

# Toward Automated Nucleic Acid Enzyme Selection

**Letha J. Sooter, Timothy Riedel, Eric A. Davidson, Matthew Levy, J. Colin Cox and Andrew D. Ellington\***

Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA

\*Corresponding author

**Methods for automation of nucleic acid selections are being developed. The selection of aptamers has been successfully automated using a Biomek 2000 workstation. Several binding species with nanomolar affinities were isolated from diverse populations. Automation of a deoxyribozyme ligase selection is in progress. The process requires eleven times more robotic manipulations than an aptamer selection. The random sequence pool contained a 5' iodine residue and the ligation substrate contained a 3' phosphorothioate. Initially, a manual deoxyribozyme ligase selection was performed. Thirteen rounds of selection yielded ligators with a 400-fold increase in activity over the initial pool. Several difficulties were encountered during the automation of DNA catalyst selection, including effectively washing bead-bound DNA, pipetting 50% glycerol solutions, purifying single strand DNA, and monitoring the progress of the selection as it is performed. Nonetheless, automated selection experiments for deoxyribozyme ligases were carried out starting from either a naive pool or round eight of the manually selected pool. In both instances, the first round of selection revealed an increase in ligase activity. However, this activity was lost in subsequent rounds. A possible cause could be mispriming during the unmonitored PCR reactions. Potential solutions include pool redesign, fewer PCR cycles, and integration of a fluorescence microtiter plate reader to allow robotic 'observation' of the selections as they progress.**

*Key words:* Deoxyribozyme/*In vitro* Selection/Ligase/Robot/SELEX.

## Introduction

Functional nucleic acid species can be selected from random sequence libraries, and in many instances have been shown to have attributes that rival those of proteins. For example, selected nucleic acid binding species (aptamers) frequently interact with their cognate targets with affinities and specificities that rival those of monoclonal

antibodies (Osborne and Ellington, 1997; Famulok and Jenne, 1998; Famulok and Mayer, 1999), while selected allosteric nucleic acid catalysts (aptazymes) have been shown to have activation parameters far in excess of those normally observed in allosteric protein enzymes (Soukup and Breaker, 2000). Moreover, because the entire selection process occurs *in vitro* the identification of functional nucleic acids is inherently more facile than the slower and more laborious *in vivo* selection of functional proteins.

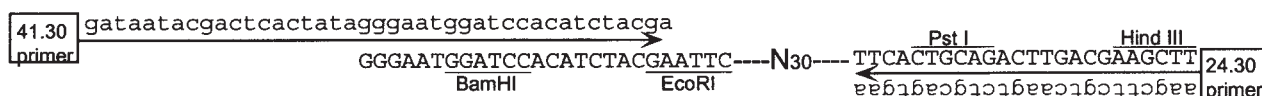
In order to develop biosensors for proteome and metabolome acquisition and apprehension, we have begun to develop automated methods for nucleic acid selection (Cox *et al.*, 1998; Cox and Ellington, 2001). As a first step, we have automated the selection of aptamers that bind to protein targets. Initially, a Biomek 2000 workstation was used to automate the final few rounds of manual selection experiments. These experiments successfully winnowed still diverse sequence populations to one or a few families or binding species. Further modifications of the protocol have allowed us to carry out complete selections against single protein targets within two days (Cox and Ellington, 2001). Selected aptamers bind their targets with nanomolar affinities.

The development of automated selections of nucleic acid catalysts has proven to be even more challenging. In order to make catalyst selections as amenable to automation as possible, we initially carried out a very simple manual selection for deoxyribozyme ligases. Our progress and the difficulties associated with adapting this manual selection protocol to an automated workstation will be discussed in detail.

## Results and Discussion

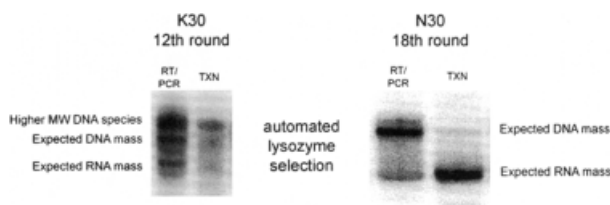
### Adapting Molecular Biology to a Robotic Workstation: Selection of Anti-Protein Aptamers

The various manual manipulations that are typically performed during *in vitro* selection experiments are straightforward, but their adaptation to a robotic workstation nonetheless proved difficult (reported more completely in Cox *et al.*, 1998, and Cox and Ellington, 2001). In particular, the workstation is currently incapable of gel isolation and ethanol precipitation, and this necessitated the development of methods that either required alternative purification or no purification. We initially chose to avoid purification entirely, and to instead finely tune amplification chemistry so as to maximize the yield of the desired products. For example, the random sequence pool that has so far been used for automated aptamer selections, N30 (Figure 1), has previously been shown to yield aptamers



**Fig. 1** The N30 Pool.

This pool has been described previously (Bell *et al.*, 1998). Residues that will be found in the transcribed RNA pool are capitalized; primer sequences are in lower case. The primer containing the T7 RNA polymerase promoter is shown at the left, while the primer containing the reverse transcription start site is at the right. Restriction enzyme recognition sites that facilitate cloning are indicated.



**Fig. 2** Amplifiability of K30 and N30 Pools.

Automated selections were carried out with either the K30 or N30 pools using lysozyme as a target. Products with higher molecular weight than the starting material are seen in both the reverse transcription/PCR (RT/PCR) and transcription (TXN) reactions following twelve rounds of automated selection. In contrast, higher molecular weight products do not accumulate during the automated selection of the N30 pool.

against a variety of protein targets (Bell *et al.*, 1998; Polard *et al.*, 2000). More importantly, the constant regions of the N30 pool were designed so that primers would hybridize poorly at their 3' ends. This strategy has previously been shown to allow 'infinite amplification' of PCR products (Cramer and Stemmer, 1993) without the accumulation of higher molecular weight products. This facet of experimental design is especially important given that the workstation cannot monitor the success of each cycle of selection and amplification, and thus the input into standard RT-PCR reactions will vary as the selection progresses. Nonetheless, despite the fact that a set number of PCR amplification cycles are used in each round (16 to 20, depending on the round) very few alternative products have been observed over a number of selections. The importance of pool design for success has been further emphasized by automated selection experiments that involved other DNA pools. An alternative pool, K30 (Marshall and Ellington, 1999, 2000), is essentially the 'inverse' of the N30 pool: all purines in the constant regions (with the exception of the T7 RNA polymerase promoter) were switched for pyrimidines, and *vice versa* (G for C, A for T, T for A, and C for G, except for the T7 RNA polymerase promoter). In automated selection experiments the K30 pool quickly accumulated higher molecular weight species; after several rounds these species dominated the selection (Figure 2).

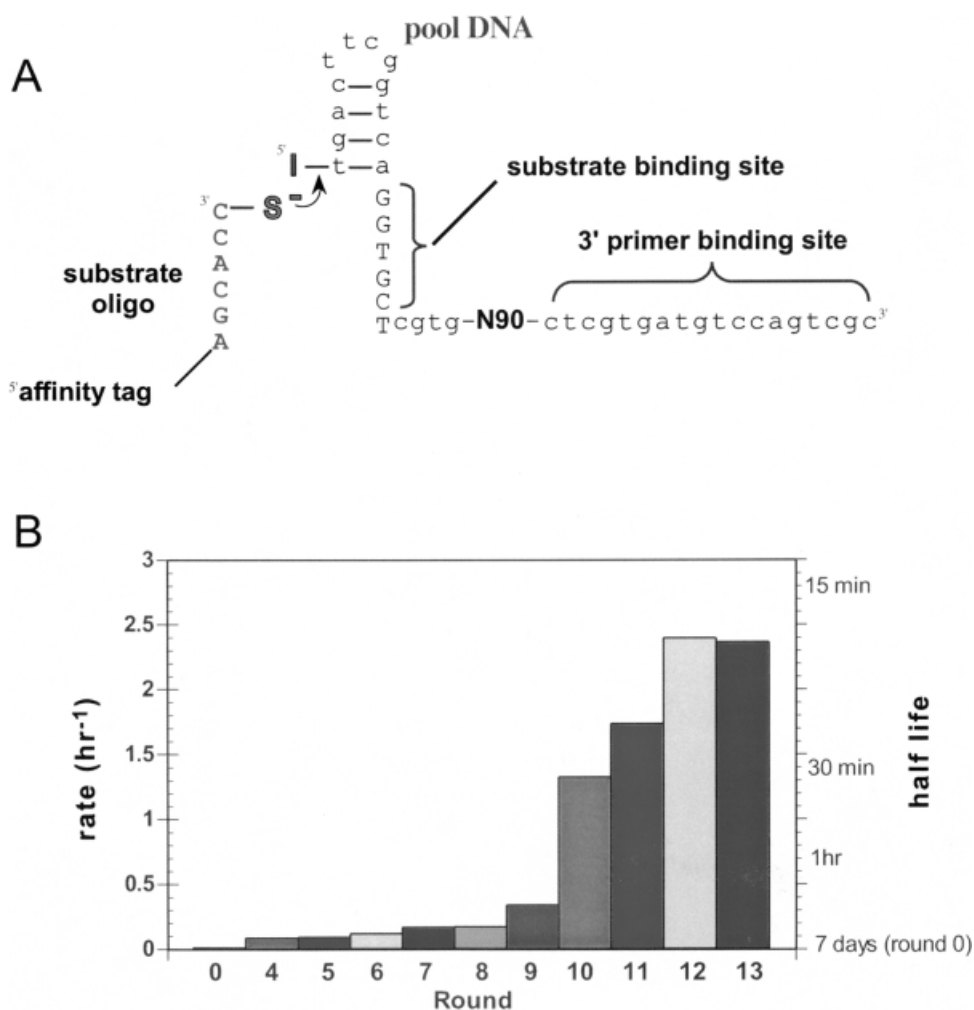
The partitioning of selected species is also different on a robotic workstation than it is for a human experimentalist. Protein:nucleic acid complexes are typically parsed from free nucleic acids by filtration through a modified cellulose filter. This procedure is amenable to automation

(via a Multiscreen filtration plate holder from Millipore, Bedford, USA). However, efficient elution of bound species from the filter proved to be difficult, primarily because manual resuspension procedures require mechanical manipulation. To obviate this problem protein targets are first immobilized on beads, and bead:protein:nucleic acid complexes are then filtered from free nucleic acids. This procedure has most of the advantages of stringent filter separation of complexes but also allows the ready manipulation of beads by robotic pipettes (Cox *et al.*, 1998; Cox and Ellington, 2001). Our familiarity with bead-based manipulations of nucleic acids proved to be especially important in the development of automated selections of nucleic acid catalysts.

Based on these advances, the robotic workstation can now be used to quickly identify aptamers against immobilized protein targets. We have successfully carried out automated selections against a number of targets, including tyrosyl tRNA synthetase from *Neurospora* mitochondria Cyt18, (Collins and Lambowitz, 1985; Guo and Lambowitz, 1992), lysozyme (Cox and Ellington, 2001), the herpes virus protein US11 (Roller *et al.*, 1996), and Rho from *Thermotoga maritima* (Liao and Dennis, 1992; Opperman and Richardson, 1994). In general, these selections were completed within two days and yielded families of aptamers with  $K_d$ 's in the nanomolar range. The one exception was the Rho protein from *T. maritima*, which yielded a series of aptamers that had high C:G ratios, similar to those previously found in a selection which targeted the Rho protein from *Escherichia coli* (Schneider *et al.*, 1993).

### Manual Selection of Deoxyribozyme Ligases as a Model System for Automation

In order to develop a model system that could be readily transferred to the robotic workstation, we initially explored the *in vitro* selection of deoxyribozyme ligases that catalyzed the formation of an unnatural internucleotide linkage (reported more fully in Levy and Ellington, 2001). In this system, a DNA pool that contained 90 random sequence positions and a 5' iodo-thymidine residue was generated (Figure 3A). This pool was successively reacted with a series of substrates that bore 3' phosphorothioate residues. After thirteen rounds of selection and amplification the catalytic ability of the pool to use a common substrate had improved by over 400-fold compared to the initial pool (Figure 3B). Several families of de-



**Fig. 3** 5IN90 Pool and Selection.

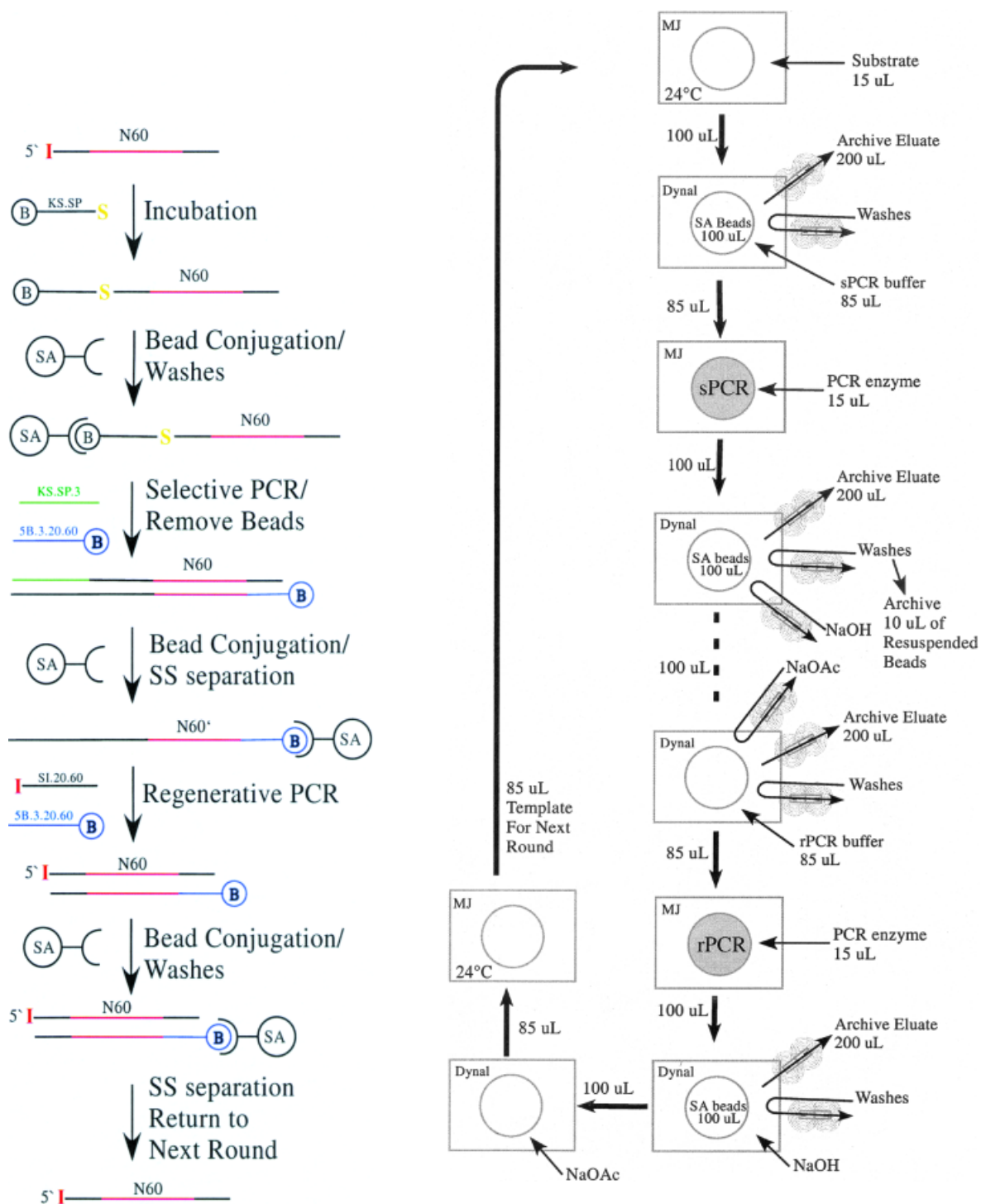
(A) This pool and selection has been described previously (Levy and Ellington, 2001). The constant sequence priming regions are shown in lowercase, except for the substrate binding site which is capitalized. The substrate oligonucleotide is shown in bold, and the 3' phosphorothioate (S) and 5' iodine (I) moieties are outlined. (B) Enhancement of ligation activity with successive rounds of selection. Rates were measured at 25 °C with 0.1  $\mu\text{M}$  single-stranded DNA, 0.15  $\mu\text{M}$  substrate, and 0.15  $\mu\text{M}$  reverse primer.

oxyribozymes could be discerned in the selected pool, and formed structures that appeared to catalyze the ligation reaction in part by the provision of an internal template.

#### Automated Selection of Nucleic Acid Catalysts

Given that it proved possible to select deoxyribozyme ligases from random sequence pools, we immediately attempted to adapt the selection procedure to the automated workstation. The automated schema is shown in Figure 4, and differs from the previously described manual selection procedure in several important ways. First, the substrate for the reaction was biotinylated. Following reaction with the DNA pool, ligated species were captured on magnetic beads bearing streptavidin. This allowed the facile separation of reacted and unreacted DNA species *via* the integrated magnetic bead separator (MPC-auto96, Dynal, Oslo, Norway). Second, in the

manual selection procedure unreacted DNA molecules were removed using 80 ml of wash buffer. To decrease the volume of washes required while maintaining the stringency of separation, the magnetic beads were washed with 7 M urea five times (200  $\mu\text{l}$ ) followed by fifteen washes with 10 mM Tris (pH 7.8, 200  $\mu\text{l}$ ). Retained, reacted DNA molecules were amplified directly on beads using a biotinylated primer (a 'selective' amplification step). The *Taq* polymerase solution (50% glycerol) proved difficult for the Biomek to pipette, and a special transfer step involving a slow aspiration was introduced. While some of the amplified DNA products likely remained bound to the streptavidin beads, the remaining amplified DNA was captured on a second set of magnetic beads and the non-biotinylated strand was removed by alkaline denaturation and neutralization. Several methods of single-strand DNA purification were initially explored: heat, urea, sodium hydroxide, heat and urea, and heat and sodium hydroxide. Double-stranded,



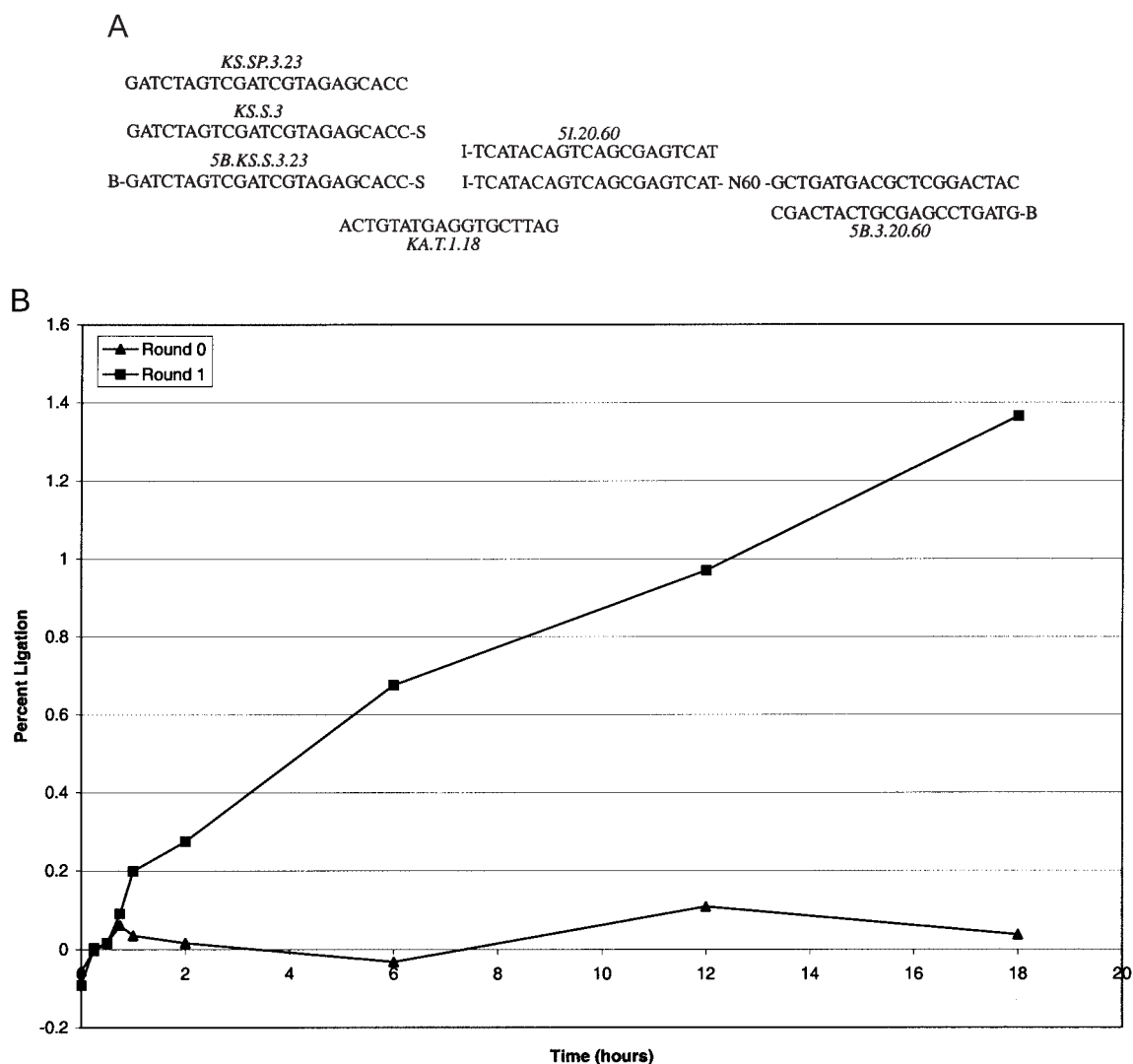
**Fig. 4** Automated Schema for Ligase Selections. For experimental details see the Materials and Methods section.

biotinylated PCR product radiolabeled with  $[^{32}\text{P}]$ -dATP was incubated with streptavidin coated magnetic beads. The beads were washed and then incubated for varying amounts of time with one of the five denaturing conditions. Afterwards, the solutions were separated from the beads and run on an acrylamide gel. Phosphoimage analysis revealed that sodium hydroxide, with or without heat, was the most effective denaturant. In ad-

ditional experiments, incubation times and temperatures were systematically varied. Following separation, single-stranded products were incubated with substrate (KS.S.3) and a splint oligonucleotide (KA.T.1.18) for 24 hours at 24 °C. The splint brings the pool and substrate into close proximity and should allow efficient ligation. Of course, if no 5' iodine moiety is present, ligation will not occur. The optimal combination of single strand

separation and 5' iodine moiety retention was achieved with a sodium hydroxide wash at room temperature. The remaining, biotinylated strand was then amplified using a second set of primers that restored the original form of the deoxyribozyme, including the 5' iodine moiety (a 'regenerative' amplification step). Single-stranded DNA was purified *via* alkaline denaturation and neutralization. The selected, single-stranded pool was then subjected to further rounds of selection and amplification. Overall, a single round of automated deoxyribozyme selection requires 11000 individual, programmed movements of the robot. To gauge the complexity of the procedure, it should be realized that the automated aptamer selections which have only recently been successfully implemented require only 1000 movements per round.

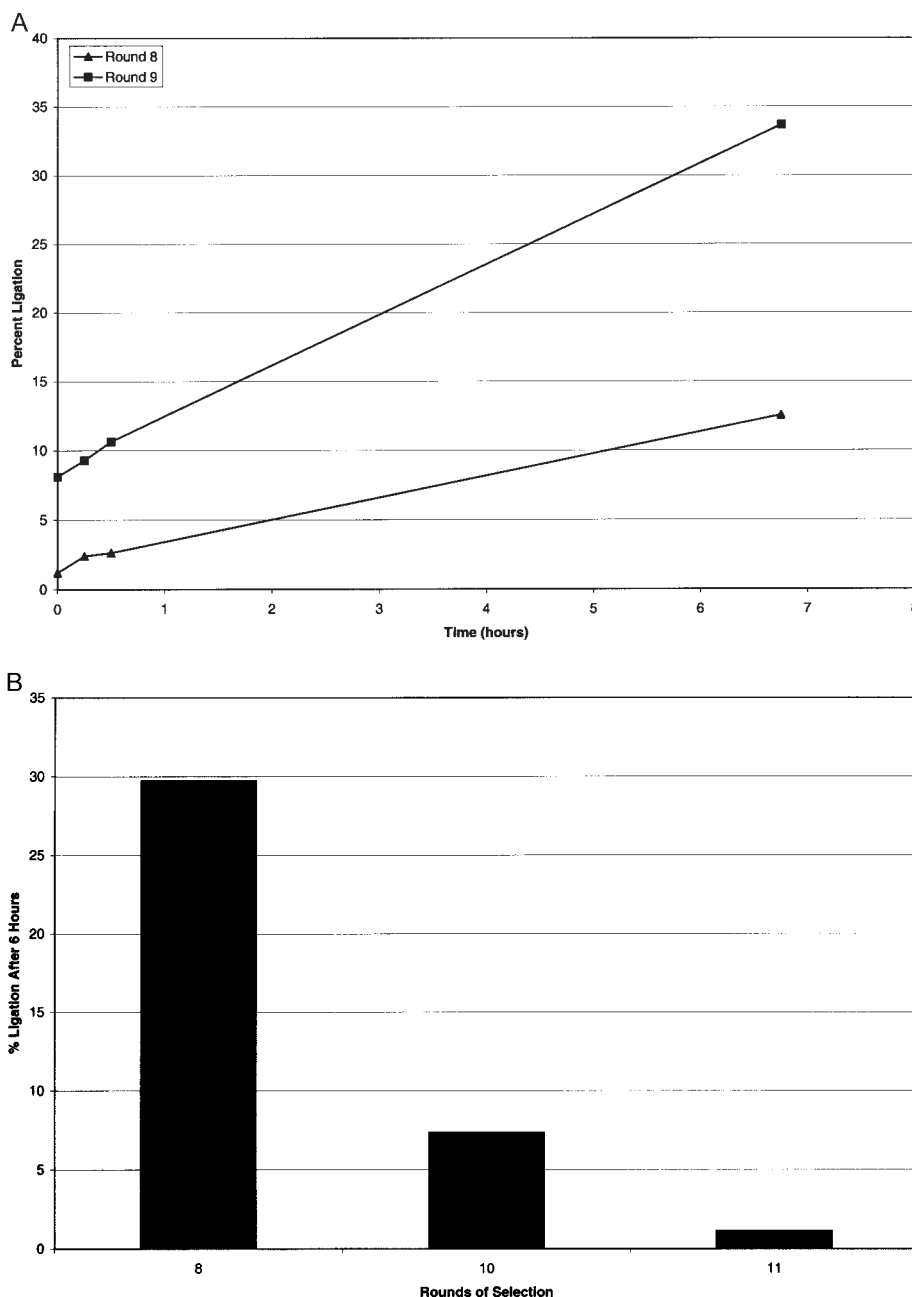
The automated selection procedure was originally attempted with a new pool, N60 (Figure 5A). After the first cycle of selection and amplification, deoxyribozyme ligase activity appeared to have accumulated (Figure 5B), but was lost in subsequent cycles. In order to better gauge whether the loss of ligase activity was due to the automated procedure or was somehow a function of the new pool, a second automated selection experiment was set up involving the deoxyribozyme pool from round 8 of the manual selection. Again, after one cycle of selection and amplification, deoxyribozyme ligase activity improved, but was lost in subsequent rounds (Figure 6). A close examination of the PCR products that accumulated during both the 'selective' and 'regenerative' steps revealed a series of bands, indicating that mis-priming was a serious problem during amplification reactions. Cycle-



**Fig. 5** Automated Selection with the N60 Pool.

(A) *KS.SP.3.23* is the forward selective PCR primer. *KS.S.3* is the same substrate with a 3' phosphorothioate (S), while *5B.KS.S.3.23* is the same substrate with a 5' biotin (B) and a 3' phosphorothioate. *5I.20.60* is the forward regenerative PCR primer with a 5' iodine (I). *5B.3.20.60* is the reverse primer with a 5' biotin.

(B) Selection rounds 0–1 of the N60 pool. Reactions were carried out at 0.1  $\mu\text{M}$  N60 pool and 0.15  $\mu\text{M}$  substrate at 24  $^{\circ}\text{C}$  in selection buffer (500 mM NaCl, 50 mM DTT, 50 mM Tris-HCl, pH 7.4).



**Fig. 6** Automated Selection with the N90 Pool.

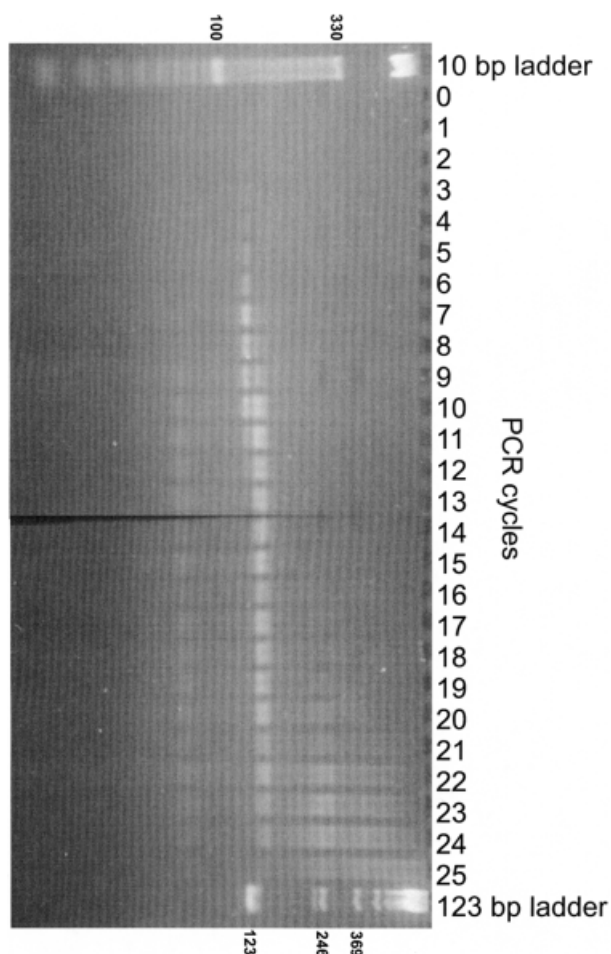
(A) Automated selection following 8 rounds of manual selection with the N90 pool. Assays were carried out at  $0.025 \mu\text{M}$  N90 pool and  $0.0375 \mu\text{M}$  substrate at  $24^\circ\text{C}$  in selection buffer (500 mM NaCl, 50 mM DTT, 50 mM Tris-HCl, pH 7.4).

(B) Selection rounds 8, 10, and 11 of the N90 pool. Reactions were carried out at  $0.1 \mu\text{M}$  N90 pool and  $0.15 \mu\text{M}$  substrate at  $24^\circ\text{C}$  in selection buffer (500 mM NaCl, 50 mM DTT, 50 mM Tris-HCl, pH 7.4). Note that these assay conditions differ from those in (A), and thus round 9 is not represented on this graph.

course reactions that mimicked the automated 'selective' and 'regenerative' amplifications also showed the accumulation of higher molecular weight bands after nine PCR amplification cycles (Figure 7).

It seems likely that the pools that were used for deoxyribozyme selection may be prone to the accumulation of amplification artefacts, just as the K30 pool was. There are several possible solutions to this problem. First and most obviously, the N60 pool can be re-designed with al-

ternative constant regions that are potentially less prone to the accumulation of amplification artefacts. Second, fewer rounds of amplification can be used during both the 'selective' and 'regenerative' PCR amplification steps. However, this latter alternative points up the single most important problem with robotic selection: in its current form it tends to be blind. Rounds of selection and amplification can be carried out, but until the machine has completed its run there is no way to ascertain how



**Fig. 7** Cycle Course of 'selective' PCR.

Two identical 100  $\mu$ l PCR reactions were carried out. Five  $\mu$ l were removed after one minute of extension time at each cycle. Samples were loaded on a 4% agarose gel, the gel was developed in  $1 \times$  TBE buffer at 60 V, and bands were visualized following staining with ethidium bromide.

well the selection is progressing. This more global problem can potentially be alleviated by adapting analytical devices to the robot that would allow it to effectively 'observe' the functionality of a selected pool. To this end, we are in the process of integrating a fluorescence microtiter plate reader with the Biomek, and in the future may be able to use fluorescently labeled pools or substrates to gauge the fraction of the population that has reacted. This value can then be used to 'set' the number of PCR amplification cycles that will be carried out in order to regenerate a deoxyribozyme population.

## Materials and Methods

### Selection and Amplification Procedures

The selection begins with the incubation of a single-stranded DNA pool and biotinylated substrate in a thermal cycler (PTC-200, MJ Research, Waltham, USA) held at 24  $^{\circ}$ C. Following the incubation, the solution is transferred to a well in the magnetic bead

separator (Dyna) containing streptavidin coated magnetic beads in  $2 \times$  binding buffer (10 mM Tris, 2 M NaCl, pH 7.5). A bead binding incubation occurs with mixing, and then the beads undergo a stringent wash with five 200  $\mu$ l washes of 7 M urea and fifteen 200  $\mu$ l washes of 10 mM Tris, pH 7.5. This co-immobilizes the species in the pool that have successfully ligated to the substrate.

The beads are resuspended in a 'selective' PCR (sPCR) mixture and transferred to the thermal cycler where the immobilized DNA molecules serve as amplification templates. Fifteen  $\mu$ l of a 50% glycerol solution containing five units of Taq polymerase (Display Systems Biotech) are transferred from a homemade enzyme cooling unit and the thermal cycler automatically performs 10 PCR amplification cycles. The amplified DNA molecules are then transferred to the Dynal magnetic bead separator and the biotinylated double strands are captured on a new set of streptavidin beads. Following capture, the beads undergo a stringent wash with 7 M urea and 10 mM Tris as before. The beads are then incubated with 0.3 N NaOH to denature the double-stranded DNA, releasing the non-biotinylated single strands. The solution containing single-stranded DNA is removed and discarded, and the beads are returned to a neutral pH with 3 M sodium acetate (NaOAc), pH 5.2.

The beads bearing biotinylated, single-stranded DNA are washed with wash buffer (5 mM Tris-HCl, 1 M NaCl, pH 7.5) and used as the template for a 'regenerative' PCR (rPCR) reaction; 10 PCR amplification cycles are performed. This solution is transferred to the magnetic bead separator where the beads are captured, and the supernatant is transferred to a fresh set of streptavidin beads. Bead binding and stringent washing are followed by a single strand separation using 100 microliters of 0.3 N NaOH. The alkaline solution is transferred to a new well on the magnetic bead separator and neutralized with 60  $\mu$ l of 3 M sodium acetate, pH 5.2. This solution contains the regenerated, iodinated, single stranded DNA pool that will be used in the next round of selection.

## Acknowledgements

This work was supported by grants from the Beckman Foundation and ARO/MURI.

## References

- Bell, S.D., Denu, J.M., Dixon, J.E., and Ellington, A.D. (1998). RNA molecules that bind to and inhibit the active site of a tyrosine phosphatase. *J. Biol. Chem.* 273, 14309–14314.
- Collins, R.A., and Lambowitz, A.M. (1985). RNA splicing in *Neurospora* mitochondria. Defective splicing of mitochondrial mRNA precursors in the nuclear mutant *cyt18-1*. *J. Mol. Biol.* 184, 413–428.
- Cox, J.C., and Ellington, A.D. (2001). Automated selection of anti-protein aptamers. *Bioorg. Med. Chem.*, in press.
- Cox, J.C., Rudolph, P., and Ellington, A.D. (1998). Automated RNA selection. *Biotechnol. Prog.* 14, 845–850.
- Crameri, A., and Stemmer, W.P. (1993). 10(20)-fold aptamer library amplification without gel purification. *Nucleic Acids Res.* 21, 4410.
- Famulok, M., and Jenne, A. (1998). Oligonucleotide libraries – variatio delectat. *Curr. Opin. Chem. Biol.* 2, 320–327.
- Famulok, M., and Mayer, G. (1999). Aptamers as tools in molecular biology and immunology. *Curr. Top. Microbiol. Immunol.* 243, 123–36.
- Guo, Q., and Lambowitz, A.M. (1992). A tyrosyl-tRNA synthetase binds specifically to the group I intron catalytic core. *Genes Dev.* 6, 1357–1372.

- Levy, M., and Ellington, A.D. (2001). Selection of deoxyribozyme ligases that catalyze the formation of an unnatural internucleotide linkage. *Bioorg. Med. Chem.*, in press.
- Liao D., and Dennis, P.P. (1992). The organization and expression of essential transcription translation component genes in the extremely thermophilic eubacterium *Thermotoga maritima*. *J. Biol. Chem.* 267, 22787–22797.
- Marshall, K.A., and Ellington, A.D. (1999). Molecular parasites that evolve longer genomes. *J. Mol. Evol.* 49, 656–663.
- Marshall, K.A., and Ellington, A.D. (2000). *In vitro* selection of RNA aptamers. *Methods Enzymol.* 318, 193–214.
- Opperman T., and Richardson J.P. (1994). Phylogenetic analysis of sequences from diverse bacteria with homology to the *Escherichia coli* rho gene. *J. Bacteriol.* 176, 5033–5043.
- Osborne, S.E., and Ellington, A.D. (1997). Nucleic acid selection and the challenge of combinatorial chemistry. *Chem. Rev.* 97, 349–370.
- Pollard, J., Bell, S.D., and Ellington, A.D. (2000). Design, synthesis, and amplification of DNA pools for *in vitro* selection. In: *Current Protocols in Nucleic Acid Chemistry*, S.L. Beaucage *et al.*, eds. (New York, USA: John Wiley & Sons, Inc.), pp. 9.2.1–9.2.23.
- Roller R.J., Monk L.L., Stuart D., and Roizman B. (1996). Structure and function in the herpes simplex virus 1 RNA-binding protein U(s)11: mapping of the domain required for ribosomal and nucleolar association and RNA binding *in vitro*. *J. Virol.* 70, 2842–2851.
- Schneider, D., Gold, L., and Platt, T. (1993). Selective enrichment of RNA species for tight binding to *Escherichia coli* rho factor. *FASEB J.* 7, 201–207.
- Soukup, G.A., and Breaker, R.R. (2000). Allosteric nucleic acid catalysts. *Curr. Opin. Struct. Biol.* 10, 318–325.

*Received April 20, 2001; accepted July 23, 2001*