In Vitro Selection Using Modified or Unnatural Nucleotides

In vitro selection is the process by which a pool of nucleic acids is enriched via iterative selection and amplification for those species that are capable of performing a particular task. Nucleic acids have been selected that bind to particular targets (aptamers), catalyze reactions (ribozymes or deoxyribozymes), or act as molecular switches (aptazymes).

Instructions for carrying out in vitro selection experiments have been detailed elsewhere in this Chapter (e.g., *UNITS 9.3 & 9.4*). This unit augments these other units by describing how modified or unnatural nucleotides can potentially be incorporated into a selection. It is strongly recommended that the researcher be conversant with a "normal" in vitro selection experiment prior to attempting selections with modified nucleotides. A normal in vitro selection experiment is already fraught with problems and pitfalls, and the addition of modified nucleotides adds an extra level of difficulty. For simplicity, this unit focuses on the transcription and amplification of RNA pools; however, similar procedures can be used for DNA pools.

CAUTION: When working with radioactivity, take appropriate precautions to avoid contamination of the experimenter and the surroundings. Carry out the experiment and dispose of wastes in an appropriately designated area, following the guidelines provided by the local radiation safety officer.

NOTE: Experiments involving RNA require careful precautions to prevent contamination and degradation by RNases (see *APPENDIX 2A*). The water used to make all solutions and buffers should be treated with diethylpyrocarbonate (DEPC; *APPENDIX 2A*). Use sterile, disposable plasticware where possible.

STRATEGIC PLANNING

When to Use Modified Nucleotides for In Vitro Selection

As discussed in other units (*UNITS 9.3 & 9.4*), the desired outcome of a selection experiment should be determined in advance and the selection strategy should be designed accordingly. The decision to include modified nucleotides in a selection experiment and the selection of which ones to use should be based on how the nucleotides might benefit the selection. There are various advantages to using modified nucleotides, the chief one being to increase the stability of the selected nucleic acid to nuclease degradation. For example, the incorporation of pyrimidines modified at the 2′ position with amino or fluoro functional groups has been shown to drastically increase the stabilities of transcribed RNA molecules. Selections from nuclease-resistant pools have yielded aptamers against human neutrophil elastase (Lin et al., 1994), vascular endothelial growth factor (Green et al., 1995; Ruckman et al., 1998), basic fibroblast growth factor (Jellinek et al., 1995), human keratinocyte growth factor (Pagratis et al., 1997), and interferon γ (Kubik et al., 1997), as well as a sequence-specific phosphodiesterase (Beaudry et al., 2000). Aptamers selected from such pools have been shown to be stable in sera or urine for as much as 2 days (Pieken et al., 1991; Lin et al., 1994; Green et al., 1995).

In addition, modified nucleotides potentially expand the chemical functionality of nucleic acids. Modified nucleotides have been included in selections for catalytic nucleic acids, and the resultant catalysts have been shown to be highly dependent upon the modifications for activity (Tarasow et al., 1997; Wiegand et al., 1997; Beaudry et al., 2000; Santoro et

Contributed by Scott M. Knudsen, Michael P. Robertson, and Andrew D. Ellington
Current Protocols in Nucleic Acid Chemistry (2001) 9.6.1-9.6.21

Table 9.6.1 Successful In Vitro Selections Using Modified Nucleotides

Modified nucleotide	Class	Target/function[a]	Citation
5-(1-Pentynyl)-deoxyuracil	Deoxy aptamer	Thrombin	Latham et al. (1994)
N-6-Aminohexyl adenosine	Ribozyme	Ligation to its 5′ end	Teramoto et al. (2000)
5-Imidizole-uracil	Ribozyme	Amide synthase	Wiegand et al. (1997)
5-Pyridylmethylcarboxamid-uracil	Ribozyme	Diels-Alderase	Tarasow et al. (1997)
C5-Imidizole-uracil analog (unnamed)	Deoxyribozyme	Sequence directed RNase	Santoro et al. (2000)
8-[2-(4-Imidazolyl)ethylamino]deoxyadenosine and 5-(3-aminoallyl)-deoxyuracil	Deoxyribozyme	RNase (internal rC with no Mg^{2+})	Perrin et al. (2001)
5-(3′-aminopropynyl)-deoxyuracil	Deoxy aptamer	ATP	Battersby et al. (1999)
aS NTPs (phosphorothioate linked RNA)	Aptamer	bFGF	Jhaveri et al. (1998)
aS dNTPs (phosphorothioate linked DNA)	Deoxy aptamer	NF-IL6	King et al. (1998)
2′-Fluoro pyrimidines	Aptamer	VEGF(165)	Ruckman et al. (1998)
2′-Amino pyrimidines	Aptamer	VPF/VEGF	Green et al. (1995)
2′-Amino pyrimidines	Ribozyme	Trans cleavage of RNA	Beaudry et al. (2000)
2′-Amino pyrimidines	Aptamer	Human neutrophil elastase	Lin et al. (1994)
2′-Amino pyrimidines	Aptamer	bFGF	Jellinek et al. (1995)
2′-Fluoro pyrimidines	Aptamer	hKGF	Pagratis et al. (1997)
2′-Amino pyrimidines	Aptamer	hKGF	Pagratis et al. (1997)
2′-Amino pyrimidines	Aptamer	IFN-γ	Kubik et al. (1997)
2′-Fluoro pyrimidines	Aptamer	IFN-γ	Kubik et al. (1997)

[a]Abbreviations: bFGF, basic fibroblast growth factor; hKGF, human keratinocyte growth factor; IFN-γ, interferon γ; NF-IL6, nuclear factor for human interleukin 6; VEGF, vascular endothelial growth factor; VPF, vascular permeability factor.

al., 2000). Interestingly, no study has yet been published in which a selection with a modified nucleotide was done side-by-side with one without the modified nucleotide to compare the relative chemical advantage imparted by the modification. In this respect, it is interesting to note that while Tarasow et al. (1997) suggested that the inclusion of the modified nucleotide was the only reason they were able to select a Diels-Alder synthetase, Seelig et al. (1999) were later able to readily select such enzymes from an unmodified pool. To the extent that a modified nucleotide will be included to hopefully enhance catalytic functionality, the choice of modification should complement the desired function. For example, an imidizole ring (with a pK_a near neutrality; Battersby et al., 1999) may be beneficial in a pool to be screened for catalysis of an acid/base reaction, while the incorporation of a thiolated residue (Jhaveri et al., 1998) could allow nucleic acids to participate in disulfide bond formation or rearrangement, reactions normally denied to them.

Overall, the most important considerations in deciding whether to use a modified nucleotide in a selection experiment are purely technical ones, i.e., can the modification be incorporated, and will it inhibit amplification of the nucleic acid pool? A number of modified nucleotides have been incorporated with varying degrees of success into selection experiments (Table 9.6.1). However, in the authors' experience, at least in some instances (e.g., the incorporation of 5-Br-dUTP with *Taq* DNA polymerase), modified nucleotides decrease the replicability of sequences into which they are incorporated, and therefore there is an inherent and unintended selection in favor of those members of a population that have fewer modified nucleotides. While most of the modifications mentioned thus far retain an unmodified Watson-Crick pairing face, this need not always

be the case. For example, Piccirilli et al. (1990) have shown that diaminopyrimidine and xanthosine can form a stable base pair. Lutz et al. (1998), using xanthosine in the template, showed that diamino pyrimidine triphosphate was incorporated by the 9°N-7 thermostable DNA polymerase (Southworth et al., 1996) in a sequence-specific fashion. However, the fidelity of this reaction was low, with only 45% incorporation of the correct nucleotide analog. This may mean that it would be difficult to carry out multiple rounds of selection with such modified nucleotides.

While it is difficult to predict in advance which combinations of polymerases and modified nucleotides will work well together (also see Critical Parameters), there is at least one strategic consideration that can be taken into account in the design of selection experiments that include modified nucleotides. While the length and composition of a random sequence pool can vary according to the type of selection that is carried out (also see *UNIT 9.2*), some attention should be paid to the choice of constant sequence regions. RNA polymerases tend to have difficulty incorporating modified nucleotides prior to positions 8 to 10 following the transcription start site (Milligan and Uhlenbeck, 1989; Kujau et al., 1997), most likely because the attempted incorporation of modified nucleotides leads to increased abortive initiation. Thus, it is wise to design a pool that lacks or limits the incorporation of the modified base within this region. Inasmuch as the first 18 or more nucleotides of a transcribed pool typically remain constant to allow for second-strand DNA synthesis, this limitation should not affect the overall diversity of the pool.

Preparation and Use of Modified Nucleotides

This unit presents procedures in two main sections: determination of the suitability of modified nucleotides for in vitro selection, followed by the use of modified nucleotides in in vitro selection. Since the main consideration is whether a given nucleotide analog can serve as a substrate and template for various enzymes and thus be selected, it is critical to determine this prior to beginning the selection. In the first set of procedures, transcription reactions are performed to show that full-length transcripts can be obtained from the modified nucleotides (see Basic Protocol 1). The prevalence of the modified nucleotide in the transcript pool is then assessed (see Basic Protocol 2). Since selected transcripts must be amplified, it is also important to verify that the modified RNAs are suitable templates for reverse transcriptase (see Basic Protocol 3). Finally, a mock selection should be performed and the products cloned and sequenced to determine the fidelity of replication (see Basic Protocol 4 for a discussion).

Once it has been determined that a given modified nucleotide(s) is suitable for use in a selection, the starting pool can be constructed and selection can be performed. In the second set of protocols, double-stranded DNA that codes for the modified pool (as generated in *UNIT 9.2*) is first transcribed and the products are gel purified (see Basic Protocol 5). A general discussion of special considerations in designing an in vitro selection procedure for modified nucleotides is then presented (see Basic Protocol 6), followed by a procedure for amplification of the selection products (see Basic Protocol 7).

BASIC PROTOCOL 1

Evaluation and Optimization of Transcription with a Modified Nucleotide

In the case of RNA selections, the ability of a modified nucleotide to be incorporated into a transcript should be evaluated. This is best done by performing a transcription reaction in which the modified nucleotide completely substitutes for its natural counterpart (e.g., 2′-fluoro-cytidine in place of cytidine). A control reaction is performed in parallel on the same template with all natural nucleotides. The comparative success of these reactions can be determined by including α-^{32}P-labeled nucleotides in the reaction mix, separating transcripts by denaturing polyacrylamide gel electrophoresis, and visualizing the full-length transcripts with a phosphorimager or via autoradiography.

Materials

Double-stranded template DNA (*UNIT 9.2*) and randomly mutagenized DNA template (*UNIT 9.4*)
3000 Ci/mmol α-^{32}P-labeled ATP (e.g., NEN Life Sciences)
Transcription mix (see recipe)
2× denaturing stop dye (see recipe)
10% (w/v) denaturing polyacrylamide gel (see recipe, *APPENDIX 3B*), 0.75 mm thick
37° to 42° and 70°C water baths or heating blocks, or a programmable thermal cycler
Gel blotting paper (Bio-Rad)
Plastic wrap
Gel dryer with vacuum (e.g., Bio-Rad)
Phosphorimager screen
Phosphorimager and image analysis software (e.g., ImageQuant)
Additional reagents and equipment for denaturing polyacrylamide gel electrophoresis (PAGE; *APPENDIX 3B*)

Perform transcription with modified and unmodified nucleotides

1. Combine the following in two separate tubes:

   ~0.5 µg double-stranded DNA template (first tube) and randomly mutagenized DNA template (second tube)
   0.5 µL 3000 Ci/mmol α-^{32}P-labeled ATP
   transcription mix for a total reaction volume of 10 µL.

   Begin to incubate at 37° to 42°C.

   It is advisable to amplify a small portion of the pool that will actually be used in the selection for these test transcriptions. As this is a test reaction, do not commit the entire pool.

   Using high-yield transcription kits (e.g., Ambion or Epicentre Technologies) instead of a homemade mix can result in increased yields with shorter reaction times, and is advised wherever possible.

2. At several time points (e.g., 2, 4, 8, and 12 hr), vortex the tubes, briefly centrifuge to bring down the liquid, and transfer 1 µL to separate tubes containing 5 µL of 2× denaturing stop dye.

Vortexing and centrifuging the reaction ensures that each sample will contain a consistent portion of the reaction. The accumulation of condensation and the precipitation of pyrophosphate during transcription can sometimes lead to variations in reaction volumes or concentrations.

If a commercial transcription kit is utilized, the time points can be much shorter (e.g., 0.25, 0.5, 1, and 2 hr).

3. Denature the time point samples by heating 3 min to 70°C and separate transcripts from aborted transcripts and unincorporated nucleotides on a 0.75-mm-thick, 10% denaturing polyacrylamide gel (*APPENDIX 3B*).

The concentration of acrylamide can be varied to efficiently separate different length products, depending on their sizes (APPENDIX 3B).

Prerunning the gel will heat it up and help to denature molecules that contain significant secondary structure. Nonetheless, pools can sometimes appear "fuzzy" on a gel due to the presence of molecules that differ slightly in length or that have different secondary structures.

Analyze transcription reactions

4. Remove one glass plate and place a sheet of gel blotting paper against the gel. Peel the gel off of the other glass plate (it will stick to the paper) and cover it with plastic wrap. Dry the gel under heat and vacuum using a commercial gel dryer.

Drying the gel can be omitted, and a wet gel can be directly used for exposure. If this is done, leave the gel against one glass plate and carefully wrap with plastic wrap to avoid contaminating the phosphorimager screen and the exposure cassette with radiation.

5. Expose the dried gel to a phosphorimager screen for 1 hr.

The exposure time may need to be increased or decreased, depending on the specific activity of the labeled nucleotide and the amount of transcript that is produced.

Gels can also be exposed to X-ray film, although this makes quantitation somewhat more cumbersome. A standard laboratory densitometer can be used for quantitation.

6. Develop the phosphorimager screen and quantitate the relative amounts of RNA transcripts produced using image analysis software such as ImageQuant.

7. Compare the intensities of product bands for the transcription reaction that includes the modified nucleotide to those for transcription with the normal nucleotide to determine the relative efficiency of incorporation for the modified nucleotide under the conditions tested.

8. Plot the amount of product that accumulates versus time to determine the optimal time for further transcriptions.

Optimize transcription reaction

9. Repeat steps 1 to 6 using different concentrations of the modified nucleotide. Take samples only at the optimal time determined in step 8.

Since it is likely that the modified nucleotide will not work as well as the normal nucleotide because it binds to the polymerase more poorly, higher modified nucleotide concentrations may be necessary to enhance transcription. Therefore, care should be taken that the amount of modified nucleotide does not unnecessarily restrict the magnesium concentration in solution. If the total nucleotide concentration approaches that of magnesium, or the concentration of modified nucleotide exceeds the normal concentration used for unmodified (i.e., 2.5 mM), then additional magnesium should be added to the reaction.

The initial sequence of the transcription product can also influence the incorporation of a modified nucleotide (see Strategic Planning for examples and references).

10. Plot the amount of product (band intensity) versus modified nucleoside triphosphate concentration.

This step is included both to maximize the yield of RNA produced with the new substrate and to efficiently use the modified nucleotide, which could potentially be expensive or difficult to produce. To conserve modified nucleotide, the concentration of modified nucleotide should be chosen at either a maximum in the above plot or at the lowest concentration at which near-maximal levels of RNA are produced.

BASIC PROTOCOL 2

Confirmation of the Presence of Modified Nucleotides

The appearance of a full-length product in the transcription reaction (see Basic Protocol 1) is encouraging, but does not indicate the level at which a modified nucleotide has been incorporated into a pool. For example, full-length RNAs could be members of the original population that incorporated the modified nucleotide at only a few positions, could have been synthesized via contaminating natural nucleotides, could have been synthesized via the misincorporation of a natural nucleotide (e.g., via G:U mismatches), or could be the result of some combination of these possibilities. In order to determine with certainty the level of modified nucleotide incorporation, transcription products should be isolated, fully digested, and separated by HPLC to identify the component nucleosides.

This protocol assumes that one has already gel-purified and quantitated full-length transcription products, isolated as below. The RNA will be digested to mononucleotides using nuclease P1, and then treated with alkaline phosphatase to generate nucleosides. Separation of individual mononucleosides will be carried out by reversed-phase HPLC. Digestions will be compared to standards containing both modified and unmodified nucleosides. Sample HPLC data are given in Figure 9.6.1.

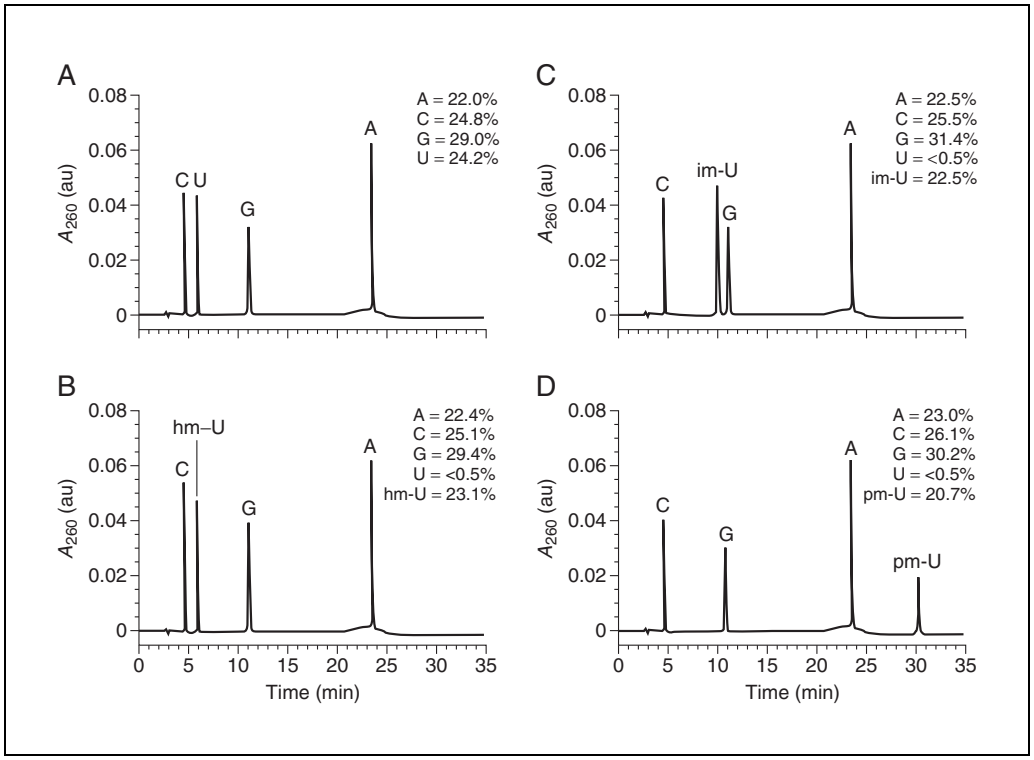

Figure 9.6.1 HPLC elution profile of pool RNA that has been treated with nuclease P1 and alkaline phosphatase as described in Basic Protocol 2. (**A**) Unmodified pool, (**B**) pool containing 5-hydroxymethyluridine, (**C**) pool containing 5-imidizolemethyluridine, (**D**) pool containing 5-phenolmethyluridine (Robertson, 2001).

Materials

Transcribed RNA pools, modified and unmodified (see Basic Protocol 1)
Nuclease P1 digestion mix (see recipe)
Alkaline phosphatase reaction mix (see recipe)
HPLC mobile phase solution: 5% (v/v) methanol in 0.1 M sodium phosphate, pH 6.0 (*APPENDIX 2A*)

37° and 50°C water baths
High-performance liquid chromatograph (HPLC) with reversed-phase C18 column (5-µm, 250 × 4.5-mm; Alltech Associates) and UV detector

1. Combine 200 pmol of each RNA pool (modified and unmodified) in separate reactions with nuclease P1 digestion mix for a total reaction volume of 5 µL and incubate 1.5 hr at 37°C.

   This step will digest the nucleic acid into its component nucleotides. Nuclease P1 will digest both RNA and DNA. It is possible that nuclease P1 will not digest after a specific modified nucleotide. If incomplete digestion is seen compared to an unmodified sample digestion, this in itself is convincing evidence of modified nucleotide incorporation.

   The unmodified transcription reaction should be run separately through the same column using the same conditions to provide a series of standard retention times for comparison. This control can also be used to obtain the extinction coefficient of the modified nucleoside, if it is not otherwise known.

2. Combine each entire nuclease P1 digestion reaction with alkaline phosphatase reaction mix in a final volume of 25 µL. Incubate 1.5 hr at 37°C, followed by 1 hr at 50°C.

3. Inject 2.5 µL of one reaction onto a reversed-phase C18 HPLC column and separate using a mobile phase flow rate of 1 mL/min and detection at 260 nm. Repeat with the other reaction.

   If the pool initially had an equimolar mix of nucleotides, then the digested RNA should contain ~25% of the modified nucleoside and no peak corresponding to the replaced nucleoside. Of course, the actual composition of pools can vary greatly, depending on the synthesis of the pool (e.g., see UNIT 9.2). Therefore, the composition of the original pool should first be determined either by sequencing individual clones (e.g., see CPMB, Chapter 7) or by first digesting and separating a transcription product made solely from natural nucleotides.

   Nucleotides are detected at 260 nm as this is close to the absorbance maxima for their aromatic bases. Nucleotides with heavily modified bases that disrupt their aromaticity may absorb at different maxima; however, A_{260} should still be used because the RNA will likely be mostly unmodified.

Evaluation of Modified RNA as a Template for Reverse Transcriptase

BASIC PROTOCOL 3

During each round of in vitro selection, the selected RNAs must be amplified. Therefore, it is necessary to determine if an RNA transcript containing modified bases will serve as a suitable template for reverse transcription, the first step of the amplification process. A control reaction can be performed on the same template with all natural nucleotides. The relative success of the reverse transcription reactions can again be visualized by incorporation of α-^{32}P-labeled nucleotides, gel electrophoresis, and analysis on a phosphorimager.

Materials

Transcribed RNA pools, modified and unmodified (prepared as in Basic Protocol 1 but without radiolabeled ATP)
3000 Ci/mmol α-^{32}P-labeled dATP (e.g., NEN Life Sciences)
RT reaction mix (see recipe)
AMV reverse transcriptase (e.g., USB)
2× denaturing stop dye (see recipe)
10% (w/v) denaturing polyacrylamide gel (see recipe; *APPENDIX 3B*), 0.75 mm thick

42° and 70°C water baths or heating blocks, or a programmable thermal cycler
Gel blotting paper (Bio-Rad)
Plastic wrap
Gel dryer with heat and vacuum (e.g., Bio-Rad)
Phosphorimager screen
Phosphorimager
Image analysis software (e.g., ImageQuant)

Additional reagents and equipment for denaturing polyacrylamide gel electrophoresis (PAGE; *APPENDIX 3B*)

Optimize RT reaction

1. Combine the following in two separate tubes:

   10 pmol transcribed RNA pool (modified in one tube, unmodified in the other)
   0.5 µL α-^{32}P-labeled dATP
   RT reaction mix for a total of 10 µL.

   Incubate 2 min at 42°C.

   Do not use radiolabeled RNA as the input. It may be difficult to differentiate product from input if both are labeled.

2. Add 10 U AMV reverse transcriptase per 10 µL RT reaction mix and incubate 50 min at 42°C.

   The success of the reverse transcriptase reaction will depend primarily on the reverse transcriptase that is used. While reverse transcriptases in general tend to be somewhat forgiving with respect to template chemistry (e.g., they can use both RNA and DNA as templates), different reverse transcriptases may have distinct abilities to utilize diverse modified bases. If one reverse transcriptase does not prove efficient in copying a particular RNA template, then another one should be used, or they should all be compared at the outset.

3. Remove 2 µL and add it to 5 µL of 2× denaturing stop dye.

4. Denature by heating 3 min to 70°C and run on a 0.75-mm-thick 10% denaturing polyacrylamide gel (*APPENDIX 3B*).

   A sample of radiolabeled RNA, as produced in the test transcriptions, can be used as a convenient size standard. RNA runs slightly slower on a gel than DNA of the same size.

   The concentration of acrylamide can be varied to efficiently separate different products (APPENDIX 3B).

Analyze RT reaction

5. Remove one glass plate and place a sheet of gel blotting paper against the gel. Peel the gel off of the other glass plate (it will stick to the paper) and cover with plastic wrap. Dry the gel in a gel dryer under heat and vacuum.

   Drying the gel can be omitted, and a wet gel can be directly exposed to the phosphorimager plate. If this is done, carefully wrap the gel with plastic wrap to avoid contaminating the screen and exposure cassette with radiation.

6. Expose the dried gel to a phosphorimager screen for 1 hr.

   The exposure time may need to be increased or decreased, depending on the specific activity of the labeled nucleotide and the amount of transcript produced.

   Alternatively, gels can be exposed to X-ray film, and the film developed to visualize bands.

7. Develop the phosphorimager screen and view the output with image analysis software such as ImageQuant.

   Full-length cDNA should have been produced from both modified and unmodified RNA templates.

   It is possible that smaller, discrete bands will be observed. These may be due to the failure of reverse transcriptase to read past secondary structures in the RNA.

   It is unlikely that discrete bands will be observed with a pool that contains many different RNA molecules, unless the secondary structures are in the constant regions. However, as a selection progresses, species that give rise to discrete stops may become fixed in a population. In this instance, the reverse transcription should be carried out at a slightly higher temperature. In extreme cases, a thermophilic DNA polymerase, Tth polymerase, has been shown to have significant reverse transcriptase activity, and can potentially be used to copy recalcitrant RNA molecules. Of course, it should be realized that even though strong stops cannot be seen early in selection, certain RNA species may drop out of a population because they cannot be copied by reverse transcriptase.

Confirmation of the Fidelity of Replication

BASIC PROTOCOL 4

If Basic Protocols 1, 2, and 3 give successful results, then it is likely that the random sequence pool can make it through the preparative steps leading to the selection. However, it is possible that the incorporation of modified nucleotides may alter the dynamics of selection experiments in one of two ways: first, modified nucleotides may in and of themselves be extremely mutagenic; second, it may be difficult to either make full-length transcripts that contain modified nucleotides or to fully copy transcripts containing modified nucleotides, in which case there will be an unseen selection against transcripts that contain modified nucleotides over the course of several cycles. In order to guard against these possibilities, it may be worthwhile to carry out at least one round of mock selection and to clone and sequence the products.

Cloning and sequencing PCR products is beyond the scope of this protocol, but can be found, for example, in CPMB Chapter 7. However, the protocols in the next section can be used to simulate a round of selection (omitting the selective steps). Two different types of selections can be carried out, depending on whether the mutagenic or replicative potential of modified nucleotides is being tested. To test the mutagenic potential, a single sequence cloned from the pool should be subjected to a round of selection. Following the selection, ~30 clones should be sequenced and the number of mutations counted.

To test the replicative potential, the pool should be subjected to a round of selection. Before and after the selection, ~30 clones should be sequenced and the base composition of the pool should be analyzed. If the base composition following selection is highly skewed relative to before selection, then selection against a given residue may prove to be a problem. It should be recalled that even small skewing can prove to be significant over many rounds of selection. For example, if a modified cytidine is present at 25% of the random sequence positions in a starting pool, but is present at only 22.5% of the positions following one round of selection, after 10 rounds of selection its frequency may fall to $0.25[(0.225/0.25)^{10}] = 0.087$, or <9% of the random sequence positions in a pool.

BASIC PROTOCOL 5

Preparation of the RNA Pool

The pool will be transcribed using the conditions determined during the trial experiments. The RNA will be gel purified, eluted from the gel, and ethanol precipitated. Alternatively, very large amounts of RNA can be purified by size-exclusion chromatography (e.g., Qiagen RNA kit). Following quantitation, it will be ready to use in the first round of selection. Transcription reactions in later rounds will be performed in much the same way, except that smaller amounts of DNA template may be used in later rounds as the diversity of the pool becomes constricted.

Materials

Double-stranded, modified template DNA (*UNIT 9.2*)
Transcription mix (see recipe)
RNase-free DNase (e.g., RQ1 DNase, Promega)
2× denaturing stop dye (see recipe)
10% (w/v) denaturing polyacrylamide gel (see recipe, *APPENDIX 3B*), 1.5 mm thick
0.3 M NaCl (*APPENDIX 2A*)
70% and 100% (v/v) ethanol
TE buffer, pH 7.5 to 8.0 (*APPENDIX 2A*)

37° to 42° and 70°C water baths or heating blocks, or a programmable thermal cycler
Fluorescent thin-layer chromatography (TLC) plate (e.g., Aldrich), wrapped in plastic wrap
0.45-µm syringe filter with disposable syringe
Spectrophotometer

Additional reagent and equipment for denaturing polyacrylamide gel electrophoresis (*APPENDIX 3B*)

Prepare RNA

1. Add ~1 µg double-stranded, modified template DNA to 20 µL of transcription mix. Incubate at 37° to 42°C for the predetermined optimal time (see Basic Protocol 1).

   If the pool contains >1 µg of template DNA, scale up the transcription reaction accordingly.

   The complexity of the pool for the initial round of selection can be determined as described in UNIT 9.3. At least 10^{13} different sequences should be used for the first round of selection. By increasing the complexity of the population to ≥10^{15} sequences, it may be possible to isolate rare binding or catalytic species (e.g., see Bartel and Szostak, 1993). In later rounds of selection, as the population is winnowed, there will be many copies of each selected species in the pool, and smaller amounts of DNA template (~1 µg/round) and RNA (~10 µg/round) can be utilized.

   Although it is not necessary to produce radiolabeled RNA at this stage, it may be useful to have a labeled sample of the original pool RNA. This can be accomplished by adding α-^{32}P-labeled ATP to the transcription reaction mix.

2. Remove template DNA by adding 5 to 10 U RNase-free DNase and incubating 15 min at 37°C. Stop the reaction by adding an equal volume of 2× denaturing stop dye.

   The DNase treatment is necessary to ensure that selected RNA molecules, rather than DNA carryover from a previous amplification reaction, is preferentially amplified.

   In some cases, amplification artifacts can overrun selection experiments (e.g., see Marshall and Ellington, 1999). In these instances, it may prove worthwhile to include dUTP in PCR amplification reactions, and to eliminate unwanted templates by the addition of uracil N-glycosylase.

3. Denature the samples 3 min at 70°C, and load the entire reaction on a 1.5-mm-thick, 10% denaturing polyacrylamide gel (*APPENDIX 3B*).

   No more than 150 μg of RNA should be loaded per 1-cm-wide lane. For a thinner gel, this amount should be proportionally decreased. This value corresponds to ~50 μL of a very efficient transcription reaction.

   Since later rounds need not contain as much template (see step 1), a 0.75-mm-thick gel may suffice to purify transcripts.

4. Visualize the product band by UV shadowing against a fluorescent TLC plate covered in plastic wrap. Excise the product band with a clean razor blade.

   Use fresh plastic wrap and razor blades every time to avoid contaminating the pool.

   It is important to only cut out the product band, as bands that migrate lower on the gel are likely to be aborted transcripts or template DNA that has not been completely digested.

   As some modified nucleotides may be especially sensitive to UV light, it is best to expose the pool for as short a time as necessary to define the bands. It may prove useful to lay an additional sheet of plastic wrap over the gel and to draw a line around the product band before excision.

Elute RNA

5. Submerge the gel slice in ~400 μL of 0.3 M NaCl per cm^2 of gel. Elute RNA overnight at 37°C with agitation.

   Increased recovery of RNA can be achieved by cutting the gel slice into smaller pieces, or by pulverizing it with a clean glass stir rod.

   Elution at higher temperatures can speed recovery time, but may lead to increased fragmentation of the pool RNA.

6. Pipet the supernatant off of the gel slice and filter it through a 0.45-μm filter attached to a syringe to remove small pieces of acrylamide. Add 2.5 vol of 100% ethanol and incubate 30 min at −20°C or 10 min at −80°C.

   Filtering out small pieces of acrylamide is especially important if the gel slice was pulverized. These pieces will precipitate with the RNA and will cause errors in determining the final concentration of RNA, as the presence of acrylamide can interfere with the determination of the sample's UV absorbance.

7. Place the precipitation reaction in a 1.5-mL microcentrifuge tube and microcentrifuge 20 to 30 min at maximum speed, 4°C.

   For larger-scale precipitations, such as purification of the original pool, 10- to 12-mL Sarstedt tubes can be used, and the reaction centrifuged 40 min at 13,000 × g, 4°C.

8. Carefully decant the supernatant without disturbing the pellet. Wash the pellet with 70% ethanol and dry the pellet.

   If the pellet is dislodged during the wash step, microcentrifuge an additional 5 min in the 70% ethanol to reattach it to the side of the tube.

   The pellet can be air dried or dried under vacuum. When air drying, cover the tube with Parafilm and poke holes in the Parafilm with a clean needle in order to guard against contamination while allowing evaporation to proceed.

9. Resuspend the RNA pellet in 25 μL TE buffer, pH 7.5 to 8.0.

   If RNA is to be used in a reaction whose volume is small relative to the volume of RNA added, it may be necessary to supplement the reaction with magnesium to avoid complete chelation by EDTA. Alternatively, RNA can be resuspended in sterile water; however, this renders the sample more susceptible to RNase degradation.

10. Determine the amount of RNA recovered by determining the absorbance of the solution at 260 nm.

The concentration of an RNA solution = (A_{260} × dilution factor)/(10,625 × n) M for a 1-cm cell, where n is the length of the RNA. A convenient dilution factor is 500, corresponding to 2 μL of RNA solution diluted in 1 mL water.

The extinction coefficient 10,625 $M^{-1}cm^{-1}$ is the average of the extinction coefficients for the four nucleobases, C = 7100, U = 8400, G = 12,000, and A = 15,000 M cm^{-1}. However, it should be noted that modification could drastically affect absorbance at 260 nm. For example, if an aromatic group is appended to a base, then the extinction coefficient of the pool would be higher than expected. In such a case, it would be wise to use the extinction coefficient for the modified base to create a new average.

The RNA is now ready for a round of selection (see Basic Protocol 6). Repeated rounds of selection are performed by generating a new (selected) DNA template (see Basic Protocol 7) and repeating this protocol to generate the new (selected) RNA pool.

BASIC PROTOCOL 6

Isolation of a Functionally Enriched Pool by In Vitro Selection

The procedure used to isolate active species from a pool is highly dependent on what function is being selected, i.e., binding versus catalysis, binding proteins versus binding small organics. Ideally, selections should be designed in which only those individuals that are capable of performing the desired function are isolated from the remainder of the pool. This is best done by a series of negative and positive selections (also see *UNITS 9.2 & 9.3*). A positive selection is performed with the intention of isolating active sequences. For example, if binding a small organic molecule is the goal, the pool may be passed over a column that presents the desired target, and binding species eluted with the target in solution (i.e., affinity chromatography).

Most methods of isolating functional sequences can also isolate nonfunctional sequences for other reasons, such as binding to the matrix rather than to the target. For this reason, a negative selection generally contains all elements of the positive selection with the exception of the specific target or substrate. Negative selection steps, in general, reduce the number of sequences that artifactually survive the positive selection steps. As rounds progress, both negative and positive selections are made more stringent by adjusting reaction concentrations or times. More detailed discussions of the design of selection experiments to yield binding or catalytic species can be found in *UNITS 9.3 & 9.4*, respectively.

Conceptually, the design of a selection using modified nucleotides is not different from any other selection. However, special consideration should be given to the prospect that a modified nucleotide can influence the relative number of molecules that will survive a negative selection step. For example, a more hydrophobic nucleotide might be bound more readily to a modified cellulose filter even in the absence of a protein target. Moreover, as previously discussed, it is possible that discrimination against modified nucleotides at any point in a replicative cycle (e.g., failure to be incorporated by RNA polymerse, failure to be utilized as a template by reverse transcriptase, failure to be utilized as a substrate or a template by *Taq* polymerase) could lead to artifactual selection of those sequences that replicate best with few or no modified nucleotides, rather than to the actual selection of binding or catalytic species containing modified nucleotides.

Amplification of a Functionally Enriched Pool

Once an RNA pool undergoes a selection function, the selected pool, presumably enriched in functional molecules, must be amplified. In the case of an RNA selection, this is best done by a combination of reverse transcription, PCR, and in vitro transcription. An unintended benefit of this convoluted process is that it makes it more difficult for molecular parasites to accumulate, as there are few molecular species that can be preferentially amplified by reverse transcriptase, *Taq* polymerase, and T7 RNA polymerase (Marshall and Ellington, 1999).

Materials

Selected RNA pool (see Basic Protocol 6)
RT reaction mix (see recipe)
AMV reverse transcriptase (e.g., USB)
PCR reaction mix (see recipe)
Taq DNA polymerase (e.g., Promega)
6× nondenaturing dye solution (see recipe)
4% (w/v) agarose gel: e.g., NuSieve in tris/borate/EDTA (TBE) electrophoresis buffer (available precast or see, e.g., *CPMB UNIT 2.6*)
Size standard
5 M NaCl (*APPENDIX 2A*)
100% and 70% (v/v) ethanol
TE buffer, pH 7.5 to 8.0 (*APPENDIX 2A*)
10 mg/ml ethidium bromide solution (*APPENDIX 2A*)

42°C water bath or heating block
Programmable thermal cycler (e.g., MJ Research)

Additional reagents and equipment for agarose gel electrophoresis (e.g., *CPMB UNIT 2.5*)

Perform reverse transcription

1. Add a sample of a selected RNA pool to RT reaction mix to a total volume of 10 µL. Incubate 2 min at 42°C.
2. Add 10 U AMV reverse transcriptase per 10 µL reaction mix. Incubate 50 min at 42°C.

In some instances, RNA samples may be attached to beads, such as streptavidin-coated agarose beads. If so, the RNA can be directly amplified from the beads, but it is helpful to first wash the support with 1× RT buffer. The RT reaction may also need to be scaled up to ensure sufficient volume to completely cover the support.

Perform amplification

3. Seed 100 µL PCR reaction mix with 2 µL of the RT reaction mix (or ⅕ the RT reaction volume). Hot start the reaction with 2.5 U *Taq* DNA polymerase.

8 to 10 cycles:	30 sec	94°C (denaturation)
	30 sec	55°C (annealing)
	60 sec	72°C (extension).

The annealing temperature is dependent on primer composition and may need to be lowered.

Adding polymerase at high temperatures (or "hot starting") helps to prevent mispriming events that can lead to the accumulation of artifacts. The easiest way to hot start a reaction is to wait until the initial 94°C step to add Taq DNA polymerase.

If the RT reaction was performed on beads, these should also be added to the PCR reaction. It is best to limit the bead volume to no more than 20 µL in a 100-mL PCR reaction.

4. Withdraw 5 µL of the reaction and mix with 1 µL of 6× nondenaturing dye solution.
5. Run on a 4% agarose gel, preferably next to a size standard. If only a faint band of the proper size is visible, run the reaction for one to two more cycles. If no band is visible, continue the amplification reaction for four more cycles and check again. Continue to amplify and monitor until a band is visible on the gel.

   Following the reaction in this manner helps to avoid overamplification, which can lead to accumulation of larger-than-normal artifactual products. This somewhat laborious method of monitoring reactions can be greatly simplified if a laboratory has access to a real-time PCR machine and has designed appropriate primer pairs in advance.

   Although the selected RNA pool can be amplified by a combined reverse transcription PCR reaction, the authors have found that separate reactions yield more consistent results. However, if effort is expended on optimizing a combined RT-PCR reaction, then the time required for a round of selection can be proportionately shortened.
6. Run four identical polymerase chain reactions with the remainder of the RT reaction. Use the optimal number of thermal cycles determined in step 5.

Purify template DNA

7. Combine the five amplification reactions. Add 30 µL of 5 M NaCl to a final concentration of 0.3 M. Add 2.5 volumes of 100% ethanol. Incubate 30 min at −20°C or 10 min at −80°C.

   If a support such as agarose beads was included in the reaction, filter these out with a Millipore filter column prior to precipitation.
8. Place the precipitation reaction in a 1.5-mL microcentrifuge tube and microcentrifuge 20 to 30 min at maximum speed, 4°C.
9. Carefully decant the supernatant without disturbing the pellet. Wash the pellet with 70% ethanol and dry the pellet.

   If the pellet is dislodged during the wash step, microcentrifuge an additional 5 min in 70% ethanol to reattach it to the side of the tube.

   The pellet can be air dried or dried under vacuum. Cover the tube with Parafilm and poke holes in the Parafilm with a clean needle in order to guard against contamination while allowing evaporation to proceed.
10. Resuspend the pellet in 25 µL TE buffer, pH 7.5 to 8.0.
11. Dilute 1 µL template DNA sample in 4 µL water and add 1 µL of 6× nondenaturing dye solution.
12. Run the sample on an agaraose gel next to a ladder or size standard of known concentration and stain with 10 µg/mL ethidium bromide solution. Visualize the bands on a UV light table. Estimate the concentration of the original template DNA stock.

    This DNA will be used to transcribe RNA for the next round of selection (see Basic Protocol 5).

    Laboratories that have access to a digital camera with quantitation software can more accurately estimate the amount of DNA in a given sample.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see ***APPENDIX 2A****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Alkaline phosphatase reaction mix

50 mM Tris·Cl, pH 8.5 (*APPENDIX 2A*)
0.1 mM EDTA
2 U alkaline phosphatase per 25 µL reaction
Store buffer without enzyme indefinitely at −20°C. Add enzyme fresh.

Denaturing polyacrylamide gel, 10% (w/v)

1× TBE electrophoresis buffer (APPENDIX 2A) containing:
10% (w/v) acrylamide
0.5% (w/v) bisacrylamide
7 M urea
Prepare fresh

See APPENDIX 3B for full details on pouring and running gels.

Denaturing stop dye, 2×

Deionized formamide containing:
0.1% (w/v) bromphenol blue
50 mM EDTA
Store indefinitely at −20°C. Daily-use aliquot may be stored up to 1 month at room temperature.

Nondenaturing dye, 6×

60% (v/v) glycerol
0.1% (w/v) bromophenol blue
Store indefinitely at −20°C. Daily-use aliquot may be stored up to 1 month at room temperature.

Nuclease P1 digestion mix

15 mM sodium acetate, pH 5.2 (*APPENDIX 2A*)
0.1 mM zinc sulfate
1 U P1 nuclease per 5 µL reaction
Store buffer indefinitely at −20°C. Add enzyme fresh.

PCR reaction mix

10 mM Tris·Cl, pH 8.4 (*APPENDIX 2A*)
50 mM KCl (*APPENDIX 2A*)
1.5 mM $MgCl_2$ (*APPENDIX 2A*)
0.2 mM each dNTP
0.5 µM each primer
2.5 U/100 µL reaction *Taq* DNA polymerase

The Tris·Cl, KCl, and $MgCl_2$ can be combined at ten times the indicated concentrations to make a 10× PCR buffer and stored indefinitely at −20°C. The remainder of the reagents should be added fresh.

RT reaction mix

100 mM Tris·Cl, pH 8.0 (*APPENDIX 2A*)
50 mM KCl (*APPENDIX 2A*)
10 mM $MgCl_2$ (*APPENDIX 2A*)
10 mM dithiothreitol (DTT)
0.4 mM each dNTP
5 µM reverse (3′) primer

continued

The Tris·Cl, KCl, and $MgCl_2$ can be combined at ten times the indicated concentrations to make a 10× RT buffer and stored indefinitely at −20°C. The remainder of the reagents should be added fresh.

Transcription mix

40 mM Tris·Cl, pH 7.9 (*APPENDIX 2A*)
26 mM $MgCl_2$ (*APPENDIX 2A*)
0.01% (v/v) Triton X-100
2.5 mM spermidine trihydrochloride
5 mM dithiothreitol (DTT)
2.5 mM each NTP
20 U/20 µL reaction RNasin (Promega)
100 U/20 µL reaction T7 RNA polymerase

The efficiency of transcription will be one of the variables that will require the greatest optimization. It is again suggested that a commercial transcription kit may give the greatest yields and best results, especially with modified nucleotides.

COMMENTARY

Background Information

Nucleic acid chemistry is constrained by the relatively limited chemistry of the five canonical nucleotides. For example, nucleic acids, especially RNA, tend to be much more chemically labile than proteins. In the case of RNA, stability is impaired by the presence of a 2′-hydroxyl group that can serve as a general base poised to attack the phosphodiester bond. To offset this problem, Eckstein and co-workers developed 2′-modified nucleotides that were resistant to chemical and enzymatic hydrolysis. Hammerhead ribozymes modified with 2′-amino or 2′-fluoro groups were shown to retain activity, and had a 1000-fold increase in stability in rabbit serum (Pieken et al., 1991). While the riboside triphosphates could be incorporated into RNA, the k_{cat}/K_M values for incorporation of 2′-modified ribonucleotides were substantially higher than for unmodified nucleotides, with a preference order of 2′-OH > 2′-NH_2 > 2′-F > 2′-H (Brieba and Sousa, 2000). In order to improve incorporation, some groups have begun to engineer or evolve polymerases that can incorporate modified nucleotides. For example, a single mutation in T7 RNA polymerase (Y639F) reduces discrimination at the 2′ position, and allows more efficient incorporation of deoxyribonucleotides into RNA (Kostyuk et al., 1995). It was subsequently shown that this mutation also allows the incorporation of NTPs with fluoro, amino, and thio groups at their 2′ positions. The Y639F mutant of T7 RNA polymerase has been shown to have up to 24-fold greater specificity for incorporation of 2′-modified NTPs, with a preference order of 2′-OH > 2′-F > 2′-H > 2′-NH_2 (Brieba and Sousa, 2000). Several 2′-modified ribonucleoside triphosphates are now commercially available (Jena BioScience, Trilink); Y639F T7 polymerase is available from Epicentre Technologies.

The replacement of the ribose 2′-OH group with other chemical moieties interferes with the primary mechanism for nuclease cleavage of RNA, attack of the 2′-hydroxyl on the bridging phosphate. However, substitutions on the backbone, such as replacing the phosphate with a phosphorothioate, have also been shown to increase oligonucleotide stability in the presence of nucleases (Zon and Geiser, 1991). An additional benefit is that phosphorothioate nucleotides have been shown to be incorporated into an elongating transcript by T7 polymerase with little or no increase in K_M (Griffiths et al., 1987). While DNA is not as vulnerable to hydrolysis as RNA, it is nonetheless susceptible to cleavage by a variety of deoxyribonucleases and phosphodiesterases. The stability of DNA can also be increased by the incorporation of phosphorothioate nucleotides, and these can be readily incorporated by *Taq* DNA polymerase (Nakamaye et al., 1988). These modified ribonucleoside and deoxyribonucleoside triphosphates are also commercially available (NEN Life Sciences, Sigma, Life Technologies).

Nucleic acid selection experiments have generated a wide variety of binding species (aptamers) and catalysts (ribozymes and deoxyribozymes). However, nucleic acids are by and large not as functional as proteins: aptamers bind their ligands less well than antibodies, for the most part, while ribozymes are slower than

protein enzymes, for the most part. While it is possible that these limitations are a function of the relative newness of aptamer and catalyst selections relative to a well-established field like protein engineering, it is also possible that functional nucleic acids are inherently limited by the diversity of the genetic alphabet. The binding interactions and chemical reactions performed by a nucleic acid biopolymer may be constrained by the functional groups they contain. The ability to expand the functional groups available to a DNA or RNA polymer through the incorporation of modified nucleotides could potentially open up realms of chemistry and binding interactions that were previously inaccessible.

Researchers are currently trying to assess this hypothesis and demonstrate the utility of modified bases in functional nucleic acids. For example, the first RNA capable of catalyzing the formation of carbon-carbon bonds utilized 5-pyridylmethylcarboxamid-UTP in place of UTP (Tarasow et al., 1997). When the most active isolate from this selection was transcribed with unmodified UTP, it was inactive. Similar results were obtained for a ribozyme that catalyzed amide bond formation (Wiegand et al., 1997) whose activity was dependent on the incorporation of 5-imidizole-UTP. Additionally, two sequence-specific RNase deoxyribozymes have been selected that were dependent on the incorporation of a 5-imidizole-TTP (Santoro et al., 2000), and both 8-2-(4-imidazolyl)ethylamino-2′-dATP and 5-(3-aminoallyl)-2′-dUTP (Perrin et al., 2001). However, none of the sequences, motifs, or activities found in these selection experiments was directly compared with ribozymes that contained canonical nucleotides and that were sieved from the same pool using the same selection conditions. Therefore, at present time, it is unclear whether modified nucleotides truly improve RNA catalysis.

Surprisingly, though, there are at least a few counterexamples that would suggest that modified nucleotides do not greatly contribute to binding or catalysis relative to natural nucleotides. Santoro and colleagues (Santoro and Joyce, 1997; Santoro et al., 2000) selected deoxyribozymes with RNA hydrolysis activity from different aliquots of the same, unamplified random sequence pool. The selection performed with natural nucleotides produced a much faster catalyst. Ultimately, it is unknown whether this indicates the superiority of natural nucleotides for this pool and this function, or whether the fraction of the original pool used for the selection of the natural catalyst just contained a "jackpot" sequence. An additional example of a catalytic activity selected for with both modified and unmodified pools is the Diels-Alder catalyzing ribozymes (Tarasow et al., 1997; Seelig and Jaschke, 1999, respectively). The modified selection yielded a catalyst with a k_{cat}/K_M of ~4 $M^{-1}sec^{-1}$, while the unmodified selection yielded a catalyst with a k_{cat}/K_M of 167 $M^{-1}sec^{-1}$. However, these selections were performed by different research groups with different pools, and thus are not directly comparable. For a final example, see Critical Parameters, discussion of multiple approaches.

Critical Parameters

Optimizing preparative reactions

A good portion of this protocol has been devoted to testing reactions and optimizing conditions. Although this process will likely take several days, it is time wisely spent. The first round of selection is by far the most important, as this is the round when selection conditions will query the greatest number of possible answers. Concomitantly, creating the largest possible number of unique sequences (10^{13} to 10^{15}) for the first round of selection is a critical task. However, because creating a large library requires a large amount of effort and large volumes of relatively expensive reagents, initial optimization and practice should always be performed on a small scale.

Similarly, it is expected that only a small number of the input molecules will survive any given round; therefore, it is essential that these successful sequences be efficiently carried to the next round. Inefficient reverse transcription or amplification reactions may lead to a selection for molecules that can be efficiently replicated (amplification artifacts) rather than to molecules that are highly functional.

Archiving reactions

In vitro selection experiments are particularly grueling because the ultimate outcome, the successful isolation of binding species or catalysts, may not be known for days, weeks, or even months. This is especially true when working with modified nucleotides, because of the strong possibility that one or more amplification reactions or selection steps will not work at some point during the course of the selection. Therefore, it is desirable to keep an archived copy of each round. After the initial round, there should be multiple copies of each winning

sequence present in the selected population. Thus, it is not necessary to use all of the RNA, cDNA, or double-stranded DNA template in a given selection or amplification step. Rather, some portion of the pool should be saved at any convenient point (e.g., as a double-stranded PCR product). In the event an unsuccessful round of selection should occur, it will not be necessary to repeat the entire experiment. Rather, selection can continue from the last successful round, with a minimal loss of time and effort.

Understanding modified nucleotide chemistry

The incorporation of additional functional groups into the context of an RNA backbone is expected to increase the diversity of its available interactions, both with itself and with its desired substrates. Thiol groups, for example, can participate directly in catalysis as nucleophiles. Additionally, disulfide bonds could be formed intramolecularly between thiols. This may add to the structural diversity of nucleic acids, normally limited to hydrogen bonding, salt bridges with metals, and stacking interactions. Charged groups, such as a lysine-like side chain, could potentially add to the structural repertoire of nucleic acids by allowing the formation of electrostatic interactions and salt bridges. The inclusion of chemical moieties with a pK_a closer to neutrality, such as an imidizole group, is also expected to benefit nucleic acid chemistry. Natural nucleotide bases contain no functional groups with unperturbed pK_a values between 3.5 and 9.2, inherently limiting the proton "push-pull" chemistry found in so many protein enzymes. The introduction of these and other functional groups can also potentially increase the abilities of nucleic acids to bind metals.

While attempts to associate nucleotide functionality with binding or catalytic properties are at best still guesswork, it is nonetheless true that failing to appreciate the chemical properties of modified nucleotides can potentially adversely affect how a selection experiment proceeds. For example, the inclusion of thiolated nucleotides can potentially lead to the unwanted formation of disulfide-bonded products if care is not taken to include a reducing agent in enzyme and/or selection reactions. For example, deoxyribozyme ligases that form an unnatural internucleotide linkage between a 5′-iodinated pool and an oligonucleotide substrate with a 3′ phosphorothioate have been selected from random sequence pools (Levy and Ellington, 2001a,b). Unless reducing agents are kept in the selection reaction, the oligonucleotide substrates will dimerize, reducing their effective concentration. Similarly, the inclusion of modified nucleotides that have altered pK_as could very easily change the pH of a concentrated stock solution, and care should be taken to make sure that all such stocks are at the desired pH and appropriately buffered. The accidental alteration of pH in enzyme or selection reactions can of course lead to decreases in product yield. Finally, the inclusion of modified nucleotides that have new metal binding or chelating properties may alter the available metal concentrations in an enzyme or selection reaction. In particular, the chelation of magnesium can lead to large changes in the efficiency of product formation by many different polymerases.

Multiple approaches

As has been apparent throughout, the in vitro selection of functional nucleic acids that contain modified nucleotides is still in its infancy, and thus there are few hard and fast rules regarding what will and will not work. Because of this, researchers need to be somewhat versatile in their approach to selection experiments involving modified nucleotides. If a given approach does not work, this does not mean that the selection experiment inherently has no chance of working, but instead indicates that it may be necessary to alter one or more parameters.

In particular, to improve the incorporation of modified nucleotides into a pool there are three variables that should be adjusted as needed. First, the buffer conditions for incorporation can be adjusted. Padilla and Sousa (1999) have systematically investigated several buffer conditions that aid the incorporation of nucleotides modified specifically at the 2′ position. These authors find better incorporation upon supplementing a transcription reaction with 0.5 mM $MnCl_2$, 1 U/μL pyrophosphatase, and either 8 mM spermidine for plasmid templates or 8 mM spermine for short DNA templates. Second, different polymerases clearly have different potentialities for the incorporation of modified nucleotides. For example, the Benner group has tested several thermostable DNA polymerases for their ability to incorporate a variety of modified nucleotides (Lutz et al., 1998, 1999). Finally, the sequence of the pool itself can have a surprisingly large effect on selection experiments. The authors originally selected a relatively fast ribozyme ligase from an RNA pool that contained 90 random

sequence positions (Robertson and Ellington, 1999; Robertson et al., 2001). A variant of this pool was generated in which only four positions in the constant region were altered, substituting a GAAA tetraloop in place of a UUCG tetraloop. Selection experiments for ribozyme ligases were initiated from amplified aliquots of this pool; one aliquot was used for the selection of ribozymes containing canonical nucleotides, other aliquots of identical complexity were used for the selection of ribozymes containing one of three different modified nucleotides (Robertson, 2001). After six rounds of selection and amplification, all of the pools had collapsed to contain a relatively few winning sequences. However, the rates of all of these winning sequences were remarkably slow (500-fold slower) compared to the ribozyme ligase originally selected from the slightly different pool. There was no discernable difference between the ribozymes that contained canonical or modified nucleotides; they were all very slow. The authors hypothesized that the pools had somehow become artificially narrowed during the course of the selection, and therefore repeated the selection experiments with new aliquots of the original, amplified pool. Again, only very slow ribozymes were obtained.

Modified nucleotides and mutations

Although a selection starts with a large number of sequences, this number is usually a small fraction of the total number of sequences possible for the length of the random region. Additionally, with each round, the number of sequences within the pool is diminished. As such, it may be useful to explore the sequence space around the selected winners in order to discover functional variants. To some extent, this occurs in any selection due to the inherent mutations that arise in the amplification process; however, this background mutation rate is small, and one may wish to increase the frequency of mutations. For these purposes, the mutagenic potential of a nucleotide analog that serves as a nonspecific template can be used to increase the diversity of a pool. Similar to mutagenic PCR (*UNIT 9.4*), these techniques can be used at the outset of a selection to mutate an existing ribozyme or aptamer for either optimization or reselection for altered specificity. For example, Kore et al. (2000) have used modified nucleotides to create a degenerate pool based on the hammerhead ribozyme, from which they selected a variant that cleaves at an alternate sequence. Additionally, a mutagenic step can be incorporated between rounds in a selection to more efficiently explore sequence space. However, it should be noted that in the latter case, one does not want the mutation rate to be so high that a few of each winning sequence from one round do not survive unchanged into the next round.

One benefit of using modified nucleotides rather than standard mutagenic PCR is that it is easy to control the amount of mutation introduced into the sample simply by adjusting the relative rate of modified to unmodified nucleotide in the amplification process. As such, a higher rate of mutation can be achieved than with mutagenic PCR, which would be particularly useful for diversifying an initial pool. Zaccolo et al. (1996) describe such a system using both a purine and pyrimidine analog in a PCR reaction, and have quantitated the frequencies of each base transition. It should be mentioned that when using mutagenic modified nucleotides, one would not want them to be included in the active pool, as their decreased fidelity would make it less likely that they would be in the same position during the next round. For example, if a DNA selection were being performed, a second amplification of the pool would be required using only natural nucleotides.

Anticipated Results

As with any selection experiments (*UNITS 9.3 & 9.4*), it is virtually impossible to anticipate the outcome of any given experiment. This is especially true when considering the incorporation of modified nucleotides, since relatively few selection experiments have so far been carried out with modified residues. However, if the protocols for the production of RNA and DNA pools that have been outlined are followed, it can be anticipated that it should be possible to generate and purify upwards of at least 10^{14} different nucleic acid sequences (~10 μg) that contain a particular modified nucleotide. For successful selection experiments, the population should be sieved by a factor of 100 to 1000 each round. That is, <1% of the total population will be recovered following a round of selection. Thus, most successful selections will show significant functional improvement or be completed within 5 to 10 rounds. However, in some selections it may take very successful functional sequences a longer period of time to overcome a great sea of mediocre functional sequences. For example, it took 18 rounds of selection to isolate cGMP-dependent aptazymes from a random sequence population

(Koizumi et al., 1999). Following selection, the population should contain a relatively small number (1 to 10) of sequence classes. A sequence class is a set of related sequences that contain a common motif. To date, the sequences of aptamers or ribozymes selected from pools containing modified nucleotides do not correspond to the sequences of aptamers or ribozymes selected from pools that contain canonical nucleotides. Thus, any selection that includes modified nucleotides is likely to produce new sequence motifs.

Time Considerations

The preparative portions of each round of selection can be expected to take several days. A transcription could take several hours to overnight, followed by ~2-hr gel purification, and an overnight elution from the excised gel pieces. The amount of time spent on the selective steps will vary greatly with each individual selection. However, in initial rounds, the reaction time alone is typically several hours to an entire day. The selective segregation of winning sequences from losing sequences will vary from a few to several hours, depending on the method used. Reverse transcription followed by PCR, isolation, and quantitation of new template DNA will take a few more hours, depending mainly on the number of cycles required for amplification. The largest variable, the number of rounds of selection required, will vary greatly depending on the activity sought, from only a few to ≥12. Cloning and assaying of individual sequences will likely also take several days. Thus, even if no problems are encountered during a selection, it is likely to take ≥1 month to complete.

Literature Cited

Bartel, D.P. and Szostak, J.W. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science* 261:1411-1418.

Battersby, T.R., Ang, D.N., Burgstaller, P., Jurczyk, S.C., Bowser, M.T., Buchanan, D.D., Kennedy, R.T., and Benner, S.A. 1999. Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog. *J. Am. Chem. Soc.* 121:9781-9789.

Beaudry, A., DeFoe, J., Zinnen, S., Burgin, A., and Beigelman, L. 2000. In vitro selection of a novel nuclease-resistant RNA phosphodiesterase. *Chem. Biol.* 7:323-334.

Brieba, L.G. and Sousa, R. 2000. Roles of histidine 784 and tyrosine 639 in ribose discrimination by T7 RNA polymerase. *Biochemistry* 39:919-923.

Green, L.S., Jellinek, D., Bell, C., Beebe, L.A., Feistner, B.D., Gill, S.C., Jucker, F.M., and Janjic, N. 1995. Nuclease-resistant nucleic acid ligands to vascular permeability factor/vascular endothelial growth factor. *Chem. Biol.* 2:683-695.

Griffiths, A.D., Potter, B.V., and Eperon, I.C. 1987. Stereospecificity of nucleases towards phosphorothioate-substituted RNA: Stereochemistry of transcription by T7 RNA polymerase. *Nucl. Acids Res.* 15:4145-4162.

Jellinek, D., Green, L.S., Bell, C., Lynott, C.K., Gill, N., Vargeese, C., Kirschenheuter, G., McGee, D.P., Abesinghe, P., Pieken, W.A., et al. 1995. Potent 2′-amino-2′-deoxypyrimidine RNA inhibitors of basic fibroblast growth factor. *Biochemistry* 34:11363-11372.

Jhaveri, S., Olwin, B., and Ellington, A.D. 1998. In vitro selection of phosphorothiolated aptamers. *Bioorg. Med. Chem. Lett.* 8:2285-2290.

King, D.J., Ventura, D.A., Brasier, A.R., and Gorenstein, D.G. 1998. Novel combinatorial selection of phosphorothioate oligonucleotide aptamers. *Biochemistry* 37:16489-16493.

Koizumi, M., Soukup, G.A., Kerr, J.N., and Breaker, R.R. 1999. Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP. *Nat. Struct. Biol.* 6:1062-1071.

Kore, A.R., Vaish, N.K., Morris, J.A., and Eckstein, F. 2000. In vitro evolution of the hammerhead ribozyme to a purine-specific ribozyme using mutagenic PCR with two nucleotide analogues. *J. Mol. Biol.* 301:1113-1121.

Kostyuk, D.A., Dragan, S.M., Lyakhov, D.L., Rechinsky, V.O., Tunitskaya, V.L., Chernov, B.K., and Kochetkov, S.N. 1995. Mutants of T7 RNA polymerase that are able to synthesize both RNA and DNA. *FEBS Lett.* 369:165-168.

Kubik, M.F., Bell, C., Fitzwater, T., Watson, S.R., and Tasset, D.M. 1997. Isolation and characterization of 2′-fluoro-, 2′-amino-, and 2′-fluoro-/amino-modified RNA ligands to human IFN-gamma that inhibit receptor binding. *J. Immunol.* 159:259-267.

Kujau, M.J., Siebert, A., and Wolfl, S. 1997. Design of leader sequences that improve the efficiency of the enzymatic synthesis of 2′-amino-pyrimidine RNA for in vitro selection. *J. Biochem. Biophys. Methods* 35:141-151.

Latham, J.A., Johnson, R., and Toole, J.J. 1994. The application of a modified nucleotide in aptamer selection: Novel thrombin aptamers containing 5-(1-pentynyl)-2′-deoxyuridine. *Nucl. Acids Res.* 22:2817-2822.

Levy, M. and Ellington, A.D. 2001a. In vitro selection of a deoxyribozyme that can utilize multiple substrates. *J. Mol. Evol.* In press.

Levy, M. and Ellington., A.D. 2001b. Section of deoxyribozyme ligases that catalyze the formation of an unnatural internucleotide linkage. *Bioorg. Med. Chem.* 9:2581-2587.

Lin, Y., Qiu, Q., Gill, S.C., and Jayasena, S.D. 1994. Modified RNA sequence pools for in vitro selection. *Nucl. Acids Res*. 22:5229-5234.

Lutz, M.J., Horlacher, J., and Benner, S.A. 1998. Recognition of a non-standard base pair by thermostable DNA polymerases. *Bioorg. Med. Chem. Lett*. 8:1149-1152.

Lutz, S., Burgstaller, P., and Benner, S.A. 1999. An in vitro screening technique for DNA polymerases that can incorporate modified nucleotides. Pseudo-thymidine as a substrate for thermostable polymerases. *Nucl. Acids Res*. 27:2792-2798.

Marshall, K.A. and Ellington, A.D. 1999. Molecular parasites that evolve longer genomes. *J. Mol. Evol*. 49:656-663.

Milligan, J.F. and Uhlenbeck, O.C. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol*. 180:51-62.

Nakamaye, K.L., Gish, G., Eckstein, F., and Vosberg, H.P. 1988. Direct sequencing of polymerase chain reaction amplified DNA fragments through the incorporation of deoxynucleoside alpha-thiotriphosphates. *Nucl. Acids Res*. 16:9947-9959.

Padilla, R. and Sousa, R. 1999. Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2′-groups using a mutantT7 RNA polymerase (RNAP). *Nucl. Acids Res*. 27:1561-1563.

Pagratis, N.C., Bell, C., Chang, Y.F., Jennings, S., Fitzwater, T., Jellinek, D., and Dang, C. 1997. Potent 2′-amino- and 2′-fluoro-2′-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor. *Nat. Biotechnol*. 15:68-73.

Perrin, D.M., Garestier, T., and Helene, C. 2001. Bridging the gap between proteins and nucleic acids: A metal-independent RNAseA mimic with two protein-like functionalities. *J. Am. Chem. Soc*. 123:1556-1563.

Piccirilli, J.A., Krauch, T., Moroney, S.E., and Benner, S.A. 1990. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* 343:33-37.

Pieken, W.A., Olsen, D.B., Benseler, F., Aurup, H., and Eckstein, F. 1991. Kinetic characterization of ribonuclease-resistant 2′-modified hammerhead ribozymes. *Science* 253:314-317.

Robertson, M.P. 2001. Ph.D. dissertation. Engineered regulation of an RNA ligase ribozyme. University of Texas, Austin.

Robertson, M.P. and Ellington, A.D. 1999. In vitro selection of an allosteric ribozyme that transduces analytes to amplicons. *Nat. Biotechnol*. 17:62-66.

Robertson, M.P., Hesselberth, J.R., and Ellington, A.D. 2001. Optimization and optimality of a short ribozyme ligase that joins non-Watson-Crick base pairings. *RNA 7:513-523.*

Ruckman, J., Green, L.S., Beeson, J., Waugh, S., Gillette, W.L., Henninger, D.D., Claesson-Welsh, L., and Janjic, N. 1998. 2′-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *J. Biol. Chem*. 273:20556-20567.

Santoro, S.W. and Joyce, G.F. 1997. A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci. U.S.A*. 94:4262-4266.

Santoro, S.W., Joyce, G.F., Sakthivel, K., Gramatikova, S., and Barbas, C.F. III. 2000. RNA cleavage by a DNA enzyme with extended chemical functionality. *J. Am. Chem. Soc.* 122:2433-2439.

Seelig, B. and Jaschke, A. 1999. A small catalytic RNA motif with Diels-Alderase activity. *Chem. Biol*. 6:167-176.

Southworth, M.W., Kong, H., Kucera, R.B., Ware, J., Jannasch, H.W., and Perler, F.B. 1996. Cloning of thermostable DNA polymerases from hyperthermophilic marine Archaea with emphasis on Thermococcus sp. 9°N-7 and mutations affecting 3′-5′ exonuclease activity. *Proc. Natl. Acad. Sci. U.S.A*. 93:5281-5285.

Tarasow, T.M., Tarasow, S.L., and Eaton, B.E. 1997. RNA-catalysed carbon-carbon bond formation. *Nature* 389:54-57.

Teramoto, N., Imanishi, Y., and Ito, Y. 2000. In vitro selection of a ligase ribozyme carrying alkylamino groups in the side chains. *Bioconjug. Chem*. 11:744-748.

Wiegand, T.W., Janssen, R.C., and Eaton, B.E. 1997. Selection of RNA amide synthases. *Chem. Biol*. 4:675-683.

Zaccolo, M., Williams, D.M., Brown, D.M., and Gherardi, E. 1996. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol*. 255:589-603.

Zon, G. and Geiser, T.G. 1991. Phosphorothioate oligonucleotides: Chemistry, purification, analysis, scale-up and future directions. *Anticancer Drug Des*. 6:539-568.

Contributed by Scott M. Knudsen, Michael P. Robertson, and Andrew D. Ellington
University of Texas
Austin, Texas