2220–2227 *Nucleic Acids Research, 1996, Vol. 24, No. 12* *© 1996 Oxford University Press*

Characterisation of antibody-binding RNAs selected from structurally constrained libraries

Jörg Hamm

Istituto di Ricerche di Biologia Molecolare, Via Pontina km 30,6, 00040 Pomezia, Roma, Italy

Received March 29, 1996; Revised and Accepted April 30, 1996

ABSTRACT

Constrained RNA libraries of limited sequence complexity were constructed and used to select RNA molecules binding to the antigen binding site of an anti-ferritin antibody. The sequences required as primer-binding sites for the selection cycle were designed to form a predictable secondary structure, which greatly facilitated the characterisation of the secondary structures of the selected RNAs. RNA–antibody interactions were studied by real-time interaction analysis to study the dynamic aspects of binding and by circular dichroism spectroscopy to search for conformational changes upon binding. The selected RNAs were analysed with a binding site sequestering assay and were shown to compete with ferritin for binding to the antigen-binding site. The experiments described here indicate that the introduction of strong structural constraints does not have to interfere with the ability to select tightly and specifically binding RNA-molecules.

INTRODUCTION

The number of building block types that can be used to generate molecules with a desired feature by *in vitro* selection from combinatorial libraries is growing rapidly. The libraries can be of very different structural classes: peptides displayed on the surface of phages, synthetic peptide libraries, antibody libraries, RNA or single-stranded DNA containing natural or modified nucleotides and chemical libraries. From any of these libraries molecules binding specifically and with high affinity to a wide variety of ligands have been isolated (for reviews see 1–3).

Although nucleic acid libraries were developed only a few years ago (4,5), a large number of RNA and DNA molecules has been selected and characterised. These libraries have been used to define binding sites of RNA-binding proteins (5–8), to generate RNAs that inhibit polymerases (9) or proteases (10), to change properties of existing ribozymes (11,12) or to create novel ribozymes (13,14). It has been shown that specifically selected RNA, also called aptamer, is capable of discriminating between molecules differing by as little as a single methyl group (15).

The structural motif responsible for the activity of an aptamer is often quite small, so it could appear that the structural elements employed by aptamers to bind to ligands which normally do not bind to RNA are similar to the elements found in naturally occurring RNA–protein interactions. These RNA elements include stem–loop structures with variable stem but fixed loop sequence, symmetric and asymmetric internal loops, pseudo-knot and G-quartet structures (16). So it might be sufficient to limit the randomised library sequences to a specific RNA motif in an otherwise constant background. This kind of strategy is used in antibodies to generate diversity of binding specificity involving only a small portion of the molecule. The variable part of the antibody is kept in loop structures that are structurally predisposed within the context of the whole antibody molecule to form binding sites.

To test the hypothesis that this concept could also be used for RNA libraries, two structurally constrained RNA libraries of limited sequence complexity, representing two different types of constraint, were designed. Both libraries were used to isolate aptamers binding to the anti-ferritin antibody H107 (35). The antibody was selected because peptides have been isolated with H107 from phage display libraries (17), allowing comparison of aptamer and peptide structures binding to the same ligand.

MATERIALS AND METHODS

In vitro selection

Template was generated by amplifying 200 pmol of single-stranded DNA-library (sequence corresponding to the RNA sequences shown in Fig. 1) for 5 cycles with 1 nmol of each primer (GQCL: primer T7P1, 5′-GGGAAGCTTAATACGACTCACTATAGGGTTGCACACCGTCTCG and primer 3′RSP1, CCCAAGCTTCCCGGGTTGCACACCGTCTCG; SSL$_{30}$: primer T7P2, GGGAAGCTTAATACGACTCACTATAGGGATGCTTCGGCATCCC and primer 3′RSP2, CCCAAGCTTACGTACCGCCGAAGCGGTAC). For all PCR reactions buffer J (50 mM NaCl; 3.75 mM MgCl$_2$; 10 mM Tris–HCl pH 8.0; 0.01% gelatine; 0.01% Triton X-100) was used. DNA was purified from 2.5% agarose gels and transcribed in a volume of 240 µl for 16 h (including 30 µCi [α-^{32}P]GTP) as described in reference 4. RNA was purified by phenol/chloroform extraction and ethanol precipitation. RNA was re-suspended in H$_2$O and DNA digested in a volume of 300 µl for 1 h with 50 U RQ DNase (Stratagene) in the recommended buffer. RNA was purified by phenol/chloroform extraction, unincorporated nucleotides removed (NAP$_5$; Pharmacia), RNA precipitated with ethanol, suspended in 52 µl NaK$_{150}$, heated to 68°C for 10 min.

RNA was always in molar excess over the antibody (2–12 nmol RNA/less than 60 pmol H107). M280 beads (coated with sheep anti-mouse antibodies; Dynal) were incubated overnight with antibody (50 µl beads + 25 µg antibody) at 4°C in PBS (+0.1% BSA), washed twice with PBS (+0.1% BSA), twice with NaK$_{150}$ (50 mM Tris–HCl pH 7.5/100 mM KCl/50 mM NaCl/5 mM

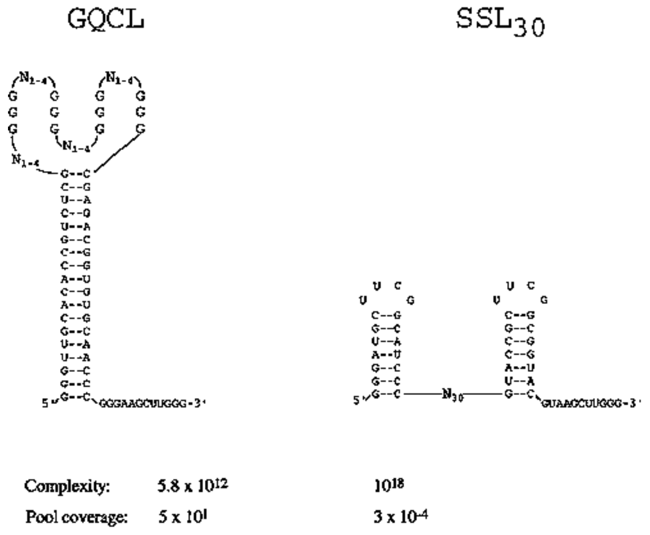

Figure 1. Library design. The potential structural constraints determined mainly by the invariant sequences selected are shown. For GQCL the 5′ and 3′ nucleotides corresponding to the primer binding sites are complementary, permitting the formation of a continuous 18 bp stem. The four groups of three G-residues are interrupted by sequences variable both in length (one to four bases) as well as in sequence. The invariant sequences of SSL_{30} have the potential to form 7 bp stems closed by a stable tetra loop sequence. The 3′-terminal 11 bases were added to increase the efficiency of the reverse transcriptase–PCR cycle. The sequence complexity and the pool coverage for the first selection cycle are indicated.

$MgCl_2$) for 30 min at room temperature. RNA (50 µl) was incubated for 30 min at room temperature with M280/α-OSM beads in a rotating wheel, the supernatant removed and beads washed once with NaK_{150}. Supernatants were combined and incubated for 30 min in a volume of 100 µl with H107 beads (50 µl beads). Beads were washed twice with 100 µl NaK_{150} or with NaK_{150}/0.1% NP-40 for 5 min at room temperature and RNA eluted with 400 µl proteinase K buffer (50 mM Tris–HCl pH 7.5; 5 mM EDTA; 1.5% SDS; 300 mM NaCl; 1.5 mg/ml proteinase K, 30 min, 37°C). RNA was extracted and re-suspended in 20 µl H_2O, 10 µl used for the reverse transcriptase reaction [200 pmol 3′RSP1 or 3′RSP2, 1 µl AMV reverse transcriptase (Stratagene); 20 min, 56°C]. RNA and DNA were extracted and re-suspended in 20 µl H_2O and 10 µl used for the PCR reaction (buffer J, 250 pmol of each primer, 20 cycles, annealing temperature 60°C). Templates were purified, suspended in 80 µl H_2O and 40 µl used in a 80 µl transcription reaction (10 µCi [α-^{32}P]GTP; 250 U T7 RNA polymerase).

Real time interaction analysis

Dextran coated sample cells were used to couple H107 (50µg/ml) covalently to the cuvette in 10 mM NaAc pH 5.0. The base line was obtained with 195 µl NaK_{150}, RNA to be analysed added in a volume of 5 µl NaK_{150} and the association monitored for 5–10 min. The sample cell was washed rapidly twice with 200 µl NaK_{150} and dissociation monitored for 5–10 min. The cell was regenerated by washing for 2 min with 200 µl 1.5 M NaCl and three times with 200 µl NaK_{150}. The same H107 coated sample cell could be used for 20–30 measurements, stored in between measurements at 4°C (in PBS + 0.1% Tween-20). Curves were fitted using the fast fit software package and association constants obtained by plotting k_{on} over aptamer concentration (error typically 1–3%).

Circular dichroism (CD)-measurements

Spectra were obtained with a JASCO J710 Spectropolarmeter at 4°C using a cuvette of 1 cm pathlength (V = 2.5 ml). Spectra were recorded from 320–220 nm at 5 nm/min and five measurements sampled. Aptamer concentrations were 2.4 µM, H107 concentration 132 µg/ml (=825 nM bivalent antibody or 1.65 µM binding sites), all samples in 100 mM KCl/10 mM Na-phosphate pH 7.5).

RESULTS

Library design

Due to the iterative nature of the *in vitro* selection process the 5′- and 3′-terminal sequences of all library molecules remain invariant because they serve as the primer binding sites both for the reverse transcriptase and for the polymerase chain reaction (PCR) step. If fixed sequences could be used to create a structural constraint, it might be easier to analyse the structure of RNA selected from libraries with short stretches of variable sequences because of the formation of a predictable secondary structure. To test the feasibility of this approach, two different libraries, GQCL and SSL_{30}, were constructed. The library GQCL should be compatible with the formation of G-quartet-like structures and contained primer binding sites of complementary sequence permitting the formation of a continuous 18 bp stem structure. The stem sequences were interrupted by four groups of three G-residues separated by four blocks of variable sequence (1–4 random nucleotides for each block). These blocks should provide sufficient structural variability for the formation of loops connecting a G-Quartet-like scaffold (Fig. 1, GQCL). The library SSL_{30} was not intended to favour a particular type of structure in the variable part of the library. Rather, each primer sequence was designed to form a stable structure independent from the rest of the molecule (Fig. 1, SSL_{30}). A single-stranded tail of 11 nt was added to both libraries to increase the sensitivity of the reverse transcriptase–PCR reaction. Although this introduced the risk that the 3′-tail could be used for base pairing interactions with bases from the variable sequence of the library, it permitted the use of stringent selection conditions. The stringency of the selection is likely to influence the number of cycles required to select an aptamer, and also the affinity for the selector molecule.

Course of selection

Since RNA binding to the antigen binding site of H107 was desired, aptamers binding to the constant region of the antibody were depleted in every cycle by using an unrelated antibody from the same species (anti-oncostatin M/α-OSM). Although the amount of RNA eluted from the initial cycles of selection increased, after four to five cycles the amount of RNA remaining decreased with each additional cycle (Table 1). This observation was reproducible, because any given selection cycle could be repeated and the previously observed amount of RNA remained after the washing steps. This is in contrast to what has been reported previously in selection experiments where the amount of selected material usually increased with each additional cycle until the majority of the selector molecules had been saturated. The decrease in the amount of RNA after five cycles of selection was accompanied by the appearance of shorter PCR products. It is possible, that once selection for binding affinity has reduced the RNA pool to a small number of individual sequences, the selective pressure of the cycle might be limited to the deletion of sequences outside the primer binding sites.

Table 1. Course of selection

Cycle	SSL$_{30}$ [pmol RNA]	GQCL [pmol RNA]
1st (wash without NP-40)	1.66	1.19
2nd (wash with NP-40)	0.15	0.15
3rd (wash with NP-40)	0.26	0.35
4th (wash with NP-40)	1.69	*1.20*
5th (wash with NP-40)	*1.69*	4.21
6th (wash with NP-40)	2.68	1.30
7th (wash with NP-40)	2.66	
8th (wash with NP-40)	1.64	

The absolute amount of RNA (values in pmol) eluted from the H107 selection cycles is indicated. Individual RNAs binding to H107 were isolated from cDNA of cycles 4 (GQCL) or 5 (SSL).

Structure of GQCL-4.7

Aptamers derived from the 4th cycle of the GQCL selection were analysed directly for binding activity by immunoprecipitation or real time interaction analysis. Specific binding to H107 was observed for 7/32 RNAs. No binding of the unselected libraries of GQCL and SSL$_{30}$ to H107 could be detected, neither by immunoprecipitation, nor by real time interaction analysis. This indicates that the scaffold structures selected for the libraries are not responsible for determining binding specificity. Sequencing showed that 6/7 were identical (clone GQCL-4.7), while the seventh clone had a single base substitution. In all clones selected the 4th group of G-residues (G^4) which was invariant in the starting library, had been mutated from GGG to AAG. Instead, new G-residues had been created from the variable regions N^3 and N^4 (Fig. 2). It is possible that the reason for the lack of recovery of the initial constrained structure is that a G-quartet like structure and a stem structure might have to be separated by loop sequences on both sides of the G-quartet. In the GQCL-design used for this selection a loop was provided only at the 5′-end of the G-quartet scaffold. The sequence outside of the primer binding sites is extremely purine-rich (20/22), and the selection of additional G-residues from the variable blocks made an unambiguous allocation of a G-quartet scaffold in GQCL-4.7 impossible (Fig. 2; 1+2). Immunoprecipitation of GQCL-4.7 with various antibodies demonstrated that it did bind specifically to H107 (Fig. 3a). Therefore, to characterise the GQCL-4.7 structure, various types of mutagenesis were employed. The 3′-tail was removed by cutting the template with a restriction enzyme (*Sma*I) and RNA transcribed from this template was analysed for binding to H107. No difference in binding was detected. Next, GQCL-4.7 derivatives with the ability to form stems of 16, 14, 12, 8 or 4 bp were constructed and tested for binding to H107 and α-OSM antibodies. Shortening the stem to 8 bp or less resulted in complete loss of binding to H107 (Fig. 3b and data not shown). This indicates that a stem is formed by the primer binding sites as anticipated with the constraint. Furthermore, the formation of a stem of a length of >8 bp is required for the formation of the active structure.

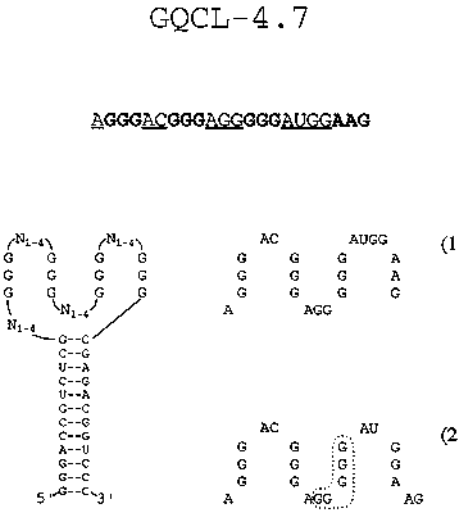

Figure 2. Structure of clone GQCL-4.7. The sequence corresponding to the initially variable part of the library is shown (5′ to 3′, from left to right). Bold letters mark the positions which contained the four groups of three G-residues in the library, underlined positions are derived from the variable positions. Two ways to fit the GQCL-4.7 sequence into a G-quartet-like structure are proposed (labelled 1+2), the dotted line around the central five G-residues indicating that there are three groups of three G-residues which could potentially be involved.

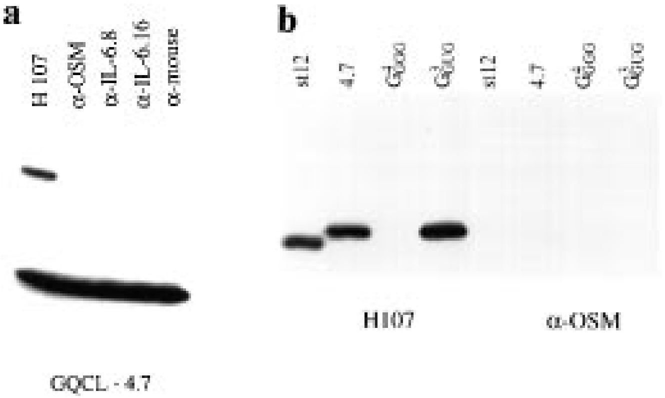

Figure 3. Binding specificity of GQCL-4.7. (**a**) Radioactively labelled RNA was immunoprecipitated with various antibodies (H107, selection antibody; α-OSM, depletion antibody; α-IL6.8/16, anti IL6 mouse monoclonal antibodies; α-mouse, magnetic beads carrying only the secondary antibody sheep anti-mouse) and RNA analysed on denaturing acrylamide gels. (**b**) Radioactively labelled RNA was immunoprecipitated with H107 or α-OSM antibodies (st12, 6 bp stem deletion as shown in Fig. 2; 4.7, original GQCL-4.7 clone; G$^4_{GGG}$, A to G substitution in the 4th group of G-residues in Fig. 2; G$^3_{GTG}$, G to T substitution in the 3rd group of G-residues in Fig. 2) and RNA analysed on denaturing acrylamide gels.

To identify residues important for interaction with H107, a mutagenised GQCL-4.7 DNA pool was synthesised. At each of the 22 positions outside of the primer binding sites (including also the four groups of G-residues) 91% of the original GQCL-4.7 nucleotide was incorporated and 3% of each of the remaining 3 nt. Several observations were made: no substitution was found for any of the A, C or U residues, the 3′-GAAG motif was absolutely conserved. No sequence carrying three G-residues in a fourth group of G (G^4) was found. The most frequent substitutions fell

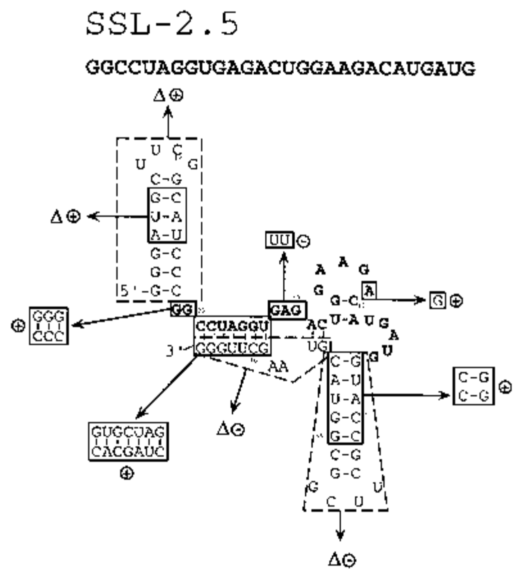

Figure 4. Structure of clone SSL$_{30}$-2.5. The sequence derived from the variable part of the library and the predicted secondary structure of the clone SSL$_{30}$-2.5 is shown (5′ to 3′, from left to right). Bold letters indicate positions variable in the original library. The result of the mutational analysis is summarised (Δ, deletion; +, binding; –, no binding; boxed sequences connected by arrows with the 2.5 sequence illustrate substitutions). The A to G change in the sequence GAAGA represents the only difference in sequence of the clone SSL-2.6, isolated together with SSL-2.5 in the initial screening. Orientation and spacing of secondary structure elements were chosen to facilitate interpretation of mutations introduced, they do not reflect a modelling attempt.

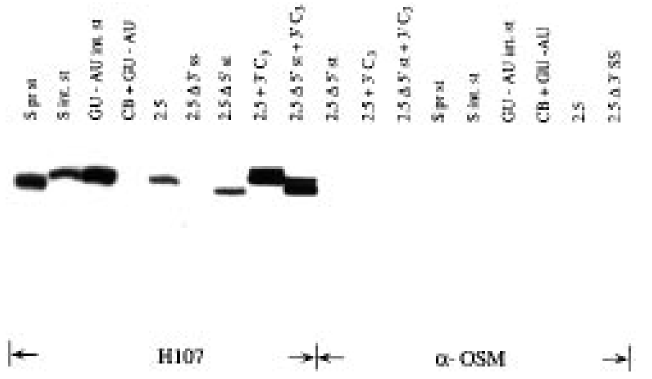

Figure 5. Binding specificity of clone SSL$_{30}$-2.5. Radioactively labelled RNA was immunoprecipitated with H107 or α-OSM antibodies (*S pr st*, substitution of 5′ and 3′ primer stems; *S int st*, substitution of internal stem; *GU-AU int st*, single base substitution in internal stem; *CB+GU-AU*, closed bubble + single base substitution; *2.5*, original SSL$_{30}$ clone; *2.5Δ3′ss*, deletion of 11 bases from the 3′-end; *2.5Δ5′st*, deletion of 5′ stem–loop; *+3′C$_3$*, addition of three C-residues at the 3′-end) and RNA analysed on denaturing acrylamide gels.

into G^2 and G^3, mainly G to U substitutions (e.g. G^3_{GUG}, Fig. 3b). The failure to isolate aptamers having three G-residues in G^4 could have been the consequence of the very high stability of a three-layered G-quartet structure. This has been observed for the riboflavin aptamer (18). To test this hypothesis, the clone G^4_{GGG} was constructed (3′-sequence: GGAAG to GGGAG) and immunoprecipitated. No binding to H107 was detectable (Fig. 3b), indicating that the A-residue might be part of the binding site, rather than having been selected for easier replication in the reverse transcriptase–PCR-cycle. Furthermore, it demonstrates that the structure of GQCL-4.7 is not a perfect three-layer G-quartet structure. Although this provides only evidence against the formation of a G-quartet structure, the only known RNA structure compatible with the selected sequence would be a G-quartet like structure with mixed G/A-quartets (19). Mixed A/G-quartets might be of similar overall structure but of reduced stability because an A-residue would be able to form only three instead of the four hydrogen bonds formed by a guanosine with neighbouring residues. However, this could be an important feature, because it might allow a conformational freedom impossible for a perfect G-Quartet structure (see below).

Structure of SSL$_{30}$-2.5

SSL-clones were analysed both by immunoprecipitation with H107 and α-OSM antibody, and with the real time interaction analyser. Three of the 12 clones bound tightly to H107 and were sequenced. Two of the clones were identical (SSL 2.5), the third clone differing only by a single nucleotide (clone SSL 2.6; Fig. 4). The sequence of the aptamers isolated from the SSL-library were complementary to the single-stranded tail added to the 3′-end of the libraries. The rest of the aptamer sequence, excluding the stable tetra loop stems formed by the primer binding sites, had no obvious secondary structure elements (Fig. 4). Various mutations were introduced into the SSL-2.5 sequence. Reducing the length of the 5′- and 3′-primer stems from 7 to 4 bp did not affect binding to H107 (Fig. 5, S pr st). While complete deletion of the 5′ stem–loop did not interfere with binding (Fig. 5, Δ5′-st), deletion of the 3′ stem–loop abolished interaction with H107 (data not shown).

The formation of the internal stem is indicated by several observations. Removal of the 3′-tail of SSL-2.5 resulted in a loss of binding (Fig. 5, 2.5Δ3′ss). Introducing a point mutation into the internal stem sequence (Fig. 6, GU to AU int st) or substituting the entire 7 bp stem by a 7 bp stem of different sequence (Fig. 5, S int st) did not affect binding. Shortening the stem to 6 bp abolished binding to H107 completely (data not shown). The two G-residues between the 5′ primer stem and the internal stem could have been involved in interactions with other regions of the aptamer. Two RNAs were generated having three G-residues at this position plus three additional C-residues at the 3′-end of the RNA, which should allow the formation of a 10 bp internal stem. This mutation was introduced into the aptamer either alone or in combination with the deletion of the 5′ primer stem. In both cases the extension of the internal stem improved binding to H107 (Fig. 5, 2.5+3′C$_3$; 2.5Δ5+3′C$_3$). Closing of the potential purine bubble on the internal stem by a substitution (GAG to UU) either in SSL-2.5 or in 2.5 (GU to AU int. stem) abolished binding completely (Fig. 5, CB+ GU to AU). This indicated an internal loop or the nucleotides by themselves were important for the formation of the active structure. For SSL-2.5 most of the sequence invariable in the library could be removed without losing binding activity. Furthermore, the 5′-variable bases 1–9 could be substituted by a different sequence without loosing binding activity, as long as it was forming a stem with the 3′-end of the aptamer. Thus only very few bases are actually forming the binding site, providing evidence in support of the RNA–antibody scenario (see Introduction and 20).

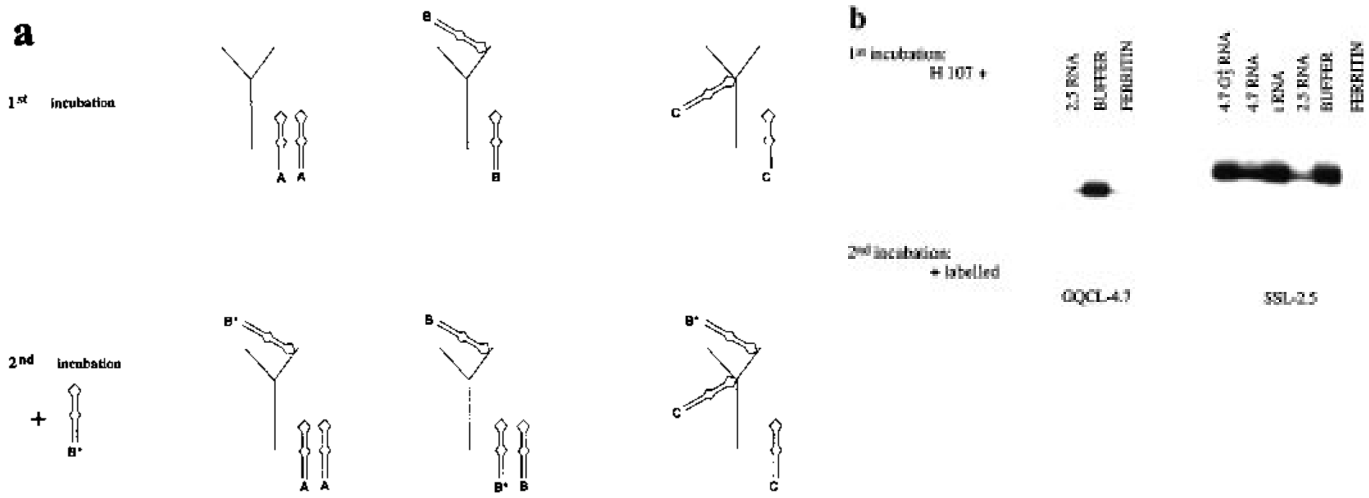

Figure 6. RNA and ferritin compete for binding to H107. (**a**) Scheme of the binding site sequestering assay: excess of unlabelled competitor (RNA or protein) is pre-incubated with a limiting amount of H107 antibody (1st incubation). A small amount of radioactively labelled RNA (marked by asterisks) is added and the immunoprecipitation is continued (2nd incubation). The radioactively labelled RNA added in the second incubation will be immunoprecipitated only if the competitor and the labelled RNA do not have overlapping binding sites. (**b**) Binding site sequestering assay: H107 (0.7 µg ≈ 7 pmol) bound to magnetic beads was incubated for 30 min with unlabelled RNA (400 pmol) or with H ferritin (28 µg; ref. 34), radioactively labelled RNA (10 pmol) was added, the immunoprecipitation continued for 15 min and RNA analysed on denaturing acrylamide gels.

The structure of SSL-2.5 bears some similarity to the structure of the hammerhead ribozyme (21,22). Three stem and stem–loop structures are present, the sequence CUGAUG connecting stems I + II of the ribozyme (the uridine turn, found also in tRNA) is similar to the sequence UGAUG found in SSL-2.5 (residues 42–46) connecting here the 3′ primer stem with the rest of the sequence derived from the variable part of the library. In the ribozyme the UG of this uridine turn is base paired with two A-residues present at the junction between stems II + III, also in SSL-2.5 two A-residues are found 3′ of the 7 bp internal stem.

RNAs and ferritin compete for binding to H107

The selection strategy used should have resulted in the isolation of aptamers binding to the antigen binding site of H107, because the library was depleted specifically in each selection cycle. To test whether this indeed was the case, a binding site sequestering assay was performed (Fig. 6a). A limiting amount of H107 (7 pmol) was coupled to magnetic beads and incubated for 30 min with 400 pmol of unlabelled RNA (specific and non-specific RNA) or with ferritin (28 µg = 1200 pmol monomeric ferritin = 50 pmol ferritin 24mer, the solution form of ferritin). Subsequently, radioactively labelled aptamer (10 pmol) was added and the immunoprecipitation was continued for 15 min.

Pre-incubation of H107 with tRNA or GQCL-4.7G^4$_{G3}$ did not interfere with binding of SSL-2.5 to H107. Pre-incubation with ferritin or cold SSL-2.5 RNA abolished binding almost completely. Pre-incubation of cold GQCL-4.7 reduced the amount of SSL-2.5 immunoprecipitated significantly but not completely, reflecting the lower affinity of GQCL-4.7 (Fig. 6b, SSL-2.5). This interpretation is supported by the observation that in the reverse experiment, pre-incubation of cold SSL-2.5 RNA excluded the immunoprecipitation of GQCL-4.7 completely (Fig. 6b, GQCL-4.7). The results of the competition experiments indicated that both aptamers bind to the antigen binding site of H107 (or to two sites very close in space), and confirmed that also binding of aptamer and ferritin is mutually exclusive.

Kinetic analysis

It was interesting to analyse the dynamic aspects of the interaction of the aptamers with H107 precisely, to gain information about on- and off-rates of aptamers. The antibody was linked covalently to a dextran-coated sample cell of a Affinity Sensors IAsys real time interaction analysis system. Unlabelled RNA was added at various concentrations and binding as well as dissociation were monitored in real time. Figure 7 shows the superimposition of binding curves obtained with various concentrations of aptamer. These measurements confirmed the observed differences in the relative affinity of GQCL-4.7 and SSL-2.5. The calculated dissociation constants were $K_D = 6.4 \times 10^{-8}$ M for SSL-2.5 and $K_D = 1.6 \times 10^{-6}$ M for GQCL-4.7. The higher affinity of SSL-2.5 compared to GQCL-4.7 seems to be due to a ~15-fold higher on-rate and a 2-fold lower off-rate of SSL-2.5 (Fig. 7). For comparison, the equilibrium signal was plotted against the concentration of aptamer to obtain a conventional binding curve. Defining K_D as the concentration at which the half-maximum response was obtained, produced numbers similar to those above [K_D-values obtained in this way were usually lower (≤ 20%) than those obtained by integrating the binding curves with the fast fit program, data not shown].

Indications for conformational changes upon binding

CD spectroscopy can be used to obtain qualitative information about a conformational change of an aptamer when binding to a ligand (23). For this reason, CD-spectra of the H107 aptamer with the highest affinity for H107 (SSL-2.5) in the presence or absence of H107 were recorded. As a control, the mutant SSL-2.5CB, which did not show detectable binding to H107, was used for comparison. The spectra of SSL-2.5 and SSL-2.5CB RNA were almost identical in the absence of H107 (Fig. 8a). Furthermore,

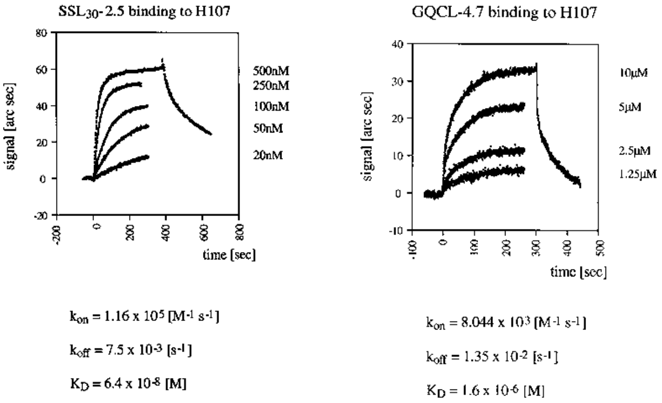

Figure 7. Real time interaction analysis. Binding of SSL_{30}-2.5 and of GQCL-4.7 to H107 was analysed with a Affinity Sensors Iasys real time interaction analysis system. H107 was covalently linked to the dextran matrix of the sample cell and association as well as dissociation of unlabelled RNA monitored (for clarity only a single dissociation curve is shown in the superimposition of the measurements). The kinetic parameters are the result of fitting the binding curves using the fast fit software package.

upon addition of H107 the spectrum of SSL-2.5CB did not change in the region around 270 nm (Fig. 8c). In contrast, in the presence of H107 the spectrum of the aptamer SSL-2.5 indicated a clear reduction of the signal around 270 nm (Fig. 8b and d). Although this is a mixed spectrum of the RNA and the antibody, at the concentrations used for these measurements the antibody is not likely to have contributed to the changes of the spectrum around 270 nm. To visualise conformational changes of aromatic side chains in proteins much higher concentrations would have been required. The observed differences of the CD-spectra of the RNA upon binding to H107 are similar to those reported for the binding sites of Rev (24), Tat (25) and arginine binding DNA aptamers (19). Although the conformational change observed upon binding does not prove that it is required for binding, it indicates that the structure of the SSL-2.5 RNA might be different in the bound and the unbound state. Consequently, detailed studies of the RNA structure in the absence of the ligand are unlikely to be sufficient to identify the reactive groups involved in ligand–RNA interaction. A dramatic change of the backbone conformation similar to that described for the Tat-binding site could not be deduced from the RNA-structure alone.

DISCUSSION

Library design

Commonly used RNA-libraries are of high sequence complexity and no attempts are made to limit diversity by introducing structural constraints. This is motivated by the reasonable assumption that the higher the sequence complexity of the library is, the higher will be the structural complexity of the library and the higher the resulting chance of obtaining tightly binding RNA. However, a problem could arise from the fact that the number of individual RNA-sequences that can be handled in a single experiment is much lower than the theoretical complexity of the library. The result is partial pool coverage. The selection of aptamers from libraries of high sequence complexity requires many rounds of selection (10–12 cycles) to enrich individual sequences sufficiently to allow their identification. This incorporates the risk of selecting against RNAs with stable secondary structures because their lower replication efficiencies might favour enrichment of less stable structures. High enrichment factors for individual sequences were observed with the motif-libraries used for this study already after four to five cycles of selection. The characterisation of the secondary structure of the H107 aptamers was greatly facilitated by the predictable structure formed by the primer binding sites. For *in vitro* selection studies aiming at a general understanding of RNA structure, the relatively easy identification of the approximate overall secondary structure of the aptamer selected from constrained libraries might not seem to be of great importance. However, if aptamers are selected with the intention of generating molecules interfering with the function of proteins, either *in vitro* or *in vivo*, a high resolution structure might not be required. In addition, the stability of the aptamer will be an important factor for applications. The secondary structure constraints used here were partly inspired by those of UsnRNAs. For example, the binding site for the common UsnRNP-proteins (Sm-binding site) is a single-stranded region flanked by stable stem–loop structures (26). So the aptamers selected from these constrained libraries might be used directly for experiments, without the need to engineer a stable structure. Pilot experiments performed with the H107 aptamers showed that both aptamers were stable for ≥ 24 h after injection into Xenopus oocytes (unpublished data).

One of the constrained libraries used for this study, GQCL, was designed to favour a three-layer G-quartet structure. GQCL-4.7, the aptamer selected from this library carried a 2 nt mutation compared to the original constraint, indicating that the chosen backbone was not compatible with the formation of the binding site for the H107 antibody. However, a constraint is likely to limit the structural complexity of a library by limiting its flexibility. The type and the right amount of constraint, sufficient for displaying the binding regions without creating a structure that is too rigid to allow binding to occur, will have to be determined experimentally by designing a constraint and analysing the outcome of the selection. Constrained libraries have been used already to select RNA (7,27). In the former selection a library with a randomised internal loop was used to obtain optimised Rev-binding elements, but it was already known that the binding site of Rev (an RNA binding protein) was an internal loop. The selection of RNA binding an antibody raised against a peptide derived from the g10 fusion protein with a library based on the stem–loop II of U1snRNA (27) is more relevant to the work presented here. However, there, the RNA structure was not characterised in detail, and the affinity of the RNA for the antibody was not determined.

Kinetic parameters

The dynamic aspects of aptamer binding were analysed with a real time interaction analyser. This method permits the determination of kinetic constants under quasi-solution conditions in real time (28). So far the affinities of aptamers have been determined only by indirect methods and no on- or off-rates were reported. The relative values for the two aptamers characterised are in good agreement with the relative affinities observed in competition assays and immunoprecipitation experiments. The affinity of GQCL-4.7 is comparable to the average affinity of aptamers isolated from unconstrained libraries (1–500 µM), the aptamer SSL-2.5 is part of the group of high affinity aptamers. This indicates that the introduction of a constraint does not necessarily

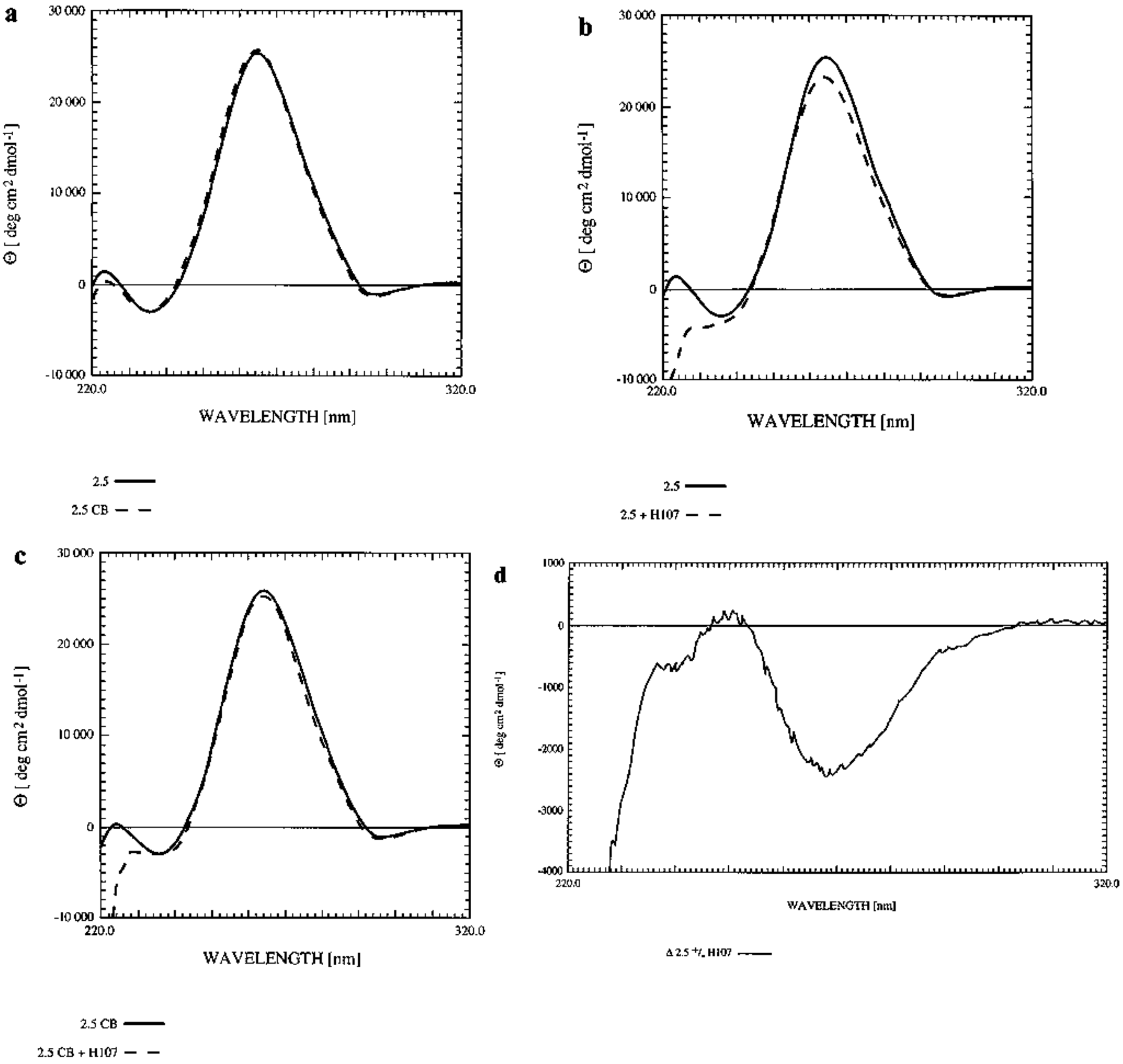

Figure 8. CD spectra of RNA or RNA–antibody complexes. Superimpositions of RNA or RNA–antibody spectra are shown. Theta values are given as mean molar ellipticity. RNA concentrations were 2.4 µM, H107 concentrations were 132 µg/ml (≈825 nM). (**a**) RNA spectra of SSL_{30}-2.5 and SSL_{30}-2.5CB. (**b**) SSL_{30}-2.5 in the presence or absence of H107. (**c**) SSL_{30}-2.5CB in the presence or absence of H107. (**d**) Difference spectrum of SSL_{30}-2.5 with/without H107.

interfere with the ability to select tightly and specifically binding aptamers.

Search for conformational changes

CD-spectroscopy has been used to predict conformational changes in the RNA-elements binding to Rev, Tat and arginine binding DNA aptamers, and at least in the case of Tat were confirmed later by the determination of the Tat-binding RNA structure in the presence and absence of ligand by NMR (29). Since the determination of RNA structure by NMR is far from trivial, CD-measurements might be a fast way to look qualitatively for conformational changes occurring in the RNA upon binding to a ligand. Although it is not clear whether it will be possible to attribute certain shapes of spectra to specific secondary structure elements of RNA, as is possible for proteins, it seems likely that the signal around 270 nm corresponds to the stacking interactions of the bases (23). If bases would move out of helical structures to interact with the ligand (as in the case of Tat) a reduction of the signal might be expected (as observed for Tat and Rev). Of course the observation of a conformational change of RNA or DNA upon binding to a ligand on its own is not sufficient to prove that the conformational change is required for binding. However, since CD measurements can be performed rapidly it could be informative to employ this type of analysis as a standard characterisation for aptamers. It might turn out that conformational changes are the rule rather than the exception.

Interestingly, the differences in affinity of SSL-2.5 and GQCL-4.7 are mainly due to different on-rates. If SSL-2.5 would simply provide more/better contact points for H107, on- and off-rates might have been affected similarly. Therefore the lower affinity of GQCL-4.7 might be the consequence of a more rigid structure, which would require more time/energy to undergo a conformational change (if it should be required for binding).

Comparison of aptamers and peptides

Although the GQCL-4.7 and the SSL-2.5 aptamers seem to have a rather different secondary structure, they have some common features. Both of them are purine rich, contain purine residues opposed to each other at the border of helical regions, and both include the motif UGGAAG. However, no common secondary structure is recognisable, and the differences in affinity for H107 suggest the formation of distinct structures.

Purine-rich sequences have been reported to be important for the activity of other aptamers, what might be the special feature of purine clusters? An attractive hypothesis would be that structural flexibility could be incorporated into an RNA structure by non Watson–Crick interactions. Hoogsteen base pairs have been found in the binding sites of ribosomal proteins, Rev, the E loop of 5SRNA, and splice sites of pre mRNA . A family of motifs involving G-A mismatches in ribosomal RNAs was studied and provided evidence for the formation of a sheared tandem structure. Since the A-residues could always maintain a similar exposure in the minor groove it was suggested that they might represent recognition or anchoring units (30). The formation of purine–purine interactions might be the RNA analogue of a hydrophobic pocket in proteins, which upon conformational change becomes available for hydrophobic interactions with a ligand. Furthermore, the functional groups of the bases used for Hoogsteen or reversed Hoogsteen interactions are different from those of the standard base pairs. This leaves for example the NH_2-group of guanine available for interactions with the ligand.

Finally, do the aptamers share some similarity to the natural H107 ligand, ferritin or the linear peptide selected from the phage display library? The exact structure of the H107 epitope is not known, however, the selected peptide has been modelled to the surface of ferritin. These simulations showed that the peptide-consensus sequence YDAxxxW could be super-imposed onto the ferritin surface residues $Y_{39}D_{42}W_{93}$. These residues are distant in the linear sequence but close in space in the three-dimensional structure (31). The ferritin residues identified to be important for binding to H107 are compatible with the recognition or the failure to do so of ferritin mutants by H107 (32). It would not be difficult to imagine that the purine cores identified in the H107 aptamers could make a spatial arrangement of contacts in a way equivalent to that of the aromatic side chains likely to represent the H107 epitope of ferritin (see fig. 4 in 31). In this case the selected aptamers might be structurally equivalent to the protein structure and not simply binding to the same site of the ligand. A similar proposal has been made for aptamers selected with an anti-human insulin receptor antibody (33). In the absence of structural data, further evidence for this hypothesis would be if an antibody could be generated against one of these aptamers and if the antibody raised against the aptamer would recognise the peptide antigen.

ACKNOWLEDGEMENTS

I would like to thank Philippe Neuner for the 10th year of high-speed/tech oligonucleotide synthesis, John Aherne of Affinity Sensors for excellent customer service and for his introduction to real time interaction analysis, Manuela Emili for artworks, Sven Behrens for helpful discussions, Elisabetta Bianchi and Raffaele DeFrancesco for advice on CD-measurements, Joe Jiricny and Iain Mattaj for valuable comments on the manuscript, and Alessandra Luzzago and Alfredo Nicosia for helpful suggestions, antibodies and encouragement when approaching local minima during the course of these experiments.

REFERENCES

1 Fedor,M.J. (1994) *Struct. Biol.* **5**, Vol.1, 267–269.
2 Joyce,G.F. (1994) *Curr. Opin. Struc. Biol.* **4**, 331–336.
3 Kenan,D.J., Tsai,D.E. and Keene,J.D. (1994) *Trends Biol. Sci.* **19**, 57–64.
4 Ellington,A.D. and Szostak,J.W. (1990) *Nature* **346**, 818–822.
5 Tuerk,C. and Gold,L. (1990) *Science* **249**, 505–510.
6 Burd,C.G. and Dreifuss,G. (1994) *EMBO J.* **13**, 1197–1204.
7 Giver,L., Bartel,D., Zapp,M., Pawul,A., Green,M. and Ellington,A. (1993) *Nucleic Acids Res.* **21**, 5509–5516.
8 Tsai,D.E., Harper,D.S. and Keene,J.D. (1991) *Nucleic Acids Res.* **19**, 4931–4936.
9 Tuerk,C., MacDougal,S. and Gold,L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6988–6992.
10 Bock,L.C., Griffin,L.C., Latham,J.A., Vermaas,E.H. and Toole,J.J. (1992) *Nature* **355**, 564–566.
11 Lehman,N. and Joyce,G.F. (1990) *Nature* **361**, 182–185.
12 Robertson,D.L. and Joyce,G.F. (1990) *Nature* **344**, 467–468.
13 Lorsch,J.R. and Szostak,J.W. (1994) *Nature* **371**, 31–36.
14 Wilson,C. and Szostak,J.W. (1995) *Nature* **374**, 777–782.
15 Jenison,R.D., Gill,S.C., Pardi,A. and Polisky,B. (1994) *Science* **263**, 1425–1429.
16 Gold,L., Polisky,B., Uhlenbeck,O. and Yarus,M. (1995) *Annu. Rev. Biochem.* **64**, 763–797.
17 Luzzago,A., Felici,F., Tramontano,A., Pessi,A. and Cortese,C. (1993) *Gene* **128**, 51–57.
18 Lauhon,C.T. and Szostak,J.W. (1995) *J. Am. Chem. Soc.* **117**, 1246–1257.
19 Harada,K. and Frankel,A.D. (1995) *EMBO J.* **14**, 5798–5811.
20 Green,L., Waugh,S., Brinkley,J.P., Hostomska,Z., Hostomsky,Z. and Tuerk,C. (1995) *J. Mol. Biol.* **247**, 60–68.
21 Pley,H.W., Flaherty,K.M. and McKay,D.B. (1994) *Nature* **372**, 68–74.
22 Tuschl,T., Gohlke,C., Jovin,T.M., Westhof,E. and Eckstein,F. (1994) *Science* **266**, 785–789.
23 Jaeger,J.A., SantaLucia,Jr,J. and Tinoco, Jr,I. (1993) *Annu. Rev. Biochem.* **62**, 255–287.
24 Daly,T.J., Rusche,J.R., Maione,T.E. and Frankel,A.D. (1990) *Biochemistry* **29**, 9791–9795.
25 Calnan,B.J., Biancalana,S., Hudson,D. and Frankel,A.D. (1991) *Genes Dev.* **5**, 201–210.
26 Mattaj,I.W. and DeRobertis,E.M. (1985) *Cell* **40**, 111–118.
27 Tsai,D.E., Kenan,D.J. and Keene,J.D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8864–8868.
28 Morton,T.A., Myszka,D.G. and Chaiken,I.M. (1995) *Anal. Biochem.* **227**, 176–185.
29 Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) *Science* **257**, 76–80.
30 Gautheret,D., Konings,D. and Gutell,R.R. (1994) *J. Mol. Biol.* **242**, 1–8.
31 Pizzi,E., Cortese,R. and Tramontano,A. (1995) *Biopolymers* **36**, 675–680.
32 Helmer-Critterich,M., Rovida,E., Luzzago,A. and Tramontano,A. (1995) *Mol. Immunol.* **32**, 101–1010.
33 Doudna,J.A., Cech,T.R. and Sullenger,B.A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 2355–2359.
34 Luzzago,A. and Cesareni,G. (1989) *EMBO J.* **8**, 569–576.
35 Luzzago,A., Arosio,P., Iacobello,C., Ruggeri,G., Capucci,L., Brocchi,E., De Simone,F., Gamba,D., Gabri,E., Levi,S. and Albertini,A. (1986) *Biochim. Biophys. Acta* **872**, 61–71.