Crystal structure of a DNA catalyst

Almudena Ponce-Salvatierra^{1,2}, Katarzyna Wawrzyniak-Turek^{1,3}, Ulrich Steuerwald², Claudia Höbartner^{1,3} & Vladimir Pena²

Catalysis in biology is restricted to RNA (ribozymes) and protein enzymes, but synthetic biomolecular catalysts can also be made of DNA (deoxyribozymes)¹ or synthetic genetic polymers². *In vitro* selection from synthetic random DNA libraries identified DNA catalysts for various chemical reactions beyond RNA backbone cleavage³. DNA-catalysed reactions include RNA and DNA ligation in various topologies^{4,5}, hydrolytic cleavage^{6,7} and photorepair of DNA⁸, as well as reactions of peptides^{9,10} and small molecules^{11,12}. In spite of comprehensive biochemical studies of DNA catalysts for two decades, fundamental mechanistic understanding of their function is lacking in the absence of three-dimensional models at atomic resolution. Early attempts to solve the crystal structure of an RNA-cleaving deoxyribozyme resulted in a catalytically irrelevant nucleic acid fold¹³. Here we report the crystal structure of the RNA-ligating deoxyribozyme 9DB1 (ref. 14) at 2.8 Å resolution. The structure captures the ligation reaction in the post-catalytic state, revealing a compact folding unit stabilized by numerous tertiary interactions, and an unanticipated organization of the catalytic centre. Structure-guided mutagenesis provided insights into the basis for regioselectivity of the ligation reaction and allowed remarkable manipulation of substrate recognition and reaction rate. Moreover, the structure highlights how the specific properties of deoxyribose are reflected in the backbone conformation of the DNA catalyst, in support of its intricate three-dimensional organization. The structural principles underlying the catalytic ability of DNA elucidate differences and similarities in DNA versus RNA catalysts, which is relevant for comprehending the privileged position of folded RNA in the prebiotic world and in current organisms.

The deoxyribozyme 9DB1 catalyses the regioselective formation of a native phosphodiester bond between the 3'-hydroxyl and the 5'-triphosphate termini of two RNA strands, using divalent metal ions (Mg^{2+} or Mn^{2+}) as cofactors^{14,15}. Ribozymes that catalyse an analogous ligation reaction are found in nature¹⁶, and have also been generated by *in vitro* selection from random RNA libraries^{17,18} or a natural ribozyme scaffold¹⁹. The 9DB1 DNA enzyme is composed of a central catalytic domain flanked by two arms that hybridize to the RNA substrates by canonical Watson–Crick base pairing (Fig. 1a, b), leaving only the two nucleotides embracing the ligation junction unpaired. The *in vitro* selection strategy used to obtain 9DB1 imposed selection pressure for base pairing outside the ligation junction and enforced 3',5' selectivity of the reaction¹⁴.

We crystallized a 44-nucleotide strand that contains the minimally required catalytic core of 31 nucleotides²⁰ in complex with a 15-nucleotide RNA strand. The structure was solved by single anomalous dispersion using data collected from hexamine cobalt (III) chloride. The model built in the experimental map was refined against a 2.8 Å resolution data set collected from a native crystal (Fig. 1c). The crystal contains one molecule in the asymmetric unit and the final model comprises all nucleotides present in the complex (Extended Data Table 1).

The structure exhibits the shape of a λ letter, consisting of the two RNA–DNA hybrid duplexes, P1 and P4, oriented at approximately 120° with respect to each other, and both tightly anchored to

the catalytic domain (Fig. 1d, e). The RNA nucleotides A–1 and G1 encompass the ligation junction and form extensive tertiary contacts with the DNA catalyst (Fig. 1d, e). The 3'-overhanging DNA and RNA nucleotides form a semi-continuous duplex in the crystal lattice

Figure 1 | Global architecture of the DNA catalyst. **a**, DNA-catalysed ligation between 3'-OH and 5'-phosphate. nt, nucleotides. **b**, Proposed secondary structure of the minimal 9DB1 deoxyribozyme²⁰. RNA nucleotides are in blue, ligation junction in light blue, DNA-binding arms in grey, and core nucleotides in black and red. Nucleotides in red do not tolerate any mutation. Blue and green lines indicate double pseudoknot interactions. **c**, Solvent-flattened electron density map contoured at 1.0σ level with orange trace of backbone and nucleobases. **d**, Secondary and tertiary structure with base pairs denoted in Leontis–Westhof presentation. DNA nucleotides in black have N-type ribose pucker, those in green have S-type conformations; grey nucleotides lie outside typical N or S conformations. P1–P5 are colour coded with respect to the structure shown in **e**. Inset: schematic overview of the double pseudoknot. **e**, Cartoon representation of the deoxyribozyme in complex with the ligated RNA product. **f**, A 60° rotation of the image shown in **e**.

¹Max Planck Research Group Nucleic Acid Chemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. ²Research Group Macromolecular Crystallography, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. ³Institute for Organic and Biomolecular Chemistry, Georg-August-University Göttingen, Tammannstr. 2, 37077 Göttingen, Germany.

Figure 2 | Tertiary contacts within the catalytic domain. **a**, Multiple stacking and base pairing contacts in the 9DB1 core. **b**, **c**, Close-up view of multiple P2 and P3 contacts. Black dashes indicate hydrogen bonds.

(Extended Data Fig. 1). Within the catalytic domain, two stacks of two and four base pairs are formed, called P2 and P3, respectively. The paired regions P1, P2 and P3 are positioned on top of one another in a non-coaxial manner (Fig. 1f). In addition, the DNA residues dT29–dT30 base pair with the RNA nucleotides A–1–G1 (denoted P5), which makes the entire DNA–RNA complex fold as a double pseudoknot (Fig. 1d).

The overall fold of the catalytic domain is stabilized by tertiary interactions that connect the regions labelled J1/2, J3/2, P2 and P3 in four successive layers of multiplets (Fig. 2a). Hydrogen bonding forms the long-range Watson–Crick base pair dC12:dG26 and the non-canonical Watson–Crick base pair dT14:dG25 in P2 (Fig. 2b, c), as well as alternative base pairs involving the sugar edges of dG8 and dG21, the Watson–Crick edges of dG8, dG21 and dA23, and the Hoogsteen edges of dA10, dG22 and dG26. Moreover, the compact DNA architecture is stabilized by stacking interactions of nucleotides in J1/2 and J3/2, which place dT11 and dG24 sandwiched between the non-canonical base pairs dG8:dA10 and dG21:dA23. These interactions are consistent with previous atomic mutagenesis data, which showed that the Watson–Crick edges of dG8, dT14, dG21 and dG25 are essential for the functional integrity of 9DB1 (ref. 21). Moreover, dG26 did not tolerate changes on the Watson–Crick or Hoogsteen edges, consistent with its interactions in a base triple with the dC12 and dG8 nucleotides (Fig. 2b).

The reactive nucleotides A–1 and G1 are positioned in the active site by a scaffold formed by four nucleotides from distinct locations of the primary sequence, corroborated by the large angular opening between the duplexes P1 and P4 (Fig. 3a, b). Thus, the junction J2/3 protrudes from the catalytic domain and extensively contacts A–1 and G1 by base pairing with dT30 and dT29, respectively, as well as by the stacking of dG27 on G1 (Fig. 3a, b). On the opposite side, a 130° kink in the backbone between dT14 and dC16 causes dA15 to form a stacking interaction with A–1 (Fig. 3b). In this way, A–1 and G1 are stacked on one another, sandwiched by dA15 and dG27 and engaged in hydrogen bonding with dT29 and dT30, resulting in a microenvironment reminiscent of a duplex.

Figure 3 | Positioning of reactive nucleotides in the catalytic domain.

a, Structure at the ligation junction. **b**, Base-pairing and stacking interactions of the ligated nucleotides G1 and A–1. Black dashes indicate hydrogen bonds. **c**, Biochemical interrogation of the base pair of nucleotides 1:29. Ligation rate or yield with original 9DB1 (dT29, black), and with mutant 9DB1 DNAs (green) complementary to the RNA substrates, that is, dT29dC for pppC1, dT29dA for pppU1 RNA. **d**, Activity assays of 9DB1 with substrate analogues modified at the 2'-position of A–1. **c**, **d**, Single-turnover kinetic assays were performed at pH 7.5 with 20 mM Mn^{2+} at 37 °C. All experiments were repeated two or three times. k_{obs} , observed rate constant, k_{rel} , relative rate constant with respect to A–1 substrate.

The identification of the previously unknown base pairs (P5) at the ligation junction was intriguing, as dT29 and dT30 were shown to tolerate mutations²⁰, and RNA substrates with uridine or guanosine at position –1 were also efficiently ligated^{14,15}. In contrast, only 5'-purine nucleotides were allowed at position 1 in the donor RNA (that is, 5'-pppG1 or 5'-pppA1)¹⁴. Therefore, the crystal structure of 9DB1 revealed an effective way of manipulating the catalytic activity by compensatory mutagenesis of the base pair at the ligation junction. Indeed, RNA substrates that were inert to the original deoxyribozyme, such as 5'-pppC1 and 5'-pppU1 RNA, are now readily ligated with mutated 9DB1 enzymes bearing G and A, respectively, in position 29 (Fig. 3c). On the basis of the crystal structure, 9DB1 deoxyribozymes can now be designed for ligation of any RNA substrates, irrespective of their sequence.

The regioselectivity of the ligation reaction is determined by the relative orientation of the reacting nucleotides. For ribozymes^{22,23} as well as deoxyribozymes²⁴, continuous Watson–Crick base pairing across the ligation junction favours nucleophilic attack by the 3'-OH group onto the 5'-triphosphate of G1. The regioselectivity of 9DB1 is dictated by tertiary contacts and specific interactions at the ligation junction, which differ markedly from a continuous double-helical arrangement. Thus, the orientation of the donor nucleotide G1 is assisted by a hydrogen bond of its 2'-OH to the minor groove of dC12:dG26 in P2 (Fig. 4a). The arrangement of the binding arms induces a shorter distance between the phosphate groups of A–1 and G1 (5.3 Å as compared to 5.9 Å in a regular A-form duplex). The hydrogen bond between the 2'-OH group of A–1 and the O4' (or O5') of G1 may sequester the 2'-OH and thereby assist the regioselective formation of 3'–5' ligated products (Fig. 3b). Consistently, removal or blocking of the 2'-OH group by introducing a 2'-deoxy or 2'-OCH₃ nucleotide at position A–1 prevented efficient ligation (Fig. 3d). Instead, replacement of the 2'-OH with 2'-NH₂ still allowed RNA ligation, although with a 30-fold

Figure 4 | Active site of 9DB1. **a**, The ligation junction and the putatively catalytic dA13 phosphate are depicted based on the crystal structure (left) and as schematic drawing of the ligation reaction (right). Red dashes indicate the proximity between *pro*-*Sp* oxygen and phosphorus atom of newly formed phosphodiester bond (red ball). **b**, Biochemical interrogation of the phosphate participation by phosphorothioate and methylphosphonate interference (pH 7.5, 20 mM Mn²⁺). Diastereomers are schematically depicted. All measurements were repeated at least twice.

slower rate (Extended Data Fig. 2). Interestingly, 2'-fluoro-modified RNA was ligated only 6–7-fold slower (Fig. 3d and Extended Data Fig. 2), suggesting that the electronegativity of fluorine and the resulting 3'-endo conformation of the ribose compensates for the loss of the 2'-OH group and orients the 3'-OH for nucleophilic attack onto the 5'-phosphate.

Catalysis of phosphoryl-transfer reactions by proteins or ribozymes often involves activation of nucleophilic groups and electrostatic stabilization of the transition state by divalent metal ions or specific chemical groups^{25,26}. The active site of an RNA enzyme that catalyses the same reaction as the DNA enzyme 9DB1, is composed of two phosphate groups coordinating a catalytic metal ion, plus an amino group from a nucleobase and a 2'-OH group that exerts its action through intermediacy of a water molecule²⁷. There are no equivalent interactions in the core of 9DB1, and we do not observe electron density for a catalytic metal ion. However, a striking feature is the phosphate group of nucleotide dA13, which lies within 3.1 Å distance from the ligation junction (Fig. 4a). To test the role of the dA13 phosphate in catalysis, we replaced either of the non-bridging oxygen atoms with a sulfur atom, and assessed the enzymatic activity of the resulting phosphorothioate-modified DNAs. In the presence of manganese ions, the ligation rate of the deoxyribozyme containing an *Sp* phosphorothioate was reduced 100-fold, whereas the *Rp* stereoisomer retained its full activity (Fig. 4b). A similar phosphorothioate interference effect was observed when magnesium was used as metal ion cofactor, and the catalytic activity could also not be rescued by the addition of cadmium ions. The absence of thiophilic metal rescue cannot prove whether metal ions are involved or not in catalysis. Upon substitution of the non-bridging phosphate oxygens with a methyl group, only one of the two resulting uncharged methylphosphonate diastereomers retained catalytic activity (Fig. 4b). Although the role of metal ions remains unclear, these data are consistent with the critical role of the non-bridging *pro*-*Sp* phosphate oxygen of dA13 in activating the 3'-OH group to promote RNA ligation.

This work provides the experimental evidence that DNA possesses the ability to fold into compact three-dimensional units, stabilized by intrinsic tertiary interactions in the absence of 2'-OH groups or

artificial hydrophobic substituents²⁸. In contrast with ribozymes, the sugar-phosphate backbone of the catalytic domain of 9DB1 exhibits much larger conformational diversity, as reflected by the broad distribution of pseudorotation phase angles found in the DNA nucleotides (Extended Data Fig. 3). Thus, of the 31 nucleotides in the catalytic domain, 8 and 20 adopt north (N)- and south (S)-type ribose puckers, respectively (Fig. 1d). The remaining three nucleotides lie outside typical N or S conformations. The less restrictive puckering allows DNA to explore a wide range of conformations, compensating for the lack of the 2'-hydroxyl group, which has an important structural role in ribozymes.

Comparison of 9DB1 with two RNA ligase ribozymes illustrates how the fold of DNA and RNA enzymes underlies the recognition and positioning of the substrates in the catalytic centre (Extended Data Fig. 4). In the class I ligase and L1 ligase ribozymes, the reactive nucleotides are positioned in extended helices (Extended Data Fig. 5)^{29,30}. In contrast, the four contributing bases in the 9DB1 ligase are provided *in trans* from different primary locations of the catalytic domain. Notably, all four DNA nucleotides adopt S-type ribose puckers, highlighting how the specific conformational propensity of deoxyribose moieties supports recognition and positioning of the substrate in the active centre.

The crystal structure of the 9DB1 deoxyribozyme demonstrates that DNA possesses the intrinsic ability to adopt complex tertiary folds that support catalysis, and unveils for the first time the active site of a DNA enzyme in the post-catalytic state. Together with mutagenic analyses, the structure elucidates the basis of substrate recognition and provides first insights for rationalization of the regiospecific bond formation. The ability of DNA to form complex three-dimensional architectures that support catalysis raises questions about the positioning of this biopolymer in the prebiotic evolution of life and may shed light on metabolic events in current organisms, in which single-stranded DNA may adopt functionally important folds.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 July; accepted 20 November 2015.

Published online 6 January 2016.

- Breaker, R. R. & Joyce, G. F. A DNA enzyme that cleaves RNA. *Chem. Biol.* **1**, 223–229 (1994).
- Taylor, A. I. *et al.* Catalysts from synthetic genetic polymers. *Nature* **518**, 427–430 (2015).
- Schlosser, K. & Li, Y. Biologically inspired synthetic enzymes made from DNA. *Chem. Biol.* **16**, 311–322 (2009).
- Coppins, R. L. & Silverman, S. K. A DNA enzyme that mimics the first step of RNA splicing. *Nature Struct. Mol. Biol.* **11**, 270–274 (2004).
- Sreedhara, A., Li, Y. & Breaker, R. R. Ligating DNA with DNA. *J. Am. Chem. Soc.* **126**, 3454–3460 (2004).
- Chandra, M., Sachdeva, A. & Silverman, S. K. DNA-catalyzed sequence-specific hydrolysis of DNA. *Nature Chem. Biol.* **5**, 718–720 (2009).
- Gu, H., Furukawa, K., Weinberg, Z., Berenson, D. F. & Breaker, R. R. Small, highly active DNAs that hydrolyze DNA. *J. Am. Chem. Soc.* **135**, 9121–9129 (2013).
- Chinnappen, D. J. & Sen, D. A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl Acad. Sci. USA* **101**, 65–69 (2004).
- Walsh, S. M., Sachdeva, A. & Silverman, S. K. DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **135**, 14928–14931 (2013).
- Chandrasekar, J. & Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl Acad. Sci. USA* **110**, 5315–5320 (2013).
- Li, Y. & Sen, D. A catalytic DNA for porphyrin metallation. *Nature Struct. Biol.* **3**, 743–747 (1996).
- Chandra, M. & Silverman, S. K. DNA and RNA can be equally efficient catalysts for carbon–carbon bond formation. *J. Am. Chem. Soc.* **130**, 2936–2937 (2008).
- Nowakowski, J., Shim, P. J., Prasad, G. S., Stout, C. D. & Joyce, G. F. Crystal structure of an 82-nucleotide RNA–DNA complex formed by the 10–23 DNA enzyme. *Nature Struct. Biol.* **6**, 151–156 (1999).
- Purtha, W. E., Coppins, R. L., Smalley, M. K. & Silverman, S. K. General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages. *J. Am. Chem. Soc.* **127**, 13124–13125 (2005).
- Büttner, L., Seikowski, J., Wawrzyniak, K., Ochmann, A. & Höbartner, C. Synthesis of spin-labeled riboswitch RNAs using convertible nucleosides and DNA-catalyzed RNA ligation. *Bioorg. Med. Chem.* **21**, 6171–6180 (2013).
- Vicens, Q. & Cech, T. R. A natural ribozyme with 3'/5' RNA ligase activity. *Nature Chem. Biol.* **5**, 97–99 (2009).

17. Bartel, D. P. & Szostak, J. W. Isolation of new ribozymes from a large pool of random sequences [see comment]. *Science* **261**, 1411–1418 (1993).
18. Ekland, E. H., Szostak, J. W. & Bartel, D. P. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **269**, 364–370 (1995).
19. Jaeger, L., Wright, M. C. & Joyce, G. F. A complex ligase ribozyme evolved *in vitro* from a group I ribozyme domain. *Proc. Natl Acad. Sci. USA* **96**, 14712–14717 (1999).
20. Wachowius, F., Javadi-Zarnaghi, F. & Höbartner, C. Combinatorial mutation interference analysis reveals functional nucleotides required for DNA catalysis. *Angew. Chem. Int. Ed. Engl.* **49**, 8504–8508 (2010).
21. Wachowius, F. & Höbartner, C. Probing essential nucleobase functional groups in aptamers and deoxyribozymes by nucleotide analogue interference mapping of DNA. *J. Am. Chem. Soc.* **133**, 14888–14891 (2011).
22. Ikawa, Y., Tsuda, K., Matsumura, S. & Inoue, T. *De novo* synthesis and development of an RNA enzyme. *Proc. Natl Acad. Sci. USA* **101**, 13750–13755 (2004).
23. Pitt, J. N. & Ferré-D'Amaré, A. R. Structure-guided engineering of the regioselectivity of RNA ligase ribozymes. *J. Am. Chem. Soc.* **131**, 3532–3540 (2009).
24. Coppins, R. L. & Silverman, S. K. Rational modification of a selection strategy leads to deoxyribozymes that create native 3'–5' RNA linkages. *J. Am. Chem. Soc.* **126**, 16426–16432 (2004).
25. Doudna, J. A. & Cech, T. R. The chemical repertoire of natural ribozymes. *Nature* **418**, 222–228 (2002).
26. Fedor, M. J. & Williamson, J. R. The catalytic diversity of RNAs. *Nature Rev. Mol. Cell Biol.* **6**, 399–412 (2005).
27. Shechner, D. M. & Bartel, D. P. The structural basis of RNA-catalyzed RNA polymerization. *Nature Struct. Mol. Biol.* **18**, 1036–1042 (2011).
28. Davies, D. R. *et al.* Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proc. Natl Acad. Sci. USA* **109**, 19971–19976 (2012).
29. Shechner, D. M. *et al.* Crystal structure of the catalytic core of an RNA-polymerase ribozyme. *Science* **326**, 1271–1275 (2009).
30. Robertson, M. P. & Scott, W. G. The structural basis of ribozyme-catalyzed RNA assembly. *Science* **315**, 1549–1553 (2007).

Acknowledgements This work was supported by the Max Planck Society. We thank J. Ludwig for a gift of 5'-triphosphorylated RNAs, J. Seikowski for assistance with RNA and DNA synthesis, F. Wachowius and B. Samanta for initial crystallization samples, P. Afonine and members of the Pena and Höbartner laboratories for discussions, and the beamline staff at the Swiss Light Source, Villigen, Switzerland for assistance with data collection.

Author Contributions Crystallographic work was performed by A.P.-S. under the supervision of V.P., biochemical experiments were performed by A.P.-S., K.W.-T. and C.H. U.S. obtained initial crystals. A.P.-S., V.P. and C.H. designed the experiments and all authors discussed the results and contributed to the manuscript.

Author Information Structural models and structure factors have been deposited in the Protein Data Bank under accession numbers 5CKI (cobalt) and 5CKK (native). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.P. (vpenna@gwdg.de) or C.H. (choebar1@gwdg.de).

METHODS

Oligonucleotide synthesis and complex formation. RNA and DNA oligonucleotides were prepared by solid-phase synthesis using phosphoramidite chemistry on polystyrene or controlled-pore glass solid supports. For RNA synthesis 2'-O-TOM-protected RNA phosphoramidites were used³¹. 3'-Modified RNA was prepared on commercially available modified (2'-OCH₃, 2'-NH₂, 2'-F-uridine) solid supports or universal support (2'-F-adenosine). All oligonucleotides were deprotected with ammonia or methyl amine, followed by tetrabutylammonium fluoride (for RNA only), desalted and purified by denaturing polyacrylamide gel electrophoresis. The phosphorothioate modification was introduced during solid-phase synthesis using phenylacetyl disulfide³². Methylphosphonate modifications were introduced using commercially available N⁴-acetyl-dC methylphosphonamidite. Phosphorothioate and methylphosphonate diastereomers were separated by RP-C18 high-performance liquid chromatography (HPLC; NH₄OAc/CH₃CN gradient) and assigned according to elution order as reported^{33,34}. Full-length deoxyribozymes were generated by splinted ligation of two fragments. 5'-Triphosphorylated donor RNA with 5'-purine nucleotides were prepared by *in vitro* transcription with T7 RNA polymerase. Synthetic triphosphorylated RNAs were prepared by solid-phase synthesis following published procedures^{35,36} or provided by J. Ludwig. Crystallization complexes were formed by mixing RNA and DNA strands in a 1:1 molar ratio, heated to 95 °C, and slowly cooled down to room temperature before MgCl₂ was added. The final concentration of the complex was 0.8 mM in 10 mM HEPES pH 8.0, 150 mM NaCl, 2 mM KCl and 20 mM MgCl₂.

Crystallization and diffraction experiments. Crystals were grown at 4 °C from solutions containing 35% 2-methyl-2,5-pentanediol (MPD), 0.5 mM spermine, 20 mM MgCl₂, 100 mM NaCl and 100 mM sodium cacodylate (pH 6.5) using the sitting-drop vapour-diffusion method. Sitting drops with volumes of 2 µl produced crystals after 5–10 days, reaching full size after 30 days. The derivatized crystals were prepared by soaking of the crystals in a solution containing the mother liquor and hexammine cobalt (III) chloride for 1 min.

Prior to data collection, crystals were cryoprotected with perfluoropolyether (PFPE) and flash frozen. The crystallographic experiments were performed on the PXII beamline at the Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland. The data sets were processed with XDS³⁷.

Crystallographic analyses. The structure of the complex was determined by single anomalous dispersion (SAD) using a hexammine cobalt (III) chloride derivative. The determination of the heavy atom substructure was done with SHELXD³⁸. Initial SAD phase calculation and automatic model building were carried out with AutoSol wizard³⁹. The resulting model from AutoSol was manually edited with COOT⁴⁰. The edited partial model was used along with the experimentally determined cobalt sites for MR-SAD phase calculation within Phaser⁴¹. The resulting map was solvent flattened using RESOLVE⁴². This map was used for tracing the entire phosphodiester backbone. The RNA strand of the model was built using Rcrane⁴³ and the DNA strand was built from individual nucleotide monophosphates. Structure refinement was carried out with phenix.refine⁴⁴. Analysis of pseudorotation phase angles was performed on the PROSIT webserver (<http://cactus.nci.nih.gov/prosit/>)⁴⁵.

Kinetic assays of DNA-catalysed RNA ligation reactions. Single-turnover assays were performed as described previously^{14,15}. Briefly, 5'-³²P-trace-labelled acceptor

RNA (3 pmol), 9DB1 deoxyribozyme, and 5'-triphosphorylated donor RNA were mixed in a 1:5:10 ratio and annealed by heating to 95 °C for 2 min, followed by slowly cooling for 15 min to room temperature (for RNA/DNA sequences see Extended Data Table 2). The assays were performed in 10 µl volume at 37 °C in buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl and 20 mM MnCl₂, or 50 mM Tris buffer, pH 8.5, 150 mM NaCl, 2 mM KCl and 80 mM MgCl₂ including 0, 0.1, 1, or 10 mM CdCl₂. Aliquots of 1 µl were removed at desired time points and quenched in stop solution (80% formamide, 1 × TBE, 50 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol). The samples were analysed by PAGE (15% polyacrylamide), and band intensities were quantified by Phosphorimaging. The yield versus time data were fit to (fraction ligated) = $Y(1 - e^{-kt})$, where $k = k_{\text{obs}}$ and Y = final yield. No statistical methods were used to predetermine sample size.

31. Pitsch, S., Weiss, P. A., Jenny, L., Stutz, A. & Wu, X. Reliable chemical synthesis of oligoribonucleotides (RNA) with 2'-O-[(Triisopropylsilyl)oxy]methyl(2'-O-TOM)-protected phosphoramidites. *Helv. Chim. Acta* **84**, 3773–3795 (2001).
32. Kumar, R. K., Olsen, P. & Ravikumar, V. T. An alternative advantageous protocol for efficient synthesis of phosphorothioate oligonucleotides utilizing phenylacetyl disulfide (PADS). *Nucleosides Nucleotides Nucleic Acids* **26**, 181–188 (2007).
33. Frederiksen, J. K. & Piccirilli, J. A. Separation of RNA phosphorothioate oligonucleotides by HPLC. *Methods Enzymol.* **468**, 289–309 (2009).
34. Koch, M. et al. Role of a ribosomal RNA phosphate oxygen during the EF-G-triggered GTP hydrolysis. *Proc. Natl Acad. Sci. USA* **112**, E2561–E2568 (2015).
35. Ludwig, J. & Eckstein, F. Rapid and efficient synthesis of nucleoside 5'-O-(1-thiotriphosphates), 5'-triphosphates and 2',3'-cyclophosphorothioates using 2-chloro-4h-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.* **54**, 631–635 (1989).
36. Goldeck, M., Tuschl, T., Hartmann, G. & Ludwig, J. Efficient solid-phase synthesis of pppRNA by using product-specific labeling. *Angew. Chem. Int. Ed. Engl.* **53**, 4694–4698 (2014).
37. Kabsch, W. X. D. S. *Acta Crystallogr. D* **66**, 125–132 (2010).
38. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D* **66**, 479–485 (2010).
39. Terwilliger, T. C. et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D* **65**, 582–601 (2009).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
41. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
42. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
43. Keating, K. S. & Pyle, A. M. RCrane: semi-automated RNA model building. *Acta Crystallogr. D* **68**, 985–995 (2012).
44. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
45. Sun, G., Voigt, J. H., Filippov, I. V., Marquez, V. E. & Nicklaus, M. C. PROSIT: pseudo-rotational online service and interactive tool, applied to a conformational survey of nucleosides and nucleotides. *J. Chem. Inf. Comput. Sci.* **44**, 1752–1762 (2004).
46. Ren, A. et al. In-line alignment and Mg²⁺ coordination at the cleavage site of the env22 twister ribozyme. *Nat. Commun.* **5**, 5534 (2014).

Extended Data Figure 1 | Semi-continuous helix in the crystal lattice. Nucleotides important for the crystal contacts are labelled accordingly.

Extended Data Figure 2 | Influence of 2'-modifications on mutant acceptor RNA U-1.

Extended Data Figure 3 | Distribution of pseudorotation phase angles of 9DB1 deoxyribozyme in comparison with twister⁴⁶ and L1 ligase ribozyme³⁰.

Extended Data Figure 4 | Active sites of RNA-ligating nucleic acid enzymes. Left: deoxyribozyme 9DB1; middle: class I ribozyme; right: L1 ligase ribozyme. Figure was generated using Protein Data Bank (PDB) accession numbers 5CKK (9DB1), 3R1L (class I) and 2OIU (L1).

Extended Data Figure 5 | Positioning of the reactive nucleotides in the active centre of ligase deoxyribozyme 9DB1 in comparison to ribozymes. A schematic of the three enzymes is shown before and after the crystal structure determination (left and central columns, respectively). Strands that were deliberately chosen are depicted in red and blue, while the ones that resulted from random selection are shown in

black. Nucleotides at the ligation junction are coloured in cyan, stacking nucleotides are depicted in light red. Nucleotides shown as white-filled circles are indicated for orientation purposes. Recognition between reactive and pairing nucleotides is shown as dashed lines (left column). Structures of product-bound nucleic acid enzymes are compared (PDB accessions 5CKK (9DB1), 3HHN (class I) and 2OIU (L1), right column).

Extended Data Table 1 | Data collection and refinement statistics

	Native	Co-Hex
Data collection		
Space group	P 43 21 2	P 43 21 2
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	83.1, 83.1, 55.9	80.8, 80.8, 55.8
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength	1.0	1.6
Resolution (Å)	41.58 – 2.80 (2.90 – 2.80)	40.42 – 2.98 (3.09 – 2.98)
<i>R</i> _{merge} (%)	5 (>100)	8 (>100)
<i>I</i> / σ <i>I</i>	33.87 (0.43)	25.23 (0.86)
CC _{1/2}	0.99 (0.90)	0.99 (0.52)
Completeness (%)	97 (97)	100 (98)
Redundancy	24.8 (25.8)	26.1 (24.0)
Refinement		
Resolution (Å)	2.80	2.98
No. reflections	5012	4053
<i>R</i> _{work} / <i>R</i> _{free}	0.27/ 0.29	0.25/ 0.31
No. atoms		
Nucleic acid	1232	1232
Ligand/ion	3	6
B-factors		
Nucleic acid	148.07	133.72
Ligand/ion	119.43	132.84
R.m.s deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.66	0.60

One crystal was used for each data set. Highest-resolution shell values are shown in parentheses.

Extended Data Table 2 | Oligonucleotides for kinetic experiments

description	5'-sequence-3'
acceptor RNA 3'-A	GGAAGUCUCAUGUACUA
acceptor RNA 3'-(2'dA)	GGAAGUCUCAUGUACU (dA)
acceptor RNA 3'-(2-OMeA)	GGAAGUCUCAUGUACU (OMeA)
acceptor RNA 3'-(2'-F-A)	GGAAGUCUCAUGUACU (FA)
acceptor RNA 3'-U	GGAAGUCUCAUGUACUU
acceptor RNA 3'-(2'-OMe-U)	GGAAGUCUCAUGUACU (OMeU)
acceptor RNA 3'-(2'-NH ₂ -U)	GGAAGUCUCAUGUACU (NH ₂ U)
acceptor RNA 3'-(2'-F-U)	GGAAGUCUCAUGUACU (FU)
donor RNA1 5'pppG	pppGAUGUUCUAGCGCCGA
donor RNA2 5'pppG	ppp G ACGCUGACCCUGAAGUUCAUCUU
donor RNA 5'pppA	ppp A ACGCUGACCCUGAAGUUCAUCUU
donor RNA 5'pppC	ppp C ACGCUGACCCUGAAGUUCAUCUU
donor RNA 5'pppU	ppp U ACGCUGACCCUGAAGUUCAUCUU
9DB1 DNA original*	CAAGGCGCTAGAACATGGATCATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGACTTCC
9DB1 DNA*,†	GAACTTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGACTTCC
9DB1 dT29C	GAACTTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGTCTGCCGTTTAAGTACATGAGACTTCC
9DB1 dT29A	GAACTTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGTATGCCGTTTAAGTACATGAGACTTCC
9DB1 dT29G	GAACTTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGTGTGCCGTTTAAGTACATGAGACTTCC
9DB1 DNA (Fig. 4b)	GCGCTAGAACATGGATCATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGAC
9DB1 DNA (ps) (Fig. 4b)‡	GCGCTAGAACATGGATC ps ATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGAC
9DB1 DNA (pCH3) (Fig. 4b)‡	GCGCTAGAACATGGATC (pCH3) ATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGAC
splint for ligation	CCCCGCCGACCGTATGATCCATGTTC

*Catalytic core is underlined.

†dT29 is in bold.

‡Ligation site is indicated with |.