Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

An efficient method for solving RNA structures: MAD phasing by replacing magnesium with zinc

E. Ennifar, P. Walter and P. Dumas*

UPR 9002 du CNRS, IBMC, 15 Rue R. Descartes, 67084 Strasbourg CEDEX, France

Correspondence e-mail: p.dumas@ibmc.u-strasbg.fr

Received 17 July 2000
Accepted 14 November 2000

The structure of a 46-nucleotide RNA complex has been successfully solved using multi-wavelength anomalous dispersion (MAD) at the zinc *K* edge. Taking advantage of the eight magnesium-binding sites, it has been shown that for five of them magnesium could be replaced by zinc. This resulted in an excellent 2.0 Å MAD electron-density map. Zinc, in common with some other transition metals, is able to replace magnesium in RNA structures, but zinc has the advantage of its *K* edge being ideally located at 1.284 Å. As most RNA molecules contain magnesium-binding sites, it is suggested that this method could be a valuable alternative to the use of bromo derivatives of bases, which is limited to chemically synthesizable and thus rather short RNA sequences.

1. Introduction

Multi-wavelength anomalous dispersion (MAD) is now recognized as a powerful method for solving crystal structures (Hendrickson, 1991). It provides decisive advantages over the classical multiple isomorphous replacement (MIR) approach. Firstly, thanks to cryocooling, one crystal is in most cases sufficient for data collection and, secondly, the resulting electron-density map is of high quality owing to the absence of lack of isomorphism (LOI). There is no longer any need for quite random screening of several heavy-atom compounds which mostly result in a significant LOI. As a consequence, solving a structure can now be performed in a very short time (Walsh *et al.*, 1999).

Nowadays, MAD using selenomethionine-modified proteins has become the most common technique for solving protein crystal structures. Many recent RNA–protein or RNA–ligand complexes have also been solved by using a seleno-substituted protein or intrinsic anomalous scatterer of the ligand (Sussman *et al.*, 2000). For RNA-only structures, large quantities of oligoribonucleotides (up to ∼35 nucleotides) can now be chemically synthesized, allowing introduction of bromo-modified bases for MAD phasing at the Br *K* edge. In favourable cases, bromouridines can also be introduced during T7 transcription by substituting Br-UTP for UTP (Su *et al.*, 1999; Wild *et al.*, 1999). However, for larger RNA structures, the tedious method of soaking crystals with appropriate heavy-atom compounds (*i.e.* with an absorption edge accessible to a tunable synchrotron beamline) remains the only method for MAD phasing.

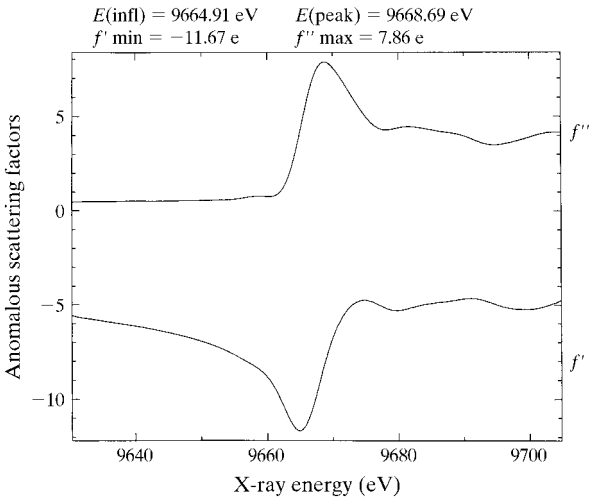

Figure 1
X-ray fluorescence spectrum for the DIS-RNA crystal soaked in zinc chloride (upper curve) and its Kramers–Krönig transform (lower curve). The upper curve has been obtained after numerical smoothing of the experimental points collected on beamline BM30 at the ESRF.

Most of these heavy atoms bind to RNA at magnesium sites, such as $Os(NH_3)_6$ in the P4-P6 ribozyme domain (Cate & Doudna, 1996), Tb^{3+} to the hammerhead ribozyme (Feig *et al.*, 1998), or Ba^{2+} and Pb^{2+} to the leadzyme (Wedekind & McKay, 1999). Although magnesium-binding sites are commonly found along RNA major grooves containing either G-residues, non-Watson–

Crick base pairs or bulged residues, only some of these sites are able to bind heavier atoms with a larger radius and different coordination geometry. In contrast, some transition metals such as Mn^{2+} or Co^{3+}, with a smaller radius, are able to replace all or most magnesium sites (Cate *et al.*, 1997; Ennifar *et al.*, 1999; Murray *et al.*, 1998). However, the energy needed to reach their *K* edge is too low to obtain a strong and stable beam intensity on many synchrotron beamlines (this is particularly true for manganese, with a *K* edge at 1.896 Å). Furthermore, absorption becomes important for large-size crystals at such wavelengths. Zinc does not have these limitations since its *K* edge is ideally located (9659 eV, 1.284 Å; Teplyakov *et al.*, 1998) and furthermore, its coordination geometry and its radius make it suitable for replacing magnesium with minimum risks of crystal alteration. Zinc is therefore an ideal candidate for MAD phasing and has been used successfully already for zinc-finger proteins (*e.g.* Zhang *et al.*, 1995). Here, we show that it is possible to solve a 46-nucleotide RNA complex with this method.

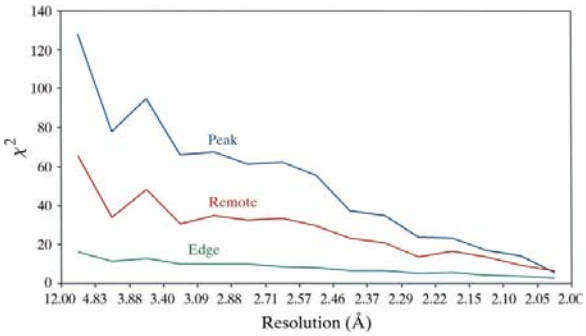

Figure 2
Assessment of experimental anomalous data quality by monitoring the χ^2 values. For each wavelength, χ^2 is calculated as the average of $(F^+ - F^-)^2/[\sigma^2(F^+) + \sigma^2(F^-)]$ in resolution shells. Each shell included ~500 independent acentric reflections.

2. Material and methods

2.1. Sample preparation and crystallization

The RNA molecule used in this study was a chemically synthesized 23-nucleotide fragment that corresponds to the dimerization initiation site (DIS) of HIV-1 genomic RNA (Yusupov *et al.*, 1999). It crystallizes as a dimer, forming a 46-nucleotide duplex containing two G·A mismatches and two looped-out adenine residues (Ennifar *et al.*, 1999). One DIS crystal that was obtained and stabilized in a magnesium-containing solution as described previously (Ennifar *et al.*, 1999; Yusupov *et al.*, 1999) was then extensively washed with a magnesium-free solution containing 100 m*M* $ZnCl_2$, 50 m*M* sodium cacodylate pH 7.0, 300 m*M* KCl, 50% methylpentanediol (MPD) for several days. Direct co-crystallization of the DIS-RNA with zinc instead of magnesium was not attempted.

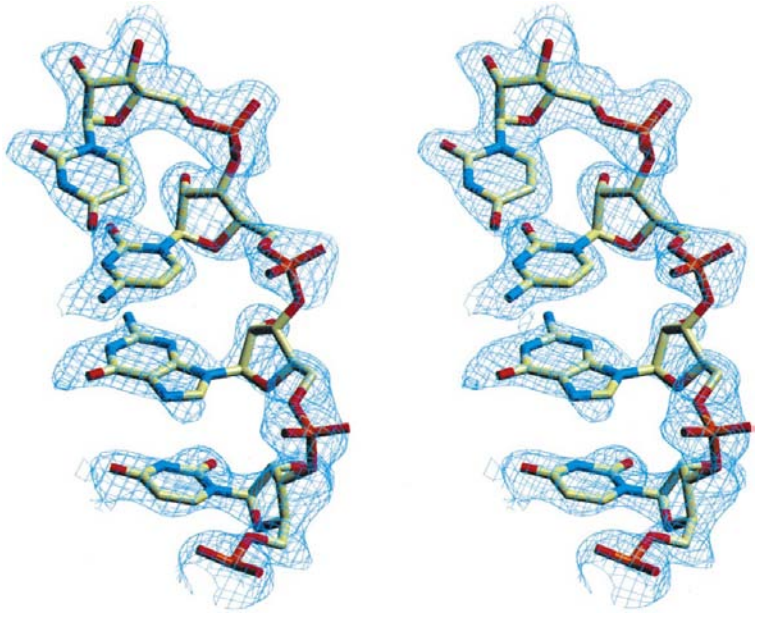

Figure 3
Experimental electron-density map calculated at 2.0 Å resolution. The map is contoured at 1.2 e.s.d. above mean level and shows the high quality of phases obtained from MAD phasing with *SHARP* and solvent-flattening with *SOLOMON*.

2.2. X-ray data collection and processing

The crystal was flash-frozen in liquid ethane prior to data collection at 100 K. Because of 'real-life' practical problems, experimental data for a three-wavelength MAD experiment were collected at two different synchrotron-radiation sources. However, the same crystal could be kept frozen and used in both cases. The data set at the remote wavelength was first collected at beamline BW7A of DESY (Hamburg, Germany) on a MAR-CCD detector, whereas the data sets at the peak and the inflection point wavelengths were collected later at beamline BM30 of the ESRF (Grenoble, France) on a MAR 345 imaging plate (Table 1). Fluorescence measurement showed an interesting anomalous signal, with $f' = -11.67$ e at the inflection point and $f'' = 7.86$ e at the peak (Fig. 1). A significant percentage of reflections at low resolution were overloaded for the two data sets at BM30 (~50% in the 15.0–4.5 Å resolution range). Ideally, a special data collection for low-resolution reflections should have been performed for more accurate measurements. As the aim of this work was not to obtain the best quality result, but rather to show the practical feasibility of the method, the overloaded reflections were simply taken into consideration and not discarded during the final scaling procedure. This resulted in an increased R_{sym}, but led to nearly complete low-resolution shells (Table 1). The anomalous signal quality at the three wavelengths is shown in Fig. 2 by the variation of the χ^2 index with resolution (see legend). This is by far the most reliable statistical index to judge the quality of such an experimental signal. It is recalled that a unit value of a χ^2 means that any differences between experimental data are accounted for by errors. The very large values found, particularly at the 'peak' wavelength, are the sign of a very valuable anomalous signal. Data were processed using the *HKL* package (Otwinowski & Minor, 1996). The space group is $P3_121$, with unit-cell parameters $a = b = 59.2$, $c = 63.7$ Å, and the zinc-soaked crystal was found to be isomorphous with the 'native' crystal (*i.e.* the one stabilized in 100 m*M* $MgCl_2$).

2.3. Localization of zinc sites and MAD phasing

Three sites were readily located with the program *LOCHVAT* (Dumas, 1994*a*,*b*) in the 15–2.7 Å resolution range using anom-

Table 1
Data collection and structure determination.

Space group	$P3_121$		
Unit-cell parameters (Å)	$a = 59.24$, $c = 63.68$		
Wavelength (Å)	1.26000 (Zn remote†, BW7A)	1.28233 (Zn peak, BM30)	1.28281 (Zn edge, BM30)
Resolution range (Å)	12–1.85	15–2.0	15–2.0
Unique reflections	21273	16722	16703
Redundancy	8.9	3.6	3.6
Average I/σ‡	32.3 (5.1)	22.9 (5.6)	25.4 (7.2)
Completeness‡ (%)	98.9 (98.7)	96.0 (62.2)	95.8 (69.5)
R_{sym}‡§ (%)	4.3 (18.8)	5.3 (19.7)	4.6 (17.2)
Phasing power¶			
Centric reflections	—	1.9	1.7
Acentric reflections (iso/ano)	—/5.6	3.7/5.9	3.1/3.5
Refined value for f'/f'' (e)	−3.37††/4.82	−8.86/9.50	−10.11/2.90
R_{cullis}‡‡	—	5.9	6.6
Mean figure of merit before/after solvent flattening	0.72 (0.39)/0.81 (0.58)		

† Taken as reference for structure solution. ‡ Values for last shell are in parentheses. § $R_{sym} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I is the measured intensity of each reflection and $\langle I\rangle$ is the intensity averaged from multiple observations of symmetry-related reflections. ¶ Phasing power = $\langle F_H/\mathrm{LOC}\rangle$, where LOC is the lack of closure. †† Used as reference and fixed during the refinement. ‡‡ $R_{cullis} = \big||F_{PH} \pm F_P| - F_H\big|/|F_{PH} - F_P|$ for centric reflections.

Table 2
Zinc-binding statistics.

Zinc-site number	1	2	3	4	5
Peak height in anomalous difference map†	32.48	23.07	17.16	14.04	11.53
Relative occupancy‡§	1.00	0.57	0.42	0.38	0.15
B factor§ (Å^2)	35.5	30.7	34.1	30.6	17.1

† The anomalous difference map was calculated with coefficients $(F^+ - F^-)\exp[i(\varphi + \pi/2)]$ at the peak wavelength in a 12–2.0 Å resolution range. Peak heights are expressed in σ units. ‡ Occupancy of the first site was arbitrarily set to one. § Values after refinement in *SHARP*.

alous differences at the peak wavelength for the acentric reflections. It is to be noted that the prior knowledge of this structure and of the magnesium-binding sites was never used during this new structure solution. The three zinc sites were then introduced in the *SHARP* program (de La Fortelle & Bricogne, 1997) for initiating MAD phasing. Heavy-atom refinement clearly revealed two additional sites in residual maps for both the peak and remote wavelengths. After a new phasing refinement with five zinc sites, a solvent-flattening procedure was applied with *SOLOMON* (Abrahams & Leslie, 1996) using a solvent content of 50% (Tables 1 and 2). The resulting 2.0 Å resolution map was of excellent quality, as judged by the clearly identifiable sugar pucker in many cases (Fig. 3). The importance of low-resolution completeness appeared clearly during this study, since a first attempt at solving the structure without including overloaded reflections led to scaling and refinement problems with *SHARP* and a non-interpretable electron-density map.

3. Conclusions

Because of the weakness of the signal, a MAD experiment usually requires great accuracy in data measurements. However, in the present case, despite some unfavourable data-collection conditions (data collection made on different beamlines and half the low-resolution reflections being overloaded for two of the wavelengths), MAD phasing with zinc was successful and produced an excellent quality electron-density map. Unlike cobalt and manganese *K*-edge wavelengths, those used in this experiment are accessible to all synchrotron beamlines and represent a good choice taking into account the flux of synchrotron radiation, the scattering efficiency and the absorption effects (Teplyakov *et al*., 1998). Furthermore, the zinc anomalous signal is fully suitable for MAD phasing as it is comparable to the anomalous signal of selenium in its oxidized form.

Evidently, the major limitation of the proposed method is the need for zinc binding. In the present case, only five of the eight magnesium sites were successfully replaced. The two strongest sites (1 and 2 in Table 2) bind to G·A mismatches, replacing partially dehydrated magnesium cations, whereas sites 3, 4 and 5 bind to G residues into the major groove of the duplex, replacing hexahydrated magnesium cations. Manganese would be a better candidate for replacing magnesium (Cate *et al.*, 1997; Ennifar *et al.*, 1999; Scott, 1995), but is much less convenient for MAD phasing as noted previously. Of course, successful soaking with zinc greatly depends on crystallization conditions. If the latter include a high concentration of monovalent cations, as with ammonium sulfate used in the molar range, the proposed method might well be of little use. Thus, *a priori*, precipitating agents such as methylpentanediol and polyethylene glycol are the most favourable. Fortunately, the latter reagents are quite often used efficiently for crystallizing RNA molecules.

We are grateful to the staff of beamline BW7A at DESY (Hamburg) and to R. Kahn at beamline BM30 (ESRF, Grenoble). This work was supported by grants from the Agence Nationale de Recherche sur le SIDA. We thank B. and C. Ehresmann for their constant support.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Cate, J. H. & Doudna, J. A. (1996). *Structure*, **4**, 1221–1229.
Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997). *Nature Struct. Biol.* **4**, 553–512.
Dumas, P. (1994*a*). *Acta Cryst.* A**50**, 526–537.
Dumas, P. (1994*b*). *Acta Cryst.* A**50**, 537–546.
Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C. & Dumas, P. (1999). *Structure*, **7**, 1439–1449.
Feig, A. L., Scott, W. G. & Uhlenbeck, O. C. (1998). *Science*, **279**, 81–84.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Murray, J. B., Terwey, D. P., Maloney, L., Karpeisky, A., Usman, N., Beigelman, L. & Scott, W. G. (1998). *Cell*, **92**, 665–673.
Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.
Scott, W. G. (1995). *Cell*, **81**, 991–1002.
Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. (1999). *Nature Struct. Biol.* **6**, 285–292.
Sussman, D., Nix, J. & Wilson, C. (2000). *Nature Struct. Biol.* **7**, 53–57.
Teplyakov, A., Oliva, G. & Polikarpov, I. (1998). *Acta Cryst.* D**54**, 610–614.
Walsh, M. A., Dementieva, I., Evans, G., Sanishvili, R. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1168–1173.
Wedekind, J. E. & McKay, D. B. (1999). *Nature Struct. Biol.* **6**, 261–268.
Wild, K., Weichenrieder, O., Leonard, G. A. & Cusack, S. (1999). *Structure*, **7**, 1345–1352.
Yusupov, M., Walter, P., Marquet, R., Ehresmann, C., Ehresmann, B. & Dumas, P. (1999). *Acta Cryst.* D**55**, 281–284.
Zhang, G., Kazanietz, M. G., Blumberg, P. M. & Hurley, J. H. (1995). *Cell*, **81**, 917–924.