COMMUNICATION

Origins of the enhanced affinity of RNA–protein interactions triggered by RNA phosphorodithioate backbone modification†

Xianbin Yang,*[a] N. Dinuka Abeydeera,[a] Feng-Wu Liu[b] and Martin Egli [iD] *[c]

The well-characterized interaction between the MS2 coat protein and its cognate RNA hairpin was used to evaluate changes in affinity as a result of phosphorodithioate (PS2) replacing phosphate by biolayer interferometry (BLI). A structure-based analysis of the data provides insights into the origins of the enhanced affinity of RNA–protein interactions triggered by the PS2 moiety.

Oligoribonucleotides (RNAs) such as RNA aptamers,[1–3] siRNAs,[4–6] and miRNAs[7,8] show tremendous potential as therapeutics against viral infections, cancer, genetic disorders, and neurological diseases.[9,10] Beyond their therapeutic potential, aptamers are of high value as tools for biological research, such as target validation,[11–13] and as biosensors in diagnostics.[14] One of the limiting factors in the success of RNA-based therapeutics is the typically rapid degradation of unmodified RNAs in serum and within cells. This problem has been greatly diminished by modifications that render RNAs resistant to the action of cellular nucleases.[15] One of the most commonly employed modifications is the replacement of a non-bridging oxygen with sulfur in phosphate linkages to form phosphoromonothioate (PS)-modified RNA.[16] However, PS-modified RNAs afford only limited protection against hydrolysis by nucleases[16] and the PS modification results in a mixture of diastereomeric RNAs that raises the potential of variable biophysical properties. Stereo-controlled synthesis of P-chiral PS-RNAs represents one possible solution to this problem,[17,18] but another lies in the

synthesis of modifications that are achiral at phosphorus such as phosphorodithioate (PS2)-modified RNAs.[19]

PS2-RNA is a very attractive RNA analog because it closely mimics natural RNA. It has been shown that dinucleoside phosphorodithioates (PS2-dimers) have high nuclease resistance.[20] In addition, a hammerhead ribozyme with a single PS2 substitution at the cleavage site maintained activity and cleavage resulted in the expected product.[21] Moreover, PS2-dimers have shown great resistance to alkaline degradation when compared to natural RNA derivatives.[20] Further, we have demonstrated that the greatly improved gene silencing activities *in vitro* and *in vivo*[19,22] as a result of introducing two PS2 modifications at the 3′-end of sense strand siRNAs were a consequence of the higher affinity of PS2-RNA for Ago2 protein, presumably caused by a hydrophobic effect.[23] Recent studies have shown that combined 2′-OMe-PS2 or 2′-F-PS2 substitution experiments with *in vitro* selected RNA aptamers[24,25] led to either a reduction or increase in binding affinity.[26] A destabilizing effect by 2′-modification could be due to altered sterics (2′-OMe *vs.* 2′-OH), loss of H-bonding (2′-OMe *vs.* 2′-OH), or arise as a consequence of modifying a residue that adopts a C2′-*endo* pucker, *i.e.* shifting the conformational equilibrium to C3′-*endo* compared to the native ribose (as seen with the 2′-F analog).

The question of how individual PS2 modifications will alter the binding between bacteriophage MS2 coat protein and a 19mer stem-loop RNA (Fig. 1A) provided the starting point for the present study. The MS2 system is well suited for three reasons. (1) The X-ray crystal structure of the complex reveals protein-phosphate contacts of a very diverse nature within a small RNA molecule.[27] (2) Experiments with PS modification have been completed in this system,[28] such that these data can be compared to the effects of PS2 modification obtained here. (3) The MS2 model system entails a relatively short RNA that is ideal for a backbone walk to generate all 18 stem-loop constructs that feature a single PS2 moiety. Substitution of two non-bridging oxygen atoms by sulfur is fairly conservative in terms of the changes in the nucleotide geometry, as revealed by our recent PS2-RNA X-ray crystal structures.[23] However, differences

[a] *AM Biotechnologies, LLC, 12521 Gulf Freeway, Houston, TX 77034, USA.*
  E-mail: xianbin.yang@am-biotech.com, xianbin@hotmail.com;
  Fax: +1-832-479-0294; Tel: +1-832-379-2175

[b] *School of Pharmaceutical Sciences, Zhengzhou University, Science Avenue 100,*
  Zhengzhou 450001, Henan, China

[c] *Department of Biochemistry, Vanderbilt University, School of Medicine, Nashville,*
  TN 37232, USA. E-mail: martin.egli@vanderbilt.edu; Fax: +1-615-343-0704;
  Tel: +1-615-343-8070

† Electronic supplementary information (ESI) available: General materials, RNA and modified RNA synthesis methods, modified RNA sequence information (ST-1), affinity information (ST-2 and SF-1), structure of the MS2 coat protein–RNA complex (SF-2), and overall view of interactions between MS2 protein and RNA hairpin (SF-3). See DOI: 10.1039/c7cc05722a
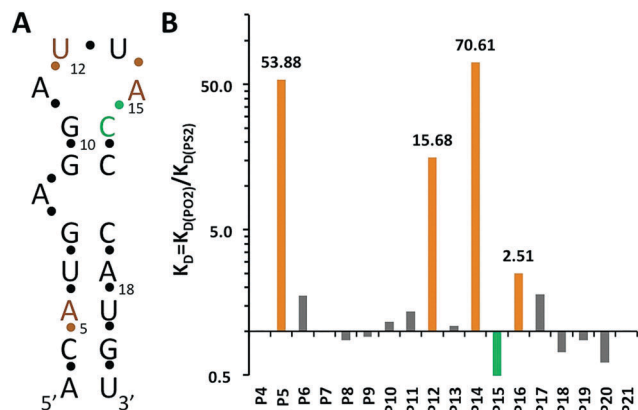
Fig. 1 (A) Secondary structure of the MS2 RNA hairpin. The residue numbering corresponds to that in crystal structures of coat protein:RNA complexes.[27] (B) Affinity changes (K_D ratios in logarithmic scale) measured for RNA hairpins with single PS2 modifications at the indicated phosphates.

between PO2 and PS2 in terms of polarizability, hydrophobicity and other factors are expected to significantly change the contribution of individual amino acid-phosphate contacts to the overall binding affinity upon PS2 substitution in a site-specific fashion.

To evaluate the effect of a PS2 substitution at a given position on the binding affinity, a 5′-biotinylated 19-nucleotide hairpin RNA (Fig. 1A)[27] and PS2 variants with one phosphate per sequence substituted by a PS2 moiety were synthesized via standard solid phase phosphoramidite chemistry (ESI,† ST-1). Kinetic characterization of each PS2 variant by BLI allowed affinity ranking of modified hairpins with respect to the native RNA. Evaluation consisted of using a serial dilution of MS2 protein screened against individual hairpin RNA variants that were immobilized onto streptavidin coated BLI sensors.[26,29] Binding and dissociation rates of native hairpin RNA were determined in parallel to the variants. ESI,† Fig. S1 (SF-1) shows the characterization of native RNA and all of its PS2 modified variants. ST-2 (ESI†) shows the calculated K_D values and relative K_D ratio for each variant relative to the native hairpin RNA. As is evident from ST-2 (ESI†), the effects of single PS2 modifications on binding are variable. Of the 18 PS2-modified RNAs tested, three displayed significantly increased binding affinity (at least 15-fold compared to the corresponding control RNA). PS2 substitutions at the remaining sites either did not affect binding or led to slightly increased/decreased binding affinities (less than 3-fold) (Fig. 1B).

The MS2 system was used previously to assess the consequences of single PS substitutions in the hairpin RNA for the tightness of the RNA-coat protein interaction.[28] The effect of a single PS substitution on the overall protein binding affinity is relatively small (2- to 5-fold). In contrast, the affinity increases of hairpin variants with P5, P12, or P14 replaced by PS2 were significantly larger (54-, 16- and 71-fold, respectively) (Fig. 1B). Unlike in a PS moiety where the negative charge resides on the sulfur, either sulfur atom in PS2 can be neutral or charged, thus rendering the latter more hydrophobic. The more dramatic increases in binding affinity for certain PO2 → PS2 substitutions are consistent with our recent results based on PS2-modified

aptamers targeting VEGF or thrombin.[26] What are the underlying reasons for affinity increases upon PS2 substitution? The fact that the P–S bond is longer than the P–O bond (by ca. 0.5 Å; 1.98 vs. 1.50 Å, respectively[30,31]) cannot be the most important factor in this regard. Instead, the increased hydrophobicity and polarizability of sulfur relative to oxygen (both PO2 and PS2 are negatively charged), potentially in combination with local flexing of the backbone to allow for an optimal fit between PS2 moiety and protein surface,[26] may be key contributors to 50- to 100-fold or even more drastic gains in affinity seen for certain PS2 variants. It is crucial that there is room to accommodate the larger PS2 moiety and an already tightly bound PO2 is unlikely to be conducive to affinity gains.

MS2 coat protein forms an icosahedral capsid (homo-180mer) that encloses RNA hairpins (SF-2, ESI†). Individual hairpins are spread across an antiparallel β-sheet such that each RNA interacts with residues from two adjacent MS2 protein molecules (SF-3, ESI†). In order to gain insight into the altered regio-specific interactions as a result of PS2 modification that give rise to affinity increases, we inspected crystal structures of MS2 coat-protein RNA complexes [PDB ID codes 1zdh, 1zdi].[27] Both complexes contain native RNA, without PS2 modification.

PS2 modification at the P14 phosphate group that is located between loop residues U13 and A14 exerts the most favorable effect on K_D among all tested modification sites. Compared to the native RNA hairpin, the PS2-14 modification results in 71-fold improvement of K_D, a gain that is primarily based on the increased association rate (ST-2, ESI†). Inspection of the crystal structure shows that C8-H of A14 and the Nζ groups of K43 (salt bridge) and K61 are in the vicinity of the pro-R_p sulfur (Fig. 2A). The Cγ2 ring carbon of Y85 is somewhat farther away (5.8 Å). By contrast, based on the conformation observed in the crystal structure, the pro-S_p sulfur is not in close contact with protein residues or RNA atoms. However, PS-modification studies demonstrated that both R_p- and S_p-PS modification affect the affinity between RNA and MS2 coat protein.[28] A sizable shift of this PS2 moiety along with a rotation could generate favorable electrostatic and hydrophobic interactions with the two lysines (including lysine methylene groups as seen in a PS2-DNA:protein complex[32]) and C8-H of A14 and the Y45 ring, similar to PS2 contacts with C8-H of G and phenylalanine, respectively, in our crystal structure of the 2′-F-PS2-RNA:thrombin complex.[26]

PS2-5 near the base of the RNA stem lies in vicinity of the R49, S51 and K57 side chains (Fig. 2B). The pro-R_p sulfur is positioned somewhat closer to the two basic residues than the pro-S_p one, but the latter has the potential to move closer to methylene groups from S51 and K57. S_p-PS modification appears to have a small beneficial effect on affinity,[28] but is clearly no match for the 54-fold improvement in K_D as a consequence of PS2 modification. The analysis here based on the structure of the complex suggests that the PS2 pro-S_p sulfur may be more important in terms of improving hydrophobic interactions. Similar to the situation in the 2′-F-PS2-RNA:thrombin interaction,[26] hydrophobic contributions are enhanced by the negative charge of the PS2 moiety that result in electrostatically favorable interactions, i.e. PS2 pro-R_p sulfur with R49/K57.
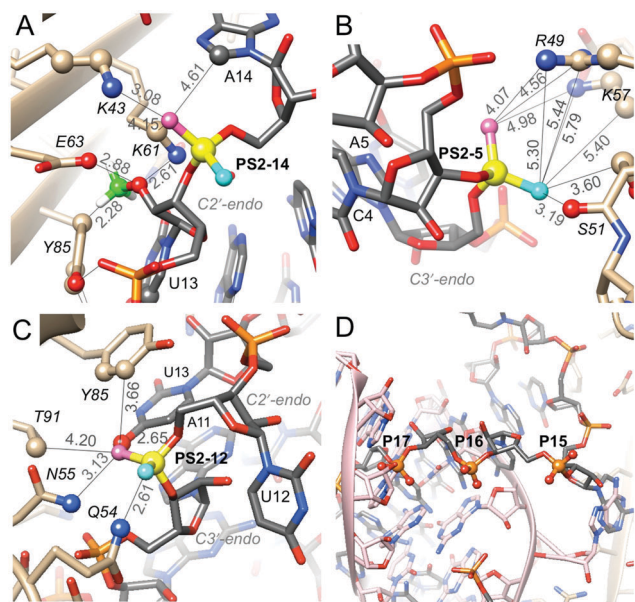
Fig. 2 Environments in MS2 protein complexes of RNA phosphates for which PS2 modification results in significant gains of affinity. (A) P14. (B) P5. (C) P12. PS2 phosphorus atoms are highlighted as filled circles in yellow along with selected side chains. Rose- and cyan-colored spheres indicate the pro-R_p and -S_p sulfurs, respectively, and the calculated position of the 2′-OMe carbon (panel A) is green. Distances in Å are based on a P–S distance of 1.9 Å. (D) P15–17 are directed away from the protein. The panel depicts hairpins in two different orientations (gray and pink carbon atoms) bound to an MS2 dimer (right-hand border) positioned on a crystallographic dyad.

Fig. 3 Interactions between phosphates P8, P10 and P11 and MS2 coat protein residues. Distances are indicated by thin lines and are in Å.

The pro-R_p and pro-S_p moieties of PS2-12 feature very different environments (Fig. 2C). The nucleobase of U13 stacks onto the Y85 side chain (supposedly seeding the RNA:MS2 interaction[27]) and the pro-R_p sulfur is situated in close vicinity (van der Waals contacts) of O4 and C5 of U13 as well as Cδ2 and Cε2 (slightly longer distances) of tyrosine. By contrast the pro-S_p sulfur is pointing away from protein and RNA and is not engaged in any interaction. This view is confirmed by the results of the PS-interference study that showed that the R_p-PS modification improves the affinity whereas S_p-PS modification plays a negligible role.[28] U13 continues the stack between G10 and A11, but U12 points away and does not engage in stacking. A slight movement of the PS2–12 moiety could lead to a favorable edge-on interaction with Y85 similar to an edge-on interaction between PS2 and phenylalanine seen in the 2′-F-PS2-RNA:thrombin complex.[26] The distances in the model for the pro-R_p sulfur are 3.94 Å and 3.66 Å (to Cδ2 and Cε2, respectively).

The three phosphates at positions P8, P10 and P11 are engaged in relatively tight interactions that are dominated by electrostatics (P8 and particularly P10, Fig. 3). Both OP1 and OP2 of P8 form H-bonds to K61 and in the case of P10, one of the oxygens (pro-S_p) forms two H-bonds to R49 and the other is H-bonded to K57. All three interactions by phosphate 10 are rather tight (distances < 3.0 Å) and there are no nearby residues that could be used to improve hydrophobic contacts to a PS2 moiety. P11 is wedged between S51, S52 and Q54 (pro-S_p), whereby the distances are rather tight and this observation
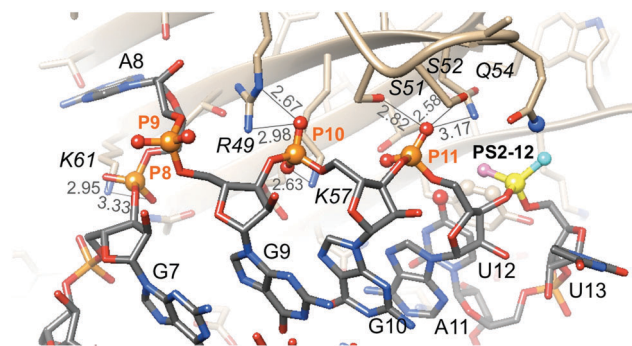
and the lack of hydrophobic moieties in the vicinity render this phosphate also a suboptimal environment for a PS2 substitution. Indeed, inspection of the affinity data for PS2 variants at these sites (ST-2, ESI†) demonstrates that the changes are negligible and similar in magnitude to PS2 substitution at P9, a moiety that lacks interactions with protein (Fig. 3).

Conversely, P15, P16 and P17 phosphates that map to a stretch of three Cs in the RNA hairpin are not engaged in direct interactions with MS2 coat protein in the crystal (Fig. 2D), and yet they lower (PS2-15, 2-fold, ST-2, ESI†) and increase affinity (2.5-fold, PS2-16, and 2-fold, PS2-17). In the crystal structure, symmetrical MS2 dimers feature two bound RNAs of different orientation and partial occupancy (Fig. 2D). However, it is unclear if such a situation exists in solution, where hairpins are supposedly bound to discrete protein dimers, and how PS2-15, -16 and -17 could indirectly affect affinity.

The structure-based analysis of the results of the PS2 walk along the backbone of the RNA hairpin that binds to MS2 coat protein using crystallographic models of complexes provides helpful insights into the observed changes in K_D. Taking into account not just the static models of MS2:RNA complexes in crystal structures but considering potential local conformational changes, significant improvements in K_D as a result of PS2 modification at three backbone sites can be rationalized by potentially favorable hydrophobic and electrostatic environments of individual PS2 moieties. Hydrophobic interactions might include edge-on contacts between a PS2 sulfur and an aromatic side chain (Fig. 2C),[26] contacts between PS2 sulfur and methylene groups of lysine (Fig. 2A and B; seen in the crystal structure of the PS2-DNA:homeodomain complex[32]), and contacts between PS2 sulfur and methylene and methyl groups of serine and threonine, respectively (Fig. 2). Such interactions would be supplemented by electrostatically favorable contacts between PS2 and lysine and arginine head groups because the PS2 group still carries a negative charge. The higher association rates for the PS2-14 and PS2-5 hairpins (ST-2, ESI†) might imply that modification there facilitates conformational selection required for complex formation.

Taking into account dynamic aspects to help explain the observed changes in K_D as a result of the PS2 backbone walk is justified in the case of the RNA hairpin interacting with MS2

coat protein. This is because close contacts between RNA and protein are limited to two regions: residues A3, C4 and A5 from one strand in the lower stem portion of the hairpin and A11, U12, U13 and A14 in the loop region (SF-2, ESI†). The nature of these contacts render it entirely possible that PS2 modification could result in more or less subtle conformational changes in the RNA that affect the interactions with the protein (we don't expect the protein to undergo any conformational changes, consistent with our findings in the structures of PS2-RNA:thrombin complexes[26]). A structural analysis of a 2′-F-PS2-modified RNA in complex with thrombin demonstrated that the RNA can locally undergo conformational changes (a shift of an individual residue of up to 3 Å and rotation of the PS2 moiety and a neighboring phosphate by more than 90° relative to the orientations of the native phosphates).[26] This RNA-induced fit resulted in enhanced hydrophobic contacts – while maintaining favorable Coulombic interactions with lysine and arginine – and calculations provided evidence that the higher polarizability of sulfur compared to oxygen is a contributing factor.[26]

Combined 2′-F-PS2[26] or 2′-OMe-PS2[22] modification of an aptamer[26] or siRNA,[22] respectively, can boost the affinity for protein targets beyond the improvements with PS2 modification alone. However, in cases where the 2′-OH moiety is engaged in a H-bond with protein, 2′-O-methylation can potentially also destroy the affinity gained by the 3′-PS2 substitution. This is seen with 2′-OMe-3′-PS2 (MS2) modified U13 that displays a K_D of 3.77 nM (ST-2, ESI:† AF151UB) compared with 1.54 nM for PO2 (ST-2, ESI:† AF147-1), and 0.028 nM for PS2 (ST-2, ESI:† AF151-12), respectively. Methylation there destroys the H-bond to E63, creates a steric conflict with Y85, and likely destabilizes the C2′-endo sugar pucker of U13 (Fig. 2A).

The PS2-mediated affinity gains achieved here for the RNA:MS2 interaction reach almost two orders of magnitude. By comparison, the affinity changes in the siRNA:Ago2, RNA:thrombin and RNA:VEGF complexes were up to 1000-fold.[23,26] Both anti-thrombin and -VEGF RNA aptamer are more complex than the MS2 RNA hairpin. Rather than an induced fit, the latter motif may merely undergo PS2-triggered flexing to promote higher affinity. Perhaps a simple RNA stem-loop motif precludes a 1000-fold boost in K_D in response to a single PS2 moiety. Nevertheless, the observed gains in affinity are impressive and dwarf those seen for duplex DNA that is conformationally more rigid by comparison.

Conflicts of interest

X. Y. is an employee of AM Biotechnologies LLC. The remaining authors declare no competing financial interests.

Notes and references

1 A. D. Keefe, S. Pai and A. Ellington, *Nat. Rev. Drug Discovery*, 2010, **9**, 537–550.
2 S. E. Osborne and A. D. Ellington, *Chem. Rev.*, 1997, **97**, 349–370.
3 C. Tuerk and L. Gold, *Science*, 1990, **249**, 505–510.
4 P. N. Pushparaj, J. J. Aarthi, J. Manikandan and S. D. Kumar, *J. Dent. Res.*, 2008, **87**, 992–1003.
5 M. A. Behlke, *Mol. Ther.*, 2006, **13**, 644–670.
6 S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl, *Nature*, 2001, **411**, 494–498.
7 S. A. Melo and M. Esteller, *Cell Cycle*, 2011, **10**, 922–925.
8 E. M. Small and E. N. Olson, *Nature*, 2011, **469**, 336–342.
9 J. Soutschek, A. Akinc, B. Bramlage, K. Charisse, R. Constien, M. Donoghue, S. Elbashir, A. Geick, P. Hadwiger, J. Harborth, M. John, V. Kesavan, G. Lavine, R. K. Pandey, T. Racie, K. G. Rajeev, I. Rohl, I. Toudjarska, G. Wang, S. Wuschko, D. Bumcrot, V. Koteliansky, S. Limmer, M. Manoharan and H. P. Vornlocher, *Nature*, 2004, **432**, 173–178.
10 P. R. Bouchard, R. M. Hutabarat and K. M. Thompson, *Annu. Rev. Pharmacol. Toxicol.*, 2010, **50**, 237–257.
11 K. K. Jain, *Drug Discovery Today*, 2004, **9**, 307–309.
12 P. S. Pendergrast, H. N. Marsh, D. Grate, J. M. Healy and M. Stanton, *J. Biomol. Tech.*, 2005, **16**, 224–234.
13 M. Blank and M. Blind, *Curr. Opin. Chem. Biol.*, 2005, **9**, 336–342.
14 T. Mairal, V. C. Ozalp, P. Lozano Sanchez, M. Mir, I. Katakis and C. K. O'Sullivan, *Anal. Bioanal. Chem.*, 2008, **390**, 989–1007.
15 G. F. Deleavey and M. J. Damha, *Chem. Biol.*, 2012, **19**, 937–954.
16 P. M. J. Burgers and F. Eckstein, *Biochemistry*, 1979, **18**, 592–596.
17 W. J. Stec and A. Wilk, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 709–722.
18 T. Wada, S. Fujiwara, T. Sato, N. Oka and K. Saigo, *Nucleic Acids Symp. Ser.*, 2004, 57–58.
19 X. Yang, M. Sierant, M. Janicka, L. Peczek, C. Martinez, T. Hassell, N. Li, X. Li, T. Wang and B. Nawrot, *ACS Chem. Biol.*, 2012, **7**, 1214–1220.
20 K. H. Petersen and J. Nielsen, *Tetrahedron Lett.*, 1990, **31**, 911–914.
21 W. B. Derrick, C. H. Greef, M. H. Caruthers and O. C. Uhlenbeck, *Biochemistry*, 2000, **39**, 4947–4954.
22 S. Y. Wu, X. Yang, K. M. Gharpure, H. Hatakeyama, M. Egli, M. H. McGuire, A. S. Nagaraja, T. M. Miyake, S. Rupaimoole, C. V. Pecot, M. Taylor, S. Pradeep, M. Sierant, C. Rodriguez-Aguayo, H. J. Choi, R. A. Previs, G. N. Armaiz-Pena, L. Huang, C. Martinez, T. Hassell, C. Ivan, V. Sehgal, R. Singhania, H. D. Han, C. Su, J. H. Kim, H. J. Dalton, C. Kovvali, K. Keyomarsi, N. A. McMillan, W. W. Overwijk, J. Liu, J. S. Lee, K. A. Baggerly, G. Lopez-Berestein, P. T. Ram, B. Nawrot and A. K. Sood, *Nat. Commun.*, 2014, **5**, 3459.
23 P. Pallan, X. Yang, M. Sierant, N. Abeydeera, T. Hassell, C. Martinez, M. Janicka, B. Nawrot and M. Egli, *RSC Adv.*, 2014, **4**, 64901–64904.
24 R. White, C. Rusconi, E. Scardino, A. Wolberg, J. Lawson, M. Hoffman and B. Sullenger, *Mol. Ther.*, 2001, **4**, 567–573.
25 P. E. Burmeister, S. D. Lewis, R. F. Silva, J. R. Preiss, L. R. Horwitz, P. S. Pendergrast, T. G. McCauley, J. C. Kurz, D. M. Epstein, C. Wilson and A. D. Keefe, *Chem. Biol.*, 2005, **12**, 25–33.
26 N. D. Abeydeera, M. Egli, N. Cox, K. Mercier, J. N. Conde, P. S. Pallan, D. M. Mizurini, M. Sierant, F. E. Hibti, T. Hassell, T. Wang, F. W. Liu, H. M. Liu, C. Martinez, A. K. Sood, T. P. Lybrand, C. Frydman, R. Q. Monteiro, R. H. Gomer, B. Nawrot and X. Yang, *Nucleic Acids Res.*, 2016, **44**, 8052–8064.
27 K. Valegard, J. B. Murray, N. J. Stonehouse, S. van den Worm, P. G. Stockley and L. Liljas, *J. Mol. Biol.*, 1997, **270**, 724–738.
28 D. Dertinger, L. S. Behlen and O. C. Uhlenbeck, *Biochemistry*, 2000, **39**, 55–63.
29 X. Lou, M. Egli and X. Yang, *Curr. Protoc. Nucleic Acid Chem.*, 2016, **67**, 7.25.1–7.25.15.
30 D. E. Volk, T. D. Power, D. G. Gorenstein and B. A. Luxon, *Tetrahedron Lett.*, 2002, **43**, 4443–4447.
31 A. Okuniewski and B. Becker, *Acta Crystallogr., Sect. E: Struct. Rep. Online*, 2011, **67**, 1749–1750.
32 L. Zandarashvili, D. Nguyen, K. M. Anderson, M. A. White, D. G. Gorenstein and J. Iwahara, *Biophys. J.*, 2015, **109**, 1026–1037.