Metal-binding sites in the major groove of a large ribozyme domain

Jamie H Cate and Jennifer A Doudna*

Background: Group I self-splicing introns catalyze sequential transesterification reactions within an RNA transcript to produce the correctly spliced product. Often several hundred nucleotides in size, these ribozymes fold into specific three-dimensional structures that confer activity. The 2.8 Å crystal structure of a central component of the *Tetrahymena thermophila* group I intron, the 160-nucleotide P4–P6 domain, provides the first detailed view of metal binding in an RNA large enough to exhibit side-by-side helical packing. The long-range contacts and bound ligands that stabilize this fold can now be examined in detail.

Results: Heavy-atom derivatives used for the structure determination reveal characteristics of some of the metal-binding sites in the P4–P6 domain. Although long-range RNA–RNA contacts within the molecule primarily involve the minor groove, osmium hexammine binds at three locations in the major groove. All three sites involve G and U nucleotides exclusively; two are formed by G·U wobble base pairs. In the native RNA, two of the sites are occupied by fully-hydrated magnesium ions. Samarium binds specifically to the RNA by displacing a magnesium ion in a region critical to the folding of the entire domain.

Conclusions: Bound at specific sites in the P4–P6 domain RNA, osmium (III) hexammine produced the high-quality heavy-atom derivative used for structure determination. These sites can be engineered into other RNAs, providing a rational means of obtaining heavy-atom derivatives with hexammine compounds. The features of the observed metal-binding sites expand the known repertoire of ligand-binding motifs in RNA, and suggest that some of the conserved tandem G·U base pairs in ribosomal RNAs are magnesium-binding sites.

Address: Deptartment of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA.

*Corresponding author.
E-mail: doudna@csb.yale.edu

Key words: group I intron, MAD phasing, osmium hexammine, RNA

Received: 5 August 1996
Revisions requested: 29 August 1996
Revisions received: 5 September 1996
Accepted: 5 September 1996

Structure 15 October 1996, 4:1221–1229

 ISSN 0969-2126

Introduction

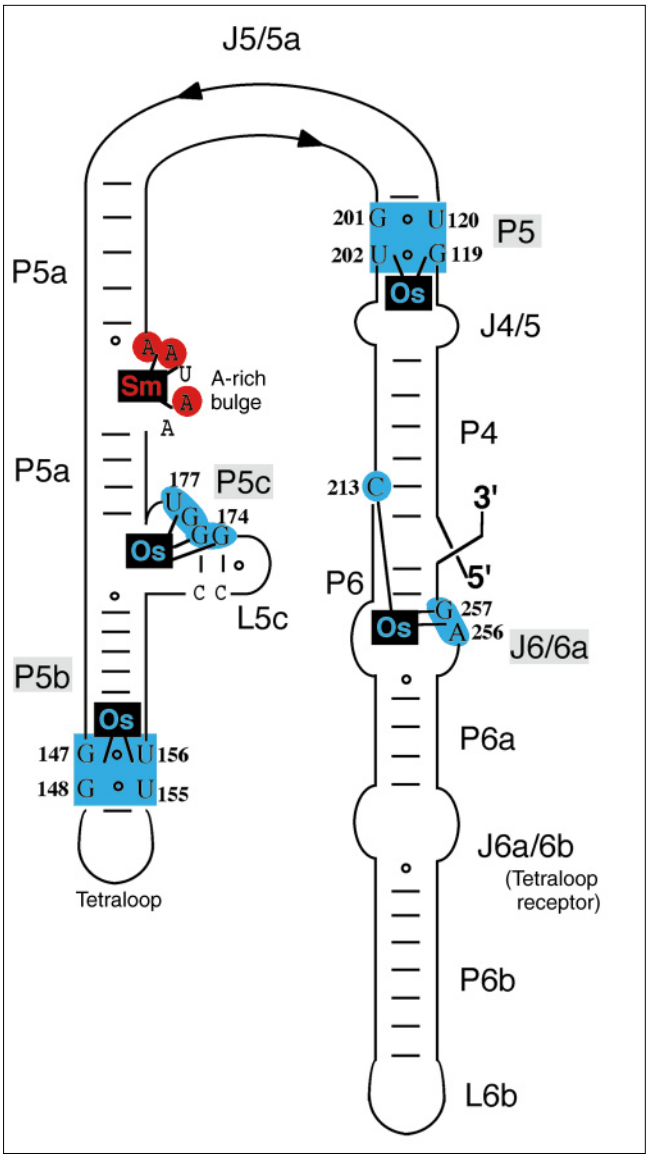
Like protein enzymes, catalytic RNAs probably use a defined set of structural motifs to build the complex architectures necessary for function. Divalent metal ions are essential components of RNA tertiary structure and, in the case of large ribozymes, are required for both stability and catalytic activity [1–6]. A key to understanding and engineering ribozyme function lies in the determination of how metals are bound specifically by RNA and positioned for a variety of uses.

In the well studied group I self-splicing introns, magnesium-dependent folding of the RNA was first detected using Fe(II)-EDTA to probe the accessibility of the phosphodiester backbone in solution [4,7]. This experimental approach revealed that half of the conserved catalytic core of the *Tetrahymena thermophila* intron is part of an independently folding domain consisting of the base-paired (P) regions P4 to P6 (P4–P6; Fig. 1) [8]. The P4–P6 domain, modeled as part of the active site by Michel and Westhof [9], folds rapidly in solution [10] and assembles *in trans* with the other half of the ribozyme to form a functional complex [11].

The 2.8Å crystal structure of the P4–P6 domain [12] provides the first detailed view of an RNA large enough to reveal structural motifs directly involved in higher order RNA folding. In the domain, two sets of stacked helices lie parallel to one another, connected by a sharp bend at one end (Fig. 2). Two key interactions stabilize the overall fold of the molecule. An adenosine-rich bulge docks in the minor groove of the P4 helix and a GAAA tetraloop binds in the minor groove of its receptor. In addition to base-specific hydrogen bonding and base stacking, these long-range contacts are further secured by pairs of interdigitated riboses termed ribose zippers [12].

Most of the contacts that stabilize internal domain structure, as well as the packing of arrays of molecules in the crystal lattice, involve the minor groove [12,13]: the wide and shallow minor grooves of A-form helices are generally more accessible than the deep major grooves [14]. However, non-canonical base pairings, loops and bulges in RNA can perturb the geometry of a helix, affecting the shape and electrostatic potential in localized regions [15,16]. Here, we show that such local perturbations within the P4–P6 domain give rise to specific metal-binding

Figure 1



Osmium- and samarium-binding sites in the P4–P6 RNA. Secondary structure of the P4–P6 domain; in red, samarium-binding site (Sm); in blue, osmium-binding sites. The sites in P5, P5b and P5c bind osmium hexammine ions, whereas that in J6/6a binds a putative osmium pentammine ion.

pockets in the major groove and near a helical junction. Osmium (III) hexammine, the metal ion used to determine the RNA structure by multiwavelength anomalous diffraction (MAD) phasing, binds at three locations in the major groove where non-standard base pairs create pockets of negative electrostatic potential. A putative osmium pentammine ion binds to a fourth site in the major groove. Samarium binds between phosphate oxygens in an adenosine-rich corkscrew structure at the junction of three helices. In three cases, the heavy atoms occupy sites normally bound by magnesium in the native RNA. The binding sites expand the known repertoire of ligand-binding motifs in RNA and may provide a general means of derivatizing appropriately designed RNAs for x-ray crystallographic analysis. Phylogenetic sequence comparisons show that one class of these motifs occurs frequently in ribosomal RNAs [17], suggesting a mechanism for metal binding in the ribosome.

Results and discussion

Crystallization conditions suggest a rationale for heavy-atom derivatives

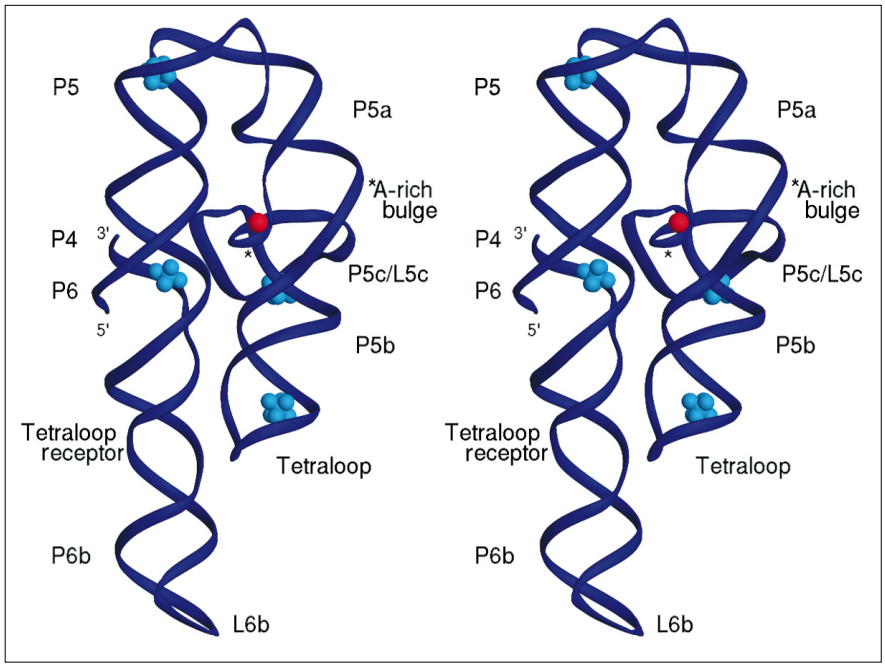
The original optimized crystallization conditions for the P4–P6 domain [18] did not produce crystals reproducibly (on average, one crystal per 15 vapor diffusion drops in 1–2 months). To generate the large number of crystals necessary to find heavy-atom derivatives, we developed a protocol using microseeding and cobalt (III) hexammine chloride. When included in crystallization solutions, low concentrations of cobalt hexammine chloride (0.1–1.0 mM) dramatically increased the number, size and growth rate of P4–P6 crystals. Crystallization drops pre-equilibrated with cobalt hexammine were microseeded with finely crushed P4–P6 crystals, consistently yielding 2–3 usable crystals per crystallization drop within 2–3 weeks. The diffraction properties of crystals obtained by this procedure were comparable with those of crystals obtained by the earlier method.

The dramatic effects of cobalt hexammine on crystal growth occurred in spite of a large excess of magnesium ions (25–50 mM) in the crystallization conditions, suggesting that specific binding sites for hexammines might be present in the RNA. To take advantage of potential sites, crystals were soaked in stabilizing solutions containing the chemically-related osmium (III) hexammine ion [19]. Initial difference Patterson maps revealed four osmium sites in the asymmetric unit; four additional sites were found by difference Fourier syntheses.

In parallel with these experiments, we systematically tested metals in the lanthanide series (from Ce^{3+} to Lu^{3+}) as potential derivatives, as lanthanides proved useful in the determination of the tRNA and hammerhead RNA crystal structures [20–22]. Diffraction experiments on lanthanide-soaked crystals led us to focus on samarium chloride as a potential heavy-atom derivative for two reasons. First, the unit-cell dimensions changed as a function of increasing ionic radius for lanthanides in the series from Lu^{3+} to Sm^{3+}, after which they remained constant (Sm^{3+}–Ce^{3+}). Second, the mosaic spread of the diffraction pattern increased as a function of increasing ionic radius for all lanthanides except for Sm^{3+}, for which the mosaicity was only slightly worse than that of native crystals (0.7° versus 0.5°–0.6°). A full anomalous data set from a samarium-derivatized crystal yielded two strong peaks in an anomalous difference Patterson map. Samarium-soaked and Sm^{3+}/Os (III) hexammine double-soaked crystals, compared using difference

Figure 2

Locations of metal-binding sites in the crystal structure of the P4–P6 domain. Stereo representation of the crystal structure of P4–P6, rotated 180° from the view in Figure 1 (with the samarium-binding site in red and osmium-binding sites in blue). The figure was produced using the program RIBBONS [51].



Fourier methods, confirmed the osmium sites and provided the first clear indications of a solvent boundary in electron-density maps. Due to poor isomorphism and weak diffraction (to 4.5 Å), the samarium derivative was abandoned in favor of the osmium hexammine derivative.

The osmium derivative ultimately provided all of the phasing power necessary to solve the structure of the P4–P6 domain. Anomalous data from an osmium hexammine-soaked crystal were measured at two wavelengths near the osmium LIII edge (peak and first inflection point). In conjunction with data from a cobalt hexammine-containing crystal and density modification, the MAD data were used to calculate an electron-density map to 3.0 Å resolution [12] which was readily interpretable.

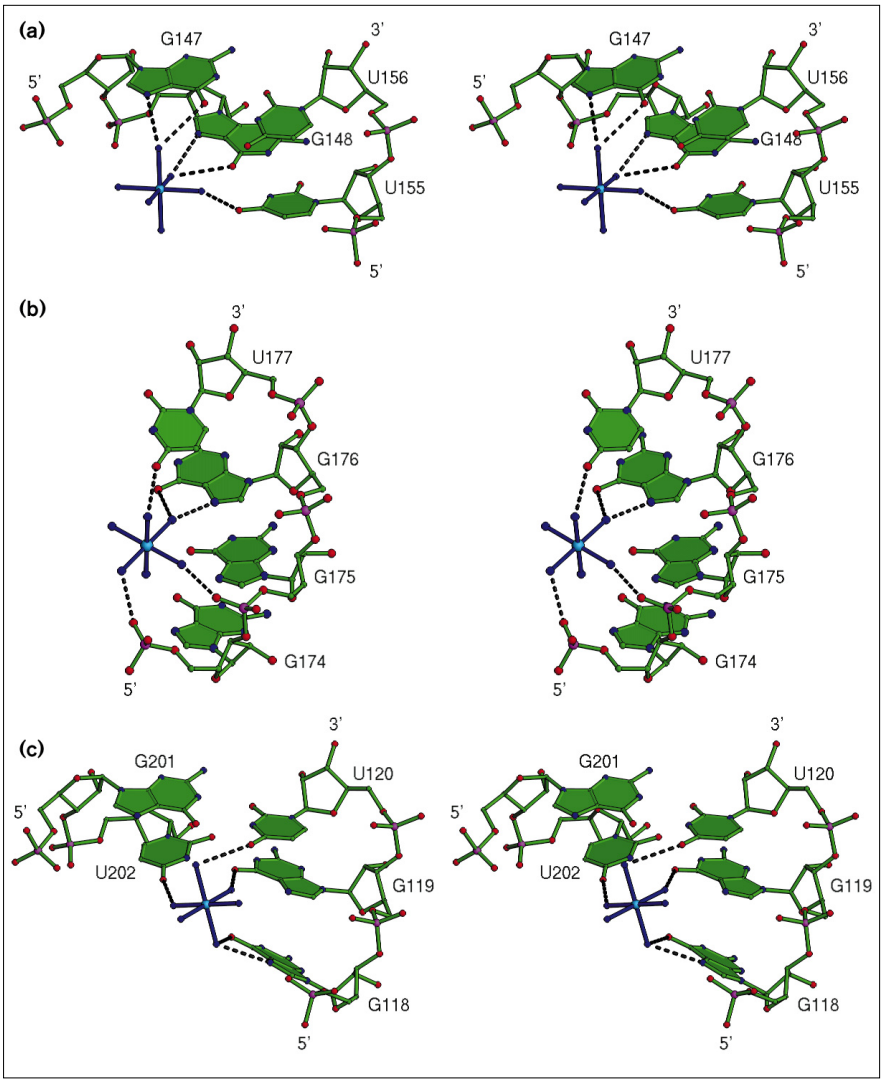
Samarium bridges phosphate oxygens in the A-rich bulge

There are two non-crystallographically related P4–P6 molecules in the asymmetric unit of the crystal, and samarium binds at the same site in each. As observed in tRNA [23,24], samarium substitutes for a magnesium ion bound by closely positioned phosphate oxygens in the native tertiary structure. Additional sites in tRNA also involve inner-sphere coordination to phosphate oxygens [25]. In the P4–P6 domain, binding occurs in the phosphate backbone of a corkscrew motif called the A-rich bulge. Phosphate oxygens from three residues in the A-rich bulge coordinate Mg^{2+} in the native RNA [12], and may coordinate Sm^{3+} in a similar manner. The low resolution diffraction limit (4.5 Å) and lack of isomorphism between the Sm^{3+} data and the native data attests to the structural perturbation that results from lanthanide binding to the RNA. Biochemical evidence supports the importance of this region for the overall folding of the domain, as well as for ribozyme activity [26,27].

Osmium hexammine binds in the major groove

Osmium hexammine binds at three sites in each P4–P6 molecule in the asymmetric unit. Unlike lanthanides, hexammines tend to replace weakly bound magnesium ions defined by outer-sphere coordination of the metal [25]. In $tRNA^{Phe}$ as well as A-form DNA, cobalt hexammine directly coordinates to 5′-GG-3′ in the major groove [25,28]. The binding sites in the P4–P6 RNA occur in the major groove near noncanonical base pairs at the ends of helices P5, P5b and P5c (Figs 1,2). The osmium hexammine site in P5 consists of the wobble base pairs 5′-GU-3′/3′-UG-5′, and that in P5b involves the wobble base pairs 5′-GG-3′/3′-UU-5′ (Fig. 1). In both of these sites, the hexammine ion is tucked between phosphate oxygens on one side and the major groove hydrogen-bond donors of G and U on the other; however, direct hydrogen bonds appear to be formed only to the bases (Fig. 3). In P5c, single-stranded G and U nucleotides hydrogen bond to the phosphate oxygens of nucleotides 163–164, creating a binding pocket at the base of the helix. Here, the osmium hexammine binds directly to hydrogen-bond donors of the bases as well as to the phosphate oxygens on the other side of the pocket (Fig. 3). A fourth osmium site is located in the major groove above the internal loop between helices P6 and P6a (joining region J6/6a; Fig. 1). In this case, the experimental electron density supports direct coordination between the osmium

Figure 3



Stereo drawings of the osmium hexammine binding sites: (a) site in P5b; (b) site in P5c; (c) site in P5. Dashed lines indicate probable hydrogen bonds; osmium hexammine is shown in blue. Although the individual ammine ligands are not visible at the resolution of this structure, the interactions were inferred by placing the osmium atom at the center of the hexammine density and optimizing hydrogen bonding parallel to the bases. This figure was produced using the program RIBBONS [51].

and a non-bridging phosphate oxygen. This coordination geometry is consistent with the binding of osmium (III) pentammine, a possible contaminant in the osmium (III) hexammine triflate preparation [19]. The refined occupancy of this site is approximately one third that of the other three (see Materials and methods section).

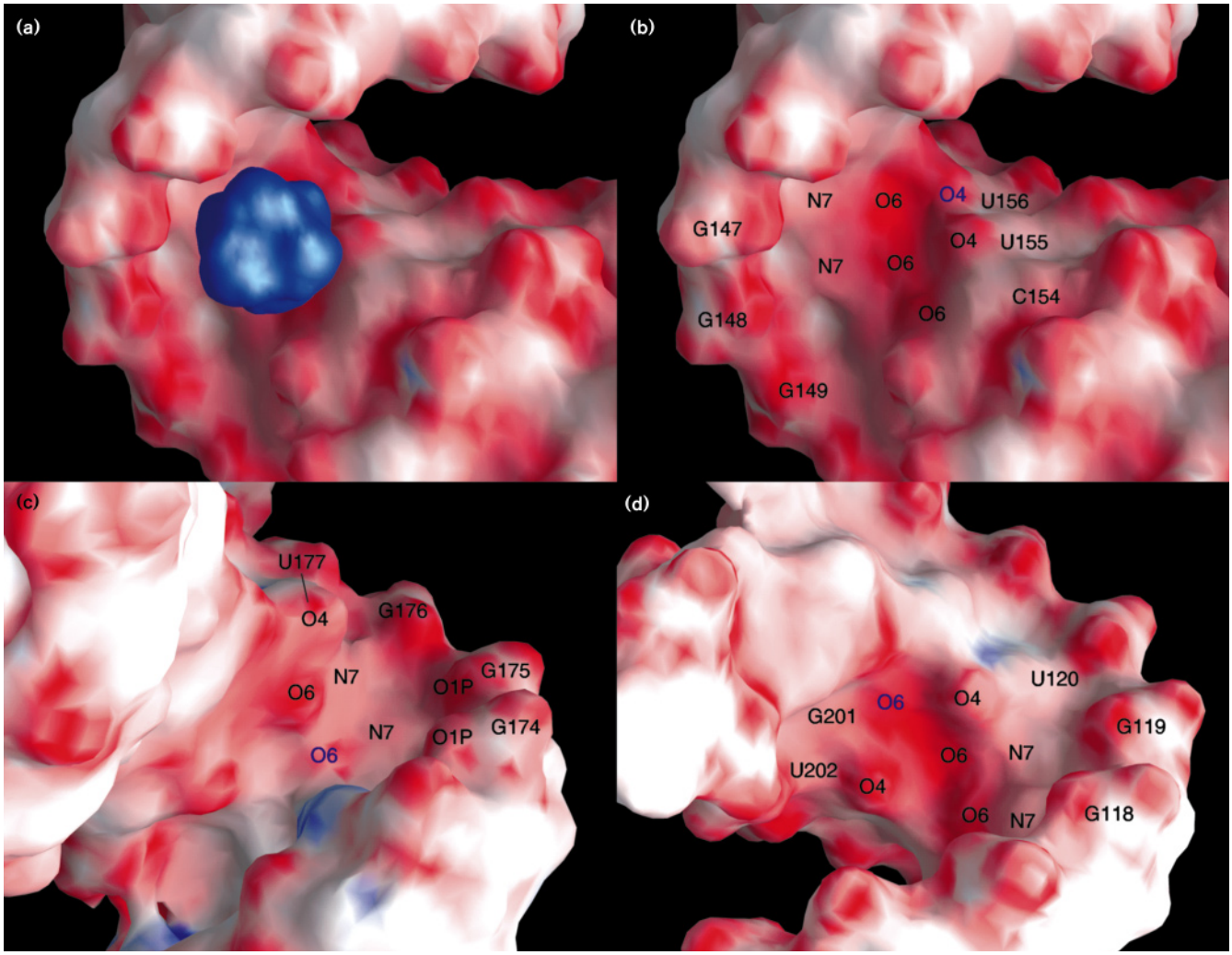
Hexammines displace magnesium ions in the native RNA

Hexammine ions and hexa-hydrated magnesium ions have comparable geometries and van der Waals radii, which may enable them to bind at similar sites in nucleic acids. Diffraction data measured from native, cobalt hexammine- and osmium hexammine-containing P4–P6 crystals were analyzed to determine whether magnesium and hexammine ions bind interchangeably. In the P4–P6 structure derived from crystals stabilized in 50 mM Mg^{2+} and 50 μM cobalt hexammine, only the two hexammine sites in P5b and P5c (Figs 1,2) show appreciable density in simulated-annealing omit maps (data not shown). Density (4σ) for an ion bound in P5 has become apparent in the most recent difference Fourier maps.

To determine whether the sites in P5, P5b and P5c bind magnesium ions, data from a native crystal was used to calculate new electron density maps (see Materials and methods section). Strong density (8–9σ) consistent with an hydrated magnesium ion occurs in the P5b and P5c sites, but not in the P5 site. On the basis of this observation, the sites in P5b and P5c are likely to be intrinsic hydrated Mg^{2+} sites in the native RNA.

Although the three hexammine-binding sites share certain features (Fig. 3), they clearly have differing affinities for metal ligands. To investigate this further, we examined

Figure 4



Electrostatic potential surfaces of the hexammine-binding sites: red indicates negative electrostatic potential and the hexammine ligand is shown in blue. Hydrogen-bond donors that have their solvent-accessible surface located within 0.5 Å of the solvent-accessible surface of the hexammine ligand are labeled in black; donors within 1.0 Å are labeled in blue. The solvent-accessible surfaces for the RNA and ligand were calculated separately; this analysis revealed that the ligands bury roughly equal solvent-accessible surface area in each site (260–270 Å^2). **(a)** P5b site with bound hexammine; **(b)** P5b site, no ligand bound; **(c)** P5c site, no ligand bound; **(d)** P5 site, no ligand bound. The electrostatic potential was calculated using the program DelPhi [49]. The figure was produced using the program GRASP [52].

the shape and electrostatic potential of each site (Fig. 4). At the resolution of our data, comparison of each site with and without the hexammine ligand bound shows that the same conformation is retained (see Materials and methods section). The three binding sites each consist of a concave surface, inside the major groove, with a highly negative electrostatic potential. These binding pockets are lined with hydrogen bond acceptors including N7 and O6 of guanosine, O4 of uridine and adjacent phosphate oxygens.

At two of the three sites in the P4–P6 RNA, hexammines displace magnesium ions involved strictly in outer-sphere coordination. As the electrostatic potential of the three hexammine-binding sites is similar, subtle differences in the geometry of the sites may account for discrimination between hexammine and magnesium ligands in P5. Particularly intriguing is the fact that the 5′-GG-3′/3′-UU-5′ wobble base pairs bind hexammines and magnesium, while the 5′-GU-3′/3′-UG-5′ site in P5 appears to bind only hexammines. The solvent-accessible surface of the 5′-GG-3′/3′-UU-5′ pairs in P5b may fit an octahedral polycation like hydrated magnesium more closely (compare Fig. 4b and 4d). Osmium hexammine may overcome a less optimal geometry of the 5′-GU-3′/3′-UG-5′ base pairs in P5 due to the additional positive charge and additional hydrogen-bond donors on the ion [29,30].

A new structural role for the G·U base pair?

G·U base pairs are often conserved in RNA phylogenies, suggesting their involvement in structural and functional interactions. For example, G·U pairs are the basis for recognition of alanyl-tRNA by its cognate synthetase [31,32], and for splice-site selection in group I self-splicing introns [33,34]. In group I introns, the G·U pair binds to the intron active site [35] and, in an NMR structure of the P1 substrate duplex [36], coordinates a magnesium ion. Binding of hexammines, and in one case a hydrated magnesium ion, to consecutive G·U base pairs in the P4–P6 domain now reveals another potential role for these non-standard pairings. Analysis of the conservation of consecutive G·U pairs in 16S and 23S ribosomal RNA showed that 5′-GG-3′/3′-UU-5′ pairs occur more frequently than 5′-GU-3′/3′-UG-5′ pairs [17]. In addition to their higher thermodynamic stability [37], 5′-GG-3′/3′-UU-5′ sites may be preferred because of their affinity for hydrated magnesium ions. It should be noted that 5′-UG-3′/3′-GU-5′ is the most frequently occurring conserved G·U tandem in ribosomal RNA [17], a tandem that does not exist in the P4–P6 domain. It will be interesting to determine whether this motif also binds metal ions.

Designing metal-binding sites in RNA

Hexammine ions bind with surprising selectivity to the P4–P6 RNA, given that hexammines can bind in multiple modes to nucleic acids [23,29]. The consecutive G·U wobble motifs may provide a general means of introducing osmium-hexammine-binding sites into RNA molecules, with obvious advantages for derivatizing RNA crystals. In the P4–P6 molecule, the G·U tandems are one base pair removed from the end of the helix (Fig. 1). As the major groove is narrow in A-form RNA, we wondered how far from the ends of a helix the G·U pairs could be placed and still bind hexammine ions. Modeling studies using the tandem G·U pairs in helix P5b indicate that the tandems would be accessible in any location within an A-form duplex (see also [14]). In addition, models of the electrostatic potential near the tandem G·U pairs suggest that the negative potential is intrinsic to the tandem, regardless of its position in the helix. The sensitivity of RNA crystal growth to the presence of low amounts of cobalt (III) hexammine may provide a quick test for the efficacy of any engineered or naturally-occurring hexammine-binding site.

Biological implications

Metal ions stabilize RNA structure and play a direct role in catalysis by ribozymes. The recently determined crystal structure of the P4–P6 domain from the *Tetrahymena thermophila* group I intron provides the first detailed view of metal-binding motifs in an RNA with extensive higher order folding. Although the overall fold of the domain is characterized by minor groove RNA–RNA contacts, the osmium hexammine ligand used for heavy atom derivatization binds at three unique sites in the major groove. All three sites involve G and U nucleotides exclusively: two are formed by tandem G·U wobble base pairs. In the native RNA, two of the sites are occupied by fully-hydrated magnesium ions. These results suggest a possible explanation for the abundance of conserved tandem G·U pairs in ribosomal RNAs: these positions may correspond to structurally-important magnesium-binding sites in the ribosome. For example, the neutralization of negative charge upon metal binding might facilitate subsequent steps in higher order RNA folding or the association of proteins.

Our results suggest a general method for heavy-atom derivatization of RNA crystals. Tandem G·U base pairs engineered within a helix would probably provide a specific binding pocket for osmium hexammine. The anomalous scattering properties of osmium, which were extremely useful for phasing the 52 kD P4–P6 domain, provide an excellent signal for phase determination by multiwavelength anomalous diffraction (MAD) and single isomorphous replacement with anomalous scattering (SIRAS) techniques.

Materials and methods

Crystals

P4–P6 RNA was prepared by using T7 RNA polymerase runoff transcription and purified by denaturing polyacrylamide gel electrophoresis, essentially as described in [18]. Diffraction-quality crystals were obtained by microseeding pre-equilibrated drops. The drops were set up using a 2:8:5 ratio of solutions a, b, and c, respectively (see below) and equilibrated against 17% (w/v) methylpentanediol (MPD) and 170 mM NaCl. Solution a contained 250 mM potassium cacodylate pH 6.0, 200 mM $MgCl_2$, and 2.5 mM spermine. Solution b contained 4 mg ml^{-1} P4–P6 RNA, annealed in the presence of 63 mM potassium cacodylate pH 6.0, 13 mM $MgCl_2$, and 160 μM Co (III) hexammine (2 equivalents per RNA), plus 0.1 mM EDTA pH 8.0 to chelate contaminating transition metals [38]. The RNA was annealed by heating to 65° for 10 min, followed by slow cooling to room temperature. Solution c contained 20–25% (w/v) MPD. Microseeds, stabilized as described below, were added to each drop by streak seeding with a cat whisker. After 2–3 weeks at 30° C, each drop typically contained at least one usable crystal.

Crystals ranging in size from 0.2–0.4 mm in each dimension were stabilized directly in the sitting drops by addition of 25% MPD (w/v), 100 mM potassium cacodylate, 50 mM $MgCl_2$, 0.5 mM spermine, and 50–100 μM cobalt hexammine. The first stabilizer was exchanged for a second containing an additional 10% isopropanol. Crystals were stored in this stabilizer until flash-frozen in liquid propane at liquid N_2 temperatures prior to data collection. For heavy-atom soaks, the cobalt hexammine was exchanged for osmium (III) hexammine or was supplemented with another metal (i.e. lanthanides). Samarium chloride was added to the crystals over the course of several days to a final concentration of 0.3 mM. Crystals grown in the absence of cobalt hexammine were obtained without using the seeding protocol, and were stabilized without using cobalt hexammine.

Data reduction and phase determination

All data sets were measured at –160° C (Table 1), processed using the programs DENZO and SCALEPACK [39], and scaled using SCALEIT in the CCP4 program suite [40]. Heavy-atom sites were initially located using standard difference Patterson methods. Parameters for the heavy atoms and initial phases were calculated using the program MLPHARE

(CCP4). After solvent flattening (CCP4 program DM), the heavy-atom parameters were refined against the solvent-flattened phases. This cycling between solvent flattening and heavy-atom parameter refinement was repeated with iterative improvements to the RNA mask. The solvent content was increased from 40 % to 62 % during the course of the iterations. An electron-density map calculated using the final solvent-flattened phases (20.0–3.0 Å) was of sufficient quality to obtain the correct register of the RNA chain (see below). Substitution of cobalt hexammine data collected at CHESS for cobalt hexammine data collected from a rotating anode source improved the map only slightly, consistent with the osmium hexammine MAD data providing the bulk of the phase information.

Model building

The initial MAD/SIR map with no solvent flattening showed many regions of the RNA backbone with connected density. RNA helical segments were built into the density without regard to sequence. These segments were used to make an improved mask (over the automatic Wang mask, [41]). Subsequent rounds of model building connected the helical segments and revealed the location of the two molecules in the asymmetric unit. We obtained the register of the sequence by a combination of three factors. The GAAA tetraloop, previously seen in the hammerhead crystal structure [22] and determined by NMR spectroscopy [42], provided the first landmark in the sequence. In addition, the size of the density of each base was used as a binary code (big versus small) to distinguish purines from pyrimidines throughout most of the sequence. Regions with weak or ambiguous base density were interpreted on the basis of the strong density of phosphates relative to the rest of the chain. In the solvent-flattened experimental map, the backbone density is continuous at 1.0 σ except for a small number of breaks in the ribose portion and in the regions discussed below. These breaks in the density were readily resolved by the location of bases extending from the backbone density, the connectivity of adjacent chains, and the packing contacts in the crystal. All model building was performed using the program O [43].

The experimental map shows density for the 5′ ends of each molecule (residues 103–105, [12]), though that for molecule A is weak. There is no density in the map for the last residue on the 3′ end of either molecule. The only other region of poor density is for loop L6b, which is consistent with the fact that many P4–P6 molecules in the lattice are hydrolyzed in this loop (JAD, unpublished data). The backbone for helix P5a is connected in the experimental map, but the base density is poorly defined in some regions (molecule A, nucleotides 129–132 and 193–195; molecule B, nucleotides 130–132, [12]). Helix P5a is near L6b from an adjacent molecule in the lattice, which may partly explain the disorder in this region.

Refinement

Three models of the P4–P6 structure have been refined against corresponding data sets: native, with cobalt hexammine bound or with osmium hexammine bound (Table 1). The initial model described above was refined against the osmium data set (Table 1). Subsequent refinement of this model against a higher-resolution cobalt hexammine data set was carried out using X-PLOR [44] with noncrystallographic symmetry (NCS) restraints applied during positional and tightly-restrained individual temperature-factor refinement. A new RNA parameter set was used [45], with some modifications. We constructed a ribose rotamer library to use in both least-squares refinement [44] and real-space fitting [43]. The ribose rotamers were designed by using ideal bond lengths and angles from the new RNA dictionary and idealized endo-ribose dihedrals [46]. Several of the ribose puckers in the P4–P6 domain are C2′-endo or O4′-exo, mainly in the nonhelical regions of the structure (to be described in detail elsewhere). The current model includes nucleotides 103–260 in each molecule in the asymmetric unit and a total of 28 metal ions and six waters. Note that the numbering scheme is based on the intact group I intron, so that nucleotides 103 and 260 are the second and penultimate nucleotides in the P4–P6 RNA, respectively [12].

This model was subsequently refined against the native and osmium hexammine data sets (Table 1). For the native data, simulated annealing

Table 1

Data sets and refinement statistics.

Parameters	Native*	Cohex†	Oshex (λ1)††
Space groups	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, b, c (Å)	74.8, 128.7, 145.9	74.8, 128.7, 145.9	74.6, 128.1, 145.8
Reflections			
Working set	28 567	34 551	26 510
Test set	1411	1850	1324
Resolution (Å)	16–2.8	18–2.5	20–2.8
Completeness $F > 2\sigma$ (%)			
Data to 3.0 Å / last shell	90/56	91/35	90/39
Redundancy	3.8	4.0	2.4§
R_{sym} (%)	5.6	4.3	4.5
Resolution (Å)	8–2.8	8–2.5	8–2.8
R_{cryst} $F > 2\sigma$ (%)	22.8	24.2	23.0
R_{free} $F > 2\sigma$ (%)	27.2	28.5	27.7
Rms bond length (Å)	0.009	0.010	0.010
Rms bond angle (°)	1.27	1.26	1.21
Rms improper angle (°)	1.29	1.19	1.31
Mean temperature factor (Å^2)	42	44	43
Rms temperature factor (Å^2)	3.6	3.4	4.1

*Native data were measured at Beamline X-4A at the National Synchrotron Light Source, Brookhaven, from a crystal grown and stabilized in the absence of cobalt hexammine (see Materials and Methods section). †Data measured from a cobalt hexammine-containing crystal at CHESS Beamline A-1. ††Data measured at the osmium LIII absorption edge (as determined by a fluorescence scan of the crystal) at Brookhaven Beamline X-4A from a crystal stabilized in osmium hexammine (see [1]). §For anomalous data, |F| was taken as $(|F^+| + |F^-|)/2$. $R_{sym} = \Sigma(|I_{obs} - \langle I \rangle|)/\Sigma \langle I \rangle$; R_{cryst} and $R_{free} = \Sigma \||F_{obs}| - |F_{calc}\||/\Sigma|F_{obs}|$, where R_{free} includes the 5 % of the amplitudes (test set) omitted from the refinement.

(with the cobalt hexammines deleted) was used to remove model bias [47]. In the resulting σ_A-weighted F_o–F_c map [48] calculated from 8–2.8 Å resolution, significant density (8–9 σ above mean peak height) was observed in the P5b and P5c hexammine-binding sites. For refinement against the osmium hexammine data, cobalt hexammines were deleted from the refined (Cohex) model and osmium hexammines were positioned based on their locations in the experimental map. Refinement of the resulting model involved rigid-body minimization followed by simulated annealing. Comparisons between the three (native, cobalt-containing, and osmium-containing) models using all-atom superpositions showed that the hexammine-binding sites were identical within coordinate error (rms deviation = 0.2–0.5 Å, coordinate error 0.4–0.5 Å [49].

Electrostatic potential calculations

The electrostatic potential of the P4–P6 domain was calculated using the non-linear Poisson–Boltzmann equation [15] in the program DelPhi [50]. Parameters for the calculation were chosen to match the crystal stabilization conditions. The dielectric constants (ε) were set to $\varepsilon_{RNA} = 4$, $\varepsilon_{solvent} = 60$ [15]; ionic strength = 0.55 M. Ions bound at the sites of interest were excluded from the calculation.

Accession numbers

Coordinates have been deposited in the Protein Data Bank.

Acknowledgements

We thank Tom Cech, in whose laboratory the P4–P6 crystallographic project was initiated; Henry Taube of Stanford University for the gift of osmium hexammine triflate; Tom Steitz, Axel Brünger, Paul Sigler, Fred Richards, Carl Correll, Adrian Ferré-d'Amaré, Alan Friedman, Virginia Rath, Scott Strobel, Jimin Wang and the staff and members of the Center for Structural Biology at Yale for invaluable helpful advice and encouragement; Virginia Rath and Pfizer Inc. for use of x-ray data collection equipment; Craig Ogata and Wayne Hendrickson for useful discussions and data collection time at beamline X-4A at the National Synchrotron Light Source, Brookhaven; the staff at Cornell High Energy Synchrotron Source (CHESS) beamline A-1; Andy Thompson and the staff at beamline BL-19 at the European Synchrotron Radiation Facility (ESRF), Grenoble; Tom Cech, Adrian Ferré-d'Amaré, Virginia Rath and Scott Strobel for comments on the manuscript. JHC was supported by NIH predoctoral training grant number 5T32GM08283-07; JAD is a Lucille P Markey Scholar in Biomedical Science, a Young Investigator of the Donaghue Medical Research Foundation, a Searle Scholar and a Beckman Young Investigator. This work was supported by grants from the Lucille P Markey Charitable Trust, the Donaghue Medical Research Foundation and the NIH.

References

1. Söll, D., Abelson, J.N. & Schimmel, P.R. (eds) (1980). *Transfer RNA: Structure, Properties and Recognition, Monograph Series 9A*. Cold Spring Harbor Press, NY, USA.
2. Guerrier-Takada, C., Haydock, K., Allen, L. & Altman, S. (1986). Metal ion requirements and other aspects of the reaction catalyzed by M1 RNA, the RNA subunit of ribonuclease P from *Escherichia coli*. *Biochemistry* **25**, 1509–1515.
3. Grosshans, C.A. & Cech, T.R. (1989). Metal ion requirements for sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme. *Biochemistry* **28**, 6888–6894.
4. Latham, J.A. & Cech, T.R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science* **245**, 276–282.
5. Piccirilli, J.A., Vyle, J.S., Caruthers, M.H. & Cech, T.R. (1993). Metal ion catalysis in the *Tetrahymena* ribozyme reaction. *Nature* **361**, 85–88.
6. Pyle, A.M. (1993). Ribozymes: a distinct class of metalloenzymes. *Science* **261**, 709–714.
7. Celander, D.W. & Cech, T.R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science* **251**, 401–407.
8. Murphy, F.L. & Cech, T.R. (1993). An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry* **32**, 5291–5300.
9. Michel, F. & Westhof, E. (1990). Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585–610.
10. Zarrinkar, P.P. & Williamson, J.R. (1994). Kinetic intermediates in RNA folding. *Science* **265**, 918–924.
11. Doudna, J.A. & Cech, T.R. (1995). Self-assembly of a group I intron active site from its component tertiary structural domains. *RNA* **1**, 36–45.
12. Cate, J.H., *et al*., & Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, in press.
13. Cate, J.H., *et al*., & Doudna, J.A. (1996). RNA tertiary structure mediation by adenosine platforms. *Science*, in press.
14. Alden, C.J. & Kim, S.-H. (1979). Solvent-accessible surfaces of nucleic acids. *J. Mol. Biol.* **132**, 411–434.
15. Sharp, K.A., Honig, B. & Harvey, S.C. (1990). Electrical potential of transfer RNAs: Codon-anticodon recognition. *Biochemistry* **29**, 340–346.
16. Weeks, K.M. & Crothers, D.M. (1993). Major groove accessibility of RNA. *Science* **261**, 1574–1577.
17. Gautheret, D., Konings, D. & Gutell, R.R. (1996). G–U base pairing motifs in ribosomal RNA. *RNA* **1**, 807–814.
18. Doudna, J.A., Grosshans, C., Gooding, A. & Kundrot, C.E. (1993). Crystallization of ribozymes and small RNA motifs by a sparse matrix approach. *Proc. Nat. Acad. Sci. USA* **90**, 7829–7833.
19. Lay, P.A., Magnuson, R.H. & Taube, H. (1989). Syntheses of pentaammineosmium (III) and -osmium(II) complexes: Pentaammine(trifluoromethanesulfonato-O)osmium(III). *Inorg. Chem.* **28**, 3001–3007.
20. Kim, S.H., *et al*., & Rich, A. (1973). Three-dimensional structure of yeast phenylalanine transfer RNA: folding of the polynucleotide chain. *Science* **179**, 285–288.
21. Robertus, *et al*., & Klug, A. (1974). Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* **250**, 546–551.
22. Pley, H.W., Flaherty, K.M. & McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature* **372**, 68–74.
23. Jack, A., Ladner, J.E., Rhodes, D., Brown, R.S. & Klug, A. (1977). A crystallographic study of metal-binding to yeast phenylalanine transfer RNA. *J. Mol. Biol.* **111**, 315–328.
24. Westhof, E., Dumas, P. & Moras, D. (1985). Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol.* **184**, 119–145.
25. Flor, P.J., Flanegan, J.B. & Cech, T.R. (1989). A conserved base pair within helix P4 of the *Tetrahymena* ribozyme helps to form the tertiary structure required for self-splicing. *EMBO J.* **8**, 3391–3399.
26. Murphy, F.L. & Cech, T.R. (1994). GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* **236**, 49–63.
27. Hingerty, B.E., Brown, R.S. & Klug, A. (1982). Stabilization of the tertiary structure of yeast phenylalanine tRNA by cobalt hexammine. *Biochem. Biophys. Acta* **697**, 78–82.
28. Nunn, C.M. & Neidle, S. (1996). The high resolution crystal structure of the DNA decamer d(AGGCATGCCT). *J. Mol. Biol.* **256**, 340–351.
29. Gessner, R.V., Quigley, G.J., Wang, A.H.-J., van der Marel, G.A., van Bloom, J.H. & Rich, A. (1985). Structural basis for stabilization of Z-DNA by cobalt hexammine and magnesium cations. *Biochemistry* **24**, 237–240.
30. Ho, P.S., Frederick, C.A., Saal, D., Wang, A.H.-J. & Rich, A. (1987). The interactions of ruthenium hexammine with Z-DNA: Crystal structure of a ruthenium hexammine salt of d(CGCGCG) at 1.2 Å resolution. *J. Biomol. Struct. Dynam.* **4**, 521–534.
31. Hou, Y.M. & Schimmel, P. (1988). A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* **333**, 140–145.
32. Park, S.J., Hou, Y.M. & Schimmel, P. (1989). A single base pair affects binding and catalytic parameters in the molecular recognition of a transfer RNA. *Biochemistry* **28**, 2740–2746.
33. Davies, R.W., Waring, R.B., Ray, J.A., Brown, T.A. & Scazzocchio, C. (1982). Making ends meet: a model for RNA splicing in fungal mitochondria. *Nature* **300**, 719–724.
34. Michel, F. & Dujon, B. (1983). Conservation of RNA secondary structure in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. *EMBO J.* **2**, 33–38.
35. Strobel, S.A. & Cech, T.R. (1996). Exocyclic amine of the conserved G–U pair at the cleavage site of the *Tetrahymena* ribozyme contributes to 5′-splice site selection and transition state stabilization. *Biochemistry* **35**, 1201–1211.
36. Allain, F.H.T. & Varani, G. (1995). Solution structure of the P1 substrate duplex from Group I self-splicing introns. *J. Mol. Biol.* **250**, 333–353.
37. He, L., Kierzek, R., SantaLucia, J.J., Walter, A.E. & Turner, D.H. (1991). Nearest neighbor parameters for G–U mismatches. *Biochemistry* **30**, 11124–11132.

38. Martell, A.E. & Smith, R.M. (1974). *Critical Stability Constants*. Plenum Press, NY, USA.
39. Otwinowski, Z. (1993). Oscillation Data Reduction Program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*. (Sawyer, L., Isaacs, N., & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
40. Collaborative Computational Project No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
41. Wang, B.C. (1985). Resolution of phase ambiguity in macromolecular crystallography. *Meth. Enzymol.* **115**, 90–112.
42. Heus, H.A. & Pardi, A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **253**, 191–194.
43. Jones, T., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
44. Brünger, A.T. (1993). *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
45. Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A.T. & Berman, H.M. (1996). New research for the refinement of nucleic acid containing structures. *Acta Cryst. D* **52**, in press.
46. Levitt, M. & Warshel, A. (1978). Extreme conformational flexibility of the furanose ring in DNA and RNA. *J. Am. Chem. Soc.* **100**, 2607–2613.
47. Hodel, A., Kim, S.-H. & Brünger, A.T. (1992). Model bias in macromolecular crystal structures. *Acta Cryst. A* **48**, 851–859.
48. Read, R.J. (1986). Improved Fourier coefficient for maps using phases from partial structures with errors. *Acta Cryst. A* **42**, 140–149.
49. Luzzati, V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Cryst.* **5**, 802–810.
50. Sharp, K.A. & Nicholls, A. (1989). *DelPhi, A Macromolecular Electrostatics Modelling Package*. Columbia University, New York.
51. Carson, M. (1991). *RIBBONS 2.0 Manual*. University of Alabama at Birmingham, AL 35294, USA.
52. Nicholls, A. (1993). *GRASP: Graphical Representation and Analysis of Surface Properties*. Columbia University, New York.