5. Discussion

5.1 Optimization of *In Vitro* Selection Conditions

The experimental determination of the relationship between structure and activity can be achieved by virtue of *in vitro* selection: functional molecules can thus be isolated from sequence space through repeated cycles of selection and amplification. The original distribution of functional molecules may however be distorted by biases in replication and selection efficiency that accumulate during the course of the selection. Therefore the use of optimal selection conditions is crucial.

As more difficult tasks require more complicated solutions, longer, more diverse sequence libraries (Sabeti et al., 1997; Bartel and Szostak, 1993a; Vaughn et al., 1996; Griffiths and Duncan, 1998; Roberts and Ja, 1999; Plückthun and Schaffitzel, 2001) are likely to be beneficial to avoid that simpler, less active, but more abundant solutions emerge first in an *in vitro* selection (Salehi-Ashtiani and Szostak, 2001). Since the four highest-affinity aptamers out of eleven GTP aptamers all originated from the pre-engineered part of the pool, the use of partially-structured libraries can contribute to increasing the likelihood of finding functional sequences. This occurs by nucleating folding and preventing the formation of many alternative, however inactive, tertiary structures, thereby focusing the search on regions of sequence space that are well populated by molecules with high information content. RNA is able to solve the problem of binding to GTP with high affinity in a large number of ways, but the emergence of a particular high-affinity aptamer in an *in vitro* selection as well as in natural evolution is unpredictable and ultimately a matter of chance.

5.1.1 Conclusions from Optimizing Selection Conditions

Any number of biochemical techniques can be applied in an *in vitro* selection to increase binding stringency such as using an excess amount of RNA over the target so that sequences can compete with each other for binding. Non-amplifiable competitor

molecules like tRNA are also often used to saturate non-specific binding sites. In addition, selection for slower off-rates can be achieved by extensive washing.

To ensure that all active species survive the initial round of selection, early rounds of selection are however typically performed under relatively permissive binding conditions since a very stringent initial selection would impose a bottleneck that could eliminate some active sequences that are initially present at very low frequencies. As functional sequences are enriched and replace non-functional sequences during the course of the selection, the stringency can be safely increased. In fact, once the pool has reached at least marginal activity, high selection stringency is required to efficiently discriminate between the best and sub-optimal sequences. Irvine et al. (1991) provide a detailed mathematical description of the optimal ligand and pool concentrations ensuring maximum enrichment without significant loss of the initially rare but best-performing molecules.

From the series of experiments described above, it could be concluded that pool RNA anneals to aptamer RNA and prevents a large fraction of it (over 95%, see blue curve, Fig. 8) from binding to GTP immobilized on thiopropyl sepharose beads. The following selection protocol was therefore developed and tested in order to establish optimal conditions for the selection of high-affinity aptamers:

Annealing and incubation of aptamer and pool RNA together was performed in a large volume (10 ml) of 1x binding buffer. Subsequent reduction of the volume by spinfiltration to 200 µl was followed by incubation with GTP immobilized on thiopropyl sepharose beads. The amount of RNA used in the first round of selection has to be adjusted to the fact that a large fraction (approximately 40%) is lost during this process by adhering to the membrane of the spinfilter. Therefore, this procedure is only to be employed in the first three rounds of selection until the diversity of the pool and the RNA concentration are sufficiently reduced that the benefit of this method - reduction of undesired interaction of RNA molecules - is outweighed by the large loss of material.

In addition, the pool should be incubated with only a small amount of GTP immobilized on thiopropyl sepharose beads such that binding sequences are in

approximately 10-fold molar excess over the ligand, which forces pool molecules to compete for binding to the ligand.

The use of a pre-column consisting of thiopropyl sepharose beads without immobilized GTP and the incubation of both uncharged and charged beads with tRNA prior to incubation with RNA also has proven beneficial for improving the specificity of the selection.

Incubation of the RNA on the beads with immobilized GTP should occur for at least 24 h, followed by 1-column volume washes until RNA molecules are only released from the column at a background level. Elution with elution buffer should then be performed for at least 24 h.

In order to increase the stringency of the selection, the amount of washes can be gradually increased over the course of the selection

Pre-elution with GTP after moderate enrichment (approximately after round 4) has also proven useful for increasing selection stringency (Davis and Szostak, 2002).

Finally, PCR conditions can be adjusted to aid in increasing selection stringency by using a hot-start protocol, a high-fidelity polymerase and high annealing temperature.

When combining these conditions established in the experiments described above and, 9-4 aptamer RNA can be significantly and specifically enriched compared to pool RNA (see Fig. 23).

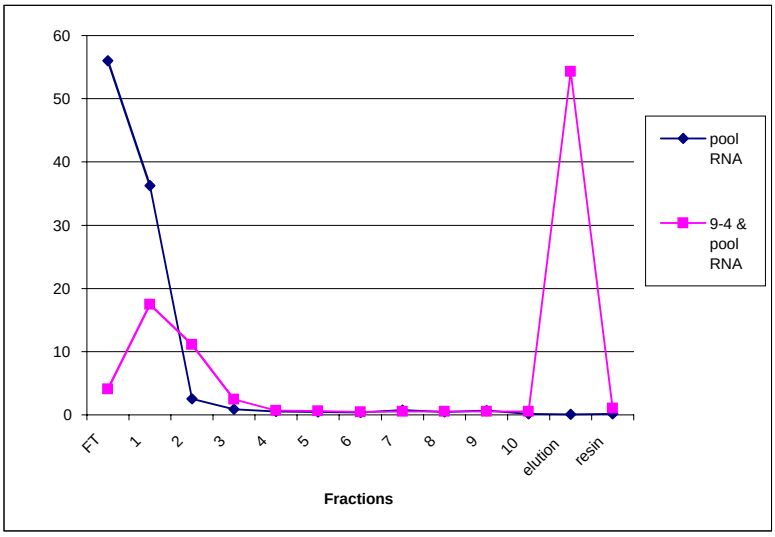

Fig. 23: Elution Profile of Aptamer RNA with Pool RNA (pink) vs. Pool RNA Alone (blue); FT: flow-through fraction; Fractions 1-10: 1-column volume wash fractions; elution: elution fraction; resin: RNA remaining bound to the column material.

5.2 Mechanisms of Aptamer Binding

5.2.1 Aptamer Stability

Results of our analysis indicate that aptamer size, information content and free energy of secondary structure formation are positively correlated and appear to be the key factors for aptamer quality, since all of these factors enhance aptamer stability. A major part is contributed by stem formation, but larger aptamers also form several loops whose stacking enhances their stability.

5.2.2 Correlation of Aptamer Structure and Quality

Free energy of secondary structure formation is correlated with K_d, thus contributing to aptamer stability. Experiments performed by Steffens and Digby (1999) and Schultes et al. (1999) suggest a correlation between aptamer fitness and free energy of secondary structure formation. The structure of a folded RNA molecule is low in energy because it has a large number of favorable interactions (Saven and Wolynes,

1997), and in at least one instance, improved stability of the secondary structure has been related to increased tertiary stability (Juneau and Cech, 1999).

RNA secondary structure is known to account for a large portion of the overall structural free energy (Schuster, 2002). Nucleic acids form secondary structures so readily because their nitrogen bases are hydrophobic, and the charged phosphate residues are hydrophilic. The bases therefore tend to turn inward, forming a base-paired duplex, which twists about a central axis to exclude water. There is a strong bias towards complementary Watson-Crick base pairs because the stacking is precise and planar and excludes water very well, generally exhibiting high negative free energy (Puglisi and Tinoco, 1989). Single-stranded regions that are constrained from forming base pairs also form planar stacks and these twist to exclude water from the spaces between them, but without an external constraint there is always a strong thermodynamic impetus for the chain to form a duplex.

Base-pairing interactions are also highly redundant within a given structure since the primary structure of nucleic acids is less well equipped to instruct unique folding pathways than this is the case for proteins. While RNA secondary structure is generally stable, its tertiary structure is however more labile than this is the case for proteins, therefore more prone to explore misfolded states (Thirumalai and Woodson, 1996; Wolynes et al., 1995) and is plastic in that sense. This inherent plasticity entails a fitness cost, because the more alternative shapes an RNA molecule can adopt, the less time it will spend in each shape, including advantageous ones (Ancel and Fontana, 2000), as has been observed in the group I intron (Pan et al., 1997; Pan and Sosnick, 1997). Low-information content aptamers are therefore likely to adopt more shapes, i.e. a lower fraction is folded in the active conformation.

On the other hand, conformations that constitute energy minima are less free to explore steric possibilities. This is reflected in some of the high-performance aptamers that show a high percentage of correctly folded RNA. Once a particular path of folding is followed, an energy barrier is formed, separating it from alternative pathways.

Transitions from the biologically active state of the molecule to a misfolded state will not generally be feasible unless energy is supplied. Folding is therefore directional.

Even though more complex structures are expected to require a more complicated folding pathway and therefore should have more opportunities to misfold than simpler structures, the correlation of K_d and fraction of RNA that is correctly folded shows that only complex structures with a high degree of correct folding had survived the selective pressure of *in vitro* selection.

5.2.3 Aptamer Binding Mechanisms

High-resolution structures of RNA aptamers have shown that most aptamers have not assumed their correct tertiary structure prior to binding to their target, but rather (re)structure during this event in an induced-fit manner, enveloping their target upon binding and unfolding upon dissociation (Dieckmann et al., 1997). Kinetic data that could for instance be obtained by surface plasmon resonance using Biacore technology are therefore expected to show comparatively slow association and dissociation rates as have been observed for aptamer 9-4 (Davis and Szostak, 2002). Even though it is difficult to obtain information about the free energy of tertiary structure formation for correctly folded RNA aptamers, the observed relationship between K_d and free energy of secondary structure formation and fraction correctly folded (Fig. 19) underline the importance of aptamer stability for achieving high-affinity binding. These results also suggest that a stable secondary structure might contribute significantly to nucleating folding into the correct tertiary aptamer structure.

5.2.4 Similar Interaction Patterns of Different Aptamers

The greatest similarities overall could be observed with the lowest-affinity aptamers, Class III and Class IV, which might in part be due to the fact that only competitor concentrations of up to 0.5 mM could be tested because of the limited solubility of GTP and its analogs. While this permits examination of the effect of the

competitor over a range of up to five orders of magnitude for higher-affinity aptamers, for weaker binders such as Class III merely data for three orders of magnitude can be made available experimentally. The experimental limitations of the method employed here for the examination of aptamer binding mechanisms therefore mandate caution when interpreting data for weaker-binding aptamers.

Aptamers Class II and 10-10 show some similarities in their interaction pattern. Crosslinking experiments might reveal the similarity of their interaction with GTP in more detail, possibly even with a region that is similar in both aptamers (see Fig. 24, marked in yellow).

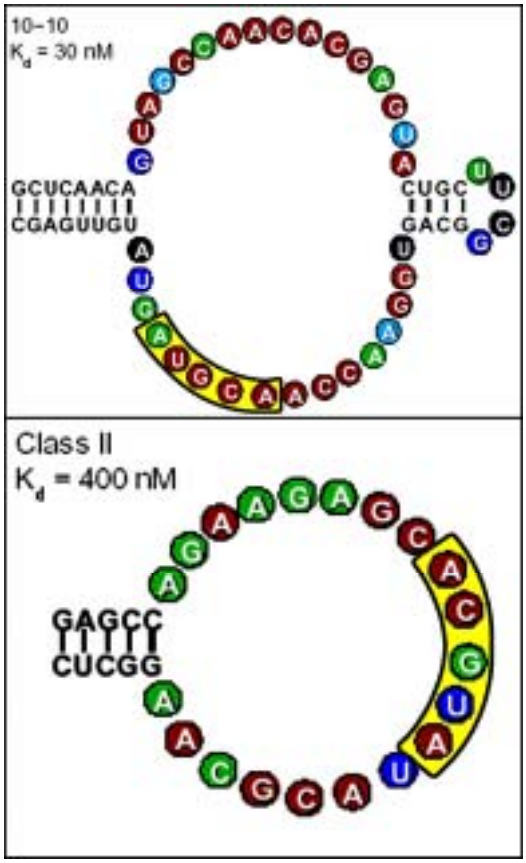

Fig. 24: Proposed Secondary Structures of Aptamers 10-10 and Class II and Possible GTP Binding Region. Common region is marked in yellow (aptamer structures courtesy of James Carothers).

5.3 Evolutionary Potential of Aptamers

Interestingly, aptamers 10-6 and 10-24 both exhibit equivalent K_d values (170 nM and 195 nM respectively), but aptamer 10-6 shows a comparatively high information content of 71 bits, whereas 10-24 only exhibits 50 bits of information content. Aptamer 10-24 differs significantly from all other aptamers in its low specificity and high variability, i.e. poorly constrained loop sequence. Together with the tendency to weak interactions and thus low specificity of this aptamer, these results suggest that 10-24 still retains ample possibility for evolving to a more highly functional aptamer with improved information content and affinity to GTP, while 10-6 seems to have exhausted most of its evolutionary potential.

In addition, aptamer 10-24 could be used as a starting point when attempting to evolve a GTP aptamer into an ATP aptamer. This aptamer with low specificity could be evolved into binding to ATP by conducting several doped re-selections with different GTP analogs that gradually bridge all mutational differences from GTP to ATP, one at a time.

5.4 Future Studies

Due to the diverse aptamer binding mechanisms, it is challenging to disentangle the entropic and enthalpic contributions in aptamer binding. Thermodynamic data from calorimetry for instance would in the first case reflect only data from the interaction of aptamer RNA and GTP, whereas in the second case a component for the correct folding of the aptamer has to be added. The magnitude of both contributions is difficult to determine, and therefore a comparison of the thermodynamic patterns for aptamers with different folding behaviors is problematic.

The lack of a uniform binding pattern suggests that each aptamer forms a structure that recognizes GTP in a unique manner through a combination of different individual interactions. Therefore multiple different solutions exist for binding to GTP, as opposed to binding to ATP, for which several independent *in vitro* selections in different

laboratories have yielded one dominant solution (Sassanfar and Szostak, 1993; Burgstaller and Famulok, 1994; Burke and Gold, 1997). It would therefore be interesting to explore the interaction of the dominant ATP aptamer and some of the few additional ATP aptamers in the same way as this has been done here for GTP aptamers.

In order to obtain a comprehensive picture of aptamer interaction, in addition to kinetic and thermodynamic factors contributing to the quality of aptamer binding, high-resolution structures are eventually needed to elucidate the tertiary structure of aptamers, both separately and when interacting with their target molecules.