Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Preliminary crystallographic characterization of an *in vitro* evolved biotin-binding RNA pseudoknot

Jay C. Nix, Angela R. Newhoff† and Charles Wilson*

Department of Biology and Center for the Molecular Biology of RNA, Sinsheimer Laboratories, University of California, Santa Cruz 95064, USA

† Current address: Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143-0448, USA.

Correspondence e-mail: wilson@biology.ucsc.edu

Received 29 September 1997
Accepted 1 April 1998

A biotin-binding RNA pseudoknot developed through *in vitro* selection has been crystallized using the hanging-drop vapor-phase diffusion method. The X-ray diffraction data indicate that the crystals belong to the space group $P4_222$ with unit-cell parameters $a = b = 55.2$, $c = 62.7$ Å and $\alpha = \beta = \gamma = 90^\circ$. The crystals are $120 \times 80 \times 40$ µm and $V_M = 2.17$ Å^3 Da^{-1}. The crystals diffract to 2.8 Å.

1. Introduction

Atomic resolution structures are required for fully understanding the structural basis for RNA function, yet only a handful of RNA molecules with tertiary structure have been successfully analyzed by X-ray crystallography (Pley *et al.*, 1994; Scott *et al.*, 1995; Cate *et al.*, 1996). The absence of bound ligands in previously determined X-ray structures (with the important exception of aminoacyl-tRNA synthetases complexed with tRNAs, *e.g.* Goldgur *et al.*, 1997) has greatly limited our understanding of how RNA structure is coupled to molecular recognition and ultimately biological function. Recent *in vitro* selection (SELEX) experiments starting with pools of random-sequence molecules have yielded a variety of RNA with highly specific ligand-binding activities (aptamers) and catalytic activities (ribozymes) (Bock *et al.*, 1992; Latham *et al.*, 1994; Lorsch & Szostak, 1994*a*,*b*; Huizenga & Szostak, 1995; Geiger *et al.*, 1996; Wang *et al.*, 1996; Burke & Gold, 1997). A combination of factors, including small size, efficient folding and opportunities for *in vitro* phylogenetic analysis, make these RNAs ideal model systems for correlating structure and function at high resolution.

Several previous NMR studies of aptamers have revealed novel base–base interactions and wide variations in backbone conformation, suggesting that there remains a broad range of unelucidated tertiary structures for RNA (Kelly *et al.*, 1996; Jiang *et al.*, 1996, 1997; Fan *et al.*, 1996; Dieckmann *et al.*, 1996; Yang *et al.*, 1996). X-ray crystallography can extend our understanding of RNA structures by adding to the body of structural information available. Additionally, crystallographic studies can provide direct information about the importance of solvent cofactors, especially divalent cations, important to structure stabilization and molecular recognition.

The current report describes a preliminary crystallographic analysis of an *in vitro* evolved RNA selected to bind biotin, a common enzyme cofactor used in carboxylation reactions (Wilson & Szostak, 1995). Recognition by the aptamer is highly specific in that only very closely related ligand analogs are capable of competing with biotin for binding. Phylogenetic analysis and chemical probing studies have shown that the aptamer folds as a purine-rich pseudoknot (Wilson *et al.*, 1998), but have not revealed how molecular recognition is achieved in the complex. Pseudoknots are one of the most common RNA tertiary-structure motifs, also appearing within 16S ribosomal RNA, the group I self-splicing intron and a variety of regulatory elements in viral transcripts (Matsufuji *et al.*, 1995; Jaeger *et al.*, 1996; Poot *et al.*, 1996; Hemmings-Mieszczak *et al.*, 1997; Rijnbrand *et al.*, 1997). The biotin aptamer pseudoknot is remarkably similar to that of the retroviral frame-shifting elements characterized by Chen *et al.* (1995). NMR experiments have suggested that the angle between the two pseudoknot helices in these elements may be a critical determinant of frame-shifting efficiency, although conflicting studies disagree as to whether the helices are, in fact, coaxially stacked (Shen & Tinoco, 1995; Chen *et al.*, 1996; Du & Hoffman, 1997). High-resolution crystallographic analysis of the biotin aptamer pseudoknot would reveal the constraints imposed upon folding into this common topology. In addition, it should explain how a ligand-binding pocket is formed from such a simple structure.

2. Methods

2.1. RNA preparation and purification

DNA templates for transcription and the T7 promoter top-strand were synthesized on a Millipore Expedite 8909 nucleic acid synthesizer and purified by denaturing 12% polyacrylamide gel electrophoresis. Template and promoter strands were pre-annealed by heating to 338 K prior to transcription. Transcription reactions were performed with 0.5 µ*M* template strand, 0.5 µ*M* T7 promoter

top-strand, 5 m*M* of each NTP, 5 m*M* DTT, 35 m*M* $MgCl_2$, 0.015%(*v*/*v*) spermidine, 0.1%(*v*/*v*) Triton X-100, 40 m*M* Tris–HCl (pH 8.1) and an optimized quantity of recombinant T7 RNA polymerase. Transcription was allowed to proceed overnight at 298 K. Small-scale reactions (100 µl) were performed with α-^{32}P-UTP to facilitate measurements of biotin-agarose-binding efficiency. A large-scale (20 ml) transcription reaction to provide material for crystallization trials was worked up as follows. Reaction was stopped by the addition of solid urea to an approximate final concentration of 4 *M* and purification was performed immediately on a 20% 14 × 16 × 0.3 cm denaturing polyacrylamide gel, run at 75 W for 5.5 h. A single large band of RNA visualized by UV shadowing was extracted from gel bits by electroelution using the Elutrap electrophoresis chamber (Schleicher & Schuell) and subsequently precipitated directly from TBE buffer by the addition of 2.5 volumes of ethanol. The resulting pellet was washed in ice-cold 70% ethanol, lyophilized, resuspended in 1 ml Milli-Q H_2O and stored at 277 K. The final yield determined spectrophotometrically was approximately 16 mg.

2.2. Affinity purification

1 ml biotin agarose (Sigma) was applied to a 1.5 ml Bio-spin column (BioRad) and equilibrated with 20 column volumes of binding buffer (0.1 *M* KCl, 10 m*M* Na-HEPES, pH 7.4, 5 m*M* $MgCl_2$). RNA was pre-equilibrated in a total volume of 1 ml binding buffer at 338 K for 5 min (annealing protocol) and then for 30 min at 298 K prior to loading onto the column. After loading with RNA, the column was washed with 15 × 1 ml binding buffer. Specifically bound RNA was then eluted with 5 × 1 ml elution buffer (binding buffer containing 5 m*M* biotin). Elution fractions were pooled and concentrated to 5–8 mg ml^{-1} using a Centricon 3 ultraspin column (Amicon Inc., Beverly, MA) and stored at 277 K until use.

2.3. Crystallization

Crystallization trials were carried out by the hanging-drop vapor-diffusion method using siliconized cover slips (Hampton Research) and 24-well Linbro plates. Drops containing 5–8 mg ml^{-1} RNA solution in biotin elution buffer were mixed with reservoir solution in 1:2, 2:2 and 2:1 µl RNA solution:reservoir ratios. Drops were mixed by pipetting and allowed to equilibrate over 1 ml reservoir at either 298 or 277 K. Initial crystallization screens were based on an expansion of the Doudna sparse matrix (Doudna *et al.*, 1993). This screen is based on conditions used previously for the crystallization of tRNA and relies primarily on small organics as precipitating agents. Early crystallization results with another *in vitro* selected RNA indicated that high concentrations of LiCl could improve crystallization (Sussman *et al.*, 1998) and we prepared, in parallel, a modified Doudna screen in which 1 *M* LiCl was added to every hanging drop and reservoir. In addition, the methyl-pentanediol-based Berger nucleic-acid miniscreen (Berger *et al.*, 1996) and a simple PEG-based screen were also tested.

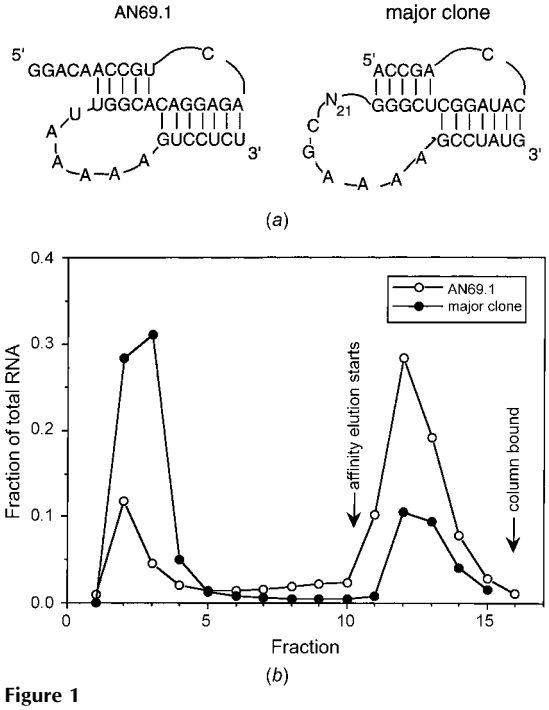

Figure 1
Secondary structures and affinity-elution profiles of (*a*) the biotin aptamer major clone and (*b*) AN69.1 used for crystallization trials. Fractions 11–15 in (*b*) were pooled and concentrated for crystallization experiments.

2.4. X-ray analysis

10% glycerol was added to the reservoir solution as a cryoprotectant. Crystals equilibrated for several minutes in this solution were mounted in a 0.2 mm nylon loop (Hampton Research) and flash-frozen under a stream of boiling nitrogen provided by an MSC fixed-tube liquid-nitrogen-based low-temperature system. Cu $K\alpha$ X-rays were produced by a Rigaku RU-200 rotating-anode X-ray generator with graphite monochromator operating at 180 kV, 50 mA. Data were collected as a series of 1° oscillation frames on a Rigaku R-AXIS IIc imaging-plate detector. The crystals were exposed for 1500 s at a crystal-to-detector distance of 100 mm. Collection at 100 K eliminated radiation-induced crystal decay.

3. Results

The original pool of biotin-binding RNAs was dominated by a single major clone whose pseudoknot core was interrupted by an extended loop of 27 nucleotides. This loop could be shortened to six nucleotides but further deletions significantly reduced its binding efficiency. We predicted that the presence of such an extended unstructured loop within this short RNA would complicate crystallization attempts and thus analyzed other minor species within the selected pool. In contrast to the major clone, all other biotin aptamers lack this extended loop. Minimal domains corresponding to each minor clone were prepared and assayed for binding to biotin agarose. As indicated in Fig. 1, one of these sequences (AN69.1) binds with an efficiency equivalent to that of the major clone yet is defined by a minimal structure of 31 nucleotides (five additional nucleotides at the 5′ end ensure high-efficiency transcription). The crystallographic trials presented in this study have focused on this version of the biotin aptamer.

A fundamental problem in macromolecular crystallization is the tendency for contaminants to poison crystal growth. This problem is likely to be extreme for RNA crystallization since chemically homogeneous molecules can readily adopt alternate stable but misfolded conformations. To minimize these problems, we have relied upon purification by ligand-affinity chromatography with the assumption that functional molecules are likely to be trapped in a single correct conformation. This process is especially straightforward for aptamers, since the protocol used to isolate them from the initial random pool may also be used to separate folded and misfolded molecules in a clonally pure mixture (Fig. 1*b*). Because the isolated RNA is obtained by competitive elution with the ligand, the material used for crystallization trials exists as a specific complex of RNA aptamer together with its ligand.

Several different sparse-matrix screens biased towards conditions used previously for RNA or DNA crystallization have proven useful in previous studies (Doudna

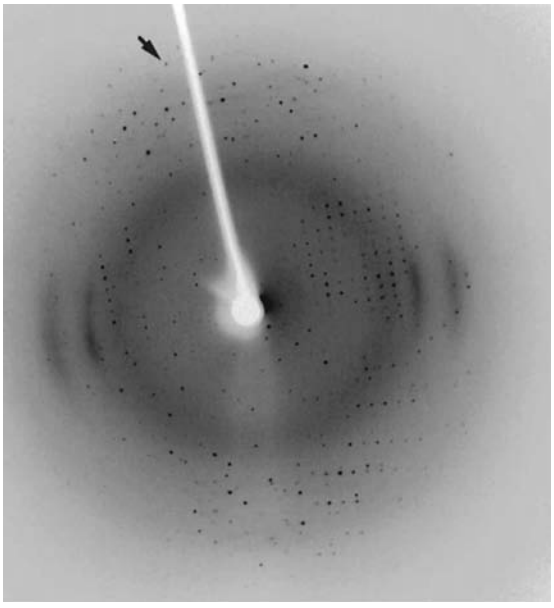

Figure 2
A 1° oscillation image of an AN69.1 crystal of dimensions 120 × 80 × 40 µm exposed for 1500 s. The arrow at the edge of the image is at 2.8 Å resolution.

et al., 1993; Scott *et al.*, 1995; Berger *et al.*, 1996). All of the initial screens attempted with AN69.1 yielded crystals of poor-to-moderate quality resulting in starburst-like multicrystals of varying size. Conditions containing polyethylene glycol yielded the largest crystals and were chosen for optimization. Improved crystals were obtained after adjusting polyethylene glycol concentration, polyamine concentration and various divalent-cation additives. Rod-shaped crystals of dimensions 120 × 80 × 40 µm were obtained in drops set up with 4% PEG 8000, 20 m*M* $MgCl_2$, 20 m*M* K-HEPES (pH 7.5) and 1.0 m*M* spermine. Staining with ethidium bromide or toluidine blue confirmed that these were RNA crystals. Crystals with similar morphology were also grown in related conditions (5% PEG 4000, 20 m*M* potassium cacodylate buffer, pH 6.5, 15 m*M* $MgCl_2$, 1.5 m*M* spermine and 2 m*M* cobalt hexamine trichloride). Crystals were grown at 277 K and took longer than three months to form. To date, these crystals have been difficult to reproduce.

A preliminary data set was collected with an average I/σ of 11 in the resolution shell to 3.4 Å, although in Fig. 2 one can see that diffraction extends out to 2.8 Å. These results indicate that while small, the crystals should be useful for determination of the structure at high resolution. A pseudoprecession plot of unmerged data indicates that the diffraction is anisotropic, extending to beyond 3.0 Å along the l axis, but to approximately 3.4 Å along the symmetric h and k axes. This anisotropy extends beyond the $h = 0$ plane to all zones. The crystals have unit-cell parameters $a = b = 55.2$, $c = 62.7$ Å, $\alpha = \beta = \gamma = 90^\circ$, and may be assigned on the basis of Laue symmetry and systematic absences to space group $P4_222$. Assuming a single RNA molecule per asymmetric unit, V_M is calculated to be 2.17 $Å^3$ Da^{-1}. We are continuing our attempts to increase the size of the crystals and to determine whether the anisotropy is a function of the freezing procedure or inherent to the crystal form itself.

References

Berger, I., Kang, C., Sinha, N., Wolters, M. & Rich, A. (1996). *Acta Cryst.* D**52**(3), 465–468.

Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. (1992). *Nature (London)*, **355**(6330), 564–566.

Burke, D. H. & Gold, L. (1997). *Nucleic Acids Res.* **25**(10), 2020–2024.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996). *Science*, **273**(5282), 1678–1685.

Chen, X., Chamorro, M., Lee, S. I., Shen, L. X., Hines, J. V., Tinoco, I. Jr & Varmus, H. E. (1995). *EMBO J.* **14**(4), 842–852.

Chen, X., Kang, H., Shen, L. X., Chamorro, M., Varmus, H. E., Tinoco, I. Jr (1996). *J. Mol. Biol.* **260**(4), 479–483.

Dieckmann, T., Suzuki, E., Nakamura, G. K. & Feigon, J. (1996). *RNA*, **2**(7), 628–640.

Doudna, J. A., Grosshans, C., Gooding, A. & Kundrot, C. E. (1993). *Proc. Natl Acad. Sci. USA*, **90**(16), 7829–7833.

Du, Z. & Hoffman, D. W. (1997). *Nucleic Acids Res.* **25**(6), 1130–1135.

Fan, P., Suri, A. K., Fiala, R., Live, D. & Patel, D. J. (1996). *J. Mol. Biol.* **258**(3), 480–500.

Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A. & Famulok, M. (1996). *Nucleic Acids Res.* **24**(6), 1029–1036.

Goldgur, Y., Mosyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O. & Khodyreva, S. (1997). *Structure*, **5**(1), 59–68.

Hemmings-Mieszczak, M., Steger, G. & Hohn, T. (1997). *J. Mol. Biol.*, **267**(5), 1075–88.

Huizenga, D. E. & Szostak, J. W. (1995). *Biochemistry*, **34**(2), 656–665.

Jaeger, L., Westhof, E. & Michel, F. (1996). *Biochimie*, **78**(6), 466–473.

Jiang, F., Kumar, R. A., Jones, R. A. & Patel, D. J. (1996). *Nature (London)*, **382**(6587), 183–186.

Jiang, L., Suri, A. K., Fiala, R. & Patel, D. J. (1997). *Chem. Biol.* **4**(1), 35–50.

Kelly, J. A., Feigon, J. & Yeates, T. O. (1996). *J. Mol. Biol.*, **256**(3), 417–422.

Latham, J. A., Johnson, R. & Toole, J. J. (1994). *Nucleic Acids Res.* **22**(14), 2817–2822.

Lorsch, J. R. & Szostak, J. W. (1994*a*). *Nature (London)*, **371**(6492), 31–36.

Lorsch, J. R. & Szostak, J. W. (1994*b*). *Biochemistry*, **33**(4), 973–982.

Matsufuji, S., Matsufuji, T., Miyazaki, Y., Murakami, Y., Atkins, J. F., Gesteland, R. F. & Hayashi, S. (1995). *Cell*, **80**(1), 51–60.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). *Nature (London)*, **372**(6501), 68–74.

Poot, R. A., Pleij, C. W. & van Duin, J. (1996). *Nucleic Acids Res.* **24**(19), 3670–3676.

Rijnbrand, R., van der Straaten, T., van Rijn, P. A., Spaan, W. J. & Bredenbeek, P. J. (1997). *J. Virol.* **71**(1), 451–457.

Scott, W. G., Finch, J. T., Grenfell, R., Fogg, J., Smith, T., Gait, M. J. & Klug, A. (1995). *J. Mol. Biol.* **250**(3), 327–332.

Scott, W. G., Finch, J. T. & Klug, A. (1995). *Nucleic Acids Symposium Series*, pp. 214–216.

Shen, L. X. & Tinoco, I. Jr (1995). *J. Mol. Biol.* **247**(5), 963–978.

Sussman, D., Reilly, K., Greensides, D. & Wilson, C. (1999). *Acta Cryst.* D**54**, 327–329.

Wang, Y., Killian, J., Hamasaki, K. & Rando, R. R. (1996). *Biochemistry*, **35**(38), 12338–12346.

Wilson, C., Nix, J. C. & Szostak, J. W. (1998). *Biochemistry*. In the press.

Wilson, C. & Szostak, J. W. (1995). *Nature (London)*, **374**, 777–782.

Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. & Famulok, M. (1996). *Science*, **272**(5266), 1343–1347.