surface at the concentrations from 50 nM to 5 µM. While SYBR green (**D**) was injected at the concentrations from 50 pM to 5 µM. The SA chip surface was prepared by immobilization with 300 RU of the 12-mer primer sequence, followed by subsequent annealing (500 RU) of the 18-mer template at a flow rate of 5 µl/min. The compounds were injected to the surfaces including a reference cell at a flow rate of 30 µl/min.

In another examination using SPR, HIV-1 RT was coupled on the chip surface to observe interactions with the inhibitors. As shown in Figure 15, the injection of **3E4** over the RT immobilized resulted in a concentration dependent signal enhancement. This observation demonstrated the direct interaction between HIV-1 RT and **3E4**.
However, inherent poor solubility seems common to the identified compounds and

hampered to obtain reproducible binding signals with **28F6** and **2E10**. The accurate dissociation constants (K_d) was not be accessible due to the same reason.
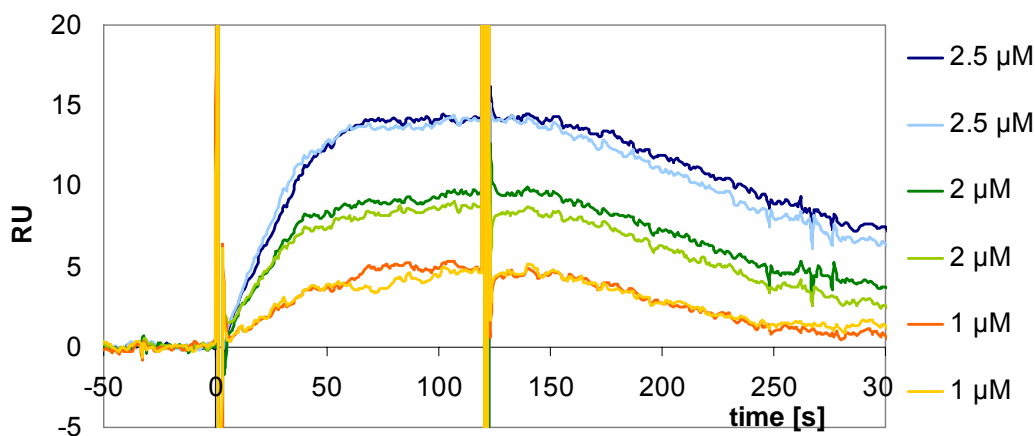
Figure 15. BIAcore sensorgrams showing reproducible binding curves of **3E4** to HIV-1 RT immobilized on the chip surface. **3E4** was injected onto the surface at the concentrations at 1, 2, and 2.5 µM at a flow rate of 30 µl/min in running buffer including 1% DMSO. The chip surface was prepared by immobilization with 2000 RU of HIV-1 RT at a flow rate of 5 µl/min. The curves were normalized with the reference cell.

2.3.9. A Cell-Based Assay using a Self-Inactivating HIV Vector System

Conventional phenotypic analysis of HIV to antiviral therapy is time-consuming and requires to culture infectious viruses. The so-called self-inactivating virus vector system allows to determine the sensitivity of HIV to RT, protease (PR) and integrase inhibitors rapidly (approximately in 10 working days) at reduced biosafety levels (Figure 17).[84] Having shown the *in vitro* specific inhibition of HIV-1 RT with the three inhibitors **28F6**, **3E4**, and **2E10**, their initial characterization *in vivo* was carried out by the group of Prof. Dr. T. Restle, University of Luebeck, using a self-inactivating HIV vector system.

The design of this assay requires two plasmids; one chimeric vector pGJ3 and the second plasmid containing the vesicular stomatitis virus G glycoprotein (VSV-G) gene. The pGJ3 vector plasmid includes the minimum essential sequences of a HIV laboratory virus, including the gag-pol and rev sequences as well as the rev responsive element (RRE), and a reporter gene. The plasmid lacks the HIV env, vif, vpu, vpr, tat

and nef genes and contains a deletion in the 3' U3 region (Figure 16). Cotransfection of the plasmid pGJ3 with VSV-G resulted in the production of replication incompetent virus vectors that contain a rudimentary HIV RNA genome with a marker gene, RT and PR proteins from the virus surrounded by capsid and matrix protein of the HIV. The vector nucleocapsid is enveloped by a VSVG-containing membrane. Infection of susceptible cells with the vectors leads to marker gene expression. The active inhibitors for RT, PR, or integrase represent reduced marker gene expression in a dose-dependent manner (Figure 17). In addition, RT and PR gene segments of the plasmid could be replaced by the corresponding region of an HIV test specimen, which permits to determine the sensitivity of drug resistant HIV isolate to inhibitors using the same system.
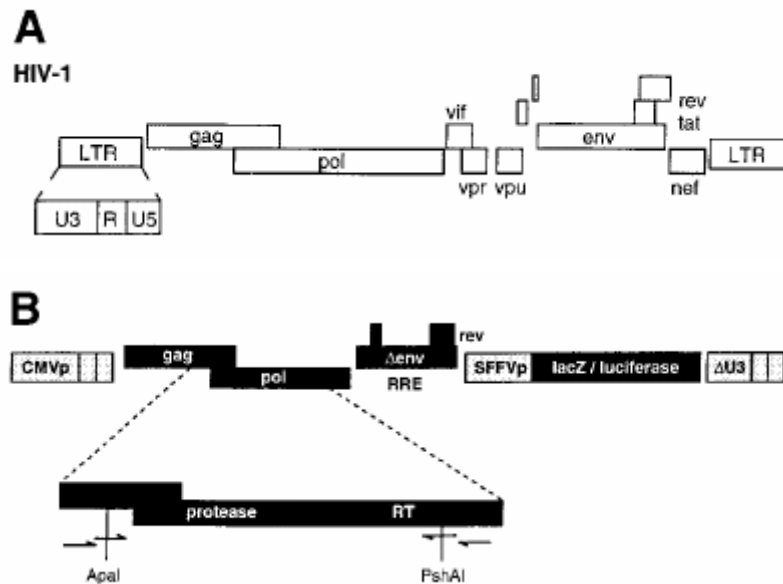
Figure 16. The self-inactivating replication-defective test vector. Plasmids pGJ3 (**B**) were constructed from a proviral clone of HIV-1 (**A**). The vector plasmid contains a chimeric cytomegalovirus promoter (CMVp) that regulates expression of the vector genome after transfection. The consecutive deletion processes of the env, nef, vpu, tat, vif, and vpr genes resulted in plasmid pGJ3 (**B**). Reporter genes driven by an internal promoter (SFFVp) were inserted in the region of the deleted nef gene. As reporter, either the E. coli lacZ or the firefly luciferase genes were selected. RT and PR sequences from laboratory or patient isolates, or from proviral clones can be alternatively inserted in the vector plasmid at the sites indicated to obtain chimeric test vector plasmids (**B**).
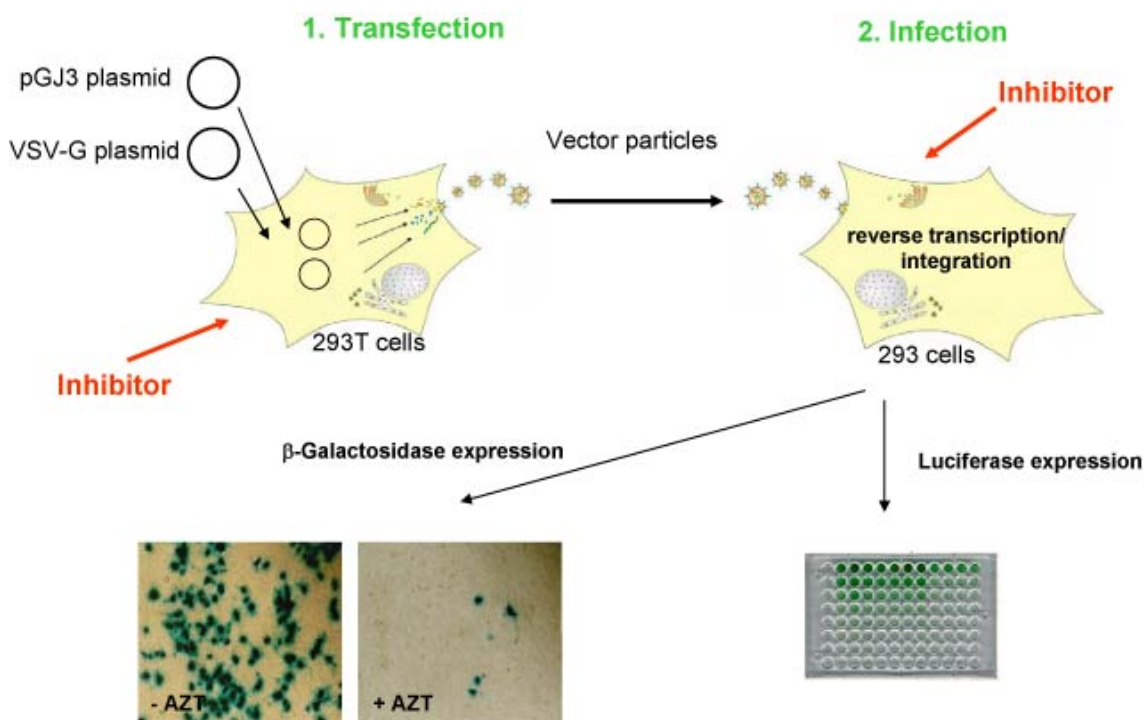
Figure 17. Cotransfection of a plasmid containing minimal HIV genome sequences (pGJ3) with a plasmid providing the vesicular stomatitis virus glycoprotein (VSV-G) leads to production of replication incompetent chimeric virus vectors which are then used to infect susceptible cells, resulting in marker gene expression. Inhibitors of the HIV-1 reverse transcriptase, protease and integrase can be analyzed in this cell-based assay system. Their inhibitory potential is reflected in a decrease of marker gene expression.

Figure 18 shows the result of the susceptibility of HIV to the identified inhibitors in the cell-based assay, where 293 cells were infected with the chimeric vector particles in the presence of **28F6**, **3E4**, and **2E10** as well as the negative control (**20H3**) and the positive control (AZT). The control compound **20H3** was identified to require significantly higher amounts to observe the inhibition of HIV-1 RT *in vitro*, compared to the other three inhibitors. One of the NRTIs, azidothymidine (AZT) was used as a positive control. Among the three compounds identified in the screening, **3E4** showed substantial reduction of reporter gene expression in concentration dependent manner. However, the other two inhibitors **28F6** and **2E10** kept the level of luciferase activity as observed with the negative control compound. The value of IC_{50} for **3E4** was determined at 5.3 μM. Nearly complete suppression of marker gene expression was achieved at 25 μM of **3E4**. It is remarkable that the inhibitor **3E4**, that is active in the cell-based assay, has exhibited the most promising characteristics accompanied with the lowest IC_{50} and the highest selectivity *in vitro*.
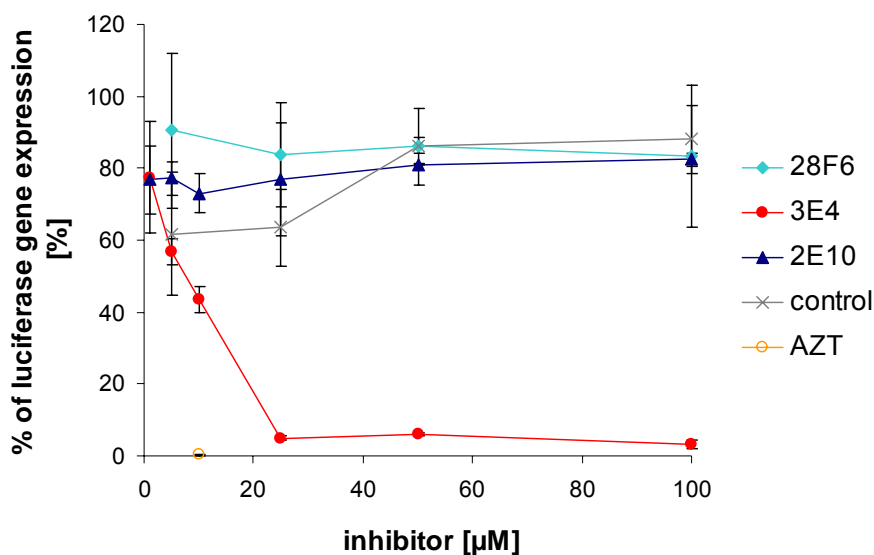
Figure 18. Inhibitory activity of small molecule RT inhibitors in the self-inactivating HIV vector system. Percent luciferase activity of 293 target cells transduced in the presence of the indicated amounts of each molecules with the chimeric vector is shown. The luciferase activity is expressed as percentage of the luciferase activity of cells cultured in the absence of inhibitors. The compound **28F6**, **3E4**, and **2E10** are the inhibitors identified from the screening, while the control corresponds to the compound **20H3** which was selected arbitrarily from the library. Positive control was performed with azidothymidine (AZT), one of the approved NRTIs. The mean and the standard deviation of duplicates are given.

2.3.10. Action of the Identified Inhibitors on *in vivo* HIV-1 Replication

Encouraged by the promising data for **3E4** from the study using the self-inactivating HIV system, further *in vivo* analysis of the inhibitors using HIV wild type was performed. The effects of the compounds to inhibit HIV-1 replication *in vivo* was investigated by the group of Prof. Dr. H.-G. Kraeusslich, University of Heidelberg, and the results are shown in Figure 19.

TZM cells infected with HIV-1 NL43 exhibit luciferase gene expression under the control of the HIV-1 LTR promoter.[85] As described in Figure 19, luciferase expression was reduced in considerable extent with **3E4**, in comparison to **28F6** and **2E10**. The assay includes the same negative and positive controls used in self-inactivating HIV vector system described above (Figure 18). In this assay, the control **20H3** resulted in some reduction of HIV-1 replication at high concentration (75 µM). Although AZT

showed extremely enhanced reduction even at low µM concentrations, **3E4** was able to reach the approximately equivalent level of reduction at 30 µM, which was also comparable to the result obtained from Figure 18.
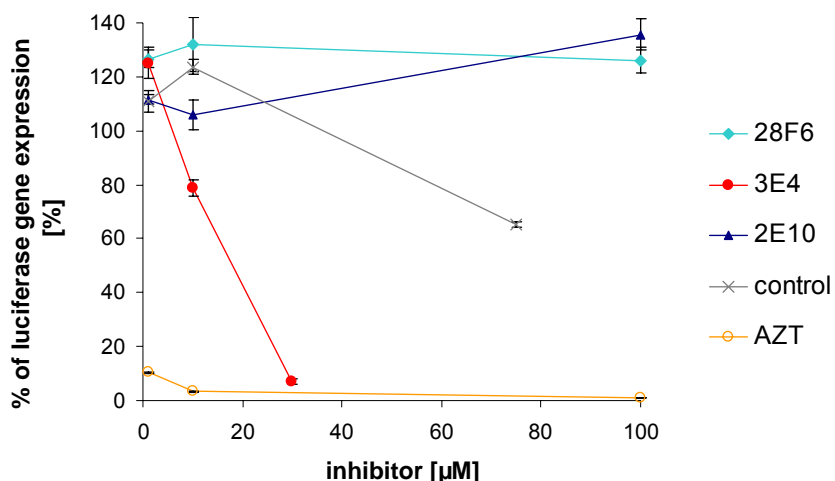
Figure 19. Effect on inhibition of HIV-1 replication by small molecule inhibitors. HIV-1 replication for TZM cells infected with HIV-1 NL43 wild-type virus in the presence of **28F6** (light blue diamonds), **3E4** (red-filled circles), **2E10** (dark blue triangles) or the negative control (gray crosses) and the positive control with the NRTI AZT (non-filled orange circle) was investigated by monitoring luciferase activity. The mean percentage of luciferase marker gene expression compared with untreated control cells from at least two replicate infections are shown.

HIV-1 virus particles were produced in MT4 co-culture in the presence of different inhibitors and used to infect TZM cells, expressing β-galactosidase under the control of the HIV-1 LTR promoter, thus enabling colorimetric detection and enumeration of the HIV-1 infected cells. The infectivity of the HIV-1 was examined with those compounds tested above (Figure 20). The infectivity of intact HIV-1 NL43 virus was arbitrarily set at 100 %. As controls, **20H3** and AZT were included. Again, **3E4** reduced the number of infected cells distinguishably from **28F6, 2E10** and the negative control. The complete reduction was observed in the presence of **3E4** at 30 µM.
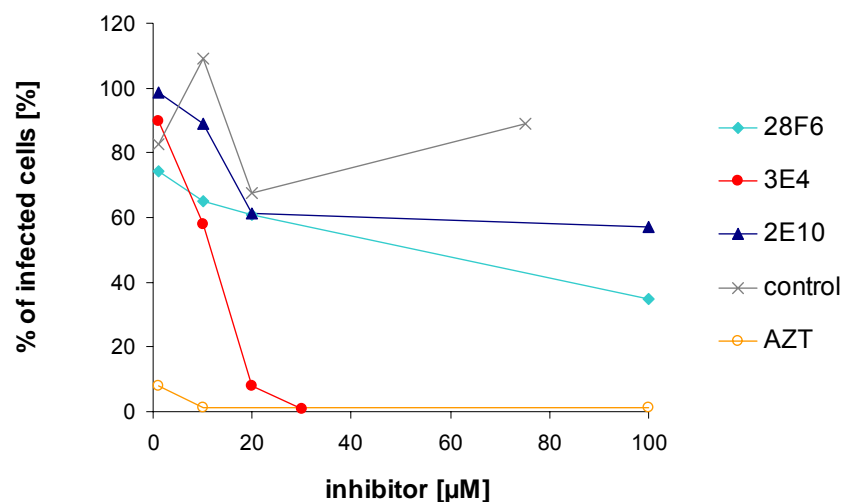
Figure 20. Effect on inhibition of HIV-1 infectivity. The supernatant from the MT4 co-culture with NL43 virus particles produced in the presence of inhibitors was used to infect TZM cells. The infected TZM cells were counted by staining with β-galactosidase substrate after 48 h. Values are expressed as percentage of the number of infected cells treated without inhibitors. The symbols for each compound are listed in the right. The mean of duplicates are given

To evaluate the cell toxicity of the inhibitors, escalating doses of the compounds alone were applied to TZM cells, and the cell viability was measured using a standard MTT assay (Figure 21). The significant reduction in transduction efficiency and infectivity could not be attributed to cytotoxic effects of the compound. Figure 21 showed that TZM cell viability was not significantly affected by the addition of **3E4** up to 20 µM as observed with AZT. Unfortunately, the cell viability steeply dropped down to 50 % level, when the concentration reached to 30 µM. The complete reduction of HIV-1 replication and infectivity showed before at the concentration more than 30 µM might have been caused in some extent by the effect of the cell toxicity of **3E4**. However, employing 20 µM of **3E4** enables to reach more than half reduction of HIV-1 replication as well as the infectivity with no cell toxicity. Hence, **3E4** is a highly potential lead structure for further anti-HIV drug development. The toxic effect in regard to **28F6** could be expected, because its possible metabolite *o*-aminoazotoluene is reasonably anticipated to be a human carcinogen.[86]
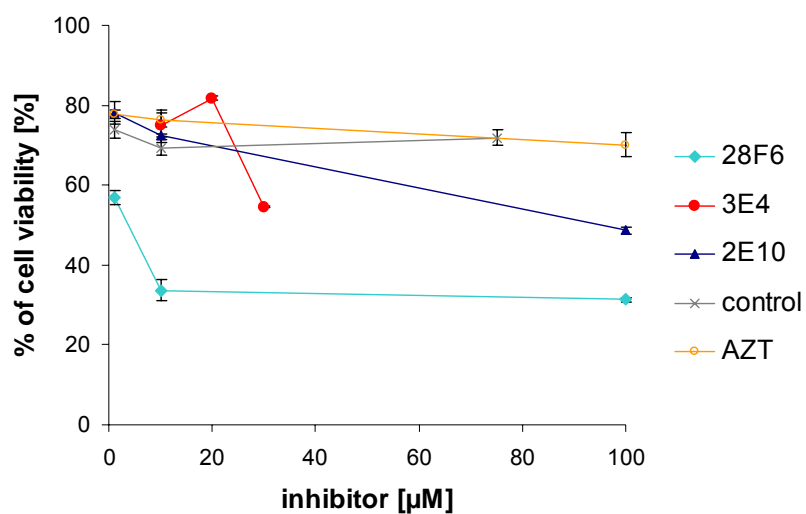
Figure 21. Cytotoxicity of small molecule inhibitors on TZM cells. To test for small molecule toxicity, 28F6, 3E4, 2E10, a negative control compound, and NRTI ATZ in concentrations ranging from 1-100 μM were added to TZM cells at 37° C for 24 hours, and the cell viabilities were monitored by MTT assay. Values are expressed as percentage of the values obtained from cultures without inhibitors. The symbols for each compound are listed in the right. The mean and the standard deviation of triplicates are given.

2.4. Discussion and Perspective

2.4.1. Design and General Application of Reporter Ribozymes to HTS

Despite the considerable number of functional allosteric ribozyme invention,[31, 87] and their great potentiality for multiple applications,[16, 48, 49, 88-92] there are only a few examples which have demonstrated applications of allosteric ribozymes to HTS formats.[45, 46] Hartig et al. screened antibiotics to identify HIV-1 Rev inhibitors using two rationally designed hammerhead allosteric ribozymes which can be regulated by Rev proteins. In one construct of the ribozyme, the anti-Rev aptamer sequence was inserted, and the conversely regulated ribozyme construct contains the Rev binding element (RBE).[46] Alternatively, using a combinatorial approach, Srinivasan et al. constructed an ADP-dependent allosteric ribozyme (RiboReporter) in a two step-*in vitro* selection, and applied it to screen protein kinase inhibitors.[45] Since these systems resulted in the identification of specific inhibitors, allosteric ribozymes should serve a useful basis to design multiple screening assays in a general way. The concept was further validated by the herein depicted work where employing an allosteric ribozyme (reporter ribozyme) as HTS assay format resulted in the identification of a highly promising HIV-1 RT inhibitor which suppresses HIV function both *in vitro* and *in vivo*.

The design of the screening assay was based on the reporter ribozyme FK-1 investigated by Hartig et al..[32] The characteristic features of FK-1 are the highly selective detection of HIV-1 RT and the domain specific sensing of HIV-1 RT binding partners. These important properties are required for the highly sensitive performance of HTS to search for specific small molecule inhibitors of HIV-1 RT. Successfully, the FK-1 ribozyme-based HTS assay was employed to identify three potential inhibitors (**28F6, 3E4, and 2E10**) of HIV-1 RT with *in vitro* inhibitory concentration that gave half-maximal activity (IC_{50}) of 2-5 μM on DNA-dependent DNA polymerase activity (Table 2).

However, the evaluation of the primary screen resulted in relatively frequent false positive hit compounds. The reliability of the assay could be increased by decreasing the concentration of library compounds (< 50 μM), and further optimization of assay conditions and data normalization.

Furthermore, the approach using a rationally engineered allosteric ribozymes (like FK-1) as the reporter basis holds the potential to be broadly applicable as screening format by constructing new fusion ribozymes. In addition to anti-RT aptamers, several aptamers targeting other components in the virus life cycles, e.g. protease, integrase, Tat,

and nucleocapsid, have been isolated via SELEX procedure.[68, 69] Some of them have shown effective suppression of HIV-1 replication *in vivo*. In principle, those aptamer sequences can be fused to ribozyme sequences in rational design approach to engineer new assay formats.

2.4.2. Functional Conversion from Aptamer into Small Molecule

One of the major tasks to use the assay based on the FK-1 ribozyme construct was to convert the aptamer inhibitory properties to the potential HIV-1 RT inhibitors. The pseudoknot aptamer sequence recognize the primer/template binding site of HIV-1 RT.[76] Endogenous expression of this aptamer by generating a tRNA^{Meti}/pseudoknot RNA chimera in the cell culture could suppress replication of HIV-1 by > 75%.[93] In another design of aptamer expression vector, in which the aptamer sequence was flanked by self-cleaving ribozymes, HIV infection and replication were strongly blocked again. In addition, replication of variety of drug resistant viruses as well as several subtypes of HIV were suppressed by this aptamer.[94] The results from different expression system support the pseudoknot aptamer as an efficacious agent in a gene therapy approach for anti-AIDS treatment. Like the pseudoknot aptamer, nucleic acid macromolecules such as ssDNA inhibitors seems to inhibit HIV-1 RT in the same way, that is to make extensive interacts with primer/template cleft of RT.[95] Considering those excellent inhibitory effect of nucleic acids aptamers, inhibitors that act with a primer/template competing mechanism might be able to suppress HIV superior to the mechanisms of NRTIs and NNRTIs, with respect to the drug resistance of RT, that always appears after drug administration.

Specific small molecule inhibitors that interact with the primer/template binding site are not known yet. (-)-Epigallocatechin gallate, which was used to demonstrate the utility of the FK-1 ribozyme-based assay (Figure 8), has been reported to be a HIV-1 RT inhibitor by Nakane et al..[79] The compound is suspected to inhibit DNA polymerase activity by interfering with binding of the primer/template, however, the detailed inhibitory mechanism and exact binding site are not yet elucidated. Later cell-based analysis of inhibitory effects of (-)-epigallocatechin gallate showed that anti-HIV activity of (-)-epigallocatechin gallate may result from an interaction with several steps in the HIV-1 life cycle.[96] Concerning only RT inhibition by (-)-epigallocatechin gallate, data from FK-1 ribozyme competition assay could support inhibitor binding to the primer/template site (Figure 8).

The pseudoknot aptamer can bind selectively to HIV-1 RT, with the closely related HIV-2 RT showing a binding affinity about 4 orders of magnitude lower.[75] In the *in vitro* selectivity assay of the identified compounds, **3E4** exhibited moderate selectivity to HIV-1 RT over HIV-2 RT (5-fold) with respect to DDDP inhibition (Figure 13, Table 4). While the other two inhibitors, **28F6** and **2E10,** gave approximately the same IC_{50} for HIV-1 and HIV-2 RTs. The factor of the differentiation obtained for **3E4** was not as efficient as that for the pseudoknot aptamer, however, the result is logically adequate, presuming the possible huge interface of RT interacting with nucleic acids macromolecules. Nevertheless, the inhibitory properties of the pseudoknot aptamer could be acceptably converted to the small molecule **3E4** in terms of protein selectivity. However, a compound that inhibits viral RTs more effective, even in a promiscuous way, than human enzymes holds great potential for further developments.

2.4.3. Inhibitory Mechanism of 3E4 on HIV-1 RT

Due to the discovered *in vitro* specific inhibition of HIV-1 RT accompanied with *in vivo* suppression of HIV replication, **3E4** justifies further characterization to understand its inhibitory mechanism.

SPR binding analysis to a dsDNA derivatized chip confirmed that HIV-1 RT was not inhibited by a mechanism, where **3E4** binds to the primer/template complex through intercalation or groove recognition (Figure 14). This results led to the hypothesis that **3E4** (also **28F6** and **2E10)** binds to HIV-1 RT or HIV-1 RT-primer/template ternary complex. Further characterization of **3E4** using HIV-1 RT derivatized chip in SPR supports the direct interaction between **3E4** and HIV-1 RT.

To know the mode of inhibition is of equal interest. HIV-1 RT recognizes both primer/template and dNTP as its substrates. The mode of inhibition of the inhibitors could be further analyzed by kinetic studies in the presence of saturated concentration of one substrate and different concentrations of the other substrate. Double reciprocal plots of initial velocity versus substrate concentrations will give information about the type of inhibition: competitive, non-competitive, or mixed type.[97, 98] The gel mobility shift assay could test if the compounds interfere with RT-primer/template complex formation.[99, 100] The most direct interpretation will be accessible from a crystal structure of the RT in complex with **3E4**.

3E4 has been investigated for inhibitory effect on DDDP and RDDP. In terms of the three enzymatic functions of HIV-1 RT, it is not known yet if **3E4** inhibits the RNase H

activity of HIV-1 RT. Specific inhibitors designed to target the RNase H domain of HIV-1 RT have been previously described.[101]

Probably the most distinctive feature of **3E4** is to inhibit DDDP preferentially to RDDP with 70-fold lower IC_{50} value. Most of the HIV-1 RT inhibitors have been supplemented with approximately equivalent IC_{50} values for both DDDP and RDDP or with the lower IC_{50} of RDDP than that of DDDP. However, the amount of information is insufficient to conclude that this observation is a general tendency. Additionally, only the IC_{50} values for RDDP have been determined for several inhibitors. The origin of **3E4** to specific inhibition of DDDP might derive from a novel inhibition mechanism. Crystal structures of HIV-1 RT in complex with DNA/DNA (primer/template)[71, 72] and DNA/RNA (primer/template)[102] have been determined. The authors who published the structure of HIV-1 RT-DNA/RNA complex made an effortful comparison with the RT-DNA/DNA complex. They observed very similar protein conformation as well as the nucleic acid helix conformation in both complexes. Slight differences were i) numerous interactions of RT with the 2'-OH groups of the RNA template in addition to sugar phosphate backbone interactions observed in common for both primer/template complexes and ii) increased involvement of the p51 subunit in binding the RNA/DNA relative to the DNA/DNA template/primer complex. It is difficult to interpret, however, **3E4** might recognize subtle structural difference between the HIV-1 RT-DNA/DNA and HIV-1 RT-DNA/RNA complexes, and inhibits only DNA polymerization derived from the HIV-1 RT-DNA/DNA complex. Another biochemical study concerning the effects of mutations in HIV-1 RT that affect residues contacting the template/primer provide some insight into this specialized property. Fisher et al. identified the N265D substitution, adjusting in the minor groove-binding track and making hydrogen bonds with the template strand (DNA and RNA), which led to a loss of processive polymerization on DNA templates but not on RNA templates.[103] This differential template usage was accompanied by a rapid dissociation of the N265D variant from DNA but not RNA templates. **3E4** might cause similar structural disruption to RT, with affecting DNA template/primer binding but without affecting RNA template/primer binding.

2.4.4. *in vivo* Inhibition Study of HIV by 3E4

The different performances among the identified inhibitors *in vitro* and *in vivo*, regarding specificity and toxicity, illustrates the importance of a secondary evaluation of

the hits resulting from the initial screening.

The cell-based assay employing the self-inactivating HIV vector system identified **3E4** as an active inhibitor in an cellular environment (Figure 18). The effect of **3E4** was further confirmed with the similar reduction of HIV-1 replication *in vivo* using a HIV wild type strain (Figure 19, Figure 20). The results from both assays fully correlate. Again, it is notably to say that the cellular activity of the pseudoknot aptamer was maintained in **3E4**.

While the anti HIV drugs targeting HIV-1 RT currently used in therapy can suppress viral replication to undetectable level, the emergence of drug resistant variants is still a unsolved problem. To test whether the drug resistant strains could be suppressed with **3E4**, further *in vivo* analysis is currently under examination.

As discussed above, **3E4** as a HIV-1 RT inhibitor may act at the primer/template binding site. Recently, Fisher et al. provided an interesting discussion concerning the emergence of resistance to template/primer competing inhibitors.[104] The authors discovered that the mutations caused by DNA aptamer treatment, that overlaps with the binding surface of the primer/template complex, result in replication defective HIV by affecting RT-primer/template interactions. In addition, the aptamer-resistant mutations retain wild-type susceptibilities to all NRTIs and NNRTIs tested *in vitro*.[105] If such a DNA aptamer-like (also pseudoknot aptamer-like) inhibitory mechanism would act in case of **3E4**, this compound would be highly promising for further evaluations. A small molecule drug like **3E4** providing aptamer-like functions will not only supersede burdensome procedures such like the preparation of expression vectors but also overcome the instability of nucleic acid-derived drugs under cellular conditions.

2.4.5. 3E4 as a Drug Lead Structure

It is not yet elucidated which stereoisomer of **3E4** is active in the assay performed. Both isomers of **4** are commercially available, thus in future both antipodes of **3E4** should be synthesized by the same strategy shown in Scheme 2. First it has to be elucidated whether the configuration at **3E4** has an effect on its activity. If one isomer turns out to be more active on HIV, the architecture of the corresponding chiral **3E4** structure can be considered as a lead compound to synthesize a number of derivatives. Comprehensive analysis using structure-activity relationship (SAR) is important for the optimization and prioritization of compound series. The required profile for **3E4** is to find derivatives with enhanced selective inhibitory effect on HIV-1 RT as well as to improve the

intrinsic poor solubility of **3E4**. Indeed, the low solubility of **3E4** was the major problem in the assays for *in vitro* and *in vivo* characterization with respect to the preparation of the samples or the administration to the cell culture. Specially the percent of DMSO, in which the compound is suspended, was reduced to 1 % in all the assays performed *in vivo*, compared to the 5 % DMSO used in the RT assays. It was confirmed that 5 % DMSO haves no interference to the activity of RT employing the *in vitro* RT assays. It implies that **3E4** might have a effect to suppress the replication of HIV-1 RT at less concentrations than those that observed. Additionally, cell toxicity has to be addressed in these endeavors.

Several anti-HIV inhibitors which share structural similarity with **3E4** are listed in Figure 22. By determining the structural nucleus within the diphenylurea group as observed similarly in Figure 12, the following three compounds have been selected. LY30046[106] was discovered by rational investigations to find the minimum molecular architecture of the tetrahydro imidazobenzo diazepinthiones (TIBO),[107] providing a new class of potent NNRTIs series derived from *N*-(2-phenylethyl)-*N'*-(2-thiazolyl)thiourea (PETT).[108] One of the PETT series inhibitor shows, besides high potency against HIV-1, marked activity against HIV-2 in a noncompetitive manner with respect to dGTP and poly(rC)/oligo(dG) as substrates, respectively. A novel molecular mode of action and mechanism of PETT series have been implied with interaction close to he active site but distal to the putative NNRTI-binding site by mutational mapping. Another RT inhibitor called KM-1 was reported as NNRTI that binds at a unconventional site to HIV-1 RT.[109, 110] Interestingly, KM-1 is believed to bind to RT in both the absence and presence of dsDNA but weakens the affinity for dsDNA so that it favors DNA dissociation. The data suggests that KM-1 distorts RT conformation and misaligns the dsDNA at the active site. One CCR antagonist, CCR is utilized for chemokine receptors in the initial stage of the infection, showed inhibitory activity of the HIV-1 envelope mediated membrane fusion during HIV-1 cell entry with extremely low IC_{50}.[111]

Taking the intriguing characteristics observed with **3E4** together with the possible speculation from the structurally similar compounds, **3E4** has been discovered as a novel type of potent inhibitor for anti-HIV therapy.
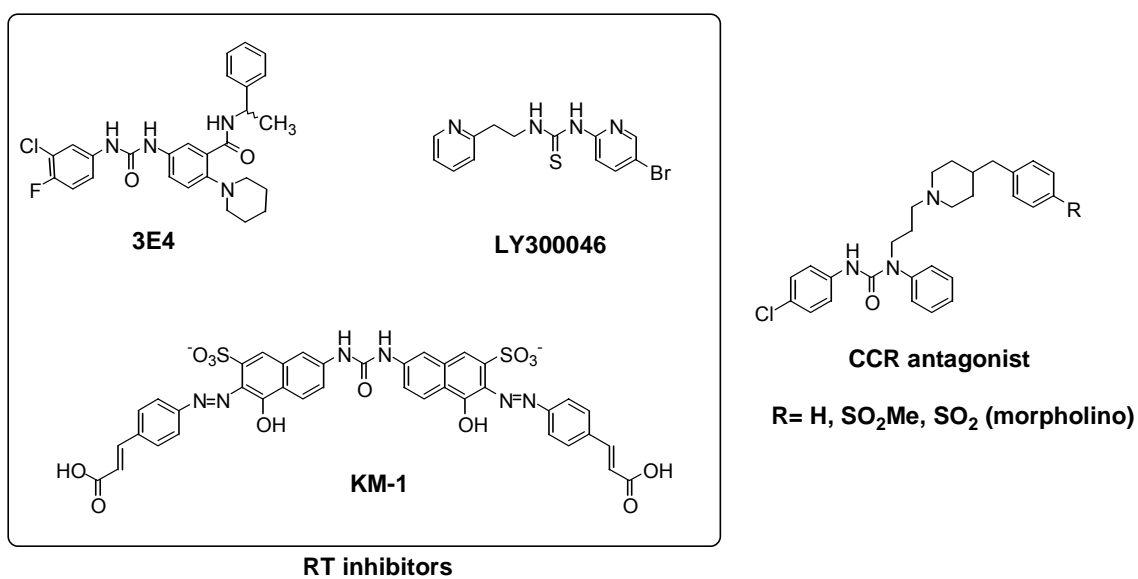
Figure 22. Structural comparison of **3E4** in other anti-HIV-1 drugs. LY300046[106] and KM-1[109, 110] of HIV-1 RT inhibitor, and one CCR antagonist[111] are shown.

3. *in vitro* Selection of Small Molecule Binding Proteins using mRNA-Display

3.1. Introduction

3.1.1. Protein Engineering by *in vitro* Selection

In the past decade, *in vitro* selection technologies have provided essential tools to evolve proteins (or peptides) with functional properties based on combinatorial chemistry. As discussed in chapter 2, engineering of proteins with new activities and rapid assignment of their functions are of fundamental interest from both biological and pharmaceutical point of views. Advances in molecular biology have made great enhancement to generate high complexity of protein variants at the DNA level. In general, the evolutional strategies applied consist of iterative rounds of selection and amplification to identify desirable members from the molecular ensembles (library). Since proteins can not be amplified directly like DNA, all selection approachs necessitate the linkage of genetic information to the phenotype (Figure 23).[5] To perform *in vitro* selection, the linkage should be tolerant to technical manipulations. The methods currently used for *in vitro* selections are either physical linkage or compartmentalization.[112-115] Specially, the techniques employing the generation of physical linkages are also referred as "display" systems.[116-119]
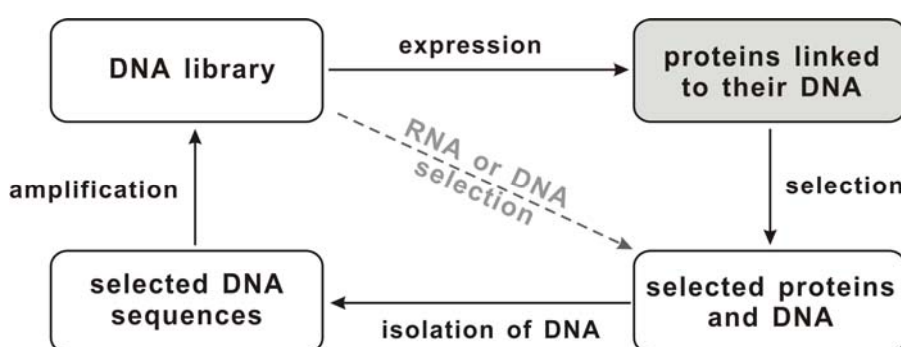
Figure 23. General selection strategy for isolation of functional proteins. Gray dotted line represents RNA or DNA selection. The figure is modified from Lin et al..[5]

Phage display is one of the first methods, that introduced indirectly the physical linking of a protein with its DNA sequence (Figure 24).[120-122] The encoded proteins are expressed on the surface of bacteriophage. A major drawback of this technology is the limitation of library size by the transformation efficiency of bacteria (approximately 10^{10}). Another approach described is ribosome display, that can be prepared totally *in vitro*, and reduce biases due to expression, resulting in a library complexity of $>10^{12}$ (Figure 24).[123-125] By eliminating the stop codon in the template mRNA, the ribosome stalls on the template after translation and generates ribosome-mRNA-peptide ternary complex. However, this ternary complex needs to be stabilized against dissociation by low temperature and high salt concentrations. The inherent instability of the complex makes it difficult to keep it intact during the selection steps and restricts the choice of conditions. A more robust method of physical linkage is to introduce a covalent bond. The modern technology called mRNA-display forms the covalent mRNA-protein conjugate by using puromycin, an antibiotic that mimics the aminoacyl end of tRNA (Figure 24). The mRNA-display system is apparently the most attractive for directed evolution among the above mentioned methods for a number of reasons. First, an expanded range of conditions can be exploited for the selection step. Second, unnatural amino acids can be incorporated directly into the proteins through suppressor codons. Finally, libraries of much greater complexity and size can be generated in the range of 10^{11}-10^{13}, depending on the transcription/translation scale.
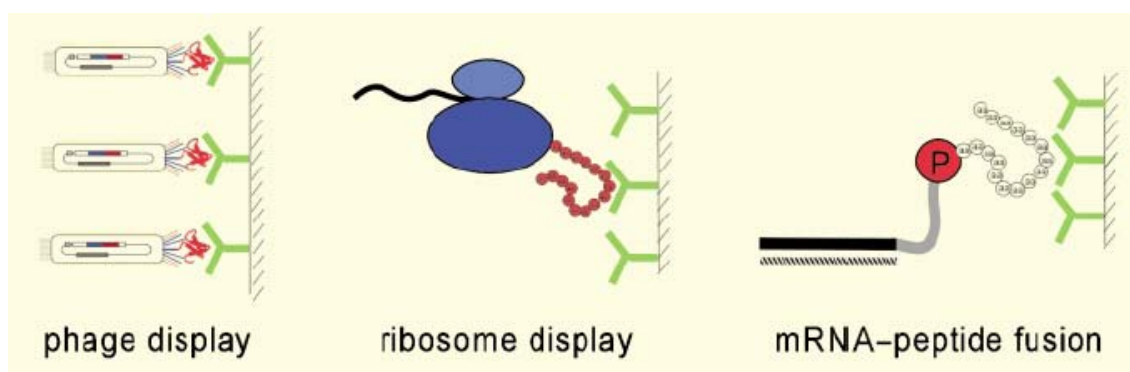
Figure 24. Illustration of representative methods for *in vitro* display technologies. The figure is reproduced from Lin et al..[5]

3.1.2. *in vitro* Selection using mRNA-Display

Initial papers on the principle of mRNA-display were published by Szostak[126] and Yanagawa[127] independently. A Selection using mRNA-display system provides an alternative method that can be applied to ligand discovery. In this approach, an encoded protein library is covalently fused to their own mRNA (Figure 25). mRNA-protein conjugates are able to be synthesized by using mRNA molecules that act as both template and protein acceptor on the surface of the ribosome. For this purpose, the mRNA contains a puromycin molecule at its 3' end. Puromycin is an antibiotics that serves as a chemically stable small molecule mimic of aminoacyl tRNA (Figure 25). Initially, the ribosome binds and transits the template in 5'→3' direction. When the ribosome reaches the 3' end of the mRNA template, puromycin enters the peptidyl transferase site to form a covalent mRNA-protein fusion.
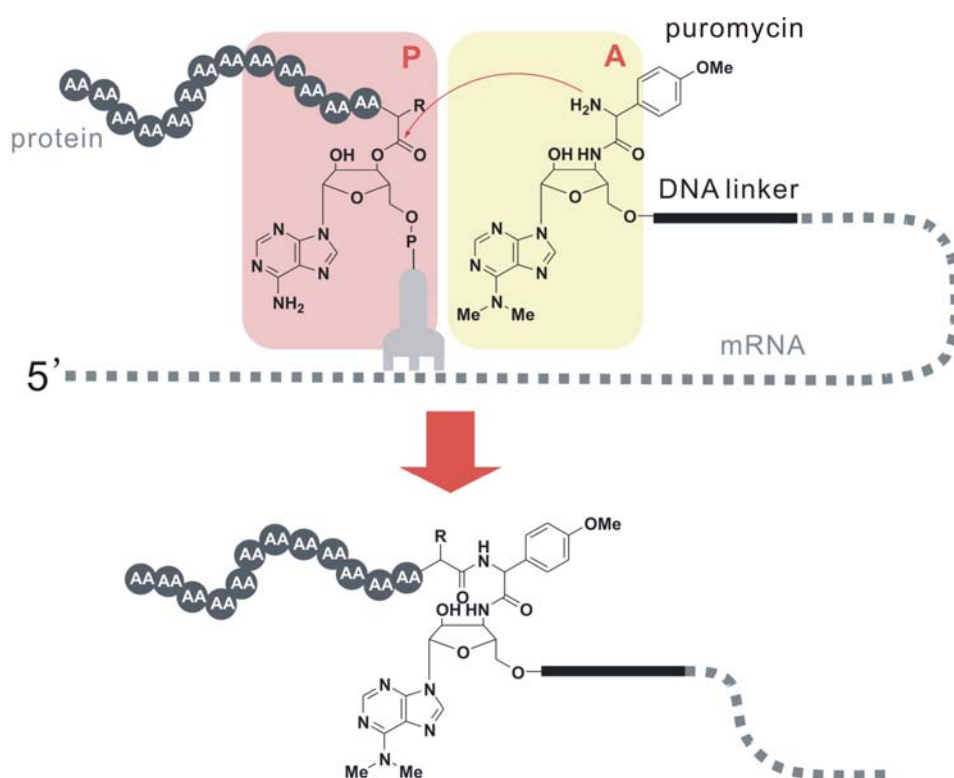
Figure 25. The formation of a mRNA-protein fusion. The ribosome reads the template in 5'→3' direction. tRNA (light gray) and the ribosome P- (pink) and A- (yellow) sites are shown. Puromycin enters the ribosome, attaching the template to the C-terminus of the nascent protein.

To perform *in vitro* selection cycles (Figure 26), a synthetic oligonucleotide containing a 3' puromycin is ligated to the 3' end of transcribed mRNA and the resulting product is

translated *in vitro* using rabbit reticulocyte lysate. The mRNA-protein fusions generated are then isolated by affinity chromatography based on the oligonucleotide tag and peptide tags. After reverse transcription, mRNA/DNA-protein fusions are selected against an immobilized antigen or ligand. Selected complexes can be PCR amplified to serve as the input material of next cycle. The functional proteins will be enriched in the fraction via successive cycles of selection and amplification.
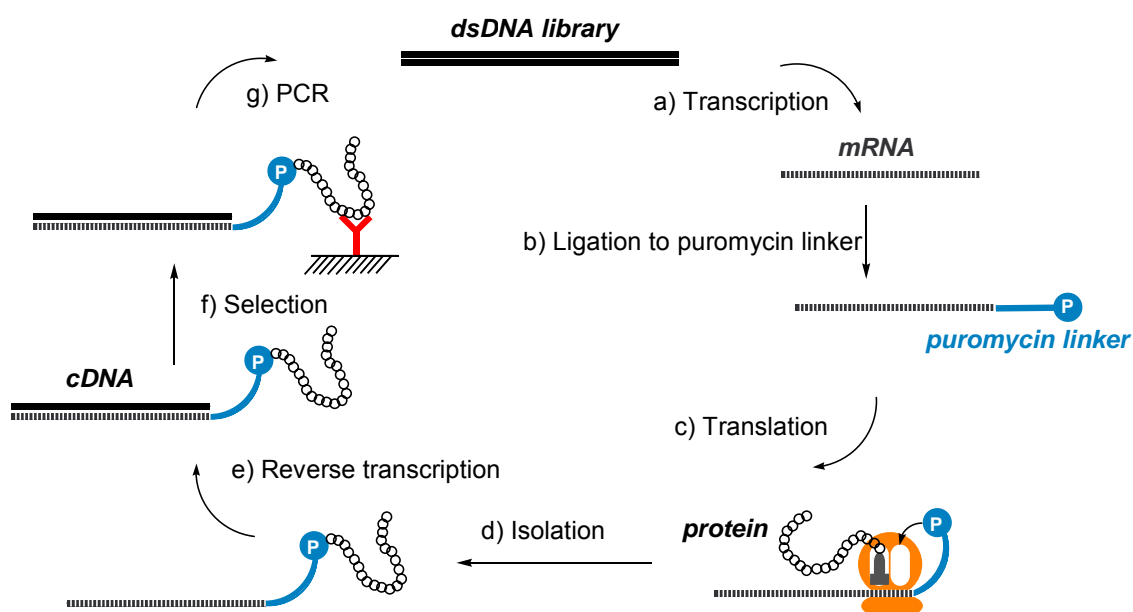
Figure 26. Schematic of mRNA-display *in vitro* selection cycle. a) A library of dsDNA (black line) is transcribed to generate mRNA (dotted line). b) The mRNA is ligated to a puromycin linker (blue). c) *In vitro* translation yields mRNA-displayed proteins. d) The complex is purified from the lysate mixtures using affinity chromatography. e) Reverse transcription generates mRNA/DNA double helix. f) The fusion library is sieved using the target of interest in the selection step. g) PCR amplification of selected molecules serves dsDNA sequences for input of the subsequent cycle.

3.1.3. Applications of mRNA-Display

Since the emergence of the mRNA-display system in 1997,[126, 127] a variety of new peptides and proteins have been discovered via *in vitro* selection. The libraries derived from mRNA-display technologies have demonstrated their ability to target versatile molecules including small molecule, nucleic acid, and protein. Keefe et al. evolved ATP binding proteins from the library containing 80 contiguous random amino acids, as simulating the evolution of primordial protein in the prebiotic era.[128] The selected

proteins had conserved CXXC sequence and functioned in zinc ion-dependent fashion. The authors estimated the frequency of occurrence of functional proteins in totally random sequence libraries, that is roughly one in 10^{11} which is similar to that observed with random RNA libraries. The peptide sequences capable of the recognition of a particular RNA loop motif called boxBR RNA have been identified from a library which has a randomized construct based on the corresponding RNA-binding domain of a λ N protein.[129] The *in vitro* selection resulted in peptides with high affinity (K_d = 0.5-5.0 nM) and better specificity than the wild-type sequence. The first protein as binding target for mRNA-displayed protein library was streptavidine reported by Wilson et al..[130] The isolated peptides contained a HPQ consensus sequence, showing nanomolar affinity. This strong and specific interaction permitted to apply them to use as peptide tag for the purification of recombinant proteins.[131] Raffler et al. exploited mRNA display to select human α-thrombin binders.[132] In addition, clot formation of human plasma was inhibited by the isolated peptides. Further approach has been performed to construct protein libraries engineered from an antibody mimic scaffold.[133] Three exposed loops of an immunoglobulin-like protein, the tenth fibronectin type III domain (^{10}Fn3), were randomized and used for a selection against tumor necrosis factor-α (TNF-α).

mRNA-display potentially facilitates proteomic analysis. Not only random sequence libraries but also cellular protein libraries can be generated by mRNA-display, and those cellular derived libraries have been selected for specific targets such as v-Abl tyrosine linase substrate,[134] anti-apoptotic protein Bcl-X$_L$,[135] and Ca^{2+} sensor calmodulin (CaM).[136] The sequence characterization of the identified proteins enables to elucidate cellular interaction partners and protein-protein interaction networks. The mentioned proteome scanning using mRNA-display resulted in all three experiments in the finding of both known and previously unknown binding partners.

Another application in the proteomic field is to establish addressable protein microarrays using mRNA-protein fusion. Weng et al. demonstrated that a standard DNA chip could be converted into a protein chip by hybridization of the informational content of the nucleic acid component of the mRNA-protein fusions.[137] The system was sensitive enough to detect sub-attomole quantities of the proteins.

mRNA-display technologies are even compatible with unnatural amino acids incorporation via amber codon suppression[138, 139] or four-base anticodon.[140] Furthermore, by reprogramming the universal genetic code for 20 naturally occurring amino acids, the unnatural peptide library containing different unnatural amino acids are able to be generated.[141] This technique greatly expands chemical and functional

diversity.

As discusssed above, various examples have demonstrated that mRNA-display is a powerful tool for both ligand discovery and interaction analysis. The major advantage of this technique is that manipulations occur *in vitro* starting from the generation of the library to selection, making it more tunable for desired investigations.

3.1.4. Coenzymes and Small Molecule Dyes as Selection Targets

Coenzymes are fundamental component in biology and are required for the function of many essential proteins common to all living organisms. At the same time, coenzyme-dependent proteins constitute a very large group of proteins that play central roles in the metabolism of the cell. The adenine binding motif shared among ATP-, coenzyme A (CoA)-, nicotinamide adenine dinucleotide (NAD)-, nicotinamide adenine dinucleotide phosphate (NADP)-, and flavine adenine dinucleotide (FAD)-dependent proteins indicates that these coenzyme-dependent proteins are likely to be "ancient proteins" in the evolutionary process.[142] The simple description for the requirement of common adenine-binding motifs are 1) backbone polar interactions that are not dependent on the protein sequence or particular properties of amino acid side chains, and 2) nonspecific hydrophobic interactions.

Proteins that bind dinucleotides (NAD and NADP, Figure 27) function pivotally in energy production, exchange and consumption. The classical dinucleotide binding fold is referred as Rossmann fold that contains βαβαβ unit.[143, 144] The core topology locates in the first 30-35 amino acids, so-called "fingerprint" region and is characteristic for 1) a phosphate binding consensus sequence GXGXXG (where X is any other amino acids), 2) six positions usually occupied by small hydrophobic amino acids, 3) a conserved negatively charged residue (Glu or Asp) at the end of the second β strand of the fingerprint, and 4) a conserved positively charged residue (Arg or Lys) at the beginning of the first β strand of the fingerprint. Although most of the dinucleotide binding proteins are thought to contain the classical Rossmann fold, still numerous nonclassical folding patterns, less observed than the classical fold, are used by proteins that bind NAD(P).

CoA is involved in various essential pathways, such as fatty-acid degradation and the citric acid cycle (Figure 27). Acetyl-CoA is a central intermediate in the metabolism of almost all kinds of biological compounds, including amino acids, carbohydrates and fats. On the contrary to dinucleotide binding fold, the CoA-dependent proteins cover a wide

variety of different topologies, including Rossmann fold, TIM barrel, or four helical up-and-down bundle.[145, 146]

Cibacron blue 3GA is a water-soluble triazine dye that binds to the nucleotide binding fold of many proteins. The similarity of its three dimensional structure with NAD(P) enables to purify NAD(P) binding proteins on affinity chromatography using immobilized cibacron blue 3GA[147, 148] as well as to analyze the structures of those proteins by co-crystallization.[149, 150]

In vitro selections of protein libraries which target coenzyme binding allow comparison with known biological coenzyme-binding motifs and might bring insight to understand protein evolution, e.g. as proteins have evolved from relatively random sequence scrambles in the ancient period.
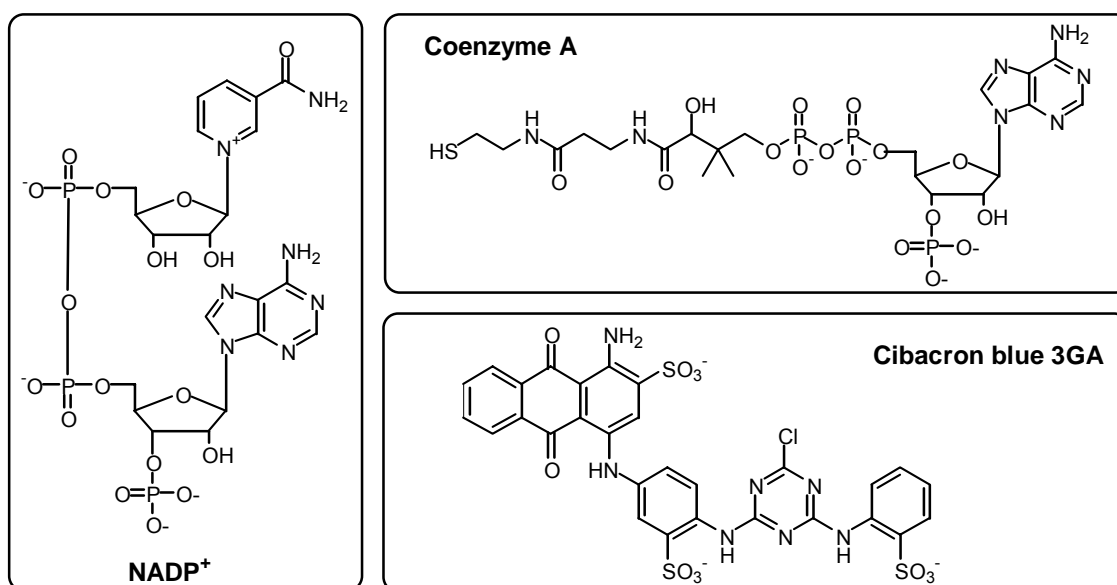
Figure 27. Structures of target molecules used for *in vitro* selections.

3.2. Aims of Study

To isolate functional proteins with desired properties, a number of methods are available in current days. As discussed in chapter 3.1, mRNA-display has turned out to be an improved generic strategy for *in vitro* selection and holds extensive potential for future applications. The aim of this study is to establish a *in vitro* selection scheme using mRNA-display to identify the proteins that can bind to small molecule targets. There is only one report which has showed the *in vitro* selection against the small molecule ATP, compared to numerous reports that describe protein targets. To enhance the possibility for small molecule recognition with folding stability, the design of secondary structural patterning such as α-helix or β-strand should be introduced to the library construction. Coenzymes (NADP and coenzyme A) have been chosen as small molecule targets, because following sequence analysis of isolated proteins allows comparison to known coenzyme binding motifs. Additionally, it is of interest to investigate which type of folding or motif can be dominantly evolved from the random sequence library. The results might provide insight to understand protein evolution in ancient era. Cibacron blue 3GA have been chosen, because it is known that the dye can interact with variety of proteins by mimicking the NAD(P) structure.

3.3. Results

3.3.1. Construction of a DNA Library of long ORFs for mRNA-Display

Chemically synthesized DNA library to construct randomized protein library is always accompanied with the problem of deletions at the DNA level due to imperfect coupling and capping efficiency during solid phase DNA synthesis. These cause frameshifts and undesired stop codons during translation, resulting in premature termination at protein library level. Accordingly, the number of intended sequences in the protein library will be greatly reduced. To address these problems, a DNA library encoding polypeptides in total length of 117 amino acids for mRNA-display was constructed using a strategy reported by Cho et al..[151] As described in Figure 28, the final DNA library consists of four short cassettes. Each cassette is designed to encode constant regions (T7, TMV, and FLAG in 5' as well as His6-tag and linker in 3') and a 22 amino acids long random region. The random region is flanked by two different restriction sites, BbsI and BbvI. Using the short DNA cassettes which have been enriched for those sequences lacking stop codons and frameshifts through protein tag purifications, the assembly proceeded by dividing the dsDNA library into two aliquots, one of which was cut with BbsI and the other with BbvI. The purified fragments were then ligated together with T4 DNA ligase. Two successive restrictions and ligations yielded the final DNA library in which the random region is 88 amino acids long.

Furthermore, not only a totally random (unpatterned: R) cassette but patterned cassettes (α, β), in which polar and nonpolar residues were arranged so as to have peptide domains with propensity for forming amphiphatic secondary structures like α-helices and β-strands, were incorporated into a final DNA library. Thus, 36% of the incorporated cassettes encode α-helices; 46% were β-strands; 18% were unpatterned random sequences in the full length DNA library. These patterning increases the fractions in the library with the ability of constituting a well-defined folding pattern.
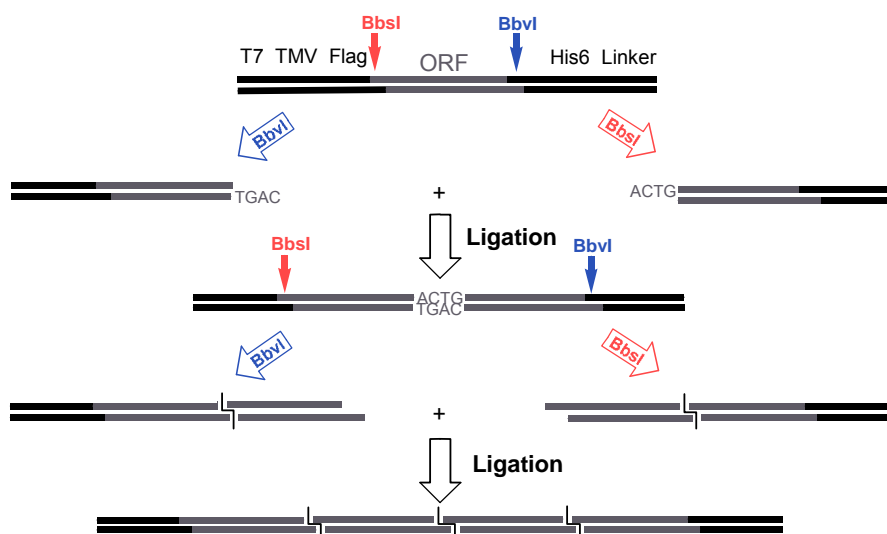
Figure 28. The strategy to construct a DNA library of long ORFs for mRNA-display.[151] Four DNA short cassettes are assembled into the full-length library. BbsI and BbvI are the restriction enzymes, which divide each cassette into two aliquots by different restriction sites. The cleavage sites are shown with arrows. After two successive ligation steps the flanking regions for the constant sequences remain the same. T7: T7 promoter sequence, TMV: tobacco mosaic virus translation enhancer sequence, Flag: Flag-tag sequence, His6: His6-tag sequence, Linker: linker annealing site.

The final full length DNA library resulted from the assembly of four cassettes was designed to encode a 88 amino acids long random sequence and several constant regions at both ends (Figure 29). At the 5'-end, the transcription promoter and translation enhancer sequences were located. The random sequence is flanked by two peptide tags for the purification of mRNA-displayed proteins: FLAG-tag at the N-terminal and His6-tag at the C-terminal position. A oligonucleotide sequence complementary to a linker used for photoligation[152, 153] was included at 3'-end. This DNA library was used to prepare mRNA-displayed protein library to perform *in vitro* selection against small molecule targets as shown in the following chapters.

Figure 29. DNA library design for mRNA-display. Untranslated sequences are in gray color and translated sequences are in white and black. White boxes encode constant regions and black box encodes random amino acids sequence (88-aa). TMV: tobacco mosaic virus translation enhancer sequence.

3.3.2. Generation of m-RNA-Displayed Protein Library

First, the DNA library was transcribed by T7 RNA polymerase and thereafter a photolinker oligonucleotide bearing the puromycin moiety was ligated at its 3' end. The mRNA linked to puromycin was translated *in vitro*, yielding the proteins displaying their mRNA via the puromycin linkage. The resulting mRNA-protein fusion library was then purified on oligo-dT cellulose and peptide tag affinity chromatography, and reverse transcribed. One of the affinity chromatography Ni-NTA agarose captures the proteins expressing a His6-tag at the C-terminus, so that the mRNA-displayed proteins that include internal stop codons or frameshifts can be excluded. Optionally anti-FLAG affinity chromatography was introduced to avoid proteins translated by internal initiation. The generation of RNA/DNA heteroduplex by reverse transcription prevented enrichment of RNA aptamers. The formation of mRNA-protein fusions and reverse transcribed mRNA/DNA-protein fusions was monitored by SDS gel analysis during the cycles (Figure 30). The library used for the *in vitro* selections consisted of 1×10^{11} - 5×10^{12} different displayed proteins (for details concerning affinity chromatography and the input diversity; see the following sections).
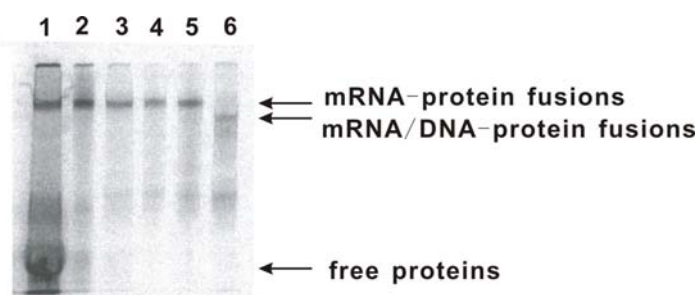
Figure 30. SDS gel analysis of the formation of mRNA-protein fusions. The loaded samples are as follows; lane 1: after *in vitro* translation, lane 2: elution from oligo-dT affinity chromatography, lane 3: elution from anti-FLAG affinity chromatography, lane 4: elution from Ni-NTA affinity chromatography, lane 5: before reverse transcription, lane 6: after reverse transcription.

3.3.3. Selection against Cibacron Blue 3GA

The random sequence library generated with 1×10^{11} different proteins was subjected to iterative cycles of *in vitro* selection and amplification. In each cycle the mRNA-displayed proteins were incubated with immobilized cibacron blue 3GA, washed and eluted with free NAD. The eluted fractions were collected and amplified by

polymerase chain reaction (PCR), and the resulting DNA was then used to generate new mRNA-displayed proteins for input into the next cycle of selection (Figure 26). The operations for each purification steps during the cycles are listed in

Table 5. After nine cycles, raising percentage (30 %) of the mRNA-displayed proteins that retain on the column after elution was observed (

Table 5, Figure 31). However, the fractions eluted with NAD remained at low amounts of the proteins even after nine cycles.

Cycle number	1	2	3	4	5	6	7	8	9
Oligo(dT)	○	○	○	○	○	○	○	○	○
FLAG	○	○	○	○	○	○	○	○	○
Ni-NTA	○	○	○	○	○	○	○	○	○
Elution [%]	10.4	0.28	0.25	0.64	1.59	0.54	0.53	0.54	0.68
Column [%]	37.7	27.6	8.3	4.5	20.7	15.0	28.1	24.0	31.0
PCR cycle	20	22	22	16	13	13	17	10	11

Table 5. The course of each cycle for cibacron blue 3GA selection using mRNA-displayed proteins. ○: the purification step performed, -: not performed. The cycle numbers required for PCR amplification after selection are depicted.

Figure 31. Progress of the selection against Cibacron blue 3GA. The bars represent the percentage of radioactivity that eluted from the column matrix (white) and remained on the column (dark gray).

Sixty-nine randomly chosen clones from the PCR product from ninth cycle were sequenced. Surprisingly 42 % of the analyzed sequences were frame-shifted from the intended frame (Figure 32). There were also sequences carrying internal stop codons (17 %), and even sequences which are composed of less or more numbers of cassettes than four (9 %). 32 % of the intact full length sequences were observed. In Figure 33, the percent of the appearance of the cassettes' type is depicted. The most frequently identified type of cassette was αβαβ. The alignment of full length sequences resulted in three sequences in different cassette' type which appeared to be duplicates with one point mutation. Other sequences exhibited to be broadly diverse. No conserved amino acids sequence motif was observed within these sequences.
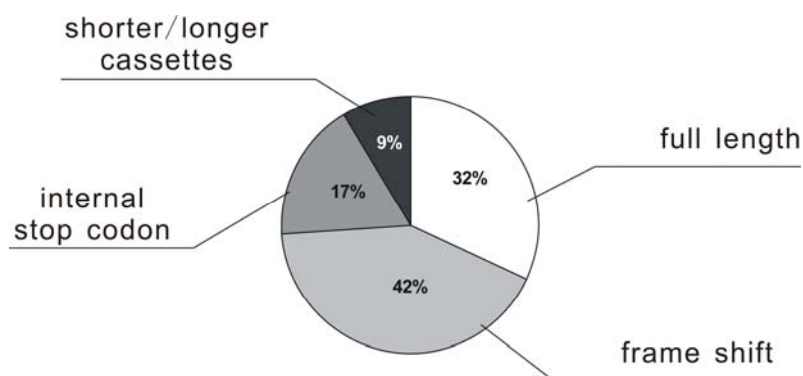
Figure 32. Analysis of cloned sequences from the elution fraction in ninth cycle.
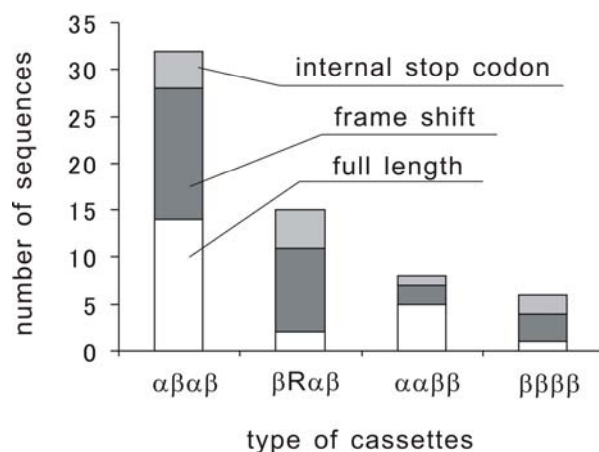
Figure 33. Percent of appeared sequences depending on the type of cassettes.

To test whether the protein fractions remained on the column are nonspecific matrix binder, the mRNA-displayed proteins from ninth cycle were incubated on cibacron blue 3GA agarose and then eluted from the column competitively by different elution conditions (Figure 34). By raising the NAD concentration to 50 mM (10-fold) in elution condition, ca. 3.0 % of the proteins were able to be eluted from the matrix. This DNA sequences were amplified by PCR and used for the subsequent selection (cycle 10). 50 mM of NAD eluted again the approximately same percentage of the proteins (3.2 %) in the tenth cycle, which implies no enrichment of NAD binders. While elution conditions including 5 mM of cibacron blue 3GA could competitively elute most of the protein fractions stayed on the column (ca. 30%). These results indicate that the proteins retained on the column are no matrix binders, and the operated selection conditions seemed to force the protein library to bind more tightly to cibacron blue 3GA than

NAD.
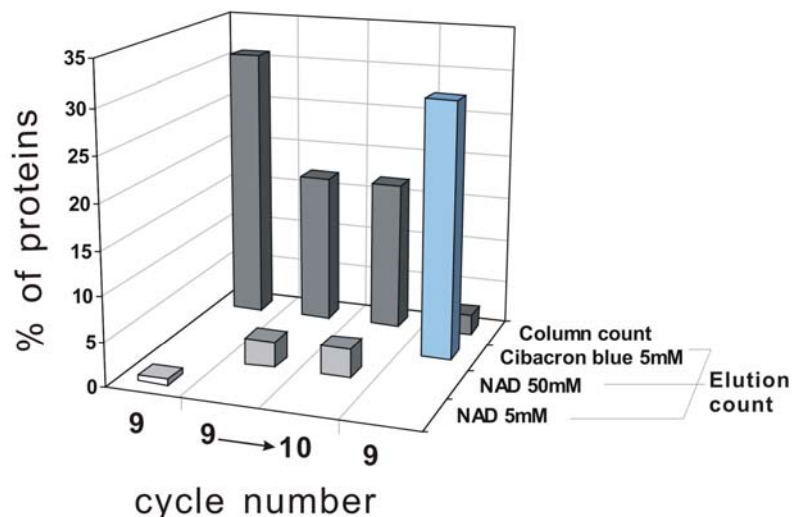
Figure 34. Different elution experiment with the mRNA-displayed proteins from ninth cycle.

3.3.4. Selection against NADP

The mRNA-displayed protein library which was composed of 5.0×10^{12} different sequences was generated for the input in the first selection cycle. This library was incubated with immobilized NADP, washed and eluted by free NADP. PCR amplification of these molecules yielded the starting material for the next round of selection. In the first cycle, a preselection against nonderivatized matrix (adipic acid hydrazide agarose) was introduced to eliminate possible matrix binders. The purification step using anti-FLAG-tag affinity chromatography was omitted. The details about the course of consecutive cycles are described in Table 6. As shown in Figure 35, no significant increase of protein amounts in both elution fractions and fractions retained on the column was observed even after seven cycles. The obtained result suggests that no enrichment of NADP binders was obtained.

Cycle number	1	2	3	4	5	6	7
Oligo(dT)	○	○	○	○	○	○	○
FLAG	-	-	-	-	-	-	-
Ni-NTA	○	○	○	○	○	○	○
Elution [%]	0.77	2.21	0.10	0.30	0.56	0.23	0.42
Column [%]	0.38	2.76	0.17	0.21	0.19	0.18	0.20
PCR cycle	21	20	19	15	17	18	17

Table 6. The course of each cycle for NADP selection using mRNA-displayed proteins. ○: the purification step performed, -: not performed. The cycle numbers required for PCR amplification after selection are depicted.
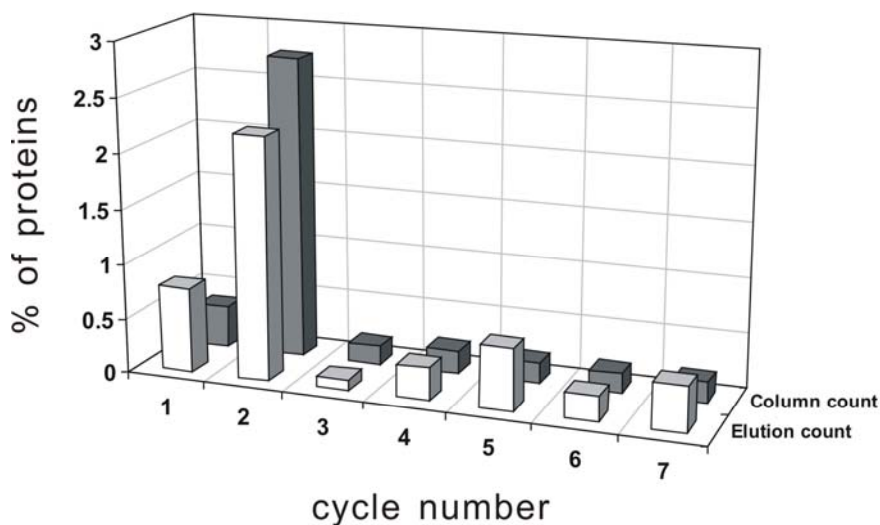
Figure 35. Progress of the selection against NADP. The bars represent the percentage of radioactivity that eluted from the column matrix (white) and remained on the column (dark gray).

3.3.5. Selection against CoA

For the selection targeting CoA, the matrix derivatized with CoA was prepared from thiopropyl sepharose as described in Figure 36. The coupling of CoA via disulphide bond enables non-competitive but specific elution of the bound proteins by covalent bond cleavage with DTT treatment. This would allow to enrich tight binders which have

low dissociation constants.
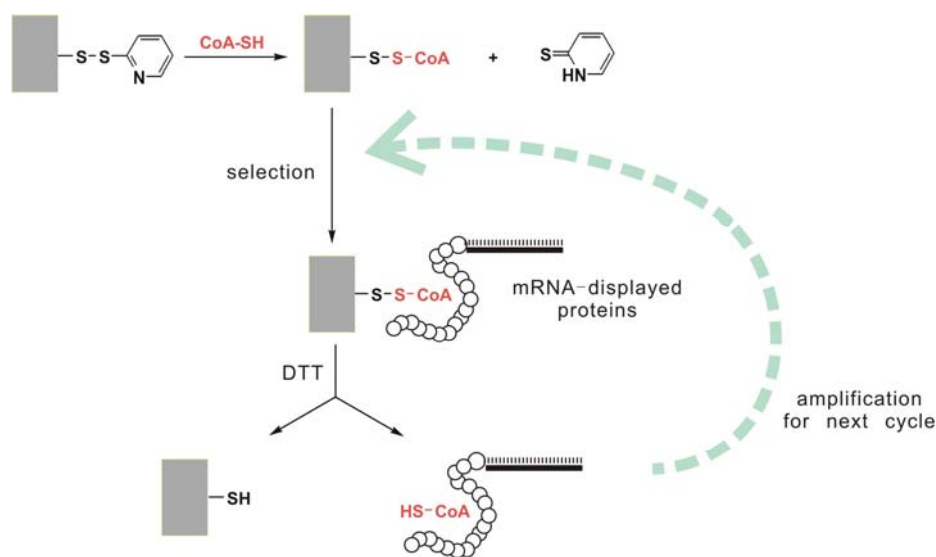
Figure 36. Affinity column preparation and selection strategy targeting CoA.

Random sequence protein library generally tends to have a intrinsic poor solubility and low potentiality for stable folding. To improve such properties of mRNA-displayed proteins, a selection strategy was designed by the addition of a small amount of denaturant in the selection condition. In comparison to phage display or ribosome display, mRNA-display is uniquely suited for this purpose because it is an entirely *in vitro* selection system based on a covalent linkage between translated proteins and their encoding mRNA. Chaput et al. re-selected ATP binding proteins, which have been isolated by using mRNA-display, to optimize protein folding stability in the presence of increasing concentrations of guanidine hydrochloride.[154] One selected sequence has shown a unique folded structure by circular dichroism, tryptophan fluorescence, and ^{1}H-^{15}N correlation NMR studies. In the following selection against CoA, two selections are performed in parallel, one in physiological condition and the other in slightly denaturing condition, to examine the effects of the selection pressure for better folding using the same target molecule.

Starting from 1.2×10^{12} unique protein sequences, several successive cycles of *in vitro* selection and amplification were performed under the respective condition. For the selection, mRNA-displayed proteins were incubated with CoA immobilized sepharose in the absence of GuHCl (physiological condition) and in the presence of 0.5 M GuHCl (slightly denaturing condition). After washing each column with the respective buffer,

the bound proteins were recovered specifically by liberating the CoA moiety from the matrix with DTT. The encoding DNA sequences were PCR amplified and used to generate a new library of mRNA-displayed proteins for input into the next selection cycle. Anti-FLAG affinity purification was skipped and after the sixth cycle under natural condition and fifth under denaturing condition, Ni-NTA purification was omitted to reduce any stringency. The details about the course of cycles are described in Table 7. As shown in Figure 37, no significant increase of protein amounts in elution fractions was observed under both physiological condition and slightly denaturing condition until the eighth and seventh cycle, respectively. The results suggest that no enrichment of CoA binders was obtained in both selections.

A Cycle number	1	2	3	4	5	6	7	8
Oligo(dT)	○	○	○	○	○	○	○	○
FLAG	-	-	-	-	-	-	-	-
Ni-NTA	-	○	○	○	○	-	-	-
Elution [%]	0.04	1.84	3.11	0.33	0.27	0.55	0.34	0.57
Column [%]	0.11	0.43	0.22	0.13	0.13	0.07	0.04	0.05
PCR cycle	20	24	25	22	22	17	16	20

B Cycle number	1	2	3	4	5	6	7
Oligo(dT)	○	○	○	○	○	○	○
FLAG	-	-	-	-	-	-	-
Ni-NTA	-	○	○	○	-	-	-
Elution [%]	0.04	0.37	0.36	0.24	0.55	0.28	0.44
Column [%]	0.10	0.21	0.10	0.08	0.08	0.05	0.04
PCR cycle	20	25	22	26	17	16	18

Table 7. The course of each cycle for Coenzyme A selection using mRNA-displayed proteins under physiological condition (**A**) and slightly denaturing condition (**B**). ○: the purification step performed, -: not performed. The cycle numbers required for PCR amplification after selection are depicted.
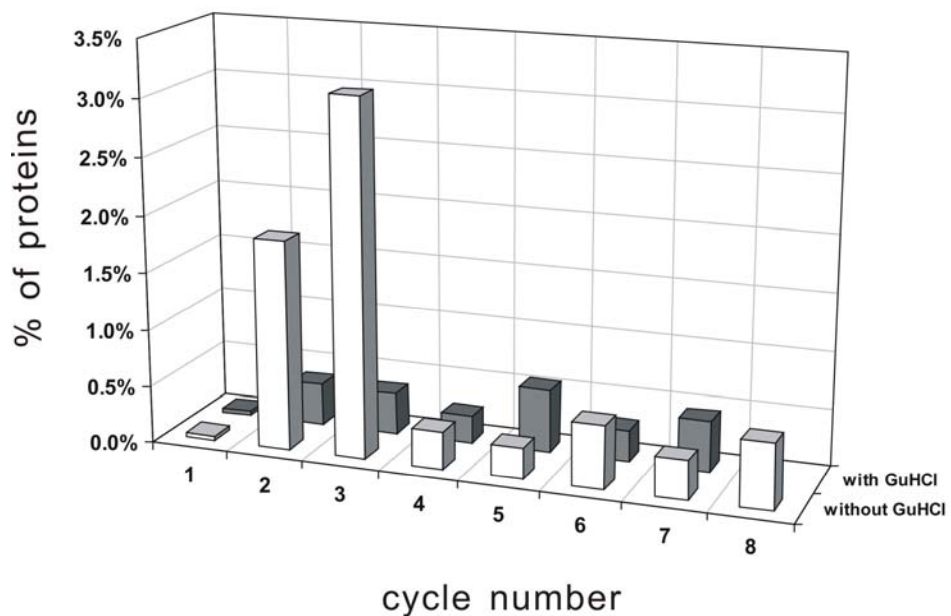
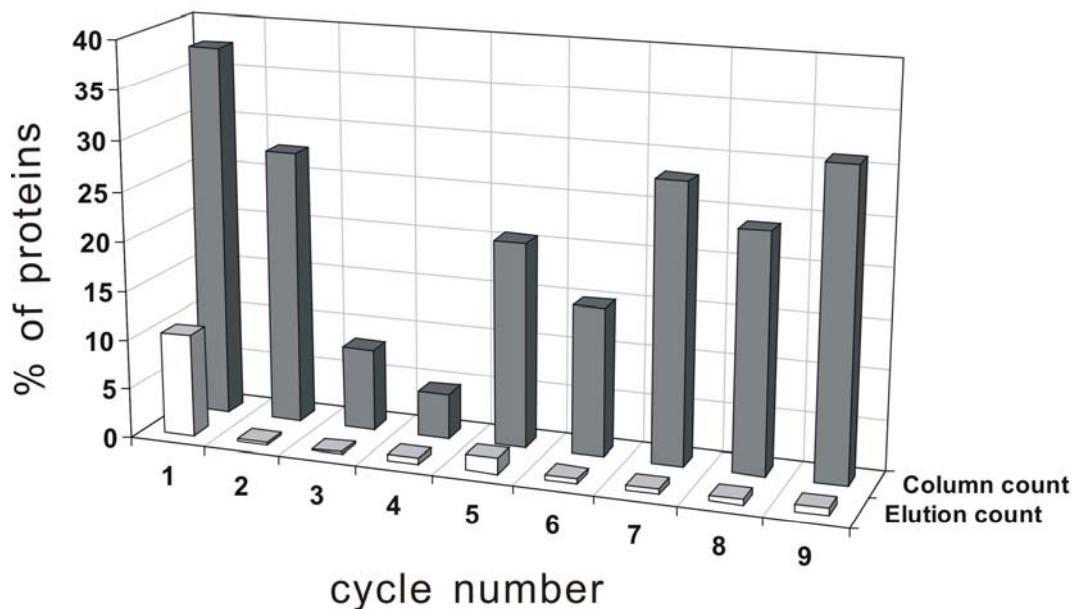**Figure 37.** Progress of the selection against Coenzyme A. The bars represent the percentage of radioactivity that eluted from the column matrix under physiological condition (white) and slightly denaturing condition (dark gray).

3.4. Discussion and Perspective

Three independent *in vitro* selections using mRNA-display have been performed to target coenzymes (NADP and CoA) and a small dye molecule (cibacron blue 3GA). The cibacron blue 3GA selection resulted in the enrichment of specific binders. However, the isolated sequences are not yet fully characterized. The sequence analysis of the fraction eluted by NAD from the ninth cycle showed unexpectedly a few numbers of full length proteins and abundant proteins derived from unintended reading frames. This argues the quality and the necessity of peptide-tag purification of the library before the selection process. A similar observation was described by Wilson et al..[130] The authors identified streptavidin binders from the patterned sequence protein library consisting of 88 amino acids. They also applied the purification step using Ni-NTA affinity chromatography to exclude such frameshifted products. However, the sequence analysis of the selected proteins turned out that their initial design of pattern of polar and nonpolar residues was discarded by frameshifts, leaving an essentially random sequence. It was mentioned that this was caused due to the higher frequency of the conserved amino acids sequence motif (HPQ) in the shifted open reading frame. This finding leads to the idea to align the majority of frameshifted sequences to search for a new consensus motifs required for target binding.

As discussed in §3.3.3, the alignment of full length sequences cloned from the ninth cycle by NAD elution showed no invariant amino acids motif and quite diverse sequences. However, the selection yielded the most frequent type of cassette for αβαβ, that include the part of important folding βαβ where the fingerprint region locates in the Rossmann fold. To characterize the cibacron blue 3GA specific binders, the fraction eluted by cibacron blue 3GA competitively (Figure 34) should be cloned and sequenced. This will provide a more detailed analysis of enriched proteins.

No enrichment of binding proteins was observed with two selections against NADP and CoA unfortunately. Considering the number (1 out of 10^{11}) for the frequency of functional proteins estimated for ATP binder,[128] the input diversity for each selection were in the same range like most other successful selections done. However, selection against ATP using mRNA-display remains the single report for selection against a small molecule target. In general, small molecule recognition by protein requires well defined binding pockets. The incidence of stably folded structures from a totally random sequence protein assembly can be easily imagined to be very low. In fact, Chaput et al. needed to improve the poor solubility and folding stability of ATP binders by re-selection under the denaturing condition.[154] Taking together, the rate of occurrence

estimated for the ATP binders might have been exaggerated for small molecule targets. Alternatively the problems arising from intrinsic poor solubility and folding instability might be addressed by the construction of scaffolded libraries. The lipocalin family[155] can be a considerably attractive candidate for such an endeavor. Lipocalins constitute a structural family of small functionally diverse proteins comprising 160-180 residues, with conserved β-barrels as their central folding motif. They exhibit great functional diversity including retinol transport, invertebrate cryptic coloration, olfaction and pheromone transport, and prostaglandin synthesis. The loop regions of the lipocalin family are remarkably dissimilar in length and conformation, reflecting the differing of shapes and chemical properties of their cognate ligands. Thus, the lipocalins show molecular recognition with the ability to bind a wide range of ligands including small hydrophobic molecules, specific cell-surface receptors, and macromolecules. The potential of small molecule binding by the lipocalins scaffold has been demonstrated by Beste et al..[156] The authors engineered the lipocalin protein by randomization in the loop region to resemble immunoglobulin. The resulting scaffold library was selected for fluorescein binding using phage display. Since it has been demonstrated with phage display, the lipocalin scaffold might be able to expand the ligand variety and property for usage in *in vitro* selections by mRNA-display due to stable folding.

4. Materials and Methods

4.1. Materials

4.1.1. Equipment

Equipment; type	Manufacturer
Agarose gel chamber	Cosmo Bio
Autoclave	Systec
Automated liquid handling system; Freedom evo	Tecan
Balance; JL-200,	Chyo
BL1500S; BP 211D	Sartorius
Centrifuge	Eppendorf
Compact reagent grade water system, Easy pure UV/UF	Barnstead
Electrophoresis power supply	Consort
EI-MS; Finnigan MAT 95XL	Thermo Finnigan
FAB-MS; Concept 1H	Kratos
Fluoroskan Ascent FL	Labsystem
Geiger-Counter; LB122	Berthold
Gel documentation system	Bio-Rad
Gel dryer	Bio-Rad
HPLC; 1100 series	Agilent technologies
Incubator shaker; 4430	Innova
Magnetic stirrer	IKA
Micropipettes	Eppendorf
Microwave	Bosch
NMR; DPX300, DPX400, DPX500	Bruker
Nucleic acid sequencing Electrophoresis apparatus	Bio-Rad
Oven	WTBbinder
PCR apparatus	Eppendorf / Biometra
pH meter	Inolab
PhosphorImager; FLA-3000	Fujifilm

Photometer; BioMate3	ThermoSpectronic
Scintillation counter; winSpectral 1414	PerkinElmer
Speed-Vac; Concentrator 5301	Eppendorf
Surface plasmon resonance (SPR); Biacore 3000	Biacore
Thermo block	Stuart Scientific
Thermo mixer	Eppendorf
UV transilluminator	Bachofer
UV/VIS spectrometer; Lambda 2S	PerkinElmer
Vertical electrophoresis apparatus	Bio-Rad
Vortexer	Neolab
Water bath	GFL

4.1.2. Consumable

Product	**Manufacturer**
Centrifuge tube (15 ml, 50 ml)	Greiner
Chromatography paper 3MM	Whatman
Disposable filter (0.2 μm)	QualiLab
Disposable syringe	Henke Sass Wolf / Braun
Glass wool	Serva
Micropipette tip	Peske / Biozym
MicroSpin column (Sephadex G-25)	Amersham
NAP-column (NAP-5, 10, 25)	Pharmacia Biotech
PolyPrep chromatography column	Bio-Rad
Prepacked micro spin column, P-30	Bio-Rad
Reaction tube	Eppendorf
Sensor chip (CM5, SA)	Biacore
Steril filteration units	Millipore
Steril surgical blade	Bayha
Ultrafree-MC centrifugal filter device	Millipore
96-well plate	Costar
384-well plate	Greiner

4.1.3. Chemicals

The water used for all biochemical experiment was deionized by compact reagent grade water system (Easy pure UV/UF), treated with diethylpyrocarbonate (DEPC), then autoclaved.

Chemical	**Manufacturer**
Acetic acid	Merck
Agarose	Invitrogen
Agarose low melting point	Sigma
Ammonium acetate	Merck
Ammonium peroxydisulfate (APS)	Roth
Ammonium sulfate	Merck
Ampiciline	Sigma
Boric acid	KMF
5-Bromo-4-chloro-3-indolyl-β-galactopyranoside (D-Gal)	Sigma
Bromophenol blue	Merck
Chloroform	Kraemer & Martin
Cibacron blue 3GA	Sigma
Coenzyme A (CoA)	Sigma
Coomasssie brilliant blue G-250	Merck
Diethyl pyrocarbonate (DEPC)	Sigma
Dimethylchlorosilane	Fluka
Dimethylformamide (DMF)	Fluka
Dimethylsulfoxide (DMSO)	Fluka
1,4-Dithiothreitol (DTT)	Acros
Ethanol	Merck
Ethidium bromide	Roth
Ethylenediaminetetraacetic acid (EDTA)	Merck
Formamide	Fluka
Glycerol	AppliChem
Glycine	AppliChem
Glycogen	Roche
Guanidine hydrochloride	Sigma
HEPES	Roth

Hydrochloric acid min. 37 %	Riedel-deHaean
Imidazole	Fluka
Isopropanol	Fisher Scientific
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Roth
Magnesium chloride	Fluka
β-Mercaptoethanol	Sigma
Methanol	Fluka
β-Nicotinamide adenine dinucleotide (NAD)	ICN
β-Nicotinamide adenine dinucleotide phosphate sodium salt (NADP)	Fluka
Potassium chloride	Acros
Potassium dihydrogen phosphate (KH_2PO_4)	Merck
Potassium hydroxide	Acros
Roti-phenol	Roth
Rotiphorese Gel 30	Roth
Rotiphorese sequencing gel concentrate	Roth
Sodium acetate	Merck
Sodium chloride	KMF
Sodium dihydrogen phosphate dehydrate ($NaH_2PO_4 \cdot 2H_2O$)	Riedel-deHaean
Sodium dodecyl sulfate (SDS)	Roth
Disodium hydrogen phosphate dihydrate ($Na_2HPO_4 \cdot 2H_2O$)	Merck
Sodium hydroxide	KMF
Spermidine	Sigma
N,N,N',N'-Tetramethylendiamine (TEMED)	Merck
Tricin	AppliChem
Tris-(hydroxymethyl)-aminomethane (Tris)	Roth
Triton X-100	Merck
Urea	AppliChem
Xylen cyanol	Merck

4.1.4. Chromatography

Chromatography	Manufacturer
Adipic acid hydrazide immobilized on cross-linked 4 % beaded agarose	Sigma
Anti-FLAG M2 affinity gel freezer-safe	Sigma
Cibacron blue 3GA immobilized on cross-linked 4 % beaded agarose	Sigma
Nicotinamide adenine dinucleotide-agarose	ICN
β-Nicotinamide adenine dinucleotide phosphate immobilized in cross-linked 4 % beaded agarose	Sigma
Ni-NTA agarose	Qiagen
Oligo(dT) cellulose type 7	Amersham
Thiopropyl sepharose 6B	Amersham
Ultro gel A4	Sigma

4.1.5. Nucleotides and Radiochemicals

Product	Manufacturer
Deoxyribonucleotides triphosphate (dNTPs)	Roche
Ribonucleotides triphosphate (NTPs)	Roche
$[\gamma\text{-}^{32}P]$-ATP (3000 Ci/mmol)	PerkinElmer
$[\alpha\text{-}^{32}P]$-GTP (3000 Ci/mmol)	PerkinElmer
$[\alpha\text{-}^{32}P]$-TTP (3000 Ci/mmol)	PerkinElmer
$[^{35}S]$-Methionine (1000 Ci/mmol)	Amersham

4.1.6. Enzymes and Proteins

Product	Manufacturer or Donor
AMV (Avian Myeloblastosis Virus) reverse transcriptase	Promega

Bbs I	New England Biolabs
Bbv I	New England Biolabs
Bovin serum albumin (BSA)	Sigma
Calf intestinal alkaline phosphatase (CIAP)	Stratagene
DNase I (RNase-free)	Roche
FLAG peptide	Sigma
HIV-1 reverse transcriptase (HIV-1 RT)	Prof. Restle, University of Luebeck
HIV-2 reverse transcriptase (HIV-2 RT)	Prof. Restle, University of Luebeck
Human DNA polymerase β	Prof. Marx, University of Konstanz
Inorganic pyrophosphatase	Roche
Klenow fragment exo-	New England Biolabs
M-MLV (Moloney Murine Leukemia Virus) reverse transcriptase	Promega
Pfu DNA polymerase	Prof. Famulok
RNasin ribonuclease inhibitor	Promega
Superscript II reverse transcriptase	Invitrogen
T4 DNA ligase	New England Biolabs
T4 polynucleotide kinease (T4 PNK)	New England Biolabs
T7 RNA polymerase	Prof. Famulok
Taq DNA polymerase	Prof. Famulok

4.1.7. Standards and Kits

Product	**Manufacturer**
Amine coupling kit	Biacore
100-bp DNA ladder	PeqLab
High Range Rainbow Marker RPN756	Amersham
Low Range Rainbow Marker RPN755	Amersham
NucleoSpin Extract II Kit	Macherey-Nagel
NucleoTrap Kit	Macherey-Nagel
QIAprep spin miniprep kit	Qiagen
Rabbit Reticulocyte Lysate	Novagen

4.1.8. Plasmid and Bacterial Strains

Product   **Manufacturer**
pGEM-T vector system   Promega
E. coli DH10B   Qiagen

4.1.9. Oligonucleotides

The synthetic DNA oligonucleotides were purchased from MWG or Metabion. The FRET-labeled RNA substrate was purchased from IBA. Other RNAs were transcribed from dsDNA synthesized by PCR using following appropriate DNA templates and primers. T7 promoter sequences are italicized in primer and template sequences.

The following oligonucleotides were used for the experiments in chapter 2.

FK-1 (RNA):
5'-GGG UCC UCU GAU GAG CUU CCG UUU UCA GUC GGG AAA AAC UGA AGC GAA ACU CGU-3'

FK-1 template:
5'-*TCT AAT ACG ACT CAC TAT* AGG GTC CTC TGA TGA GCT TCC GTT TTC AGT CGG GAA AAA CTG AAG CGA AAC TCG T-3'

T7 promoter primer19:
5'-*TCT AAT ACG ACT CAC TAT* A-3'

3'primer_FK-1:
5'-ACG AGT TTC GCT TCA GTT TTT CC-3'

HHRz-FRET-substrate (RNA):
5'-FAM-ACG AGU CAG GAU U-TAMRA-3'

anti–HIV-1 RT aptamer (RNA):
5'-GGG AGC UUC CGU UUU CAG UCG GGA AAA ACU GAA-3'

anti-HIV-1 RT aptamer template:
5'-*TAA TAC GAC TCA CTA TAG* GGA GCT TCC GTT TTC AGT CGG GAA AAA CTG AA-3'

T7 promoter primer17:
5'-*TAA TAC GAC TCA CTA TA*-3'

3'primer_anti-HIV RT aptamer:

```
5'-TTC AGT TTT TCC CGA CTG-3'
```

RT-assay template:
```
5'-GTG CGT CTG TCA TGT CTG TCA GAA ATT TCG CAC CAC-3'
```
RT-assay primer:
```
5'-GTG GTG CGA ATT TC TGA C-3'
```
PCR_template for RT-assay template:
```
5'-TCT AAT ACG ACT CAC TAT AGT GCG TCT GTC ATG TCT GTC AGA AAT TTC GCA
CCA C-3'
```

Biotin-P12II:
```
5'-Biotin-CAC ACC CCG CGC-3'
```
RT18II:
```
5'-GCG CGG GGT GTG GTG GTA-3'
```

The following oligonucleotides were used for the experiments in chapter 3. The sequences designed for protein tag, restriction enzyme recognition, and translation initiation were underlined with their explanation. The sequences for start and stop codons were shown in bold letters. The two sequences for proralen insertion were in italicized bold letters. XYZ indicates one codon with randomized nucleotide sequences.

mRNA-display library:
```
5'-TTC TAA TAC GAC TCA CTA TA GGG ACA ATT ACT ATT TAC AAT TAC A
                 T7                        TMV
ATG GAC TAC AAG GAC GAC GAC GAC AAG GGA AGA CTG AGT GGA (XYZ)17 CCA
start          FLAG tag                   Bbs I              random
GGG AGT AGT (XYZ)21 AGT (XYZ)21  AGT GGA (XYZ)17 CCA GGG AGT  AGT
            random       random           random
GCT AGC GTA GCT  GCA CAT CAC CAC CAC CAC CAT GGC AGC GGC GCT ATT TAA
out of frame stop   Bbv I         His tag                 psoralen   stop
AA-3'
```

5'primer:
```
5'-TTC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT AC-3'
```
3'primer:
```
5'-TTT TAA ATA GCG CCG CTG CCA TGG TG-3'
```

5'short primer:

`5'-TTC TAA TAC GAC TCA CTA TAG GGA C-3'`

3'short primer:

`5'-TTT CAA ATA GCG CCG CTG CCA TG-3'` RT primer:

`5'-GCC ATG GTG GTG GT-3'`

Photo linker:

`5'-(Ps)-`**`UAGCGCCGCU`**`-(dA)`$_{21}$`-(18)(18)-dCdC-(Pu)-3'`

(Ps: psoralen C-6, Pu: puromycin CPG, 18: Spacer C 18 PEG, **N**: 2'-OMe modified RNA nucleotide)

4.2. General Methods for Manipulation of Nucleic Acids

All buffers and solutions for RNA manipulation were prepared with ultra pure water treated with diethylpyrocarbonate (DEPC). To prepare DEPC-treated water, 1L of ultra pure water was mixed with 2 ml of DEPC and incubated overnight at room temperature, then autoclaved.

4.2.1. Phenol/Chloroform Extraction

Phenol/chloroform extraction was used to remove proteins from nucleic acid solutions. An equal volume of phenol was added to the aqueous nucleic acid solution, vortexed for 1 min, and centrifuged for 3 min at 14,000 rpm to separate the phases. The upper aqueous phase was transferred into to a new tube. To remove residual phenol from the nucleic acid containing aqueous solution, double volume of chloroform was added to this solution, vortexed, and centrifuged for 3 min at 14,000 rpm. After separation of the upper phase, DNA/RNA was precipitated as described in §4.2.2.

4.2.2. Ethanol Precipitation

Ethanol precipitation is performed to remove salts or reaction products. RNA or DNA was precipitated by adding 0.1 volume 3.0 M sodium acetate (NaOAc) pH 5.4, and 3.0 volumes of ice-cold absolute ethanol. After shaking thoroughly the samples were

incubated for 10 min at -80°C (for 15 ml tubes this time was elongated to 30 min) or overnight at -20°C. RNA or DNA was recovered by centrifugation for 15 min at 4°C at 14,000 rpm (30 min for 15 ml tubes). The supernatant was discarded. The resulting pellet was washed with 100 µL of 70 % (v/v) ethanol, centrifuged for 5 min at 14,000 rpm, the supernatant was taken off. After drying at room temperature, the pellet was resuspended in DEPC-treated water. To precipitate short oligonucleotides, or tiny amount of nucleic acids, 1 µL of solution of glycogen (20 mg / mL) was added to ethanol mixture.

4.2.3. Photometric Concentration Determination of Nucleic acids

The concentration of aqueous nucleic acid solutions was determined by UV absorbance at a wavelength of 260 nm. An optical density of 1 corresponds to a concentration of approx. 50 µg/ml of double stranded DNA, 33 µg/mL of single stranded DNA, 40 µg/ml of single stranded RNA, or 25 µg/mL of oligonucleotide under 40 bases. The concentration was determined by following equation:

$$c\ [mM] = (f \cdot A \cdot OD_1) / (n \cdot Mr)$$

f = dilution factor
A = measured absorption value
OD_1 = concentration by an optical density of 1
n = number of nucleotides
Mr = average molecular weight of each nucleotide
    (dNTP: 330 g/mol, NTP: 345 g/mol)

The concentration of short oligonucleotides (up to 40 nucleotides) which sequences were known was determined using the sum of the absorption coefficient ε (Table 8) of each nucleotide and the following equation:

$$c \ [\text{mM}] = (f \cdot A) / \sum \varepsilon$$

dNTP	$\varepsilon \ [\text{mM}^{-1} \cdot \text{cm}^{-1}]$	NTP	$\varepsilon \ [\text{mM}^{-1} \cdot \text{cm}^{-1}]$
dATP	15.4	ATP	15.2
dCTP	9.0	CTP	7.05
dGTP	13.7	GTP	12.1
dTTP	10.0	UTP	8.4

Table 8. Absorption coefficients of dNTPs and NTPs.

4.3. General Methods for Gel Electrophoresis

Nucleic acids were analyzed on agarose gels (non-denaturing condition) or polyacrylamide gels (denaturing condition). Proteins or fusion proteins for mRNA-display were analyzed on tricine SDS gels.

4.3.1. Agarose Gel Electrophoresis

Linear dsDNA was separated and identified using non-denaturing agarose gels at concentrations ranging from 0.7 to 2.5 % (w/v). The location of the nucleic acid within the gel was determined by ethidium bromide staining. Agarose gels were cast by melting the agarose in 0.5 x TBE buffer in a microwave. Subsequently the liquid was poured into a horizontal mold and left to harden after adding ethidium bromide solution (1 µL/10 mL). The gel well was loaded with 7 µL probes resuspended in agarose loading buffer at a 2:1 ratio. The molecular size of DNA probes was estimated using a DNA mix size marker (PeqLab) containing xylene cyanol and bromophenol blue. To achieve maximum resolution, gels were run at 170 V in 0.5 x TBE buffer. The lanes were visualized using a 366 nm UV-transilluminator.

Agarose gel loading buffer:

reagent	MW	concentration	required for10 mL
Glycerol	92.09	50% (v/v)	5 mL
1 M Tris/HCl	121.14	50 mM	500 μL
0.5 M EDTA, pH 8.0	372.2	50 mM	1 mL
Bromophenol blue	691.94	0.05% (w/v)	5 mg

4.3.2. Isolation of Nucleic Acids from Agarose Gel

DNA was separated using low melting point agarose gel from 0.7 - 2 % at 70 V in TAE buffer. The bands of DNA fragment were detected by UV shadowing and excised using sterile surgical blade. DNA fragments were purified using gel the extraction kit from Macherey-Nagel (e.g. NucleoSpin Kit) by following the instructions supplied by the manufacturer.

4.3.3. Denaturing Polyacrylamide Gel Electrophoresis (PAGE)

For the purification and/or analysis of single-stranded DNA or RNA, denaturing polyacrylamide gels were used. Gels 6-20 % (v/v) were performed under denaturing conditions using 8.3 M urea. The gel mixture was prepared from appropriate amounts of stock solutions B, C, and D. The polymerization was initiated with 800 μL 10% APS (ammonium peroxodisulfate) and 40 μL of TEMED (N,N,N',N'-tetramethylendiamine) per 100 mL of gel mixture. The gel mixture was poured between two glass plates (ca. 18 x 20 cm) held apart by spacers of 1-1.5 mm (1 mm for analytical use, and 1.5 mm for preparative use) and sealed with clamps after insertion of the comb. After polymerization (60-90 min), the vertical gels were pre-run for 20 - 30 min at 370 V in 1 x TBE, and the samples were loaded on the gel in loading buffer (1:1 ratio) after a short denaturation (1-3 min at 95ºC). The gels were then run at 370 V for the appropriate time (usually 90 min). The course of electrophoresis was monitored with a control lane containing 0.25% of bromophenol blue and xylene cyanol in loading buffer. For analytical purposes, the radioactive RNA or DNA bands were detected after exposure to a PhosphorImager screen. For preparative purposes, the RNA or DNA was visualized using film exposure or UV shadowing.

Stock solution B:	10 x TBE in 8.3 M urea
Stock solution C:	25 % acrylamide solution in 8.3 M urea
	2 % N,N'-methylene bis acrylamide
Stock solution D:	8.3 M urea

PAGE loading buffer:

reagent	MW	concentration	required for 10 mL
Urea	60.1	9 M	5.4 g
0.5 M EDTA, pH 8.0	372.2	50 mM	1 mL
Bromophenol blue	691.94	0.05 % (w/v)	5 mg
Xylene cyanol	538.62	0.05 % (w/v)	5 mg

4.3.4. Isolation of Nucleic Acids from Denaturing Polyacrylamide Gels

After electrophoretic separation, bands were visualized by UV-shadowing at 254 nm. The band of the desired size was excised using a sterile scalpel, and the gel slice was crushed and soaked in one volume 0.3 M sodium acetate solution (pH 5.4). To allow passive elution, the suspension was incubated at 65°C for 1.5 h or overnight by room temperature with strong agitation. To separate the nucleic acid containing liquid phase from the gel fragments, the sample was filtered through a syringe which was pre-packed with silanized glass wool. Finally, the nucleic acid was ethanol precipitated and collected by centrifugation as described in previous section (§4.2.2). Optionally the resuspended aqueous RNA solution was filtered through Ultrafree-MC centrifugal filter device (Millipore) to remove remaining small gel pieces.

4.3.5. Sequencing Gel Electrophoresis

For analysis of kinetic studies, 15 % sequencing gels were used (gel size 30 x 50 cm, gel thickness 0.4 mm, glass plates from Bio-Rad). 150 ml of gel mixture was prepared as described in §4.3.3. The surfaces of glass plates were silanized with 5 % of dimethylchlorosilane in dichloromethane before pouring the gel mixture. After polymerization (60 – 90 min), the vertical gels were pre-run for 30 min at 2000 V in 1 x TBE. The samples were denatured by adding two volume of STOP solution and treatment for 5 min at 95 ºC, then cooled down on ice immediately before loading. The

gels were then run at 2000 V for the appropriate time. The course of electrophoresis was monitored with a control lane containing 0.25 % of bromophenol blue and xylene cyanol in STOP solution. To visualize the radioactive nucleic acids bands the gel was transferred to a sheet of Whatman paper, dried for 40 min at 80 ºC on a gel drying apparatus (Bio-Rad) and then exposed to PhosphorImager screen.

STOP solution:

reagent	MW	concentration	required for 50 mL
Formamide	45.04	80 %	40 mL
0.5 M EDTA, pH 8.0	372.2	20 mM	2 mL

4.3.6. Tricine SDS Polyacrylamide Gel Electrophoresis

Tricine SDS polyacrylamide gels were used to separate proteins or analyze *in vitro* translated mRNA-displayed fusion proteins. They have a higher resolution than glycine SDS polyacrylamide gels, especially for lower weight peptides. The percentage of the gel was adjusted to 12 % or 8 % according to molecular weight of the proteins. The gel mixture was prepared from appropriate amount of stock solutions, and then the polymerization was initiated by addition of a 10% APS solution and TEMED. The solution was poured between two glass plates held apart by spacers of 1 mm, and the polymerization usually completed in 30 min. After the gel was placed into a gel chamber, the inner buffer tank was filled with 1 x tricine SDS electrophoresis buffer, and the outer tank was filled with 1 x tricine SDS anode buffer. The samples (up to 15 µL) with 5 µL of loading buffer were loaded into wells without heating. The gel was run at 125 V for ca. 1 h. The course of electrophoresis was monitored with a control lane containing protein markers (Amersham). The gel was removed from glass plates, placed into fixing solution for 10 min with agitation, exchanged into water for another 10 min while shaking. The gel was transferred onto a piece of Whatman paper and dried for 30 min at 80 ºC on the gel dryer. The radioactive protein samples were visualized by exposure of a dried gel to a phosphor imager screen.

Tricine SDS polyacrylamide gel (12 %):

reagent	concentration	required for 6 mL
3 x Tricine SDS gel buffer	1 x	2000 μL
Water		900 μL
30 % Bis-acrylamide (Roth)	12 %	2500 μL
Glycerol	10 %	600 μL
10 % APS		60 μL
TEMED		6 μL

Tricine SDS loading buffer:

reagent	MW	concentration	required for 10 mL
1 M Tris/HCl, pH 6.8	121.14	50 mM	500 μL
10% SDS	288	4 % (w/v)	4 mL
1 M DTT	154.2	0.1 M	1 mL
Glycerol	92.09	12 % (v/v)	1.2 mL
Water			3.3 mL

3 x Tricine SDS gel buffer:

reagent	MW	concentration	required for 100 mL
Tris	121.14	3 M	36.34 g
SDS	288	0.3 % (w/v)	300 mg

pH value was adjusted to 8.45 with HCl.

10 x Tricine SDS electrophoresis buffer:

reagent	MW	concentration	required for 1000 mL
Tris	121.14	1 M	121.14 g
Tricin	178.2	1 M	178.2 g
SDS	288	1 % (w/v)	10 g

pH value was adjusted to 8.25 with NaOH/HCl.

10 x Tricine SDS anode buffer:

reagent	MW	concentration	required for 1000 mL
Tris	121.14	2 M	242.28 g

pH value was adjusted to 8.9 with NaOH/HCl.

Fixing solution:

reagent	MW	concentration	required for 1000 mL
Methanol	32.04	40 % (v/v)	400 mL
Acetic acid	60.05	10 % (v/v)	100 mL
Water		50 % (v/v)	500 mL

4.3.7. Coomassie Staining of SDS Gels

To stain the protein bands, the gels were soaked in coomassie solution for 30 min, and then transferred into destaining solution for 30 − 120 min until the protein bands became distinct. The gel was placed onto a piece of Whatman paper and dried for 30 min at 80ºC on a gel dryer.

Coomassie solution:

reagent	MW	concentration	required for 1000 mL
Acetic acid	60.05	10 % (v/v)	100 mL
Methanol	32.04	45 % (v/v)	450 mL
Water		45 % (v/v)	450 mL
Coomassie brilliant blue		0.0125 % (w/v)	125 mg

Destaining solution:

reagent	MW	concentration	required for 1000 mL
Acetic acid	60.05	10 % (v/v)	100 mL
n-Butanol	74.12	30 % (v/v)	300 mL
Water		60 % (v/v)	600 mL

4.4. Enzymatic Reactions

4.4.1. Polymerase Chain Reaction (PCR)

dsDNA constructs encoding reporter ribozymes, aptamers, and DNA library sequences were generated and amplified by the polymerase chain reaction (PCR) using *Taq* DNA polymerase. Reactions in 100 µL final volume were performed in a PCR thermocycler. Standard conditions were programmed as follows;

operation	time	temperatur
Denaturing	50 s	94°C
Annealing	60 s	55°C
Extension	60 s	72°C
Heated lid		110°C

PCR mixture (100 µL):

reagent	amont	stock	resultant
PCR buffer	10 µL	10 x	1 x
MgCl$_2$	6 µL	25 mM	1.5 mM
5' primer	1 µL	100 µM	1 µM
3' primer	1 µL	100 µM	1 µM
dNTP mix (ACTG)	0.8 µL	25 mM	0.2 mM
DNA template	x µL		1-10 nM
Taq DNA Polymerase	0.5 µL	5 U/µL	0.025 U/µL
Water	y µL		

To monitor the amplification of desired DNA product, aliquots (5 µL) from different cycles were taken, mixed with agarose loading buffer (2.5 µL), then loaded on agarose gel as described in §4.3.1. One reaction mixture without DNA template as negative control was also operated equally to check the unspecific amplification of primers or DNA contamination.

PCR was performed with optimized cycle numbers, in large scale, if necessary. The PCR product was finally isolated and purified by phenol/chloroform extraction as written in section §4.2.1 and the dsDNA was ethanol precipitated by the method described in §4.2.2.

4.4.2. *in vitro* Transcription

For *in vitro* synthesis of RNA, DNA-dependent RNA polymerase from bacteriophage T7 was used. The DNA templates are required to contain a double stranded T7 promoter sequence. The reaction mixtures were prepared as follows, and incubated at 37°C for 4 h to overnight. When the resulting RNA needed to be body-labeled, α-^{32}P GTP was added to the mixture before T7 RNA polymerase addition.

Transcription reaction mixture (100 µL):

reagent	amount	stock	resultant
Transcription buffer	20 µL	5 x	1 x
NTP mix (ACUG)	10 µL	25 mM	2.5 mM
IPP	0.2 µL	0.2 U/µL	0.0004 U/µL
[α-^{32}P]-GTP	1 µL	10 µCi/µL	0.1 µCi/µL
RNasin	1.25 µL	40 U/µL	0.5 U/µL
DNA template	x µL		1-3 µM
T7 RNA polymerase	2 µL	50 U/µL	1 U/µL
Water	y µL		

To digest DNA templates, 0.5 µL DNase I (5 U/µL) was added to the mixture and incubated at 37°C for at least 10 min. One equivalent volume of PAGE loading buffer was added to the mixture, and subsequently heated to 95°C for 1 min. The desired RNA was purified on PAGE (see §4.3.3 and §4.3.4).

4.4.3. 5'-Dephosphorylation of RNA

Prior to labeling, the transcribed RNA was dephosphorylated at the 5′-end using calf intestine alkaline phosphatase (CIAP). The reaction mixture was composed as follows;

5'-Dephosphorylation reaction mixture (50 µL):

reagent	amount	stock	resultant
CIAP buffer	5 µL	10 x	1 x
RNA	150 pmol	10 µM	3 µM
BSA	0.5 µL	100 x	1 x
CIAP	0.3 µL	57 U/µL	0.34 U/µL
RNasin	0.5 µL	40 U/µL	0.4 U/µL
Water	y µL		

The reaction mix was then adjusted to a final volume of 50 µL and incubated for 15 min at 37°C. After addition of 0.15 µL CIAP the reaction was incubated at 55°C for an additional 15 min. The reaction was stopped by adding 0.5 µL 0.5 M EDTA solution and the enzyme inactivated for 10 min at 75°C. The volume was adjusted to 200 µL by

addition of water and the reaction mixture was extracted with chloroform and phenol (see §4.2.1). The RNA was recovered after ethanol precipitation (see §4.2.2) and the pellet was resuspended in 20 µL of DEPC treated water. 3 µL of this sample per 20 µL reaction volume can be used for 5'-radioactive end labeling.

4.4.4. 5'-Phosphorylation of Nucleic Acids

Radioactive 5'-end labeling of DNA or RNA was performed using T4 polynucleotide kinase (T4 PNK) with [α-^{32}P]-ATP. Nucleic acids synthesized on solid phase can be phosphorylated directly, while RNA transcribed *in vitro* should be dephosphorylated as described in §4.4.3 prior to labeling. The reaction mixture was incubated for 1 h at 37°C. After heat denaturation at 100°C for 2 min, the sample was immediately cooled down on ice. Unincorporated radioactive dNTP was removed by gel filtration (G-25 column). To generate a 3 µM solution of radioactive DNA primer, 20 µL of 10 µM cold primer was added to 50 µL of labeled DNA primer.

5'-Phosphorylation reaction mixture (50 µL):

reagent	amount	stock	resultant
RNA/DNA template (20 pmol)	2 µL	10 µM	0.4 µM
T4 PNK buffer	5 µL	10 x	1 x
[α-^{32}P]-ATP	5 µL	10 µCi/µL	1 µCi/µL
RNasin (added for RNA template)	0.75 µL	40 U/µL	0.6 U/µL
T4 polynucleotide kinase	5 µL	10 U/µL	1 U/µL
Water	y µL		

4.4.5. RNase Test

All buffers required for RNA manipulation were examined for RNase contamination. For this purpose, radioactive labeled RNA in 25 µL of each buffer was incubated at least for 4 h at room temperature. After addition of PAGE loading buffer, the RNA degradation of each sample was analyzed on PAGE. The radioactive bands were visualized by exposure of the gel to a phosphor imager screen.

4.4.6. Restriction Digestion of dsDNA

The digestion of dsDNA was done by restriction enzymes to construct DNA libraries of long ORFs used for mRNA display. Each DNA cassette was restricted by either Bbs I or Bbv I in the supplied buffer under the recommended condition (at 37°C overnight). Usually approximately 1 µL of restriction enzyme per 25µL of reaction volume including up to 300 pmol of dsDNA was used. The digested fragments were purified from low melting point agarose gels (§4.3.2).

4.4.7. Ligation of DNA Fragments

The digested DNA fragments by the restriction enzymes were ligated by T4 DNA ligase using the conditions recommended by the supplier. Ligation was completed at 16°C overnight. After phenol/chloroform extraction and ethanol precipitation, the ligated DNA could be amplified using appropriate primers and PCR.

4.5. DNA Sequencing

For DNA sequencing, the plasmids were sent to Qiagen for sequencing service. Otherwise, dideoxy (Sanger) sequencing method was used.

4.5.1. Ligation of PCR Product into Plasmids

For preparation of the plasmid to be transformed into *E. coli*, pGEM-T Vector System from Promega was used. This system represents a convenient system for cloning of PCR products into vectors which carry single 3'-T overhangs at the insertion sites. The ligation reactions including positive and negative controls were performed by the following protocols described in the technical instruction from the kit.

Ligation reactions (10 µL):

reagent	standard	positive	negative
2 x Rapid ligation buffer	5 µL	5 µL	5 µL
pGEM-T Vector (50 ng)	1 µL	1 µL	1 µL
PCR product	x µL	-	-
Control insert DNA	-	2 µL	-
T4 DNA ligase	1 µL	1 µL	1 µL
Water	y µL	y µL	y µL

4.5.2. Transformation

The appropriate high efficiency competent cells (DH10B) were transformed with the ligation products using the protocols provided by Promega with some modifications. For transformation, 5 µL of the respective ligation reaction was gently mixed with 100 µL of freshly thawed competent cells. After 20 min incubation on ice, the cells were warmed at 37°C for 40-45 sec, and immediately returned on ice for 2 min. Then, 750 µL SOC medium (room temperature) was added to the cells, and incubated for 1 h at 37°C with shaking (~300 rpm). 100 -250 µL of transformation culture was plated onto LB/ampicillin/IPTG/X-Gal plates. The plates were incubated overnight at 37°C.

4.5.3. Isolation of Plasmid DNA

Purification of high copy plasmid DNA from overnight cultures of E. coli was done using the QIAprep spin miniprep kit from Qiagen. All procedures were followed using the handbook of the supplier.

4.5.4. DNA Sequence Analysis

The isolated plasmids were added to sequencing reaction mixture with an appropriate primer described as follows. Primer extension using fluorescent dye-labeled terminators produced chain terminated products, which were analyzed by an automated DNA sequencer.

Sequencing reaction mixture (10 µL):

reagent	amount	stock	resultant
Premix	2 µL		
DNA-template	380 ng		
Primer	10 pmol	100 µM	1 µM
Water	y µL		

operation	time	temperatur
Initial denaturing	2 min	96°C
Denaturing	10 s	96°C
Annealing	10 s	53°C
Extension	4 min	60°C

Steps 2-4 were repeated 25 times.

4.6. Reporter Ribozyme-Based Screening Assay

4.6.1. Screening Procedure using Reporter Ribozyme

Screening to identify small molecule inhibitors of HIV-1 RT was performed in 384-well plates. Each reaction in total volume of 25 µL was carried out under multiple turnover condition with excess FRET-substrate (200 nM) over the reporter ribozyme FK-1 (10 nM) in HH buffer (50 mM Tris/HCl, pH 7.9, 25 mM NaCl) at 37°C in the presence of HIV-1 RT (200 nM) and a small molecule entity (100 µM) from the library. The library compounds and the reaction mixture were combined into 384-well plates by using an automated liquid handling system (Tecan). It was necessary to use ceramic tips and to prepare the enzyme mixture in Pyrex glass vials for robotic operation. After the plates were pre-incubated for 15 min at 37°C in Ascent Fluoroskan apparatus, the ribozyme reaction was initiated by adding $MgCl_2$ to a final concentration of 8 mM. The cleavage reaction was monitored in real-time by measuring FAM-dye fluorescent emission (520 nm) for 15 min at 60 seconds intervals at a fixed excitation wavelength (492 nm). Negative control without HIV-1 RT was included for each compound. Each plate also contained standard ribozyme reaction without compound as a positive control. Every reaction was duplicated.

4.6.2. Identification of Hit Compounds

The initial velocity (fluorescence/min) was determined by plotting the increment of fluorescence intensity, multiplied by factor 10000, versus time. For each compound, relative ribozyme activity (A_{rel}) was calculated by following equation;

$$A_{rel} = V_{RT+} / V_{RT-}$$

A_{rel} = relative activity
V_{RT+} = averaged number of initial velocity of ribozyme in the presence of HIV-1 RT
V_{RT-} = averaged number of initial velocity of ribozyme in the absence of HIV-1 RT

Standard reaction without compound usually gave the A_{rel} value of 0.2-0.3 depending on the plate. The compounds which showed the value of $A_{rel} > 0.4$ were identified as hit compounds.

4.7. Reverse Transcriptase (RT) Assay

The RNA-dependent DNA polymerase activity was assayed by monitoring the RNA template/DNA primer-directed incorporation of dNTPs, while the DNA-dependent DNA polymerase activity was assayed by DNA template/DNA primer-directed incorporation of dNTPs. As an example, the assay to examine the inhibition on DNA-dependent DNA polymerase activity of HIV-1 RT by a certain compound is described in the following.
After heat denaturing at 95°C for 5 min (RNA template/DNA primer denaturing at 70°C), the DNA template (10 µM) was annealed to a 5'-[^{32}P]-end labeled complementary DNA primer (3 µM) (see §4.4.4) by slow cooling to room temperature over 1 h. The small molecule inhibitors were pre-incubated at 37°C for 5 min at various concentrations in the enzyme mixture containing 14 nM HIV-1 RT and 50 µM dNTP. Polymerase reactions were carried out in 20 µL volumes, containing 50 mM Tris/HCl, pH 8.0, 50 mM KCl, and 10 mM MgCl$_2$. The reactions were initiated by the addition of DNA template/5'-[^{32}P]-DNA primer complex and incubated at 37°C for 15 min. The polymerase activity was deactivated by addition of 40 µL of STOP solution (80 % formamide, 20 mM EDTA) and denatured by heating at 95°C for 5 min prior to gel analysis (15 % sequencing gel). The gel was visualized by autoradiography and the

amount of DNA synthesized was quantified by phosphor imaging analysis. IC_{50} values were caluculated from the inhibition dose-dependent curves using Origin software.

Assays of HIV-2 RT, AMV, MMLV, Klenow fragment, human DNA polymerase β were carried out in the buffers supplied from the manufacturers. The reaction time was elongated to 1 h for the Klenow fragment and human DNA polymerase β. Optionally the concentration of dNTP was raised to 200 μM.

4.8. Surface Plasmon Resonance (SPR)

To elucidate the inhibition mechanism of HIV-1 RT by small molecule inhibitors, the possible interaction between double stranded DNA and a small molecule was analyzed via surface plasmon resonance (SPR) using Biacore instrumentation.

The Biacore equipment enables to monitor biomolecular interactions in real time, using a non-invasive optical detection principle based on SPR. The SPR response reflects a change in mass concentration at the detector surface as molecules bind or dissociate.

SPR is an optical phenomenon arising in thin metal films under conditions of total internal reflection. The phenomenon produces a sharp dip in the intensity of reflected light at a specific angle, the resonance angle. The position of this resonance angle depends on refractive index of the medium when other factors kept constant. The refractive index is directly correlated to the concentration of dissolved material in the medium. Therefore, SPR enables to measure changes in the concentration of molecules in a surface layer of a solution in contact with the sensor surface.

The surface immobilized with dsDNA was prepared as follows. While, HIV-1 RT immobilization was done by using amino coupling kit onto CM5 chip.

4.8.1. Immobilization of Oligonucleotides on the Sensor Surface

Oligonucleotides were synthesized with a biotinylated 5'-end (Biotin-P12II). Sensor chip SA was activated with three consecutive 1 min injections of 1 M NaCl in 50 mM NaOH prior to the immobilization. The oligonucleotide solution was prepared in a 1:1 mixture of 1 M NaCl and HBS-EP buffer supplied by the manufacturer, with a final concentration of 50 nM. Typically, 330 - 350 resonance units (RU) of oligonucleotide were immobilized on a flow cell at a flow rate of 5 μL/min.

4.8.2. Binding of Small Molecules to dsDNA

The running buffer contained 50 mM Tris/HCl pH 8.0, 50 mM KCl, 10 mM $MgCl_2$, and 0.5 % DMSO. The RT18II template sample was suspended in the running buffer by adding 1 µL of a 100 µM template solution into 199 µL of running buffer. The double stranded oligonucleotide construct (P12II/RT18II) was formed by injecting 35 µL of a 500 nM solution of RT18II at a flow rate of 5 µL/min. This annealing step usually resulted in an increase of 500 RU. The stock solutions for the small molecules were a 200-fold concentrate in 100 % DMSO. Dilutions were prepared in running buffer such that 1 µL of a respective compound was added to 199 µL of running buffer to give a final DMSO concentration of 0.5 % (v/v) in each sample. The flow rate was set to 30 µL/min and 60 µL of the dilution sample (50 pM – 5 µM) was injected simultaneously over the surfaces including the reference cell (program: KINJECT). For each experiment, a calibration for bulk differences between flow cells and buffer jump was performed. Between the sample injections, the needle and IFC were washed to ensure no contamination with compounds from the previous injection. Regeneration of the surfaces were successfully done by either 60 - 90 µL injection of running buffer containing 0.5-2.5 % DMSO, or by 5 µL injection of up to 20 mM of NaOH containing 0.5-5 % DMSO. The latter condition caused slight dissociation of the template from the surface in some experiments.

4.9. Inhibition Study of HIV-1 Replication *in vivo*

The study of the inhibition of HIV-1 replication by small molecules *in vivo* was done by usage of an assay setup termed "self-inactivating HIV vector system" in collaboration with the group of Prof. Restle at University of Luebeck. Other *in vivo* assays were performed in collaboration with the group of Prof. Klaeusslich at University of Heidelberg.

4.9.1. Self-Inactivating HIV Vector System[84]

A self-inactivating replication-defective plasmid (pGJ3) was constructed from the proviral plasmid pHXB2D by the method reported by Jármy *et al.*[84] 1 x 10^6 293T cells were co-transfected with 8 µg of plasmid VSV-G and pGJ3-Luci, using the calcium

phosphate method. After 40 h of incubation in the presence of 10 mM sodium butyrate, cell culture supernatants containing viral particles were harvested and stored at -80°C until infection.

The compounds in 100 % DMSO were diluted in PBS buffer to give a final concentration of 1 % DMSO, and then pre-incubated on 2×10^4 293 cells for 30 min in a 96-well plate. Subsequently, infection with equal amounts (50 µl) of vector-containing supernatants was performed, and luciferase activity was determined after 24 h incubation. Normalization in relation to cell viability was done with 10 µM fluorescein diacetate (FDA), which is a marker for esterase activity of living cells.

As a positive control, the assay was carried out with 10 µM AZT to see complete inhibition.

4.9.2. HIV-1 Replication Assay

2×10^4 TZM cells were grown in 48well plate. Different concentrations of compounds dissolved in 2 µL of 100 % DMSO were diluted in 150 µL of cell culture medium (DMEM), then 50 µL virus containing solution (= 10000 infectious units) was added to give a final concentration of 1 % DMSO. TZM cells were infected with 200 µL of DMEM containing virus and inhibitor after removal of old medium. After incubation for 2 days at 37°C under 5 % CO_2 atmosphere, the luciferase marker gene expression, under the control of the HIV-1 LTR promoter, was measured (Steady-Glo luciferase assay from Promega). NL43 which has been produced in a co-culture with C8166 cells was used as virus.

4.9.3. HIV Infectivity Assay

Fresh co-culture was prepared by dilution of co-culture of NL43 virus with uninfected MT4 cells in the ratio of 1:10 (1×10^6 MT4/mL). After 4 h incubation at 37°C under 5 % of CO_2 atmosphere, 500 µL of co-culture and 5 µL of different concentration of inhibitors were added to each 48 well. After 2 days at 37°C under 5 % of CO_2, cell-free supernatant of the MT4 co-culture was used to infect TZM cells, which enables to determine amount of infectious particles which have been produced in the presence of inhibitor. After 2 days postinfection, cells were fixed with 3% PFA for 30min, washed with PBS. The infected TZM cells were stained with β-galactosidase substrate (TZM

cells also contain a β-Galactosidase reporter gene under the control of HIV-1 LTR). After 2-3 hours incubation at 37°C, infected cells turned blue and could be counted under the microscope.

4.9.4. MTT Cell Toxicity Assay

The MTT-assay was performed to determine cell viability by quantifying its metabolite. MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazoliumbromid) is up taken by living cells. Mitochondrial dehydrogenases convert it to a violet formazan that cannot diffuse through cell membranes and crystallizes in vital cells. The crystals can be solubilized with organic solvents and than be measured with a colorimetric assay. Only living cells with an active metabolism, as well as cells which are in an early apoptotic phase can convert MTT. Cells in a late apoptotic phase and dead cells cannot.

1×10^4 TZM cells were seeded in 96-well plate. 100 μL of medium containing 1μL of different concentrations of inhibitors were added, and incubated for 24 h at 37°C under 5% CO_2 atmosphere. The medium was removed, and 100μl of MTT-solution (1:10 dilution of 5mg/mL MTT-stock solution in PBS into DMEM) was added to 96-well plate. After incubation for 3-4 h at 37°C, the medium was removed carefully, then 100 μL of DMSO:ethanol (1:1) was added. Absorbance both at 550-570nm (L1) and at 620-650nm (L2) was measured to determine the value of L1-L2.

4.10. Preparation and Purification of mRNA-Displayed Proteins

mRNA templates are prepared from appropriate DNA library by *in vitro* transcription. The following sections describe the preparation of mRNA display templates with DNA linker 3'-terminated with puromycin, the use of the mRNA display template to yield mRNA-displayed proteins and their subsequent purifications. In order to separate the mRNA displayed proteins from both mRNA templates which don't carry the proteins and free proteins, it is necessary to perform a purification step based on a protein affinity tag. In the initial cycles of in vitro selection, all the enzymatic reactions were scaled up to keep the library diversity.

The FLAG tag affinity chromatography was optionally used only in the selection against cibacron blue 3GA.

4.10.1. **Photoligation of RNA to Psoralen-Photolinker**

To form a covalent bond between transcribed RNA and the linker carrying puromycin at 3'-end, photoligation was conducted by usage of psoralen.[152, 153] The photolinker was designed to contain a psoralen-C6, 2'-O-methyl RNA, oligo(dA), spacer, and puromycin from 5'- to 3'-end. The ligation reaction mixture was prepared on ice, and operated as follows in a PCR cycler;

Photoligation reaction mixture (50 µL):

reagent	amount	stock	resultant
RNA	x µL	x µM	2.5 µM
Ligation buffer	5 µL	10 x	1 x
Photolinker	2.5 µL	100 µM	5 µM
Water	y µL		

operation	time	temperature
Denaturing	150 s	85°C
Annealing	20 min	4°C
Lid temperature		110°C

The solution was transferred into disposable cuvette, and irradiated with 366 nm light by a UV transilluminator for 15 min. The sample was precipitated with ethanol, and then resuspended in water.

10 x Photoligation buffer:

reagent	MW	concentration	required for 50 mL
1 M Tris/HCl, pH7.2		250 mM	12.5 mL
NaCl	58.44	1.5 M	4.38 g

4.10.2. *in vitro* **Translation**

In vitro translation of photoligated mRNA in rabbit reticulocyte lysate produced mRNA-displayed proteins via covalent bond formation promoted by puromycin. Radioactive [^{35}S]-methionine labeling of translated proteins enables to detect fusion products and free proteins on tricine SDS gels. Translation mixtures were made up as

follows by adding the lysate last;

in vitro Translation mixture (100 μL):

reagent	amount	stock	resultant
mRNA display template	x μL	x μM	400 nM
Amino acid mix without Met	8 μL	12.5 x	1 x
[^{35}S]-Methionine	8 μL	10 μM	800 nM
KCl	4 μL	2.5 M	100 mM
Mg(OAc)$_2$	5 μL	25 mM	1.25 mM
RNasin	2.5 μL	40 U/μL	1 U/μL
Rabbit reticulocyte lysate	35 μL	100 %	35 % (v/v)
Water	y μL		

After incubation at 30°C for 1 h, KCl (4 M) and MgCl$_2$ (1 M) were added to the final concentration of 340 mM and 100 mM, respectively. The sample was incubated for additional 10 min at room temperature. To release the fusion proteins from the ribosome, EDTA was added to a final concentration of 100 mM, and incubated for 10 min at room temperature. Both the concentrations of KCl and Mg(OAc)$_2$ in the translation mixture and the later addition of KCl and MgCl$_2$ after incubation were optimized for every library template used.

4.10.3. Oligo(dT) Affinity Chromatography

Oligo(dT) affinity chromatography was used for isolation of mRNA-protein fusions from the translation mixture. The cellulose resin has oligo(dT) covalently attached to it, which can form specific Watson-Crick base pairs with the target oligo(dA) tag. For this purpose, the photolinker contains oligo(dT) sequences. mRNA-protein fusions and photolinker could be purified by the column, while other proteins which lack mRNA (free proteins) were discarded. Typically, the binding capacity for the Pharmacia type 7 oligo(dT) cellulose is around 200 nmol/g of resin.

The translation mixture was diluted 1:15 or 1:20 with oligo(dT) binding buffer. The appropriate amount of oligo(dT) cellulose (2 mg per mL of diluted sample) was weighted out (up to 20 mg for PolyPrep chromatography column from Bio-Rad), washed 3 times with 1 mL of water, then equilibrated with 1 mL of oligo(dT) binding buffer. The diluted lysate was incubated with oligo(dT) cellulose at 4°C for 1 h while

shaking. Optionally, washing of cellulose and the incubation steps could be performed in 15 or 50 mL of centrifuge tubes. The mixture was filtered through the column, and the flow through was poured again into the column to collect all the remaining cellulose. The resin was washed 3 times with 1 mL of oligo(dT) binding buffer, subsequently with 1 mL of oligo(dT) wash buffer. The bound molecules were eluted with 2 times of 250 µL water at 4°C for 30 min while shaking to denature DNA hybridization. Small aliquots (1–2 % of each fraction) from wash and elution samples were used to measure the containing radioactivity by a scintillation counter.

Oligo(dT) binding buffer:

reagent	MW	concentration	required for 100 mL
KCl	74.6	1 M	7.46 g
Tris	121.14	100 mM	1.21 g
Triton X-100	643.37	0.25% (v/v)	250 µL

pH was adjusted to 8.0 with NaOH/HCl.

Oligo(dT) wash buffer:

reagent	MW	concentration	required for 100 mL
KCl	74.6	100 mM	746 mg
Tris	121.14	10 mM	121 mg
Triton X-100	643.37	0.25% (v/v)	250 µL

pH was adjusted to 8.0 with NaOH/HCl.

The amount of mRNA displayed proteins in the oligo(dT) elution was calculated with the following equation;

$$\text{mRNA-displayed proteins [mol]} = (C_{Met} \cdot V \cdot P) / N$$

C_{Met} = total concentration of hot and cold methionine in the translation reaction mixture
V = the translation reaction volume
P = the counts in the oligo(dT) elution / the counts in the translation mixture
N = the average number of methionine in a single displayed protein

4.10.4. Ni-NTA Chromatography

Ni-NTA affinity chromatography was used to isolate mRNA-protein fusions carrying His-6 tag. Ni^{2+} ions immobilized on agarose can form the complexes with the histidine residues of the proteins. Because His-6 tag was placed at C-terminus of protein library, those proteins that do not carry a His-6 tag e.g. through the presence of internal stop codons or immature translation were not captured on Ni-NTA agarose. The bound proteins were competitively eluted with the buffer including imidazol.

100–200 µL of Ni-NTA agarose was washed 3 times with 1 mL of water, and once with 1 mL of 1 x Ni-NTA binding buffer on PolyPrep chromatography column from Bio-Rad. The sample was diluted with 2 x Ni-NTA binding buffer, incubated for 60 min at 4°C, and filtered through the column membrane. The flow through was discarded. The resin was washed with following wash buffers;

2 times with 500 µL of Ni-NTA first wash buffer
1 times with 500 µL of Ni-NTA first wash buffer : Ni-NTA second wash buffer = 4 : 1
1 times with 500 µL of Ni-NTA first wash buffer : Ni-NTA second wash buffer = 3 : 2
1 times with 500 µL of Ni-NTA first wash buffer : Ni-NTA second wash buffer = 2 : 3
1 times with 500 µL of Ni-NTA first wash buffer : Ni-NTA second wash buffer = 1 : 4

Elution was done by incubating the agarose in 2 times with 250 µL of Ni-NTA elution buffer at 4°C for 30 min each while shaking. Finally 0.5 M of EDTA was added to a final concentration of 5 mM to chelate eluted Ni^{2+} ions. The eluate was desalted on NAP column twice to remove denaturant. 1–2 % samples of each fraction were analyzed by tricine SDS PAGE and scintillation counting for radioactivity.

Denaturing 2 x Ni-NTA binding buffer:

reagent	MW	concentration	required for 100 mL
Guanidine HCl	95.5	6 M	57.4 g
NaCl	58.4	1 M	5.86 g
$Na_2HPO_4 \cdot 2H_2O$	177.99	200 mM	3.56 g
Tris	121.14	20 mM	242 mg
Triton X-100	643.37	0.5 % (v/v)	500 µL
β-Mercaptoethanol	78.1	200 mM	1402 µL

pH was adjusted to 8.0 with NaOH/HCl.

Denaturing 1 x Ni-NTA binding buffer:

reagent	MW	concentration	required for 100 mL
Guanidine HCl	95.5	3 M	28.7 g
NaCl	58.4	500 mM	2.93 g
$Na_2HPO_4 \cdot 2H_2O$	177.99	100 mM	1.78 g
Tris	121.14	10 mM	121 mg
Triton X-100	643.37	0.25 % (v/v)	250 μL
β-Mercaptoethanol	78.1	100 mM	701 μL

pH was adjusted to 8.0 with NaOH/HCl.

Denaturing Ni-NTA first wash buffer:

reagent	MW	concentration	required for 100 mL
Guanidine HCl	95.5	3 M	28.7 g
Urea	60.1	8 M	48.1 g
NaCl	58.4	500 mM	2.93 g
$NaH_2PO_4 \cdot 2H_2O$	156.01	100 mM	1.56 g
Tris	121.14	10 mM	121 mg
Triton X-100	643.37	0.25 % (v/v)	250 μL
β-Mercaptoethanol	78.1	100 mM	701 μL

pH was adjusted to 6.3 with NaOH/HCl.

Ni-NTA second wash buffer:

reagent	MW	concentration	required for 100 mL
Guanidine HCl	95.5	3 M	28.7 g
NaCl	58.4	500 mM	2.93 g
Tris	121.14	10 mM	121 mg
Triton X-100	643.37	0.25 % (v/v)	250 μL
β-Mercaptoethanol	78.1	100 mM	701 μL

pH was adjusted to 8.0 with NaOH/HCl.

Ni-NTA elution buffer:

reagent	MW	concentration	required for 100 mL
Guanidine HCl	95.5	3 M	28.7 g
NaCl	58.4	500 mM	2.93 g
Tris	121.14	10 mM	121 mg
Triton X-100	643.37	0.25 % (v/v)	250 µL
Imidazol	68.1	250 mM	1.70 g
β-Mercaptoethanol	78.1	100 mM	701 µL

pH was adjusted to 8.0 with NaOH/HCl.

4.10.5. FLAG Tag Chromatography

The anti-FLAG M2 antibody immobilized on agarose binds with high affinity to the FLAG octapeptide sequences (DYKDDDDK). This amino acid sequences are encoded in N-terminus of protein library sequences, thus mRNA-protein fusion molecules with full length proteins were isolated by FLAG tag purification after Ni-NTA purification. Proteins resulting from internal translation initiation were discarded due to the lack of FLAG peptide.

100 µL of anti-FLAG M2 agarose was washed with 3 times with 1 mL of FLAG tag clean buffer, then equilibrated 2 times with 1 mL of 1 x FLAG binding buffer on a PolyPrep chromatography column from Bio-Rad. The elution from Ni-NTA purification was diluted with 2 x FLAG tag binding buffer, added to the resin, and incubated at 4°C for 1 h while shaking. The resin was washed 3 times with 1 mL of 1 x FLAG binding buffer. The bound proteins were eluted 2 times of 500 µL FLAG tag binding buffer containing 10 µM FLAG peptide at 4°C for 30 min each while shaking. 1–2 % samples of each fraction were analyzed by tricine SDS PAGE and scintillation counting for radioactivity. Eluates were desalted on NAP columns (Pharmacia) to exchange the buffer into water.

FLAG tag clean buffer:

reagent	MW	concentration	required for 100 mL
Glycine	75.1	100 mM	751 mg
Triton X-100	643.37	0.25 % (v/v)	250 µL

pH was adjusted to 3.5 with NaOH/HCl.

2 x FLAG tag binding buffer:

reagent	MW	concentration	required for 100 mL
NaCl	75.1	300 mM	1.75 g
HEPES	238	100 mM	2.38 g
Triton X-100		0.5 % (v/v)	500 µL

pH was adjusted to 8.0 with NaOH/HCl.

4.10.6. Desalting on NAP Column

For desalting or buffer exchange of the samples, NAP columns (NAP-5, 10, 25) from Pharmacia were used. The samples can be eluted from the pre-equilibrated columns with water or desired buffer. According to the principle of gel filtration, small molecules and salts tend to stay longer on the column, while the larger molecules like proteins or nucleic acids flow through the column. Particularly, the NAP column was blocked with salmon sperm DNA und BSA prior to desalting of mRNA-displayed proteins.

As an example, to a NAP-5 column 10 mL of water, 1 mL of salmon sperm DNA and BSA solutions, and further 10 mL of water were added. 0.5 mL of the sample solution was loaded on the top of the column. Finally the sample was eluted with 1 mL of water.

Salmon sperm DNA und BSA:

reagent	concentration	required for 1 mL
Salmon sperm DNA (10 mg/mL)	1 mg/mL	100 µL
BSA (10 mg/mL)	0.1 mg/mL	10 µL
H_2O		890 µL

4.10.7. Reverse Transcription

Reverse transcription creates a DNA/RNA-hybrid from a single strand RNA template and a RT primer which is necessary to give a starting double stranded region. DNA/RNA double strand formation also prevents the formation of tertiary RNA structures to avoid the selection of RNA aptamers.

The following mix I and II were prepared on ice. Mix I was incubated for 10 min at 42°C prior to addition of mix II. Mix I and II were mixed, and incubated for 50 min at

42°C. 2 % of samples before and after the reaction were analyzed by tricine SDS PAGE to confirm the DNA/mRNA-protein fusion formation. The sample was desalted on pre-equilibrated NAP column.

Reverse transcription reaction mix I (50 µL):

reagent	amount	stock	resultant
RT buffer	20 µL	5 x	1 x
DTT	10 µL	100 mM	10 mM
dNTP mix	0.8 µL	25 mM	0.2 mM
Superscript II RT	2 µL	200 U/µL	4 U/µL
Water	17.2 µL		

Reverse transcription reaction mix II (50 µL):

reagent	amount	stock	resultant
RT primer	2 µL	100 µM	2 µM
RNA template	x µL	x µM	1 µM
Water	y µL		

4.11. *in vitro* Selection using mRNA displayed proteins

The approaches for three different selections targeted the small molecules, cibacron blue 3GA, NADP, and CoA, are depicted in succeeding sections. The former two selections applied competitive elution to obtain bound species, while the latter used non-competitive elution and cleaving the target molecule from a resin. The percentage of bound proteins to the input library proteins were determined from the number of the radioactive counts measured by scintillation counting. Theoretically, cycles of selection and amplification should be repeated until the proportion of the resultant mRNA-displayed proteins in the selected fraction is no longer increasing.

Approximate input library diversity for the initial selection was estimated from the amount of mRNA-displayed proteins after oligo(dT) purification and the yield of protein tag affinity chromatography.

4.11.1. Selection against Cibacron Blue 3GA

The desalted samples of DNA/mRNA-protein fusion library were diluted at the ratio of 1:1 with 2 x cibacron blue selection buffer. In the first cycle of the selection, 200 µL of cibacron blue 3GA agarose was used. In the subsequent selections, 100 µL of cibacron blue 3GA agarose was washed 3 times with 500 µL of 1 x cibacron blue selection buffer, and then incubated with mRNA displayed proteins for 1 h at 4°C while shaking. After drain for flow through, the remaining agarose was washed with 3 to 30 column volumes of cibacron blue selection buffer. The bound molecules were eluted 4 times with 250 µL of cibacron blue elution buffer including NAD^+ at 4°C for 15 min each while shaking. Elution fractions were pooled, and desalted on pre-equilibrated NAP column. 2 % of samples from every fraction were analyzed by scintillation counter.

Desalted samples (through employment of NAP columns) were ethanol precipitated, and the resulting pellet was resuspended in 100 µL of water. 20 µL aliquots of this solution were used for test PCR (100 µL) to optimize the cycle number for amplification of the corresponding DNA fragment.

2 x Cibacron blue selection buffer:

reagent	MW	concentration	required for 1 L
Tris	121.14	40 mM	4.84 g
NaCl	58.44	100 mM	5.84 g
KCl	74.54	100 mM	7.46 g
$MgCl_2$	203.30	20 mM	4.06 g
0.5 M EDTA, pH 8.0		800 µM	1600 µL

pH was adjusted to 7.5 with NaOH/HCl.

Cibacron blue elution buffer:

reagent	MW	concentration	required for 200 mL
Tris	121.14	20 mM	484 mg
NaCl	58.44	50 mM	584 mg
KCl	74.54	50 mM	746 mg
$MgCl_2$	203.30	10 mM	406 mg
0.5 M EDTA, pH 8.0		400 µM	160 µL
NAD	663.4	5 mM	665.4 mg

pH was adjusted to 7.5 with NaOH/HCl.

4.11.2. Selection against NADP

The desalted samples of DNA/mRNA-protein fusion library were diluted at the ratio of 1:1 with 2 x NADP selection buffer. In the first cycle of the selection, pre-selection using 200 µL of adipicacid dihydrazide agarose was done to exclude the matrix binders, prior to the incubation with 200 µL of NADP agarose. In the subsequent cycles, 100 µL of NADP agarose was washed 3 times with 1 mL of 1 x NADP selection buffer, then incubated with mRNA displayed proteins for 1 h at 4°C while shaking. After drain for flow through, the remaining agarose was washed with 10 to 30 column volumes of NADP selection buffer. The bound molecules were eluted 4 times with 250 µL of NADP elution buffer including 5 mM of NADP at 4°C for 15 min each while shaking. Elution fractions were pooled, and desalted on pre-equilibrated NAP column. 2 % of samples from every fraction were analyzed by scintillation counting.

Desalted samples (through employment of NAP columns) were ethanol precipitated, and the resulting pellet was resuspended in 100 µL of water. 20 µL aliquots of this solution were used for test PCR (100 µL) to optimize the cycle number for amplification of the corresponding DNA fragment.

5 x NADP selection buffer:

reagent	MW	concentration	required for 1 L
HEPES	121.14	100 mM	23.83 g
NaCl	58.44	250 mM	14.61 g
KCl	74.54	250 mM	18.64 g
$MgCl_2$	203.30	50 mM	10.17 g
0.5 M EDTA, pH 8.0		2000 µM	4 mM
Triton X-100		0.5 %	5 mL

pH was adjusted to 7.5 with NaOH/HCl.

NADP elution buffer:

reagent	MW	concentration	required for 50 mL
HEPES	121.14	20 mM	238.3 mg
NaCl	58.44	50 mM	146.1 mg
KCl	74.54	50 mM	186.4 mg
MgCl$_2$	203.30	10 mM	101.7 mg
0.5 M EDTA, pH 8.0		400 µM	40 µM
Triton X-100		0.1 %	50 µL
NADP	787.38	5 mM	196.8 mg

pH was adjusted to 7.5 with NaOH/HCl.

2 x NADP elution buffer:

reagent	MW	concentration	required for 20 mL
5 x NADP selection buffer			8 mL
NADP	787.38	10 mM	157.5 mg

4.11.3. Immobilization of Coenzyme A to Thiopropyl Sepharose

Coenzyme A was coupled to the sepharose via the thiol group. Thiopropyl sepharose 6B purchased from Amersham contains reactive 2-thiopyridyl disulphide group attached to sepharose through a chemically stable ether linkage. The thiopropyl sepharose reacts with thiol groups under mild conditions to form mixed disulphides, and release 2-thiopyridone which can be detected by its absorbance (absorbance coefficient at 343 nm = 8.08×10^3 M^{-1}cm^{-1}, this value is approximately equivalent at 260 nm).

The gel slurry was prepared by recommended procedure from supplier. 500 µL of the gel (1000 µL of the slurry) was incubated in 500 µL of TPS buffer containing 0.01 M CoA at 4°C overnight. After drain, the gel was washed with TPS buffer until no absorbance at 343 nm was observed. Approximate amounts of released 2-thiopyridone and non-coupled CoA (absorbance coefficient at 260 nm = 16.4×10^3 M^{-1}cm^{-1}) were quantified by spectro-photometrical measuring of absorbance. The non-reacted thiopyridyl group on sepharose was removed by incubation at 4°C for 30 min in TPS buffer containing 2-mercaptoehtanol. The concentration of 2-mercaptoethanol was adjusted to the amount of the remaining thiopyridyl group. The sepharose was then intensively washed with TPS buffer to equilibrate the gel, and the column was stored in 1:1 ratio in TPS buffer. Additionally, the amount of CoA coupled to the sepharose was

quantified by eluting CoA from the small amount of the prepared gel (50 µL) with 50 mM 2-mercaptoethanol in TPS buffer.

TPS buffer:

reagent	MW	concentration	required for 500 mL
Tris	121.14	0.1 M	6.06 g
NaCl	58.44	0.5 M	14.61 g
0.5 M EDTA, pH 8.0		1 mM	1 mL

pH was adjusted to 7.5 with NaOH/HCl.

4.11.4. Selection against Coenzyme A

The selections targeted coenzyme A were operated in two fashions in parallel; one in the absence of guanidine hydrochloride, and the other in the presence of guanidine hydrochloride. The aim to make a slightly denaturing condition by adding 0.5 M of guanidine hydrochloride in the selection buffer and elution buffer was not only to confer better solubility on the proteins but to force to select proteins which are folded stably.

The desalted samples of DNA/mRNA-protein fusion library were diluted at the ratio of 1:1 with 2 x CoA selection buffer. In first cycle of the selection, pre-selection using 100 µL of 2-mercaptoehanol-blocked thiopropyl sepharose was done to exclude matrix binders, prior to the incubation with 100 µL of CoA immobilized sepharose. In the subsequent cycles, 20 µL of CoA sepharose was washed 3 times with 1 mL of 1 x CoA selection buffer, and then incubated with mRNA displayed proteins for 1 h at 4°C while shaking. After drain for flow through, the remaining sepharose was washed with 8 column volumes of CoA selection buffer. The bound molecules were eluted 6 times with 500 µL of CoA elution buffer including DTT by cleaving disulfide bond. Elution fractions were pooled, and desalted on pre-equilibrated NAP column. 1 % of samples from every fraction were analyzed by scintillation counting.

Desalted samples (through employment of NAP columns) were ethanol precipitated, and the resulting pellet was resuspended in 200 µL of water. 5 µL aliquots of this solution were used for test PCR (100 µL) to optimize the cycle number for amplification of the corresponding DNA fragment.

5 x CoA selection buffer:

reagent	MW	concentration	required for 1 L
HEPES	121.14	100 mM	23.83 g
NaCl	58.44	250 mM	14.61 g
KCl	74.54	250 mM	18.64 g
MgCl$_2$	203.30	50 mM	10.17 g
0.5 M EDTA, pH 8.0		2000 μM	4 mM
Triton X-100		0.5 %	5 mL

pH was adjusted to 7.5 with NaOH/HCl.

Denaturing CoA selection buffer:

reagent	MW	concentration	required for 50 mL
5 x CoA selection buffer			10 mL
Guanidine HCl	95.33	0.5 M	2.39 g

CoA elution buffer:

reagent	MW	concentration	required for 50 mL
5 x CoA selection buffer			10 mL
DTT	154.3	50 mM	385.8 mg

Denaturing CoA elution buffer:

reagent	MW	concentration	required for 50 mL
5 x CoA selection buffer			10 mL
Guanidine HCl	95.33	0.5 M	2.39 g
DTT	154.3	50 mM	385.8 mg

4.12. Synthesis

All reagents are commercially available and used without further purification. Solvents are purchased over molecular sieves (Fluka) and used directly without further purification unless otherwise noted. All reactions were conducted under rigorous exclusion of air and moisture. Chemical shift (δ) are given in ppm with tetramethylsilane as an internal standard, and coupling constants (*J*) are given in hertz (Hz).

NMR spectra: Bruker (DPX 300, DPX 400 und DPX 500) with the solvent peak as internal standard.
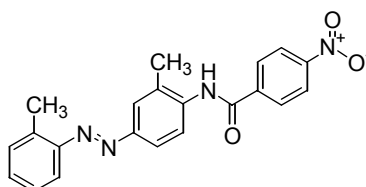
EI-MS: Finningan MAT 95XL.

FAB MS: Concept 1H using 3-nitrobenzyl alcohol (NBA) as a matrix.

Flash chromatography: Merck silica gel G60 (230–400 mesh).

Thin-layer chromatography: Merck precoated plates (silica gel 60 F_{254}).

4.12.1. Synthesis of 28F6

**(*E*)-2-(4-*p*-Nitrobenzamido-3-methylphenyl)-1-*o*-tolyldiazene (28F6)**

To a solution of o-aminoazotoluene (**1**) (50 mg, 0.222 mmol) in CH_2Cl_2 (1 mL), dried triethylamine (46.2 µL, 0333 mmol) was added at 0°C. Then, 4-nitrobenzoylchloride (**2**) (45.3 mg, 0.244 mmol) in CH_2Cl_2 (1 mL) was added to this mixture dropwise at 0°C. After stirring at 0°C for 15 min, the reaction mixture was subsequently allowed to warm to room temperature, and continued to stir for 4 h. The reaction was quenched with saturated $NaHCO_3$ solution, and extracted. The organic phase was washed with brine, dried over $MgSO_4$, concentrated, and purified by flash column chromatography (ethyl acetate/cyclohexane 1:19 → 1:3) to yield **28F6** (15 mg, 18 %) as a orange solid.

R_f = 0.54 (ethyl acetate/cyclohexane 1:3); ^{1}H-NMR (400 MHz, DMSO-d_6): δ 2.39 (s, 3H), 2.69 (s, 3H), 7.30-7.45 (m, 1H), 7.43-7.45 (m, 2H), 7.57 (d, *J* = 7.83 Hz, 1H), 7.66 (d, *J* = 8.46 Hz, 1H),   7.78 (dd, *J* = 8.46, 2.15 Hz, 1H), 7.84 (d, *J* = 1.77 Hz, 1H), 8.21-8.23 (m, 2H), 8.37-8.40 (m, 2H), 10.3 (s, 1H); ^{13}C-NMR (100 MHz, DMSO-d_6): δ

17.1, 18.0, 115.0, 120.2, 123.6, 124.8, 126.6, 126.7, 129.3, 131.2, 131.4, 134.3, 137.5, 138.8, 140.0, 149.3, 150.0, 150.1, 163.9; MS (EI): 374.1 [M]$^+$

4.12.2. Synthesis of 3E4

2-Chloro-5-nitro-*N*-(1-phenylethyl)benzamide (5)
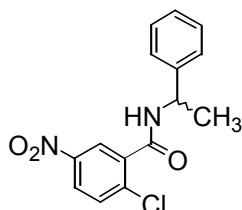
To a stirred solution of 1-chloro-5-nitrobenzoic acid (**3**) (1.01 g, 5.0 mmol) and *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (1.15 g, 6.0 mmol) in DMF (10 mL) was added (±)-α-methylbenzylamine (**4**) (645 μL, 5.0 mmol), and the mixture wasstirred at room temperature for 19 h. After the mixture was concentrated in vacuo, the residue was diluted in CH_2Cl_2, extracted with saturated $NaHCO_3$ solution. The combined extracts were washed with brine, dried over $MgSO_4$, concentrated and purified by column chromatography on silica gel (ethyl acetate/cyclohexane 1:3 → 4:1) to yield **5** (526 mg, 35 %) as a white solid.

R_f = 0.38 (ethyl acetate/cyclohexane 1:3); ^{1}H-NMR (400 MHz, CD_3OD): δ 1.57 (d, *J* = 6.19 Hz, 3H), 5.26 (m, 1H), 6.40 (d, *J* = 6.44 Hz, 1H), 7.19-7.34 (m, 5H), 7.49 (d, *J* = 8.84 Hz, 1H), 8.10 (dd, *J* = 8.72, 2.78 Hz, 1H), 8.39 (d, *J* = 2.78 Hz, 1H); ^{13}C-NMR (100 MHz, CD_3OD): δ 21.7, 50.2, 125.4, 125.8, 126.5, 128.0, 129.1, 131.6, 136.7, 137.7, 142.3, 146.8, 163.5; MS (EI): 304.1 [M]$^+$

**5-Nitro-*N*-(1-phenylethyl)-2-(piperidin-1-yl)benzamide (6)[157]**
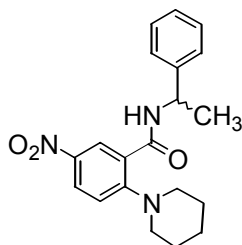
A mixture of **5** (506 mg, 1.66 mmol) and piperidine (197 μL, 1.99 mmol) in dry pyridine (5 mL) was refluxed for 24 h. Solvent was removed from the reaction mixture by evaporation, and the residue was diluted in CH_2Cl_2, extracted, and washed with 0.5

M HCl (40 mL). The combined extracts were dried over $MgSO_4$, concentrated and purified by column chromatography on silica gel (CH_2Cl_2/ethyl acetate 50:1 → 10:1) to yield **6** (518 mg, 88 %) as a yellow solid.

R_f = 0.25 (CH_2Cl_2/ethyl acetate 50:1); ^{1}H-NMR (400 MHz, CD_3OD): δ 1.32-1.52 (m, 6H), 1.57 (d, *J* = 6.82 Hz, 3H), 2.87-2.90 (m, 4H), 5.22-5.29 (m, 1H), 7.15 (d, *J* = 8.97 Hz, 1H), 7.22 (tt, *J* = 7.05, 1.66 Hz, 1H), 7.27-7.35 (m, 4H), 8.13 (dd, *J* = 8.97, 2.91 Hz, 1H), 8.81 (d, *J* = 2.91 Hz, 1H), 9.00 (d, *J* = 6.57 Hz, 1H); ^{13}C-NMR (100 MHz, CD_3OD): δ 22.0, 23.6, 25.8, 49.9, 54.6, 120.4, 126.7, 126.8, 127.5, 127.9, 128.8, 129.1, 143.1, 143.6, 157.4, 164.2; MS (EI): 353.1 $[M]^+$

5-Amino-*N*-(1-phenylethyl)-2-(piperidin-1-yl)benzamide (7)
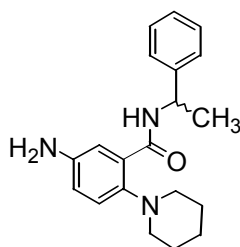
6 (493 mg, 1.40 mmol) was dissolved in methanol (7 mL) under an argon atmosphere and 10% Pd-C (250 mg) was added. The reaction mixture was subsequently floated with hydrogen applying a balloon. After stirring under a hydrogen atmosphere for 24 h, the reaction mixture was filtered through celite, and the celite was washed with ethanol, methanol, and CH_2Cl_2. The collected filtrate was evaporated and purified by column chromatography on silica gel (CH_2Cl_2/methanol 30:1 →25:1) to yield **7** (403 mg, 89 %) as a yellow solid.

R_f = 0.29 (CH_2Cl_2/methanol 25:1); ^{1}H-NMR (400 MHz, CD_3OD): δ 1.36 (bs, 4H), 1.46-1.51 (m, 2H), 1.54 (d, *J* = 6.95 Hz, 3H), 2.70 (bs, 4H), 3.62 (bs, 2H), 5.23 (tq, *J* = 7.07, 6.95 Hz, 1H), 6.63 (dd, *J* = 8.46, 2.91 Hz, 1H), 6.99 (d, *J* = 8.46 Hz, 1H), 7.14-7.34 (m, 5H), 7.53 (d, *J* = 2.91 Hz, 1H), 11.2 (d, *J* = 6.44 Hz, 1H); ^{13}C-NMR (100 MHz, CD_3OD): δ 22.2, 23.8, 26.4, 49.2, 55.1, 117.3, 117.9, 118.1, 122.9, 126.5, 127.1, 128.6, 144.0, 144.1, 144.1, 165.4; MS (EI): 323.2 $[M]^+$

1-(3-(1-Phenylethylcarbamoyl)-4-(piperidin-1-yl)phenyl)-3-(3-chloro-4-fluoro-phenyl)urea (3E4)[158]
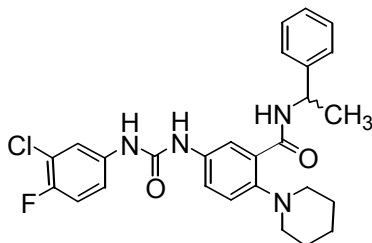
3-Chloro-4-fluorophenylisocyanate (160 µL, 1.29 mmol) and **7** (398 mg, 1.23 mmol) were dissolved in dry THF, and stirred at room temperature for 18 h. The mixture was concentrated and purified by 2 times of column chromatography on silica gel (CH_2Cl_2/methanol 1:0 →50:1, then ethyl acetate/cyclohexane 0:1 →2:1) to yield **3E4** (397 mg, 65 %) as a yellow solid.

R_f = 0.20 (CH_2Cl_2/methanol 25:1); ^{1}H-NMR (400 MHz, CD_3OD): δ 1.35-1.50 (m, 6H), 1.53 (d, J = 6.95 Hz, 3H), 2.72-2.75 (m, 4H), 5.15 (dq, J = 7.33, 7.01 Hz, 1H), 6.85 (t, J = 8.84 Hz, 1H), 7.01-7.05 (m, 1H), 7.15-7.24 (m, 6H), 7.41 (dd, J = 6.57, 2.65 Hz, 1H), 7.80 (d, J = 2.65 Hz, 1H), 7.95 (dd, J = 8.72, 2.65 Hz, 1H), 8.42 (d, J = 9.09 Hz, 2H), 11.64 (d, J = 7.33 Hz, 1H); ^{13}C-NMR (100 MHz, CD_3OD): δ 22.6, 23.7, 26.3, 49.9, 55.1, 116.3 (d, ^6J = 21.97 Hz), 119.3, 119.4, 120.7 (d, ^2J = 18.37 Hz), 121.7 (d, ^5J = 5.79 Hz), 123.1, 125.1, 126.1, 127.0, 127.6, 128.8, 135.9 (d, ^4J = 3.00 Hz), 136.6, 142.6, 147.8, 153.6, 153.9 (d, ^1J = 243.86 Hz), 166.3; MS (FAB): 495.2 [M+H]$^+$

4.12.3. Synthesis of 2E10

1-(4-(4-Nitrophenyl)piperazin-1-yl)-2-o-tolylethanone (10)[159]
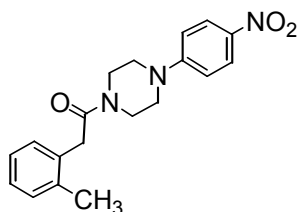
To a mixture of *o*-tolylacetic acid (**8**) (751 mg, 5.0 mmol) and *N*-hydroxybenzotriazole (1.36 g, 10.1 mmol) in DMF (7 mL) was added a solution of *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (958 mg, 5.0 mmol) in DMF (6 mL). The mixture was stirred at room temperature for 2 h, and a solution of 1-(4-nitrophenyl)piperazine (**9**) (1.04 g, 5.0 mmol) in CH_2Cl_2 (4 mL) was added. Then,

the resulting mixture was stirred at room temperature for 24 h, and concentrated under reduced pressure. The residue was distributed between CH_2Cl_2 and saturated $NaHCO_3$ solution. Phases were separated, washed with NaCl, dried over $MgSO_4$, and concentrated. The residue was purified by column chromatography on silica gel (CH_2Cl_2/methanol 1:0 →40:1) to yield **10** (1.38 g, 82 %) as a orange solid.

R_f = 0.23 (CH_2Cl_2/methanol 50:1); ^{1}H-NMR (300 MHz, CD_3OD): δ 2.23 (s, 3H), 3.24-3.78 (m, 8H), 3.66 (s, 2H), 6.69-6.74 (m, 2H), 7.04-7.13 (m, 4H), 8.02-8.07 (m, 2H); ^{13}C-NMR (75 MHz, CD_3OD): δ 19.9, 38.8, 41.3, 45.3, 47.1, 113.2, 126.1, 126.6, 127.4, 128.8, 130.7, 133.4, 136.5, 139.3, 154.5, 170.0; MS (EI): 339.1 [M]$^+$

1-(4-(4-Aminophenyl)piperazin-1-yl)-2-o-tolylethanone (11)
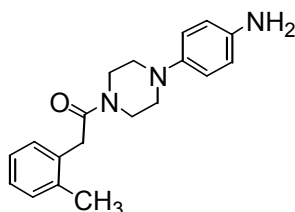
A mixture of **10** (1.02 g, 3.0 mmol) and 10% Pd-C (500 mg) in ethanol (10 mL) was stirred at room temperature for 18 h under a hydrogen atmosphere. Then, the mixture was filtered through celite, and the celite was washed with ethanol, methanol, and CH_2Cl_2. The collected filtrate was concentrated and purified by column chromatography on silica gel (ethyl acetate/cyclohexane 3:1 →1:0) to yield **11** (796 mg, 86 %) as a brawn solid.

R_f = 0.23 (ethyl acetate); ^{1}H-NMR (400 MHz, CD_3OD): δ 2.22 (s, 3H), 2.81 (t, J = 5.12 Hz, 2H), 2.93 (t, J = 4.93 Hz, 2H), 3.30 (bs, 2H), 3.64 (s, 2H), 3.75 (t, J = 5.24 Hz, 2H), 6.54-6.58 (m, 2H), 6.68-6.72 (m, 2H), 7.06-7.11 (m, 4H); ^{13}C-NMR (100 MHz, CD_3OD): δ 19.9, 38.7, 42.1, 46.3, 51.3, 51.6, 116.4, 119.4, 126.4, 127.2, 128.9, 130.5, 133.8, 136.4, 141.1, 144.0, 169.0; MS (EI): 309.2 [M]$^+$

1-(4-Nitrophenyl)-3-(4-(4-(2-o-tolylacetyl)piperazin-1-yl)phenyl)urea (2E10)[158]
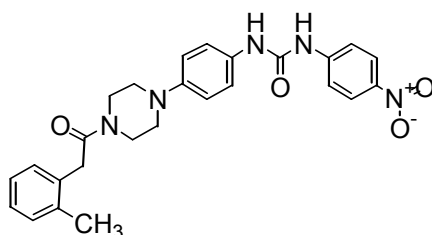
11 (808 mg, 2.61 mmol) and 4-nitrophenyl isocyanate (450 mg, 2.74 mmol) were dissolved in THF/DMSO (28 mL/1.5 mL), and stirred at room temperature for 18 h. After removal of THF under reduced pressure, the product was precipitated by addition of CH_2Cl_2, filtered off, washed with CH_2Cl_2 and ethyl acetate. Column chromatography on silica gel (CH_2Cl_2/methanol 1:0 →30:1) to yield **2E10** (1.01 g, 82 %) as a colorless solid.

R_f = 0.42 (CH_2Cl_2/methanol 25:1); ^{1}H-NMR (400 MHz, DMSO-d_6): δ 2.20 (s, 3H), 3.04 (t, J = 5.24 Hz, 4H), 3.61-3.64 (m, 4H), 3.72 (s, 2H), 6.90-6.93 (m, 2H), 7.06-7.16 (m, 4H), 7.31-7.35 (m, 2H), 7.65-7.69 (m, 2H), 8.15-8.18 (m, 2H), 8.65 (s, 1H), 9.31 (s, 1H); ^{13}C-NMR (100 MHz, DMSO-d_6): δ 19.2, 37.4, 41.1, 45.2, 49.0, 49.4, 116.6, 117.2, 120.0, 125.1, 125.6, 126.4, 129.3, 129.7, 131.1, 134.7, 136.6, 140.7, 146.6, 146.6, 152.0, 168.6; MS (EI): 473.2 $[M]^+$, 474.2 $[M+H]^+$

5. References

[1]   A. D. Ellington, J. W. Szostak. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818.

[2]   C. Tuerk, L. Gold. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **1990**, *249*, 505.

[3]   W. Sneader. *Drug Prototypes and their Exploitation*, Wiley, London, **1996**.

[4]   C. Lipinski, A. Hopkins. Navigating chemical space for biology and medicine. *Nature* **2004**, *432*, 855.

[5]   H. Lin, V. W. Cornish. Screening and selection methods for large-scale analysis of protein function. *Angew Chem Int Ed Engl* **2002**, *41*, 4402.

[6]   S. J. Klug, M. Famulok. All you wanted to know about SELEX. *Mol Biol Rep* **1994**, *20*, 97.

[7]   L. Gold, B. Polisky, O. Uhlenbeck, M. Yarus. Diversity of oligonucleotide functions. *Annu Rev Biochem* **1995**, *64*, 763.

[8]   L. Gold. Conformational properties of oligonucleotides. *Nucleic Acids Symp Ser* **1995**, 20.

[9]   D. S. Wilson, J. W. Szostak. In vitro selection of functional nucleic acids. *Annu Rev Biochem* **1999**, *68*, 611.

[10]   S. V. Taylor, P. Kast, D. Hilvert. Investigating and Engineering Enzymes by Genetic Selection. *Angew Chem Int Ed Engl* **2001**, *40*, 3310.

[11]   B. R. Stockwell. Exploring biology with small organic molecules. *Nature* **2004**, *432*, 846.

[12]   G. F. Joyce. The antiquity of RNA-based evolution. *Nature* **2002**, *418*, 214.

[13]   P. B. Moore, T. A. Steitz. The involvement of RNA in ribosome function. *Nature* **2002**, *418*, 229.

[14]   J. A. Doudna, T. R. Cech. The chemical repertoire of natural ribozymes. *Nature* **2002**, *418*, 222.

[15]   G. J. Hannon. RNA interference. *Nature* **2002**, *418*, 244.

[16]   R. R. Breaker. Natural and engineered nucleic acids as tools to explore biology. *Nature* **2004**, *432*, 838.

[17]   B. A. Sullenger, E. Gilboa. Emerging clinical applications of RNA. *Nature* **2002**, *418*, 252.

[18]   E. A. Doherty, J. A. Doudna. Ribozyme structures and mechanisms. *Annu Rev Biochem* **2000**, *69*, 597.

[19]   E. A. Doherty, J. A. Doudna. Ribozyme structures and mechanisms. *Annu Rev Biophys Biomol Struct* **2001**, *30*, 457.

[20]   D. M. Long, O. C. Uhlenbeck. Kinetic characterization of intramolecular and intermolecular hammerhead RNAs with stem II deletions. *Proc Natl Acad Sci U S A* **1994**, *91*, 6977.

[21]   T. Tuschl, F. Eckstein. Hammerhead ribozymes: importance of stem-loop II for activity. *Proc Natl Acad Sci U S A* **1993**, *90*, 6991.

[22]   M. J. Fedor, O. C. Uhlenbeck. Kinetics of intermolecular cleavage by hammerhead ribozymes. *Biochemistry* **1992**, *31*, 12042.

[23]   D. E. Ruffner, G. D. Stormo, O. C. Uhlenbeck. Sequence requirements of the hammerhead RNA self-cleavage reaction. *Biochemistry* **1990**, *29*, 10695.

[24]   H. W. Pley, K. M. Flaherty, D. B. McKay. Three-dimensional structure of a hammerhead ribozyme. *Nature* **1994**, *372*, 68.

[25]   W. G. Scott, J. T. Finch, A. Klug. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **1995**, *81*, 991.

[26]   C. Hammann, D. M. Lilley. Folding and activity of the hammerhead ribozyme. *Chembiochem* **2002**, *3*, 690.

[27]   M. Zivarts, Y. Liu, R. R. Breaker. Engineered allosteric ribozymes that respond to specific divalent metal ions. *Nucleic Acids Res* **2005**, *33*, 622.

[28]   Y. Komatsu, S. Yamashita, N. Kazama, K. Nobuoka, E. Ohtsuka. Construction of new ribozymes requiring short regulator oligonucleotides as a cofactor. *J Mol Biol* **2000**, *299*, 1231.

[29]   A. Ferguson, R. M. Boomer, M. Kurz, S. C. Keene, J. L. Diener, A. D. Keefe, C. Wilson, S. T. Cload. A novel strategy for selection of allosteric ribozymes yields RiboReporter sensors for caffeine and aspartame. *Nucleic Acids Res* **2004**, *32*, 1756.

[30]   M. Koizumi, J. N. Kerr, G. A. Soukup, R. R. Breaker. Allosteric ribozymes sensitive to the second messengers cAMP and cGMP. *Nucleic Acids Symp Ser* **1999**, 275.

[31]   J. Tang, R. R. Breaker. Rational design of allosteric ribozymes. *Chem Biol* **1997**, *4*, 453.

[32]   J. S. Hartig, M. Famulok. Reporter ribozymes for real-time analysis of domain-specific interactions in biomolecules: HIV-1 reverse transcriptase and the primer-template complex. *Angew Chem Int Ed Engl* **2002**, *41*, 4263.

[33]   D. Y. Wang, D. Sen. Rationally designed allosteric variants of hammerhead

ribozymes responsive to the HIV-1 Tat protein. *Comb Chem High Throughput Screen* **2002**, *5*, 301.

[34]   M. P. Robertson, A. D. Ellington. In vitro selection of nucleoprotein enzymes. *Nat Biotechnol* **2001**, *19*, 650.

[35]   M. P. Robertson, S. M. Knudsen, A. D. Ellington. In vitro selection of ribozymes dependent on peptides for activity. *Rna* **2004**, *10*, 114.

[36]   T. Hermann, D. J. Patel. Adaptive recognition by nucleic acid aptamers. *Science* **2000**, *287*, 820.

[37]   W. C. Winkler, R. R. Breaker. Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* **2005**, *59*, 487.

[38]   W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins, R. R. Breaker. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **2004**, *428*, 281.

[39]   K. M. Thompson, H. A. Syrett, S. M. Knudsen, A. D. Ellington. Group I aptazymes as genetic regulatory switches. *BMC Biotechnol* **2002**, *2*, 21.

[40]   S. Seetharaman, M. Zivarts, N. Sudarsan, R. R. Breaker. Immobilized RNA switches for the analysis of complex chemical and biological mixtures. *Nat Biotechnol* **2001**, *19*, 336.

[41]   N. K. Vaish, V. R. Jadhav, K. Kossen, C. Pasko, L. E. Andrews, J. A. McSwiggen, B. Polisky, S. D. Seiwert. Zeptomole detection of a viral nucleic acid using a target-activated ribozyme. *Rna* **2003**, *9*, 1058.

[42]   K. Kossen, N. K. Vaish, V. R. Jadhav, C. Pasko, H. Wang, R. Jenison, J. A. McSwiggen, B. Polisky, S. D. Seiwert. High-throughput ribozyme-based assays for detection of viral nucleic acids. *Chem Biol* **2004**, *11*, 807.

[43]   N. K. Vaish, F. Dong, L. Andrews, R. E. Schweppe, N. G. Ahn, L. Blatt, S. D. Seiwert. Monitoring post-translational modification of proteins with allosteric ribozymes. *Nat Biotechnol* **2002**, *20*, 810.

[44]   R. Penchovsky, R. R. Breaker. Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes. *Nat Biotechnol* **2005**, *23*, 1424.

[45]   J. Srinivasan, S. T. Cload, N. Hamaguchi, J. Kurz, S. Keene, M. Kurz, R. M. Boomer, J. Blanchard, D. Epstein, C. Wilson, J. L. Diener. ADP-specific sensors enable universal assay of protein kinase activity. *Chem Biol* **2004**, *11*, 499.

[46]   J. S. Hartig, S. H. Najafi-Shoushtari, I. Grune, A. Yan, A. D. Ellington, M. Famulok. Protein-dependent ribozymes report molecular interactions in real time. *Nat Biotechnol* **2002**, *20*, 717.

[47]   M. Famulok. Allosteric aptamers and aptazymes as probes for screening

approaches. *Curr Opin Mol Ther* **2005**, *7*, 137.

[48] M. Famulok, G. Mayer. Intramers and aptamers: applications in protein-function analyses and potential for drug screening. *Chembiochem* **2005**, *6*, 19.

[49] P. Burgstaller, A. Jenne, M. Blind. Aptamers and aptazymes: accelerating small molecule drug discovery. *Curr Opin Drug Discov Devel* **2002**, *5*, 690.

[50] J. S. Hartig, I. Grune, S. H. Najafi-Shoushtari, M. Famulok. Sequence-specific detection of MicroRNAs by signal-amplifying ribozymes. *J Am Chem Soc* **2004**, *126*, 722.

[51] S. H. Najafi-Shoushtari, G. Mayer, M. Famulok. Sensing complex regulatory networks by conformationally controlled hairpin ribozymes. *Nucleic Acids Res* **2004**, *32*, 3212.

[52] F. Barre-Sinoussi, J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, L. Montagnier. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **1983**, *220*, 868.

[53] R. C. Gallo, P. S. Sarin, E. P. Gelmann, M. Robert-Guroff, E. Richardson, V. S. Kalyanaraman, D. Mann, G. D. Sidhu, R. E. Stahl, S. Zolla-Pazner, J. Leibowitch, M. Popovic. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* **1983**, *220*, 865.

[54] A. D. Frankel, J. A. Young. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* **1998**, *67*, 1.

[55] B. G. Turner, M. F. Summers. Structural biology of HIV. *J Mol Biol* **1999**, *285*, 1.

[56] R. J. Pomerantz, D. L. Horn. Twenty years of therapy for HIV-1 infection. *Nat Med* **2003**, *9*, 867.

[57] L. Menendez-Arias. Targeting HIV: antiretroviral therapy and development of drug resistance. *Trends Pharmacol Sci* **2002**, *23*, 381.

[58] T. Imamichi. Action of anti-HIV drugs and resistance: reverse transcriptase inhibitors and protease inhibitors. *Curr Pharm Des* **2004**, *10*, 4039.

[59] E. De Clercq. New approaches toward anti-HIV chemotherapy. *J Med Chem* **2005**, *48*, 1297.

[60] C. F. Pereira, J. T. Paridaen. Anti-HIV drug development--an overview. *Curr Pharm Des* **2004**, *10*, 4005.

[61] Y. Pommier, A. A. Johnson, C. Marchand. Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov* **2005**, *4*, 236.

[62] M. Baba, O. Nishimura, N. Kanzaki, M. Okamoto, H. Sawada, Y. Iizawa, M. Shiraishi, Y. Aramaki, K. Okonogi, Y. Ogawa, K. Meguro, M. Fujino. A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity. *Proc Natl Acad Sci U S A* **1999**, *96*, 5698.

[63] G. A. Donzella, D. Schols, S. W. Lin, J. A. Este, K. A. Nagashima, P. J. Maddon, G. P. Allaway, T. P. Sakmar, G. Henson, E. De Clercq, J. P. Moore. AMD3100, a small molecule inhibitor of HIV-1 entry via the CXCR4 co-receptor. *Nat Med* **1998**, *4*, 72.

[64] B. Yang, L. Gao, L. Li, Z. Lu, X. Fan, C. A. Patel, R. J. Pomerantz, G. C. DuBois, H. Zhang. Potent suppression of viral infectivity by the peptides that inhibit multimerization of human immunodeficiency virus type 1 (HIV-1) Vif proteins. *J Biol Chem* **2003**, *278*, 6596.

[65] N. C. Gaddis, E. Chertova, A. M. Sheehy, L. E. Henderson, M. H. Malim. Comprehensive investigation of the molecular defect in vif-deficient human immunodeficiency virus type 1 virions. *J Virol* **2003**, *77*, 5810.

[66] A. M. Sheehy, N. C. Gaddis, J. D. Choi, M. H. Malim. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **2002**, *418*, 646.

[67] A. Lau, K. M. Swinbank, P. S. Ahmed, D. L. Taylor, S. P. Jackson, G. C. Smith, M. J. O'Connor. Suppression of HIV-1 infection by a small molecule inhibitor of the ATM kinase. *Nat Cell Biol* **2005**, *7*, 493.

[68] P. J. Joshi, T. S. Fisher, V. R. Prasad. Anti-HIV inhibitors based on nucleic acids: emergence of aptamers as potent antivirals. *Curr Drug Targets Infect Disord* **2003**, *3*, 383.

[69] D. M. Held, J. D. Kissel, J. T. Patterson, D. G. Nickens, D. H. Burke. HIV-1 inactivation by nucleic acid aptamers. *Front Biosci* **2006**, *11*, 89.

[70] H. Jonckheere, J. Anne, E. De Clercq. The HIV-1 reverse transcription (RT) process as target for RT inhibitors. *Med Res Rev* **2000**, *20*, 129.

[71] A. Jacobo-Molina, J. Ding, R. G. Nanni, A. D. Clark, Jr., X. Lu, C. Tantillo, R. L. Williams, G. Kamer, A. L. Ferris, P. Clark, et al. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc Natl Acad Sci U S A* **1993**, *90*, 6320.

[72] H. Huang, R. Chopra, G. L. Verdine, S. C. Harrison. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **1998**, *282*, 1669.

[73]   J. Ren, R. Esnouf, E. Garman, D. Somers, C. Ross, I. Kirby, J. Keeling, G. Darby, Y. Jones, D. Stuart, et al. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat Struct Biol* **1995**, *2*, 293.

[74]   C. Tuerk, S. MacDougal, L. Gold. RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proc Natl Acad Sci U S A* **1992**, *89*, 6988.

[75]   O. Kensch, B. A. Connolly, H. J. Steinhoff, A. McGregor, R. S. Goody, T. Restle. HIV-1 reverse transcriptase-pseudoknot RNA aptamer interaction has a binding affinity in the low picomolar range coupled with high specificity. *J Biol Chem* **2000**, *275*, 18271.

[76]   J. Jaeger, T. Restle, T. A. Steitz. The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. *Embo J* **1998**, *17*, 4535.

[77]   A. Jenne, W. Gmelin, N. Raffler, M. Famulok. Real-time characterization of ribozymes by fluorescence resonance energy transfer (FRET). *Angewandte Chemie-International Edition* **1999**, *38*, 1300.

[78]   A. Jenne, J. S. Hartig, N. Piganeau, A. Tauer, D. A. Samarsky, M. R. Green, J. Davies, M. Famulok. Rapid identification and characterization of hammerhead-ribozyme inhibitors using fluorescence-based technology. *Nat Biotechnol* **2001**, *19*, 56.

[79]   H. Nakane, K. Ono. Differential inhibitory effects of some catechin derivatives on the activities of human immunodeficiency virus reverse transcriptase and cellular deoxyribonucleic and ribonucleic acid polymerases. *Biochemistry* **1990**, *29*, 2841.

[80]   K. Post, J. Guo, K. J. Howard, M. D. Powell, J. T. Miller, A. Hizi, S. F. Le Grice, J. G. Levin. Human immunodeficiency virus type 2 reverse transcriptase activity in model systems that mimic steps in reverse transcription. *J Virol* **2003**, *77*, 7623.

[81]   S. Doublie, M. R. Sawaya, T. Ellenberger. An open and closed case for all polymerases. *Structure* **1999**, *7*, R31.

[82]   T. A. Steitz. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem* **1999**, *274*, 17395.

[83]   K. Singh, M. J. Modak. A unified DNA- and dNTP-binding mode for DNA polymerases. *Trends Biochem Sci* **1998**, *23*, 277.

[84]   G. Jarmy, M. Heinkelein, B. Weissbrich, C. Jassoy, A. Rethwilm. Phenotypic analysis of the sensitivity of HIV-1 to inhibitors of the reverse transcriptase, protease, and integrase using a self-inactivating virus vector system. *J Med Virol*

2001, *64*, 223.

[85] X. Wei, J. M. Decker, H. Liu, Z. Zhang, R. B. Arani, J. M. Kilby, M. S. Saag, X. Wu, G. M. Shaw, J. C. Kappes. Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* **2002**, *46*, 1896.

[86] o-Aminoazotoluene. *Rep Carcinog* **2002**, *10*, 12.

[87] G. A. Soukup, R. R. Breaker. Allosteric nucleic acid catalysts. *Curr Opin Struct Biol* **2000**, *10*, 318.

[88] J. J. Rossi. Ribozymes, genomics and therapeutics. *Chem Biol* **1999**, *6*, R33.

[89] S. K. Silverman. Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA. *Rna* **2003**, *9*, 377.

[90] G. A. Soukup, R. R. Breaker. Nucleic acid molecular switches. *Trends Biotechnol* **1999**, *17*, 469.

[91] J. Hesselberth, M. P. Robertson, S. Jhaveri, A. D. Ellington. In vitro selection of nucleic acids for diagnostic applications. *J Biotechnol* **2000**, *74*, 15.

[92] R. R. Breaker. Engineered allosteric ribozymes as biosensor components. *Curr Opin Biotechnol* **2002**, *13*, 31.

[93] L. Chaloin, M. J. Lehmann, G. Sczakiel, T. Restle. Endogenous expression of a high-affinity pseudoknot RNA aptamer suppresses replication of HIV-1. *Nucleic Acids Res* **2002**, *30*, 4001.

[94] P. Joshi, V. R. Prasad. Potent inhibition of human immunodeficiency virus type 1 replication by template analog reverse transcriptase inhibitors derived by SELEX (systematic evolution of ligands by exponential enrichment). *J Virol* **2002**, *76*, 6545.

[95] D. J. Schneider, J. Feigon, Z. Hostomsky, L. Gold. High-affinity ssDNA inhibitors of the reverse transcriptase of type 1 human immunodeficiency virus. *Biochemistry* **1995**, *34*, 9599.

[96] K. Yamaguchi, M. Honda, H. Ikigai, Y. Hara, T. Shimamura. Inhibitory effects of (-)-epigallocatechin gallate on the life cycle of human immunodeficiency virus type 1 (HIV-1). *Antiviral Res* **2002**, *53*, 19.

[97] J. E. Reardon, W. H. Miller. Human immunodeficiency virus reverse transcriptase. Substrate and inhibitor kinetics with thymidine 5'-triphosphate and 3'-azido-3'-deoxythymidine 5'-triphosphate. *J Biol Chem* **1990**, *265*, 20302.

[98] J. Balzarini, M. J. Perez-Perez, A. San-Felix, M. J. Camarasa, I. C. Bathurst, P. J. Barr, E. De Clercq. Kinetics of inhibition of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase by the novel HIV-1-specific nucleoside

analogue [2',5'-bis-O-(tert-butyldimethylsilyl)-beta-D-ribofuranosyl]-3'-spiro-5 "- (4"-amino-1",2"-oxathiole-2",2"-dioxide)thymine (TSAO-T). *J Biol Chem* **1992**, *267*, 11831.

[99] S. Loya, M. Bakhanashvili, Y. Kashman, A. Hizi. Mechanism of inhibition of HIV reverse transcriptase by toxiusol, a novel general inhibitor of retroviral and cellular DNA polymerases. *Biochemistry* **1995**, *34*, 2260.

[100] S. Loya, A. Rudi, Y. Kashman, A. Hizi. Mode of inhibition of HIV reverse transcriptase by 2-hexaprenylhydroquinone, a novel general inhibitor of RNA-and DNA-directed DNA polymerases. *Biochem J* **1997**, *324 (Pt 3)*, 721.

[101] S. R. Budihas, I. Gorshkova, S. Gaidamakov, A. Wamiru, M. K. Bona, M. A. Parniak, R. J. Crouch, J. B. McMahon, J. A. Beutler, S. F. Le Grice. Selective inhibition of HIV-1 reverse transcriptase-associated ribonuclease H activity by hydroxylated tropolones. *Nucleic Acids Res* **2005**, *33*, 1249.

[102] S. G. Sarafianos, K. Das, C. Tantillo, A. D. Clark, Jr., J. Ding, J. M. Whitcomb, P. L. Boyer, S. H. Hughes, E. Arnold. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *Embo J* **2001**, *20*, 1449.

[103] T. S. Fisher, T. Darden, V. R. Prasad. Mutations proximal to the minor groove-binding track of human immunodeficiency virus type 1 reverse transcriptase differentially affect utilization of RNA versus DNA as template. *J Virol* **2003**, *77*, 5837.

[104] T. S. Fisher, P. Joshi, V. R. Prasad. Mutations that confer resistance to template-analog inhibitors of human immunodeficiency virus (HIV) type 1 reverse transcriptase lead to severe defects in HIV replication. *J Virol* **2002**, *76*, 4068.

[105] T. S. Fisher, P. Joshi, V. R. Prasad. HIV-1 reverse transcriptase mutations that confer decreased in vitro susceptibility to anti-RT DNA aptamer RT1t49 confer cross resistance to other anti-RT aptamers but not to standard RT inhibitors. *AIDS Res Ther* **2005**, *2*, 8.

[106] C. Ahgren, K. Backro, F. W. Bell, A. S. Cantrell, M. Clemens, J. M. Colacino, J. B. Deeter, J. A. Engelhardt, M. Hogberg, S. R. Jaskunas, et al. The PETT series, a new class of potent nonnucleoside inhibitors of human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents Chemother* **1995**, *39*, 1329.

[107] R. Pauwels, K. Andries, J. Desmyter, D. Schols, M. J. Kukla, H. J. Breslin, A. Raeymaeckers, J. Van Gelder, R. Woestenborghs, J. Heykants, et al. Potent and selective inhibition of HIV-1 replication in vitro by a novel series of TIBO

derivatives. *Nature* **1990**, *343*, 470.

[108] J. Auwerx, M. Stevens, A. R. Van Rompay, L. E. Bird, J. Ren, E. De Clercq, B. Oberg, D. K. Stammers, A. Karlsson, J. Balzarini. The phenylmethylthiazolylthiourea nonnucleoside reverse transcriptase (RT) inhibitor MSK-076 selects for a resistance mutation in the active site of human immunodeficiency virus type 2 RT. *J Virol* **2004**, *78*, 7427.

[109] A. G. Skillman, K. W. Maurer, D. C. Roe, M. J. Stauber, D. Eargle, T. J. Ewing, A. Muscate, E. Davioud-Charvet, M. V. Medaglia, R. J. Fisher, E. Arnold, H. Q. Gao, R. Buckheit, P. L. Boyer, S. H. Hughes, I. D. Kuntz, G. L. Kenyon. A novel mechanism for inhibition of HIV-1 reverse transcriptase. *Bioorg Chem* **2002**, *30*, 443.

[110] L. Z. Wang, G. L. Kenyon, K. A. Johnson. Novel mechanism of inhibition of HIV-1 reverse transcriptase by a new non-nucleoside analog, KM-1. *J Biol Chem* **2004**, *279*, 38424.

[111] S. Imamura, O. Kurasawa, Y. Nara, T. Ichikawa, Y. Nishikawa, T. Iida, S. Hashiguchi, N. Kanzaki, Y. Iizawa, M. Baba, Y. Sugihara. CCR5 antagonists as anti-HIV-1 agents. Part 2: Synthesis and biological evaluation of N-[3-(4-benzylpiperidin-1-yl)propyl]-N,N'-diphenylureas. *Bioorg Med Chem* **2004**, *12*, 2295.

[112] A. D. Griffiths, D. S. Tawfik. Man-made enzymes--from design to in vitro compartmentalisation. *Curr Opin Biotechnol* **2000**, *11*, 338.

[113] D. S. Tawfik, A. D. Griffiths. Man-made cell-like compartments for molecular evolution. *Nat Biotechnol* **1998**, *16*, 652.

[114] A. D. Griffiths, D. S. Tawfik. Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *Embo J* **2003**, *22*, 24.

[115] Y. F. Lee, D. S. Tawfik, A. D. Griffiths. Investigating the target recognition of DNA cytosine-5 methyltransferase HhaI by library selection using in vitro compartmentalisation. *Nucleic Acids Res* **2002**, *30*, 4937.

[116] P. A. Lohse, M. C. Wright. In vitro protein display in drug discovery. *Curr Opin Drug Discov Devel* **2001**, *4*, 198.

[117] P. Amstutz, P. Forrer, C. Zahnd, A. Pluckthun. In vitro display technologies: novel developments and applications. *Curr Opin Biotechnol* **2001**, *12*, 400.

[118] W. J. Dower, L. C. Mattheakis. In vitro selection as a powerful tool for the applied evolution of proteins and peptides. *Curr Opin Chem Biol* **2002**, *6*, 390.

[119] R. W. Roberts. Totally in vitro protein selection using mRNA-protein fusions and ribosome display. *Curr Opin Chem Biol* **1999**, *3*, 268.

[120] C. Rader, C. F. Barbas, 3rd. Phage display of combinatorial antibody libraries. *Curr Opin Biotechnol* **1997**, *8*, 503.

[121] R. H. Hoess. Protein design and phage display. *Chem Rev* **2001**, *101*, 3205.

[122] G. P. Smith, V. A. Petrenko. Phage Display. *Chem Rev* **1997**, *97*, 391.

[123] J. Hanes, L. Jermutus, S. Weber-Bornhauser, H. R. Bosshard, A. Pluckthun. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proc Natl Acad Sci U S A* **1998**, *95*, 14130.

[124] J. Hanes, A. Pluckthun. In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **1997**, *94*, 4937.

[125] P. Amstutz, J. N. Pelletier, A. Guggisberg, L. Jermutus, S. Cesaro-Tadic, C. Zahnd, A. Pluckthun. In vitro selection for catalytic activity with ribosome display. *J Am Chem Soc* **2002**, *124*, 9396.

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

[127] N. Nemoto, E. Miyamoto-Sato, Y. Husimi, H. Yanagawa. In vitro virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett* **1997**, *414*, 405.

[128] A. D. Keefe, J. W. Szostak. Functional proteins from a random-sequence library. *Nature* **2001**, *410*, 715.

[129] J. E. Barrick, T. T. Takahashi, J. Ren, T. Xia, R. W. Roberts. Large libraries reveal diverse solutions to an RNA recognition problem. *Proc Natl Acad Sci U S A* **2001**, *98*, 12374.

[130] D. S. Wilson, A. D. Keefe, J. W. Szostak. The use of mRNA display to select high-affinity protein-binding peptides. *Proc Natl Acad Sci U S A* **2001**, *98*, 3750.

[131] A. D. Keefe, D. S. Wilson, B. Seelig, J. W. Szostak. One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expr Purif* **2001**, *23*, 440.

[132] N. A. Raffler, J. Schneider-Mergener, M. Famulok. A novel class of small functional peptides that bind and inhibit human alpha-thrombin isolated by mRNA display. *Chem Biol* **2003**, *10*, 69.

[133] L. Xu, P. Aha, K. Gu, R. G. Kuimelis, M. Kurz, T. Lam, A. C. Lim, H. Liu, P. A. Lohse, L. Sun, S. Weng, R. W. Wagner, D. Lipovsek. Directed evolution of high-affinity antibody mimics using mRNA display. *Chem Biol* **2002**, *9*, 933.

[134] T. P. Cujec, P. F. Medeiros, P. Hammond, C. Rise, B. L. Kreider. Selection of v-abl tyrosine kinase substrate sequences from randomized peptide and cellular proteomic libraries using mRNA display. *Chem Biol* **2002**, *9*, 253.

[135] P. W. Hammond, J. Alpin, C. E. Rise, M. Wright, B. L. Kreider. In vitro selection and characterization of Bcl-X(L)-binding proteins from a mix of tissue-specific mRNA display libraries. *J Biol Chem* **2001**, *276*, 20898.

[136] X. Shen, C. A. Valencia, J. W. Szostak, B. Dong, R. Liu. Scanning the human proteome for calmodulin-binding proteins. *Proc Natl Acad Sci U S A* **2005**, *102*, 5969.

[137] S. Weng, K. Gu, P. W. Hammond, P. Lohse, C. Rise, R. W. Wagner, M. C. Wright, R. G. Kuimelis. Generating addressable protein microarrays with PROfusion covalent mRNA-protein fusion technology. *Proteomics* **2002**, *2*, 48.

[138] S. Li, S. Millward, R. Roberts. In vitro selection of mRNA display libraries containing an unnatural amino acid. *J Am Chem Soc* **2002**, *124*, 9972.

[139] A. Frankel, S. Li, S. R. Starck, R. W. Roberts. Unnatural RNA display libraries. *Curr Opin Struct Biol* **2003**, *13*, 506.

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

[141] K. Josephson, M. C. Hartman, J. W. Szostak. Ribosomal synthesis of unnatural peptides. *J Am Chem Soc* **2005**, *127*, 11727.

[142] K. A. Denessiouk, V. V. Rantanen, M. S. Johnson. Adenine recognition: a motif present in ATP-, CoA-, NAD-, NADP-, and FAD-dependent proteins. *Proteins* **2001**, *44*, 282.

[143] C. R. Bellamacina. The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *Faseb J* **1996**, *10*, 1257.

[144] A. M. Lesk. NAD-binding domains of dehydrogenases. *Curr Opin Struct Biol* **1995**, *5*, 775.

[145] C. Engel, R. Wierenga. The diverse world of coenzyme A binding proteins. *Curr Opin Struct Biol* **1996**, *6*, 790.

[146] J. J. Kim, K. P. Battaile. Burning fat: the structural basis of fatty acid beta-oxidation. *Curr Opin Struct Biol* **2002**, *12*, 721.

[147] S. T. Thompson, K. H. Cass, E. Stellwagen. Blue dextran-sepharose: an affinity column for the dinucleotide fold in proteins. *Proc Natl Acad Sci U S A* **1975**, *72*, 669.

[148] Y. Li, G. Kunyu, C. Lubai, Z. Hanfa, Z. Yukui. Affinity chromatography of yeast alcohol dehydrogenase using immobilized monochlorotriazine colourless compounds. *Biomed Chromatogr* **1997**, *11*, 180.

[149] J. F. Biellmann, J. P. Samama, C. I. Branden, H. Eklund. X-ray studies of the

binding of Cibacron blue F3GA to liver alcohol dehydrogenase. *Eur J Biochem* **1979**, *102*, 107.

[150] X. Ysern, H. J. Prochaska. X-ray diffraction analyses of crystals of rat liver NAD(P)H:(quinone-acceptor) oxidoreductase containing cibacron blue. *J Biol Chem* **1989**, *264*, 7765.

[151] G. Cho, A. D. Keefe, R. Liu, D. S. Wilson, J. W. Szostak. Constructing high complexity synthetic libraries of long ORFs using in vitro selection. *J Mol Biol* **2000**, *297*, 309.

[152] M. Kurz, G. Kuang, P. A. Lohse. An efficient synthetic strategy for the preparation of nucleic acid-encoded peptide and protein libraries for in vitro evolution protocols. *Molecules* **2000**, *5*, 1259.

[153] M. Kurz, K. Gu, P. A. Lohse. Psoralen photo-crosslinked mRNA-puromycin conjugates: a novel template for the rapid and facile preparation of mRNA-protein fusions. *Nucleic Acids Res* **2000**, *28*, E83.

[154] J. C. Chaput, J. W. Szostak. Evolutionary optimization of a nonbiological ATP binding protein for improved folding stability. *Chem Biol* **2004**, *11*, 865.

[155] D. R. Flower. The lipocalin protein family: structure and function. *Biochem J* **1996**, *318 (Pt 1)*, 1.

[156] G. Beste, F. S. Schmidt, T. Stibora, A. Skerra. Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold. *Proc Natl Acad Sci U S A* **1999**, *96*, 1898.

[157] S. K. Dubey, A. K. Singh, H. Singh, S. Sharma, R. N. Iyer, J. C. Katiyar, P. Goel, A. B. Sen. Synthesis of substituted 1-hydroxy-2-naphthanilides as potential cestodicidal agents. *J Med Chem* **1978**, *21*, 1178.

[158] J. Valgeirsson, E. O. Nielsen, D. Peters, T. Varming, C. Mathiesen, A. S. Kristensen, U. Madsen. 2-arylureidobenzoic acids: selective noncompetitive antagonists for the homomeric kainate receptor subtype GluR5. *J Med Chem* **2003**, *46*, 5834.

[159] F. Zaragoza, H. Stephensen, S. M. Knudsen, L. Pridal, B. S. Wulff, K. Rimvall. 1-alkyl-4-acylpiperazines as a new class of imidazole-free histamine H(3) receptor antagonists. *J Med Chem* **2004**, *47*, 2833.

6. Appendix

6.1. Abbreviations

AA	amino acid
AIDS	acquired immunodeficiency syndrome
AMV	avian myeloblastosis virus
APS	ammonium peroxydisulfate
ATM	ataxia-talengiectasia-mutated
ATP	adenosine 5'-triphosphate
AZT	azidothymidine, *zidovudine*
bp	base pair(s)
BSA	bovine serum albumine
cDNA	complementary DNA
Ci	curie
CIAP	calf intestine alkaline phosphatase
CMVp	cytomegalovirus promoter
CoA	coenzyme A
cpm	counts per minute
CTP	cytosine 5'-triphosphate
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytosine-5'-triphosphate
DDDP	DNA-dependent DNA polymerase
DEPC	diethyl pyrocarbonate
dGTP	2'-deoxyguanosine-5'-triphosphate
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
dsDNA	double stranded DNA
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDAC	*N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride

|	
EDTA	ethylenediaminetetraacetic acid
EI	electron impact ionization
FAB	fast atom bombardment
FAD	flavine adenine dinucleotide
FAM	6-carboxyfluorescein
FRET	fluorescence resonance energy transfer
g	gram
GTP	guanosine 5'-triphosphate
HAART	highly active antiretroviral therapy
HEPES	*N*-(2-hydroxyethyl)-piperazine-*N*'-2-ethanesulfonic acid
HIV-1	human immunodeficiency virus type-1
HIV-2	human immunodeficiency virus type-2
HTS	high throughput screening
kDa	kilodaltons
KF	klenow fragment
L	liter
LTR	long terminal repeat
M	molar
min	minutes
MMLV	moloney murine leukemia virus
mRNA	messenger RNA
MW	molecular weight
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
NNRTI	non-nucleoside RT inhibitor
NRTI	nucleoside analogue RT inhibitor
NTP	nucleotide triphosphate
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PNK	polynucleotide kinase
pol β	human DNA polymerase β

PR	protease
RBE	Rev binding element
RDDP	RNA-dependent DNA polymerase
RNA	ribonucleic acid
RNase	ribonuclease
RNase H	Ribonuclease H
rpm	revolutions per minute
rt	room temperature
RT	reverse transcriptase
SDS	sodium dodecyl sulfate
SELEX	systematic evolution of ligands by exponential enrichment
SFFV	Spleen focus-forming virus
SPR	surface plasmon resonance
ssDNA	single stranded DNA
TAMRA	6-carboxytetramethylrhodamine
Taq	*thermus aquaticus*
TBE	Tris borate EDTA
TEMED	*N,N,N',N'*-tetramethylethylendiamine
THF	tetrahydrofuran
TMV	tobacco mosaic virus
Tris	tris-(hydroxymethyl)-aminomethane
tRNA	transfer RNA
TTP	thymidine-5'-triphosphate
U	unit
UTR	untranslated region
UV	ultraviolet
VSV-G	vesicular stomatitis virus G glycoprotein

6.2. Amino Acid Codes

Single letter code	**Three letter code**	**Amino acid**
A	Ala	alanine
C	Cys	cysteine
D	Asp	aspartic acid
E	Glu	glutamic acid
F	Phe	phenylalanine
G	Gly	glysine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophane
Y	Tyr	tyrosine

6.3. Vektor pGEM-T

6.3.1. pGEM-T Vector Circle Map
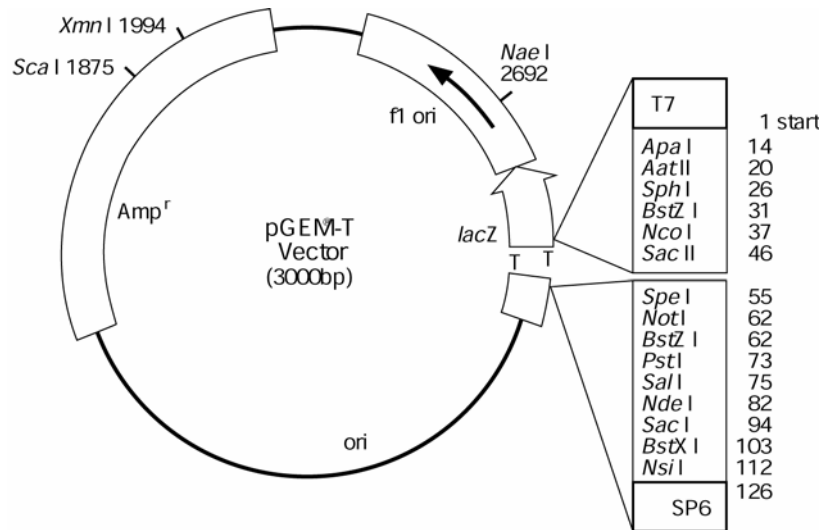
6.3.2. pGEM-T Vector Sequence

```
  1   GGGCGAATTG GGCCCGACGT CGCATGCTCC CGGCCGCCAT GGCCGCGGGA
 51   T*ATCACTAGT GCGGCCGCCT GCAGGTCGAC CATATGGGAG AGCTCCCAAC
101   GCGTTGGATG CATAGCTTGA GTATTCTATA GTGTCACCTA AATAGCTTGG
151   CGTAATCATG GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA
201   ATTCCACACA ACATACGAGC CGGAAGCATA AAGTGTAAAG CCTGGGGTGC
251   CTAATGAGTG AGCTAACTCA CATTAATTGC GTTGCGCTCA CTGCCCGCTT
301   TCCAGTCGGG AAACCTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACGC
351   GCGGGGAGAG GCGGTTTGCG TATTGGGCGC TCTTCCGCTT CCTCGCTCAC
401   TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT
451   CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG
501   AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC
551   GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCTGACGAG CATCACAAAA
601   ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC
651   CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT
701   GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC
751   TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC
801   TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC
```

130

```
 851  CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT
 901  CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA
 951  GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG
1001  AAGAACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA
1051  AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT
1101  GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA
1151  AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA
1201  ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC
1251  TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT AAAGTATATA
1301  TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA
1351  TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC
1401  GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC
1451  AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA
1501  ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC
1551  GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC
1601  GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG
1651  TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA
1701  TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC
1751  CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA GTGTTATCAC
1801  TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT GCCATCCGTA
1851  AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA
1901  GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAATA CGGGATAATA
1951  CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT
2001  TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT
2051  GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA
2101  GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA
2151  ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTTCAATA
2201  TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG
2251  AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA
2301  AAAGTGCCAC CTGATGCGGT GTGAAATACC GCACAGATGC GTAAGGAGAA
2351  AATACCGCAT CAGGAAATTG TAAGCGTTAA TATTTTGTTA AAATTCGCGT
2401  TAAATTTTTG TTAAATCAGC TCATTTTTTA ACCAATAGGC CGAAATCGGC
2451  AAAATCCCTT ATAAATCAAA AGAATAGACC GAGATAGGGT TGAGTGTTGT
2501  TCCAGTTTGG AACAAGAGTC CACTATTAAA GAACGTGGAC TCCAACGTCA
2551  AAGGGCGAAA AACCGTCTAT CAGGGCGATG GCCCACTACG TGAACCATCA
2601  CCCTAATCAA GTTTTTTGGG GTCGAGGTGC CGTAAAGCAC TAAATCGGAA
```

131

```
2651 CCCTAAAGGG AGCCCCCGAT TTAGAGCTTG ACGGGGAAAG CCGGCGAACG
2701 TGGCGAGAAA GGAAGGGAAG AAAGCGAAAG GAGCGGGCGC TAGGGCGCTG
2751 GCAAGTGTAG CGGTCACGCT GCGCGTAACC ACCACACCCG CCGCGCTTAA
2801 TGCGCCGCTA CAGGGCGCGT CCATTCGCCA TTCAGGCTGC GCAACTGTTG
2851 GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT ATTACGCCAG CTGGCGAAAG
2901 GGGGATGTGC TGCAAGGCGA TTAAGTTGGG TAACGCCAGG GTTTTCCCAG
2951 TCACGACGTT GTAAAACGAC GGCCAGTGAA TTGTAATACG ACTCACTATA
```

6.4. Sequences of Selected Clones from the Ninth Cycle of Cibacron Blue Selection

The following shows amino acid sequences cloned after ninth cycle of selection, which possessed intact ORFs. (**M**: the position for possible internal initiation, *****: internal stop codon)

<u>Type</u>	<u>No</u>	<u>Amino acid sequence</u>	
αβαβ:	**5**	SGECFFKNIHSSSYNTVRIPGG	SENLNLSLNLNIDVKITVAPGS
		SGNNPINNLKNISSGLLKVPGS	SEEFEINLSIYSTIGLKFGLKPGS
	8	SGNRALLRLNNSANDLSATPGS	SECSSITLDIPIAVA**MNM**YPES
		SGNRITKDTKNIIKDSIGSPGS	SEEHKIHFEIKLCIKLN**M**KPGS
	20	SGNDI**M**RN**M**YDLVHETFKIPGS	SENLNNALELYFRLSFTIYPGS
		SGKGLSKDIEELLGKLVVAPGS	SEYTYFEIKFK*****TLGFNLWPGS
	21	SGQKILRRVKRTTRRVLDLPGT	SEELYLTLQLNLTFNLTLWPGS
		SGHKILDGFNNSSNCLSSAPGS	SEELSFEIKLDITITINFTPGS
	29	SGNKSIGS**M**NDIIRYFTNAPGS	SEYVELGFKIKFEINFDKAPGS
		SGNR**M**VRRFKD**M**TGWTIDVPGS	SEDINFAIDFNLDVDLK**M**EPGS
	34	SGNDIVRN**M**YDLVHETFKIPGS	SENLNNALELYFRLSFTIYPGS
		SGKGLSKDIEELLRKLVGAPGS	SEYTYFEIKFKYTLGFNLWPGS
	37	SGKGSHEYLNGSLNEIFRTPGS	SEWLNTCLTLAFAFGVCVTPGS
		SGNHFSDNTGNIFQDVAGTPGS	SEKVTVKSDLTVSIELKLEPGS
	40	SGNQITDESDQSAGKILLLPGS	SEDLGLTFYLNLKLS**M**EFSPGG
		SGHSLLNKLDDALESTFGLPGS	SEDITIHIHLKFEINVRIEPGS
	45	SRHGILEESDRFIGHSLNIPGS	SEESYLKFEINID**M**TLHLNPGS
		SGQNFIKDASNFIKDTLRTPGS	SENLRVNLKL*****IALYKHLSPGS
	49	SGNHTLENAKRLTESIIDPPGS	SESIEVALDFEIKFKFTLDPGS
		SGKHLINKLSDTLSEFINLPGS	SEKLPFKLEIDIHLE**M**SLRPGS
	50	SGNN**M**LDNTSWTTHQLISLPGS	SEKLYFYLR**M**KFEINFNLKSGS
		SGSYTLKNVDET**M**SNSFRLPGS	SETVSINITVKFG*****NVSI*****PGS
	52	SGKDSLGRSKKTIKELIDAPGS	SECLSITLDIPIAVA**MNM**YPES
		SGNSVSKGLYNTSSRLTCIPGS	SDRITIKFEINFKLYLTIHPGS
	53	SGRSSIERLEKLTL**M**TLNLPGS	IEAF*****LKLYITVNVHFYITPGS
		TGQNT**M**RKARGVLHEI*****G**M**PGS	SENL**M**ISI*****I**M**FRLDLYLNPGS
	68	SGKDVSKNFQDTVDNVSDSPGS	SEPLNFSFEIKLNLKFSLSPGS
		SGQELSGNIQNTVETTLQLPGS	SEYRAIDLHFNLNFNLNLEPGS

```
          70    SGNHASKDINRFTRRSFRLPGS SEYINLDAKLYIDVTMRMNPGS
                SGQDLANNSWNSASRLLDSPGS SGTLTITLASRFTFETNVNPGS
          72    SGRQLLRSINSTSYNFVVIPGS SENMEFEINFTITINLSLYPGS
                SGNEITDNLNNLTDNMTRCQRV VSI*PSIYTYG*TSV*I*RQGV
          77    SGNRIFGDFKQATRSTSGLPGS SEELTMHLDVEITYNKSVAPGS
                SGHKLLHGINNTLNKSMGTPGS SEYLDLEIPLLFEMRLKLKPGS
          79    SEQTSGGIKRMFKNSTDSPAS SEHIDVNFEIKLNINLAFGPGS
                SGQDLINRTRSVSKGTSRLPGS SETLTFYFAFSLGLKLKLPPGS
βRαβ:     3     SEEIEVGFEIKLNLNIELKPGS SVIALYVKDNSNVIKRKAIMML
                GGHEITDEISSALSNILSLPGS SEYLHLNFRLQISFTLDFEPGS
          6     SEAIKLWFEINLELNVGFNSGG SEDGTMMIMEK*VNLSQTQNTR
                SGNSSVNNSKKAFNKLSRVPGS SEHITLELQINVGLNVAMPPGS
          10    SEEIEVGFEIKLNLNIELKPGS SVIALYVKDNSNVIKRKAIMML
                SGHEITDEISSALSNILSLPGS SEYLHLNFRLQISFTLDFEPGS
          15    SENFSFEIKPRFEMTFWIHPGS SEYFLD*DRLSCILDLNVTASL
                SGQRSSEEFGNLINRVTELPGS SEYIDPQLTLCFNVKMNFEPGS
          36    SEYLNLSINFSFEIKLDMDLES S*AEHE*LMFINHLLVHKEIKI
                SGNKLADNLRQITGYTINLPGS SENLSLTFEINVNIKLYLKPGS
ααββ:     1     SGRKSSNEVKNMFEELTNLPGS SGQDSLRKTRYSSKNIINLPGS
                SDDIDISLNFQINMDLKFWPES SEKLTVTFEINIQLNFEVSPGS
          4     SGQGSLGRIEDLIRNTINLPGS SGNETINRSWNTSNKATDMPGS
                SEKFHVGIDIRFEINLNFGLGS SEELALALKLRLTLNPPVSPGS
          42    SGRNALKKTHNTSQNSLNVPGS SGKDATNGVDNAVSKILSLPGS
                SEKLKISVEITFRFSIELGPGS SEGLTLNIKLNLNVKLTIKPGS
          62    SGRNALKKTHNTSQNSLNVPGS SGKDATNGVDNVVGKILSLPGS
                SEKLKISVEITFRFSIELGPGS SEGLTLNIKLNLNVKLTIKPGS
          65    SGNNSSGEIYHVSSELINSPGS SVHYSVGDSRDMI*STTGSPGS
                SENLALYFEIKFKLNIDLSPGS SENINLSISVTLSLSLKITPGS
          69    SGNHMMNETDKTSDSIVKIPGS IGQGLIKNISDTISSLISAPGS
                SEYFSLSITFSVDVEIGLKPGG SENIKLNFEIKFNLKFRLGPGS
ββββ:     22    SENINLYFWFYFNISFSLAPGS SEKFYLALELSFEIKFSQKPGS
                SDRLEMEISMKVYLSISFNPGS SENLKIKIKIAIEMNLNLRPGS
          33    S*TITLSLKLRFEINLKLNPGS SENLHMEIDIKIKIDMPINPGS
                SEHLWLNICLSIRLDLALSPGS SECFDFKLDLKLRMTLTLTPGS
          58    SEALKTTLRIGIHVDMGIELGS SEAEHE*LMFINHLLVHKEIKI
                SGNKLADNLRQITGYTINLPGS SVELSFEIKLNFNLAIKVNPGS
```

αααα: **35**   SGDNFTQYFKSFSNEILEIPGS SGNNIGNDSDKSSLQLIK*****PGS

                      SGKNTTHNANDILSSAIWVPRS SGNDS**M**NNLQNTSERA**M**PGS

βRRα: **39**   SETSWFEINLKFQLRLDLSPGS SVKRNLIDDR*****DGDKNPIQKV**M**

                      SAAPQVSIRGVPRYVGEKHTTV SGKGFSSEVEGSTDRLINSPGS

6.5. Acknowledgments

DAAD (Deutscher Akademischer Austauschdienst), for their financial support for me to learn the German language and to initiate and pursue my study in Germany (10/2001-09/2004).

Finally, my parents and my friends, for unconditional support, understanding and eternal encouragement to rescue my motivation.

6.6. Curriculum Vitae

Name	Satoko Yamazaki
Current Address	Wolfstr. 29, 53111 Bonn Germany
Permanent Address	Higashiuematsu 86-1, odaka, midori-ku, Nagoya, Japan, 459-8001
Date of Birth	02. October 1976
Place of Birth	Aichi, Japan
Nationality	Japanese

Education

10/2001-03/2006	**PhD of Natural Science (rerum naturalium)**
	Kekulé Institute of Organic Chemistry and Biochemistry
	University of Bonn
	Supervisor: Professor Dr. Michael Famulok
04/1999-03/2001	**Master of Science in Engineering**
	Institute of Biotechnology, Department of Engineering
	Nagoya University
	Dissertation: *Multiple Component Coupling Using Aluminum Tris(2,6-diphenylphenoxide)(ATPH)*
	Supervisor: Professor Dr. Hisashi Yamamoto
04/1995-03/1999	**Bachelor of Science in Engineering**
	Institute of Biotechnology, Department of Engineering
	Nagoya University
	Dissertation: *Novel Three Component Coupling of Ketone, Cyclic Ether and Epoxide Using Aluminum Tris(2,6-diphenylphenoxide)(ATPH)*
	Supervisor: Professor Dr. Hisashi Yamamoto (04/1998–03/1999)
04/1992-03/1995	Aichishukutoku Senior High School
04/1989-03/1992	Odaka Junior High School
04/1983-03/1989	Odaka Primary School