research papers

Novel complex MAD phasing and RNase H structural insights using selenium oligonucleotides

Rob Abdur, Oksana O. Gerlits, Jianhua Gan, Jiansheng Jiang, Jozef Salon, Andrey Y. Kovalevsky, Alexander A. Chumanevich, Irene T. Weber and Zhen Huang*

Department of Chemistry and Department of Biology, Georgia State University, Atlanta, GA 30303, USA

Correspondence e-mail: huang@gsu.edu

The crystal structures of protein–nucleic acid complexes are commonly determined using selenium-derivatized proteins *via* MAD or SAD phasing. Here, the first protein–nucleic acid complex structure determined using selenium-derivatized nucleic acids is reported. The RNase H–RNA/DNA complex is used as an example to demonstrate the proof of principle. The high-resolution crystal structure indicates that this selenium replacement results in a local subtle unwinding of the RNA/DNA substrate duplex, thereby shifting the RNA scissile phosphate closer to the transition state of the enzyme-catalyzed reaction. It was also observed that the scissile phosphate forms a hydrogen bond to the water nucleophile and helps to position the water molecule in the structure. Consistently, it was discovered that the substitution of a single O atom by a Se atom in a guide DNA sequence can largely accelerate RNase H catalysis. These structural and catalytic studies shed new light on the guide-dependent RNA cleavage.

1. Introduction

Three-dimensional structure determination and study of proteins, nucleic acids and protein–nucleic acid complexes have emerged as fields of enormous importance for bio-molecular mechanism research and drug discovery (Liu *et al.*, 2012; Nakanishi *et al.*, 2012). X-ray crystallography is one of the most direct and powerful tools for structure determination of these macromolecules and complexes. X-ray crystal structures of proteins and protein–nucleic acid complexes are commonly determined using selenium-derivatized (*i.e.* seleno-methionyl) proteins *via* multi-wavelength or single-wavelength anomalous diffraction (MAD or SAD) phasing (Elkayam *et al.*, 2012; Ferré-D'Amaré *et al.*, 1998; Hendrickson, 1991, 2000; Schirle & MacRae, 2012; Yang *et al.*, 1990). Inspired by the selenium-derivatized protein approach, our research group has pioneered and developed selenium-derivatized nucleic acids (SeNAs) for X-ray crystal structure study of nucleic acids (Carrasco *et al.*, 2001; Du *et al.*, 2002; Lin *et al.*, 2011; Sheng *et al.*, 2012, 2013; Sun *et al.*, 2012; Zhang *et al.*, 2012, 2013). Similar to the selenium substitution of the S atoms in proteins, the O atoms in nucleic acids are replaced with Se atoms, since selenium and oxygen are also in the same elemental family.

In addition to structural studies, Se atoms can be atom-specifically incorporated into nucleic acids for functional investigations as well *via* selective oxygen replacement. This Se atom-specific derivatization is also called Se atom-specific mutagenesis (SAM). In this report, SAM is used to provide new insights into both the structure and catalysis of RNase H. RNase H is a sequence-nonspecific endonuclease that selectively digests the RNA portion of the RNA/DNA duplex

(Stein & Hausen, 1969). RNase H is involved in many important biological processes (Arnold *et al.*, 1992; Green *et al.*, 1975; Hippenmeyer & Grandgenett, 1985; Wintersberger, 1990), including the removal of the RNA primers from Okazaki fragments in replication. It can also silence gene expression directly *via* the antisense mechanism (Veal *et al.*, 1998; Vickers *et al.*, 2003; Walder & Walder, 1988; Wu *et al.*, 2004). Catalytic studies on RNase H have been extensively performed (Nowotny *et al.*, 2005; Nowotny & Yang, 2006; Pallan & Egli, 2008; Reijns *et al.*, 2011; Wu *et al.*, 2004), especially investigating the roles of the metal cations near the scissile phosphate. Recently, the crystal structures of *Bacillus halodurans* RNase H and human RNase H1 complexed with RNA/DNA duplex substrates were determined at 1.5 and 2.2 Å resolution, respectively (Nowotny *et al.*, 2005; Nowotny & Yang, 2006) and revealed insights into cation-assisted RNA hydrolysis (Nowotny *et al.*, 2007; Yang *et al.*, 2006).

Here, we report the structure of the RNase H–RNA/Se-DNA complex, which was determined *via* seleno-nucleic acid MAD phasing. The RNase H–RNA/DNA complex was used as a model system to demonstrate the proof of principle of crystal structure determination using selenium-derivatized nucleic acids instead of the protein counterpart. Furthermore, it is easier and more convenient to synthesize and purify selenium-derivatized nucleic acids (SeNAs; Lin *et al.*, 2011; Sheng & Huang, 2010; Sheng *et al.*, 2012; Sun *et al.*, 2012, 2013; Zhang *et al.*, 2012) than to express and purify selenium-derivatized proteins. We also revealed that RNA cleavage by RNase H is facilitated by a local subtle unwinding of the duplex, thereby shifting the RNA scissile phosphate closer to the enzyme active site. Moreover, we observed experimentally that in the presence of RNase H the scissile phosphate formed a hydrogen bond to the water nucleophile and helped to position the water molecule in the structure. This research has successfully demonstrated the potential usefulness of the SeNA strategy in both structural and functional studies of protein–nucleic acid complexes.

2. Materials and methods

2.1. Oligonucleotide and protein preparation and crystallization

The native and modified DNA or RNA oligonucleotides synthesized in the laboratory (Salon *et al.*, 2008) were purified by HPLC twice with and without a DMTr protecting group in order to guarantee high purity. Protein expressions (Nowotny *et al.*, 2005; Nowotny & Yang, 2006) were carried out in *E. coli* BL21 (DE3) pLys cells (purchased from Invitrogen). Transformation was accomplished by the heat-shock method. The DNA portion of the DNA/RNA duplex (5′-ATGTCGp-3′/5′-UCGACA-3′; one-base overhang at both ends) was derivatized. Prior to co-crystallization with RNase H, the purified Se-DNA (5′-AT-^{Se}G-TC-^{Se}Gp-3′) and RNA (5′-UCGACA-3′) were annealed in a 1:1 molar ratio by heating the mixture to 90°C for 1 min and then allowing it to cool slowly to 25°C. The resulting Se-DNA/RNA duplex was mixed with the protein

Table 1
Data-collection, phasing and structure-refinement statistics for Se-DNA/RNA–RNase H (PDB entry 3twh).

Values in parentheses are for the highest resolution shell.

	Crystal 1	Crystal 2	
Wavelength (Å)	0.9795 [peak]	0.9797 [peak]	0.9794 [inflection]
Resolution range (Å)	50.0–1.80	50.0–1.80	50.0–1.80
	(1.86–1.80)	(1.86–1.80)	(1.86–1.80)
Space group	*C*2	*C*2	*C*2
Unit-cell parameters			
a (Å)	80.7	80.6	80.6
b (Å)	37.8	37.8	37.8
c (Å)	61.9	61.9	61.9
β (°)	96.6	96.7	96.6
Unique reflections	17321 (1704)	17220 (1702)	16858 (1641)
Completeness (%)	99.3 (98.8)	99.0 (98.7)	96.9 (95.1)
R_{merge} (%)	7.2 (26.1)	6.2 (27.0)	4.8 (20.0)
$\langle I/\sigma(I) \rangle$	15.3 (10.5)	15.4 (10.4)	14.1 (9.8)
Multiplicity	7.2 (6.2)	7.3 (6.5)	3.1 (2.9)
R factor (%)	17.5 (27.3)		
R_{free} (%)	20.8 (34.0)		
No. of amino-acid residues	134		
No. of nucleic acid residues	12		
No. of Mg^{2+} ions	2		
No. of water molecules	138		
Overall B factor (Å^2)	24.0		
R.m.s.d., bond lengths (Å)	0.026		
R.m.s.d., bond angles (°)	2.0		
Ramachandran plot, residues in (%)			
Most favored region	99.2		
Allowed region	0.8		
Disallowed region	0		

(final concentration of 8 mg ml^{-1}) in a 1:1 molar ratio in the presence of 5 m*M* MgCl$_2$. Co-crystallization of the Se-DNA/RNA duplex with RNase H was achieved by screening with the The Classics Suite kit (Qiagen). By using the sitting-drop vapor-diffusion method at 25°C, crystals were readily obtained from condition No. 96 of the crystallization screen [buffer, 0.1 *M* MES pH 6.5; precipitant, 12%(*w/v*) PEG 20 000]. Details are included in the Supporting Information.[1]

2.2. MAD data collection, phasing and structure determination

Diffraction data were collected from crystals of the Se-DNA/RNA–RNase H complex on beamlines X25 and X29 of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. A number of crystals were scanned in order to find those with strong anomalous scattering at the *K* absorption edge of selenium. 25% glycerol was used as a cryoprotectant and X-ray data were collected under a liquid-nitrogen stream at 99 K. The selected wavelengths for the selenium MAD data are listed in Table 1. Each crystal was exposed for 15 s per image with 1° rotation, and a total of 180 images were obtained for each data set. Two crystals were used to collect the MAD/SAD data sets. All data were processed using *HKL*-2000 and *DENZO/SCALEPACK* (Otwinowski & Minor, 1997). The structure was solved by the MAD method using *SOLVE/RESOLVE* (Terwilliger, 2003; Wang *et al.*, 2004). The resulting model was refined using

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: CB5033).

REFMAC5 (Murshudov et al., 2011) within CCP4 (Winn et al., 2011). The DNA/RNA duplex was modeled into the structure using Coot (Emsley et al., 2010). Metal ions and water molecules were added either automatically or manually using Coot. Data-collection, phasing and structure-refinement statistics are given in Table 1.

2.3. Catalytic hydrolysis by RNase H

The DNA with its complementary RNA was allowed to form a duplex by heating and subsequent cooling. Each RNase H hydrolysis reaction (volume 5 µl) contained DNA template (150 nM final concentration; DNA-N, DNA-S or DNA-Se) and ^{32}P-labeled RNA substrate (mixture of cold and hot RNAs; 150 nM final concentration). To each hydrolysis reaction, WT or TR RNase H enzyme (10 nM final concentration) and reaction buffer (final conditions 75 mM KCl, 50 mM Tris–HCl pH 7.8, 3 mM MgCl$_2$, 1 mM diborane) were added. The reactions were incubated at 37°C for 30 min unless mentioned otherwise. Details of these experiments are included in the Supporting Information.

3. Results

3.1. Crystallization, Se anomalous phasing and structure determination of the novel Se-DNA/RNA–RNase H complex

We synthesized selenium-modified DNAs to investigate the DNA/RNA–RNase H complex. The duplexes of the native and selenium-modified oligonucleotides (Salon et al., 2008) were complexed with RNase H (an inactive mutant with a single D132N mutation). The RNase H–RNA/Se-DNA complex (5′-UCGACA-3′/5′-AT-^{Se}G-TC-^{Se}Gp-3′) was crystallized in a buffer consisting of 0.1 M MES pH 6.5 and precipitant [12%(w/v) PEG 20 000]. The crystals were also grown using the sitting-drop vapor-diffusion method. The crystals appeared within a week and reached their maximum size within a month. Two crystals were used to collect the MAD/SAD data sets. The figure of merit of the individual SAD phasing data was relatively low and a good electron-density map for the model could not be produced. We used one SAD data set as a reference for MAD phasing of the other diffraction data set. The overall figure of merit (FOM) of the initial phases was 0.630, which produced an interpretable electron-density map. Anomalous diffraction data from the protein-complex crystals containing seleno-oligonucleotides can be collected at the Se K edge. After the SAD or MAD data set had been collected from a single crystal, the Se atoms can be located by direct methods and the phase can be determined similarly to the well established protein strategy (Elkayam et al., 2012; Ferré-D'Amaré et al., 1998; Hendrickson, 1991, 2000; Schirle & MacRae, 2012; Yang et al., 1990).

The selenium-complex structure was finally determined (PDB entry 3twh; 1.80 Å resolution; Fig. 1) via the selenium anomalous signal and MAD phasing (Table 1). Since an inactive mutant enzyme was used, the RNA substrate was not cleaved during the crystallization, which was confirmed by MS analysis of the crystals. Moreover, we have determined the selenium-complex structure by both MAD phasing and molecular replacement. The selenium complex structure determined by the MAD phasing technique is identical to the same complex structure determined by molecular replacement. Compared with the corresponding native structure (PDB entry 2g8u; 2.70 Å resolution; Nowotny & Yang, 2006) with the same sequences, the selenium-containing structure has a higher resolution (Table 1).

3.2. Subtle conformation change of the RNA/DNA substrate duplex by the Se atom on the nucleobase

By taking advantage of Se atom-specific replacement in the nucleobases (Hassan et al., 2010; Lin et al., 2011; Salon et al., 2007, 2008; Sheng et al., 2012; Sun et al., 2012; Zhang et al.,
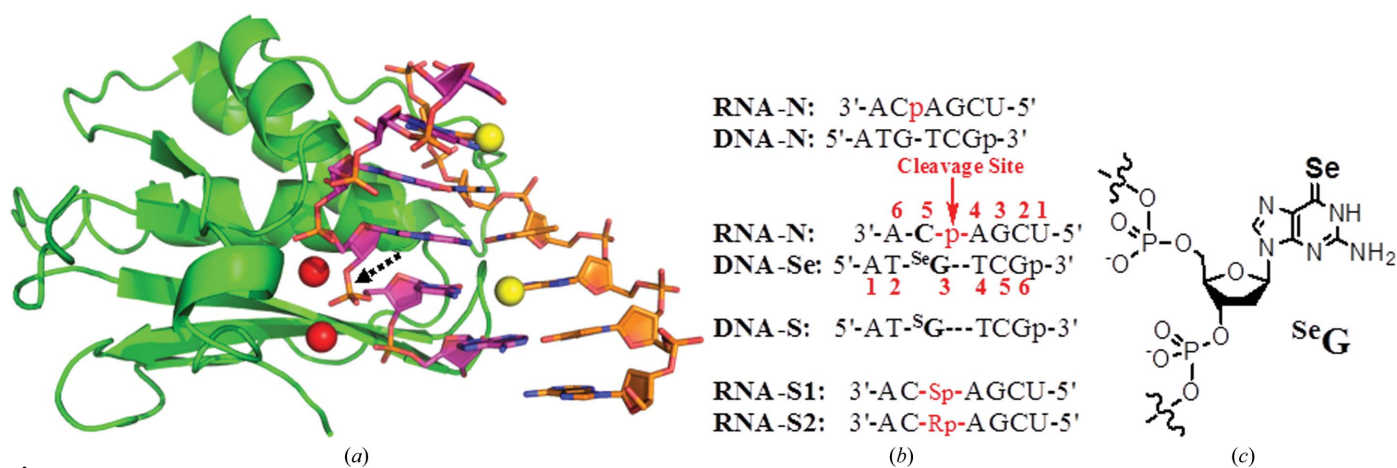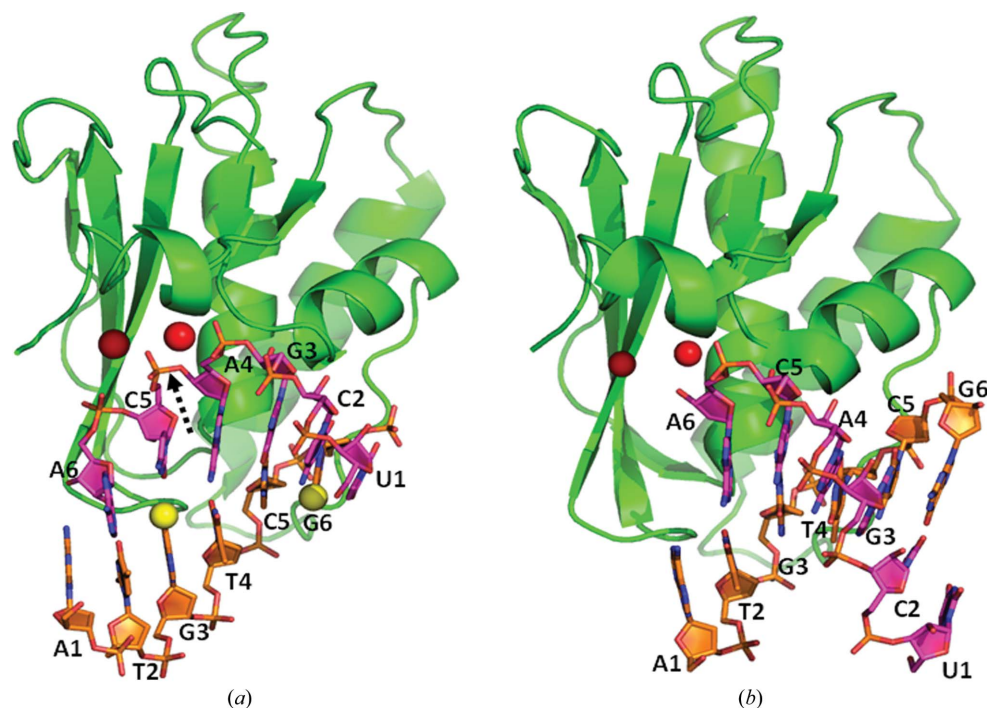

Figure 1
Crystal structure of RNase H complexed with the selenium-modified DNA and RNA duplex. (a) The overall structure of the Se-DNA/RNA–RNase H ternary complex determined at 1.80 Å resolution (D132N mutant; PDB entry 3twh). The protein is shown in green and the RNA and DNA strands in pink and orange, respectively. The red and yellow spheres are Mg^{2+} ions and Se atoms, respectively. The cleavage site is indicated by the arrow. (b) Sequences of the native and modified DNAs and RNAs. ^{Se}G and ^{S}G represent 6-Se-G and 6-S-G, respectively. (c) The Se-DNA derivatized with 6-Se-deoxyguanosine.

2012), the impact of the RNA/DNA duplex and conformational change on catalytic hydrolysis was investigated. The crystal structure (Fig. 1) of *B. halodurans* RNase H complexed with RNA/DNA modified with selenium at the 6-position of guanosine was determined at 1.80 Å resolution. RNase H is a sequence-non-specific enzyme and the binding of RNA at different locations is possible. Except RNase H binding to the

Se-modified duplex by a two-nucleotide shift with respect to the native duplex, the Se-derivatized complex structure is virtually identical to the corresponding native complex structure with the same RNA/DNA sequence (Fig. 2; Nowotny *et al.*, 2005, 2008). The active site of RNase H is positioned at the phosphate linkage between A4 and C5 of the RNA molecule in the Se-modified duplex, where C5 pairs with ^{Se}G3 of the DNA sequence (Figs. 1a and 1b). The C5 5'-phosphate appears to be the scissile phosphate. In contrast, the active site of RNase H complexed with the corresponding native RNA/DNA duplex (PDB entries 2g8v and 2g8u; Nowotny *et al.*, 2005; Nowotny & Yang, 2006) is positioned at the interface between A6 and U1 in two different RNA molecules in the pseudofiber formed *via* the stacking of the multiple RNA/DNA duplexes.

The crystal structure indicates that the Se-modified short RNA/DNA duplex (3'-ACAGCU-5'/5'-AT-^{Se}G-TC-^{Se}Gp-3') results in the formation of the substrate–RNase H complex (Figs. 1a and 2a) instead of the product–RNase H complex originally designed and formed by the corresponding native RNA/DNA substrate duplex (PDB entries 2g8v and 2g8u). The structure of the substrate–RNase H complex (PDB entry 1zbi; 1.85 Å resolution) was determined using a longer RNA/DNA duplex

Figure 2
Structures of RNase H (D132N mutant) complexed with DNA/RNA duplexes (with the same sequences). (a) The Se-DNA/RNA–RNase H structure (PDB entry 3twh; 1.80 Å resolution; this work). The cleavage site in the selenium-modified structure (PDB entry 3twh) is between the two Mg^{2+} ions and is indicated by an arrow. (b) The native DNA/RNA–RNase H structure (PDB entry 2g8u; 2.70 Å resolution; Nowotny & Yang, 2006). The sequences of the DNA (5'-ATGTCG-3') and the RNA (5'-UCGACA-3') are the same in all structures. In both structures proteins are shown in green and RNA and DNA strands in pink and orange, respectively. The red and yellow spheres are Mg^{2+} ions and Se atoms, respectively.
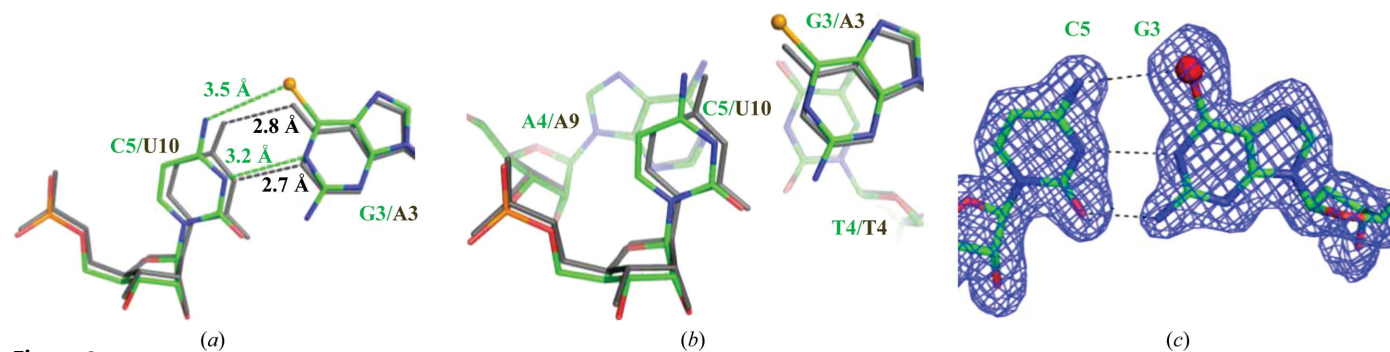
Figure 3
The subtle conformation change (or local subtle unwinding of the duplex) of the Se-DNA/RNA duplex complexed with RNase H. Residues are shown as sticks in atomic color scheme for D132N mutant RNase H–Se-DNA/RNA (PDB entry 3twh; 1.80 Å resolution; 5'-AT-^{Se}G-TC-^{Se}Gp-3'/5'-UCGA-p-CA-3', where 'p' is the cleavage site; C, green; N, blue; O, red; P and Se, orange). Residues are shown as gray sticks for D132N mutant RNase H–DNA/RNA (PDB entry 1zbi; 1.85 Å resolution; 5'-G**AATCA**GGTGTC-3'/5'-GACACC**UGA-p-UU**C-3', where 'p' is the cleavage site). (a) Local subtle unwinding of the duplex shown *via* comparison of the Se-dG3/rC5 base pair (in PDB entry 3twh) with dA3/rU10 (in PDB entry 1zbi) at the active site of RNase H. (b) Structural comparison showing the local backbone shift. The structures are shown in the same orientation as in (a). The rA/T base pairs (labeled as A4/A9 pairing with T4/T4) next to the scissile phosphate in both structures (PDB entries 1zbi and 3twh) are virtually identical. (c) Electron-density map of the C/^{Se}G base pair (rC5/dG3); the $2F_o - F_c$ map is contoured at the 1.5σ level.
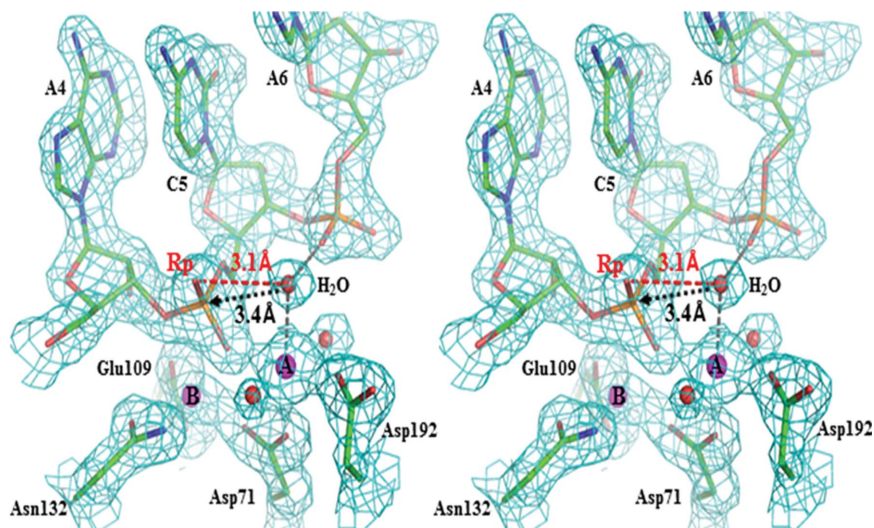
substrate (Nowotny *et al.*, 2005; Nowotny & Yang, 2006), such as a 12 bp RNA/DNA duplex (5′-G**AATCA**GGTGTC-3′/ 3′-**CUUAGU**CCACAG-5′), to which two RNase H molecules were bound. Despite the shift 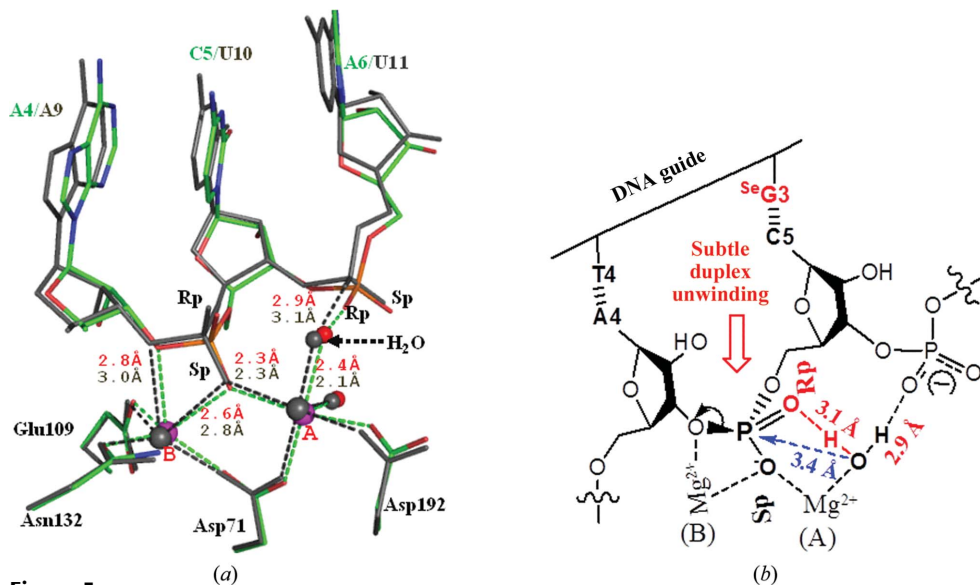of the RNase H binding site, the Se-DNA/RNA duplex retains a very similar overall structure to the corresponding native structure. However, a subtle conformation change (by 0.5–0.7 Å) of the Se-dG3/rC5 base pair was observed when comparing the selenium-modified substrate–enzyme complex (PDB entry 3twh) with the native substrate–RNase H complex (Fig. 3a; PDB entry 1zbi). The rA/T base pairs next to the scissile phosphate in these two structures are virtually identical (Fig. 3b).

Figure 4
Stereoview of the structure of the cleavage site of the Se-DNA/RNA duplex in complex with RNase H. Residues are shown as sticks in atomic color scheme for D132N mutant RNase H–Se-DNA/RNA (PDB entry 3twh; 1.80 Å resolution; 5′-AT-^{Se}G-TC-^{Se}Gp-3′/5′-UCGA-p-CA-3′, where 'p' is the cleavage site; C, green; N, blue; O, red; P, orange). Water and Mg^{2+} ions are shown as red and pink spheres, respectively. Stereoview showing the detailed interaction observed at the cleavage site of D132N mutant RNase H–Se-DNA/RNA outlined with an electron-density map ($2F_o − F_c$ map contoured at the 1.5σ level). One hydrogen bond (3.10 Å in length) was observed between the nucleophilic water molecule (labeled H$_2$O) and the R_p O atom (labeled Rp) of the scissile phosphate.

3.3. The substrate shifts its scissile phosphate to the RNase H active site

The small conformational change of the Se atom in the nucleobase causes a local subtle unwinding of the duplex. Consistently, the crystal structure (Figs. 4 and 5) indicated that this subtle unwinding leads to a small shift of the scissile phosphate closer to the RNase H active site. The 5′-phosphate of rC5 (the scissile phosphate) in the selenium-modified complex is shifted towards the active center of RNase H by approximately 0.3 Å owing to the local subtle unwinding of the duplex. The Se-DNA/RNA–RNase H complex (PDB entry 3twh) was determined at high resolution (1.80 Å) and the estimated overall coordinate error is approximately 0.1 Å. This complex (PDB entry 3twh) is considered as a substrate–enzyme complex (the selenium complex) related to another substrate–enzyme complex (the native complex, containing a D132N mutation) determined at high resolution (PDB entry 1zbi; 1.85 Å resolution; Nowotny *et al.*, 2005; Nowotny & Yang, 2006). Therefore, comparing these two structures allowed observation of the backbone shift in the sele-

Figure 5
Shift of the scissile phosphate of the Se-DNA/RNA substrate towards the RNase H active site. Residues are shown as sticks in atomic color scheme for D132N mutant RNase H–Se-DNA/RNA (PDB entry 3twh; 1.80 Å resolution; 5′-AT-^{Se}G-TC-^{Se}Gp-3′/5′-UCGA-p-CA-3′, where 'p' is the cleavage site; C, green; N, blue; O, red; P and Se, orange). Residues are shown as sticks in gray for D132N mutant RNase H–DNA/RNA (PDB entry 1zbi; 1.85 Å resolution; 5′-G**AATCA**GGTGTC-3′/5′-GACACC**UGA**-p-**UU**C-3′, where 'p' is the cleavage site). Water and Mg^{2+} ions are shown as spheres (red and pink in D132N mutant RNase H–Se-DNA/RNA and gray in D132N mutant RNase H–DNA/RNA). (a) Superposition of the cleavage-site interactions highlighted with their distances (green dashed lines and red letters for PDB entry 3twh; gray dashed lines and black letters for PDB entry 1zbi). (b) The proposed guide-dependent RNA cleavage facilitated by the scissile phosphate and the local subtle unwinding. The R_p O atom of the scissile phosphate forms a hydrogen bond to the nucleophilic water molecule. The hydrogen-bond values of the selenium-modified complex structure are in red.

nium complex with respect to the native complex. A small change in a substrate structure (such as a fraction of 1 Å) is sufficient for catalysis and can contribute to acceleration of the reaction, which has been observed in the case of isocitrate dehydrogenase catalysis (Mesecar *et al.*, 1997).
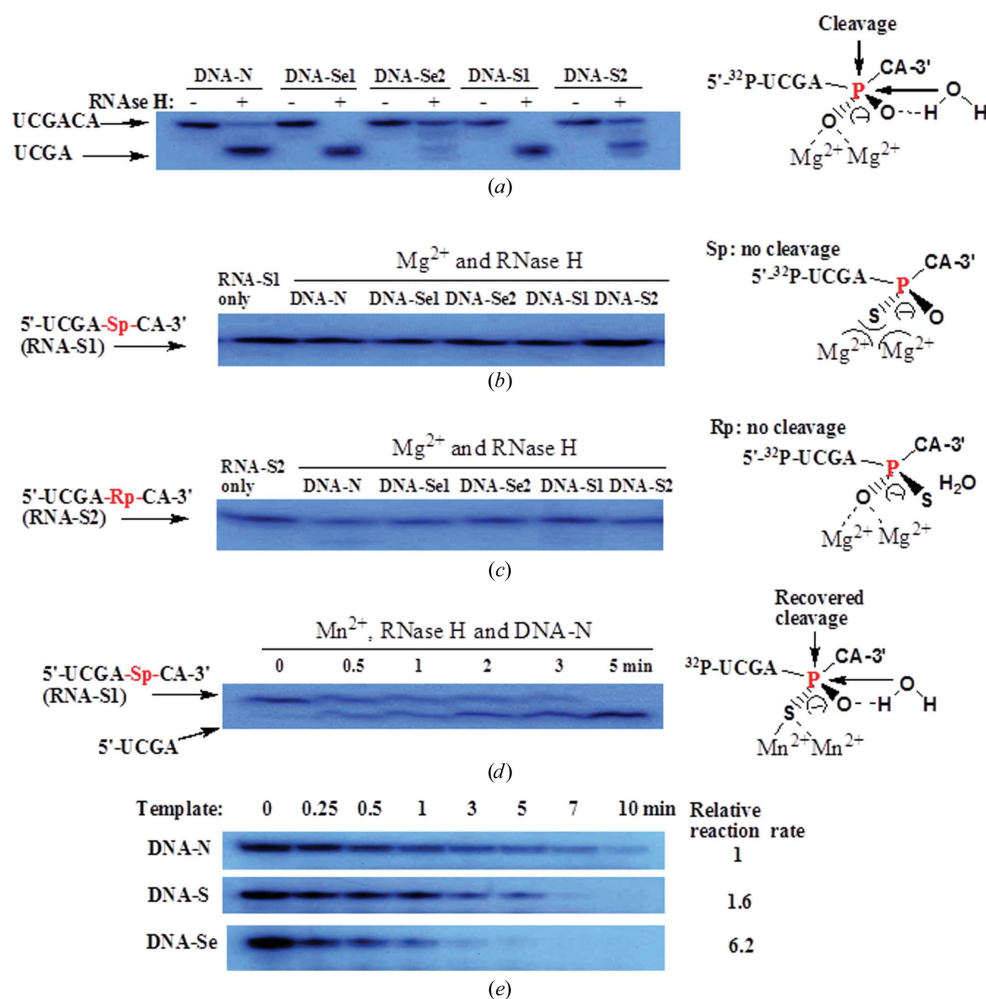
3.4. The scissile phosphate forms a hydrogen bond to the nucleophilic water

The distance from Mg^{2+}-A to the pro-S_p O atom of the scissile phosphate remains the same (2.3 Å) in both the native and the selenium-modified structures (Fig. 5a). In contrast, the distances from Mg^{2+}-B to the pro-S_p and 3′ O atoms of the scissile phosphate (2.6 and 2.8 Å, respectively) are shortened by 0.2 Å in the selenium-modified structure compared with those in the native complex (Fig. 5a). The shortened distances

between Mg^{2+}-B and these two O atoms suggest stronger interactions, which is consistent with the catalyzed RNA cleavage. The changes in the distances are also consistent with the B factors of Mg^{2+}-B, which are higher than the B factors of Mg^{2+}-A in both the native and the selenium-modified structures. This suggests that Mg^{2+}-B is more dynamic than Mg^{2+}-A in both structure and catalysis.

Similarly, the hydrogen bond (2.9 Å in length) between the nucleophilic water and the pro-R_p O atom of the 3′-phosphate next to the scissile phosphate is stronger and shorter (by 0.2 Å) compared with that in the native system. In contrast, the distance between this nucleophilic water molecule and Mg^{2+}-A (2.4 Å) is longer by 0.3 Å. Weakening the nucleophile bonding to Mg^{2+}-A and strengthening the hydrogen bond to the deprotonating pro-R_p O atom of the 3′-phosphate next to the scissile phosphate are consistent with 'in-line' attack of the nucleophilic water on the scissile phosphorus center (Figs. 4 and 5). The nucleophilic water molecule is in close proximity (3.4 Å) to the scissile phosphorus center. Interestingly, this nucleophilic water molecule is also close to the pro-R_p O atom of the scissile phosphate (3.10 Å) and they form a genuine hydrogen bond (Fig. 4). This hydrogen bond may help to position the nucleophilic water molecule in the structure.

3.5. RNA/DNA minimal substrates for RNase H catalytic cleavage

Unlike the native substrate–enzyme complex using a 12 bp RNA/DNA duplex (PDB entry 1zbi), in which RNase H can bind more than one site, the short Se-DNA/RNA substrate (with five base pairs; Fig. 1) allows only one RNase H molecule to bind owing to its minimal size. Interestingly, this duplex substrate with a short size forces RNase H to interact with the RNA/DNA duplex in a 'pseudo-specific' manner. On the basis of the crystal structure, we predict that RNase H cleaves the RNA substrate at the 5′-phosphate of C5 and only offers two fragments, *i.e.* the 5′-product (5′-UCGA-3′) and the 3′-product (5′-pCA-3′). This was confirmed by the RNA substrate digestion. RNase H generated one 5′-product (Fig. 6a) and the

Figure 6
'Pseudo-specific' cleavage of RNA substrates by RNase H. RNase H cleavage of short native and modified RNA substrates in the presence of the short native and modified DNA guides (see sequences in Fig. 1). In the reactions, the concentration of RNA substrate was equal to that of the native, single-Se, double-Se, single-S or double-S modified DNAs. Truncated RNase H (59–196) was also used in these experiments and similar results were obtained. (a) Cleavage of the native RNA by RNase H in the presence of the native and modified DNAs. (b) Cleavage inhibition by the S_p sulfur modification of the RNA substrate (RNA-S1, S_p diastereomer). (c) Cleavage inhibition by the R_p sulfur modification of the RNA substrate (RNA-S2, R_p diastereomer). (d) Recovery of S_p-RNA cleavage by replacing Mg^{2+} with Mn^{2+} cations. (e) RNA hydrolysis by RNase H in the presence of DNA guides, including native DNA-N, DNA-S and DNA-Se. The relative values of the reaction rates using these templates are DNA-N, 1; DNA-S, 1.6; DNA-Se, 6.2.
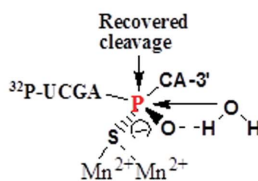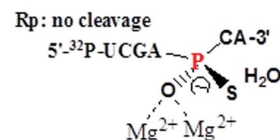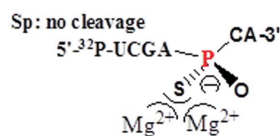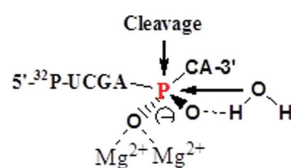
radioactive intensity of the cleaved product was equal to the intensity of the starting material.

On the basis of the crystal structure (Figs. 4 and 5), we also predicted that a sulfur modification at the C5 5′-phosphate (such as sulfur replacement of a non-bridging O atom) can significantly inhibit the RNA cleavage. As expected, sulfur modifications of the C5 5′-phosphate (replacing the pro-S_p or pro-R_p O atom) indeed prevented RNA substrate cleavage (Figs. 6b and 6c). Sulfur substitution of the pro-S_p O atom of the scissile phosphate disrupts the interaction with both Mg^{2+} ions (Fig. 6b), while sulfur replacement of the pro-R_p O atom of the scissile phosphate disrupts the interaction between the scissile phosphate and the nucleophilic water molecule. Therefore, both sulfur modifications inhibit RNA cleavage (Figs. 6b and 6c).

Excitingly, our compensation experiment by switching from Mg^{2+} to Mn^{2+} cation could effectively recover the cleavage of the S_p-modified RNA (Fig. 6d) owing to re-establishment of the interactions between the cations (Mn^{2+}) and the S_p S atom (Figs. 5 and 6d). These results confirm the conclusion from the crystal structure: the C5 5′-phosphate is indeed the cleavage site. This minimal substrate opens up the opportunity for atom-specific substitution and kinetic studies to address the impact of the duplex conformation on guide-dependent RNA cleavage. The experimental results on the RNA sulfur modifications have demonstrated that both the pro-S_p and the pro-R_p O atoms of the scissile phosphate play critical roles in RNA cleavage, which is consistent with computation study and prediction (De Vivo *et al.*, 2008; Elsässer & Fels, 2010; Rosta *et al.*, 2011).

3.6. The selenium-modified DNA accelerates RNA substrate hydrolysis catalyzed by RNase H

On the basis of our structural study, we hypothesized that owing to the facilitated subtle conformation change (or to the local subtle unwinding of the duplex), the Se-DNA guide can better assist RNase H to catalyze RNA hydrolysis than the corresponding native DNA guide. Thus, we carried out RNA hydrolysis by RNase H in the presence of native and sulfur- and selenium-nucleobase-modified DNAs (Figs. 1b and 6e). The relative reaction rates in the presence of the native (DNA-N) and modified DNAs were measured. Consistent with our hypothesis, the Se-DNA guide (DNA-Se) is much more efficient (6.2-fold faster) in RNA cleavage than the corresponding native DNA, while the S-DNA (DNA-S) is 1.6-fold more efficient than the native DNA. Our experimental results suggest that the subtle conformation change *via* the single-atom nucleobase modification can push the scissile phosphate towards the enzyme active site, thereby significantly accelerating RNA cleavage.

4. Discussion

The main structural differences between the selenium-modified and native complexes (PDB entries 3twh and 2g8u, respectively) were observed at the termini of the DNA/RNA

duplexes, owing to the flexibility of the overhung ends. Despite these differences, the native and modified complexes retain very similar structures. Alignment of the selenium-modified structure and the native DNA/RNA structure (PDB entry 1zbi; 12-mer DNA with a different sequence; Fig. 3) indicates that the selenium modification does not cause a significant perturbation. Another major difference between the native and selenium complexes is in the binding interface between the RNase H and the RNA/DNA duplex (Fig. 2). The native RNA/DNA–RNase H structure (D132N mutant; PDB entry 2g8u; 2.70 Å resolution; Nowotny & Yang, 2006) represents a mimic of the product–enzyme complex, where RNase H binds to the junction of two RNA/DNA duplexes. In contrast, RNase H (with the same D132N mutant; PDB entry 3twh; this work) recognizes the internal position of the RNA/Se-DNA duplex, which has the same sequence as in the native complex structure (PDB entry 2g8u). RNase H binds to the RNA/Se-DNA duplex in an enzyme–substrate binding fashion, in which the two Mg^{2+} ions directly interact with the scissile phosphate and are positioned individually at each site of the scissile symmetry plane of the phosphate. Again, these differences between the native and selenium-modified structures may be attributed to the subtle duplex unwinding (by 0.5–0.7 Å) caused by the Se modification. The ^{Se}G (G3) pairs with the rC5 nucleotide, the 5′-phosphate of which is the cleavage site. As a consequence of the substrate-duplex unwinding (Fig. 5b), the scissile phosphate may shift closer to the transition state of the catalytic hydrolysis.

Furthermore, we discovered in our structure that the pro-R_p O atom of the scissile phosphate forms a hydrogen bond (3.10 Å) to the water nucleophile and helps to position the water molecule in the structure. This nucleophilic water may be also activated by Mg^{2+}-A (Figs. 4 and 5) and the 3′-phosphate of the 3′-adjacent nucleotide of the scissile phosphate. This activated water molecule is in the in-line attacking orientation (Fig. 4) and is within striking distance (3.4 Å) of the scissile phosphorus center, which is consistent with the increased rate of hydrolysis of RNA (6.2-fold faster) by RNase H in the presence of the Se-DNA guide (Fig. 6e). The faster catalytic reaction observed on using the selenium-modified guide is consistent with our crystal structure, in which the scissile phosphate is shifted closer to the active center of RNase H. Our experimental results suggest that the local subtle unwinding *via* the single-atom nucleobase modification can shift the scissile phosphate towards the enzyme active site, thereby significantly accelerating the RNA cleavage.

In summary, we have reported a protein–nucleic acid complex structure determined using selenium-derivatized nucleic acids instead of the protein counterparts. We also found from the crystal structure that the substrate-duplex conformation can play a significant role in guide-dependent RNA cleavage. Furthermore, we have observed that in the presence of RNase H the RNA scissile phosphate can form a hydrogen bond to the nucleophilic water molecule and position it in the structure for potential catalysis. Our structural and catalytic investigations shed new light on RNase cleavage. Our unique Se atom-specific mutagenesis (SAM) of nucleic

acids has opened a new avenue for crystal structure determination as well as mechanistic studies of protein–nucleic acid complexes.

References

Arnold, E., Jacobo-Molina, A., Nanni, R. G., Williams, R. L., Lu, X., Ding, J., Clark, A. D. Jr, Zhang, A., Ferris, A. L., Clark, P., Hizi, A. & Hughes, S. H. (1992). *Nature (London)*, **357**, 85–89.

Carrasco, N., Ginsburg, D., Du, Q. & Huang, Z. (2001). *Nucleosides Nucleotides Nucleic Acids*, **20**, 1723–1734.

De Vivo, M., Dal Peraro, M. & Klein, M. L. (2008). *J. Am. Chem. Soc.* **130**, 10955–10962.

Du, Q., Carrasco, N., Teplova, M., Wilds, C. J., Egli, M. & Huang, Z. (2002). *J. Am. Chem. Soc.* **124**, 24–25.

Elkayam, E., Kuhn, C. D., Tocilj, A., Haase, A. D., Greene, E. M., Hannon, G. J. & Joshua-Tor, L. (2012). *Cell*, **150**, 100–110.

Elsässer, B. & Fels, G. (2010). *Phys. Chem. Chem. Phys.* **12**, 11081–11088.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Ferré-D'Amaré, A. R., Zhou, K. H. & Doudna, J. A. (1998). *Nature (London)*, **395**, 567–574.

Goedken, E. R. & Marqusee, S. (1999). *Protein Eng.* **12**, 975–980.

González, A., Pédelacq, J.-D., Solà, M., Gomis-Rüth, F. X., Coll, M., Samama, J.-P. & Benini, S. (1999). *Acta Cryst.* D**55**, 1449–1458.

Green, M., Grandgenett, D., Gerard, G., Rho, H. M., Loni, M. C., Robins, M., Salzberg, S., Shanmugam, G., Bhaduri, S. & Vecchio, G. (1975). *Cold Spring Harb. Symp. Quant. Biol.* **39**, 975–985.

Hassan, A. E., Sheng, J., Zhang, W. & Huang, Z. (2010). *J. Am. Chem. Soc.* **132**, 2120–2121.

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Hendrickson, W. A. (2000). *Trends Biochem. Sci.* **25**, 637–643.

Hippenmeyer, P. J. & Grandgenett, D. P. (1985). *J. Biol. Chem.* **260**, 8250–8256.

Lin, L., Sheng, J. & Huang, Z. (2011). *Chem. Soc. Rev.* **40**, 4591–4602.

Liu, Q., Dahmane, T., Zhang, Z., Assur, Z., Brasch, J., Shapiro, L., Mancia, F. & Hendrickson, W. A. (2012). *Science*, **336**, 1033–1037.

Mesecar, A. D., Stoddard, B. L. & Koshland, D. E. Jr (1997). *Science*, **277**, 202–206.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. (2012). *Nature (London)*, **486**, 368–374.

Novogrodsky, A., Tal, M., Traub, A. & Hurwitz, J. (1966). *J. Biol. Chem.* **241**, 2933–2943.

Nowotny, M., Cerritelli, S. M., Ghirlando, R., Gaidamakov, S. A., Crouch, R. J. & Yang, W. (2008). *EMBO J.* **27**, 1172–1181.

Nowotny, M., Gaidamakov, S. A., Crouch, R. J. & Yang, W. (2005). *Cell*, **121**, 1005–1016.

Nowotny, M., Gaidamakov, S. A., Ghirlando, R., Cerritelli, S. M., Crouch, R. J. & Yang, W. (2007). *Mol. Cell*, **28**, 264–276.

Nowotny, M. & Yang, W. (2006). *EMBO J.* **25**, 1924–1933.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). *Protein Sci.* **4**, 2411–2423.

Pallan, P. S. & Egli, M. (2008). *Cell Cycle*, **7**, 2562–2569.

Reijns, M. A., Bubeck, D., Gibson, L. C., Graham, S. C., Baillie, G. S., Jones, E. Y. & Jackson, A. P. (2011). *J. Biol. Chem.* **286**, 10530–10539.

Richardson, C. C. (1965). *Proc. Natl Acad. Sci. USA*, **54**, 158–165.

Rosta, E., Nowotny, M., Yang, W. & Hummer, G. (2011). *J. Am. Chem. Soc.* **133**, 8934–8941.

Salon, J., Jiang, J., Sheng, J., Gerlits, O. O. & Huang, Z. (2008). *Nucleic Acids Res.* **36**, 7009–7018.

Salon, J., Sheng, J., Jiang, J., Chen, G., Caton-Williams, J. & Huang, Z. (2007). *J. Am. Chem. Soc.* **129**, 4862–4863.

Schirle, N. T. & MacRae, I. J. (2012). *Science*, **336**, 1037–1040.

Sheng, J., Gan, J., Soars, A. S., Salon, J. & Huang, Z. (2013). *Nucleic Acids Res.* **41**, 1723–1734.

Sheng, J. & Huang, Z. (2010). *Chem. Biodivers.* **7**, 753–785.

Sheng, J., Zhang, W., Hassan, A. E., Gan, J., Soares, A. S., Geng, S., Ren, Y. & Huang, Z. (2012). *Nucleic Acids Res.* **40**, 8111–8118.

Stein, H. & Hausen, P. (1969). *Science*, **166**, 393–395.

Sun, H., Jiang, S., Caton-Williams, J., Liu, H. & Huang, Z. (2013). *RNA*, **19**, 1309–1314.

Sun, H., Sheng, J., Hassan, A. E., Jiang, S., Gan, J. & Huang, Z. (2012). *Nucleic Acids Res.* **40**, 5171–5179.

Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N., Kuhara, S. & Horikoshi, K. (2000). *Nucleic Acids Res.* **28**, 4317–4331.

Terwilliger, T. C. (2003). *Methods Enzymol.* **374**, 22–37.

Veal, G. J., Agrawal, S. & Byrn, R. A. (1998). *Nucleic Acids Res.* **26**, 5670–5675.

Vickers, T. A., Koo, S., Bennett, C. F., Crooke, S. T., Dean, N. M. & Baker, B. F. (2003). *J. Biol. Chem.* **278**, 7108–7118.

Walder, R. Y. & Walder, J. A. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 5011–5015.

Wang, J. W., Chen, J. R., Gu, Y. X., Zheng, C. D., Jiang, F., Fan, H. F., Terwilliger, T. C. & Hao, Q. (2004). *Acta Cryst.* D**60**, 1244–1253.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Wintersberger, U. (1990). *Pharmacol. Ther.* **48**, 259–280.

Wu, H., Lima, W. F., Zhang, H., Fan, A., Sun, H. & Crooke, S. T. (2004). *J. Biol. Chem.* **279**, 17181–17189.

Yang, W., Hendrickson, W. A., Crouch, R. J. & Satow, Y. (1990). *Science*, **249**, 1398–1405.

Yang, W., Lee, J. Y. & Nowotny, M. (2006). *Mol. Cell*, **22**, 5–13.

Zhang, W., Hassan, A. E. & Huang, Z. (2013). *Sci. China Chem.* **56**, 273–278.

Zhang, W., Sheng, J., Hassan, A. E. & Huang, Z. (2012). *Chem. Asian J.* **7**, 476–479.