

Cite this: *Chem. Sci.*, 2024, 15, 13068

All publication charges for this article have been paid for by the Royal Society of Chemistry

Chemo-enzymatic production of base-modified ATP analogues for polyadenylation of RNA[†]

Rachel M. Mitton-Fry,^{ID}*^{ab} Jannik Eschenbach,^{ID}^a Helena Schepers,^a René Rasche,^{ID}^a Mehmet Erguven,^{ID}^{acd} Daniel Kümmel,^{ID}^{ad} Andrea Rentmeister,^{ID}*^{acd} and Nicolas V. Cornelissen,^{ID}*^a

Base-modified adenosine-5'-triphosphate (ATP) analogues are highly sought after as building blocks for mRNAs and non-coding RNAs, for genetic code expansion or as inhibitors. Current synthetic strategies lack efficient and robust 5'-triphosphorylation of adenosine derivatives or rely on costly phosphorylation reagents. Here, we combine the efficient organic synthesis of base-modified AMP analogues with enzymatic phosphorylation by a promiscuous polyphosphate kinase 2 class III from an unclassified *Erysipelotrichaceae* bacterium (EbPPK2) to generate a panel of C2-, N⁶-, or C8-modified ATP analogues. These can be incorporated into RNA using template independent poly(A) polymerase. C2-halogenated ATP analogues were incorporated best, with incorporations of 300 to >1000 nucleotides forming hypermodified poly(A) tails.

Received 8th June 2024

Accepted 12th July 2024

DOI: 10.1039/d4sc03769c

rsc.li/chemical-science

Introduction

The production of mRNA therapeutics requires modified nucleoside-5'-triphosphates (NTPs), e.g., pseudouridine (Ψ) and N1-methylpseudouridine (m¹Ψ) replacing uridine (U), which increase the stability of mRNA and lower the resulting immune response.¹ In contrast to replacing nucleotides in the entire mRNA,² post-transcriptional modification of the poly(A) tail has been shown to improve protein production in cells independent of the encoded protein of interest.³ Additionally, base-modified nucleotides play critical regulatory roles in a variety of non-coding RNAs, and incorporation of non-natural nucleotides allows the study or manipulation of RNA localization and function.⁴

In spite of recent improvements in NTP synthesis,⁵ low yields in the last 5'-triphosphorylation steps (typically 4–33%) limit the number of base-modified NTPs available.⁶ For instance, the synthesis of base-modified ATP analogues is considered “troublesome”⁷ and often low yielding (8–30%)⁸ with few exceptions,⁹ rendering it an underexplored chemical space.

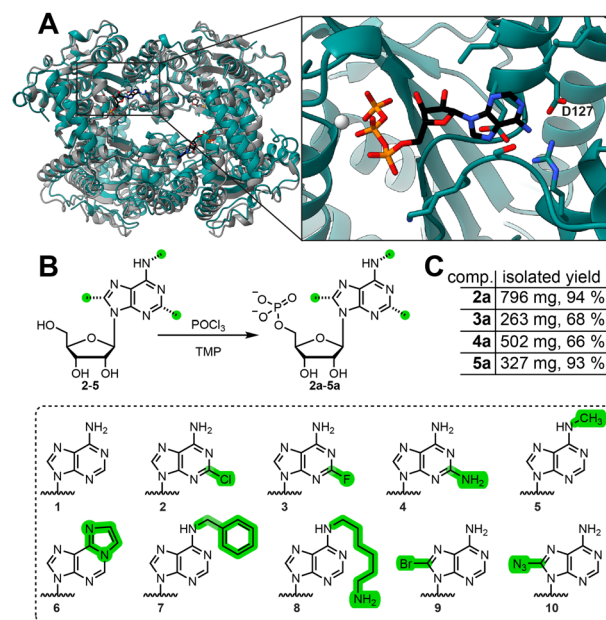


Fig. 1 Model of EbPPK2 and substrate panel of AMP analogues. (A) AlphaFold2 model of tetrameric EbPPK2 (teal), overlaid with the tetrameric PPK2 from *Meiothermus ruber* (gray) bound to ATP (PDB: 5LD1). Protein tetramers aligned with a root mean square deviation (RMSD) of 0.8 Å over 231 C_α atom pairs. Inset shows the predicted active site from a single subunit of EbPPK2, with all sidechains found within 4 Å of the nucleobase of ATP represented as sticks. D127 (labelled) coordinates the N⁶-position of the base. Alignment, RMSD calculation and figure made using Chimera.¹⁸ (B) Scheme of the chemical synthesis of base-modified AMP analogues (top) and the nucleobase portion of substrates tested for EbPPK2 (bottom). Green highlighting represents potential modifications, where H in all positions corresponds to natural adenosine 1. (C) Isolated yields of the base-modified AMP analogues 2a–5a.

^aInstitute of Biochemistry, University of Münster, Corrensstr. 36, D-48149 Münster, Germany. E-mail: mittonfryr@denison.edu; andrea.rentmeister@cup.lmu.de; cornelissen@uni-muenster.de

^bDepartment of Chemistry and Biochemistry, Denison University, 100 W. College St., Granville, Ohio 43023, USA

^cInstitute of Chemical Epigenetics, Ludwig-Maximilians-University Munich, Butenandtstr. 5-13, Haus F, D-81377 Munich, Germany

^dCells in Motion Interfaculty Centre, University of Münster, Waldeyerstraße 15, D-48149 Münster, Germany

[†] Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4sc03769c>

While chemical 5'-monophosphorylation of adenosine is known to proceed efficiently (>90%),¹⁰ enzymatic strategies have emerged as viable alternatives for 5'-triphosphorylation.¹¹

Polyphosphate kinase family 2 (PPK2) enzymes are particularly attractive, as they use inexpensive and non-toxic Graham's salt (sodium polyphosphate) as a cofactor to phosphorylate AMP or ADP and derivatives thereof.¹² For example, base-modified ATP analogues have been prepared using a two enzyme PPK2-I/PPK2-II mixture.^{11a} Recently, Tavanti *et al.* screened 92 PPKs and identified a PPK2 class III enzyme (PPK2-III) from an unclassified *Erysipelotrichaceae* bacterium (EbPPK2) which served as a robust and efficient ATP recycling catalyst.¹³ As observed for other PPK2-III enzymes,¹⁴ EbPPK2 accepts Graham's salt as a phosphate source and efficiently catalyses production of both ADP and ATP starting from AMP. The reported equilibria of AMP (~3%), ADP (~27%) and ATP (~70%) are consistent with thermodynamic calculations for other PPK2 enzymes.¹⁵

In this work, base-modified adenosine analogues were phosphorylated chemically, and the resulting AMP analogues were converted to the corresponding ATP analogues by EbPPK2. We applied the chemo-enzymatically generated ATP analogues to *in vitro* reactions with poly(A) polymerase (PAP) to prepare 30-mer RNAs with hypermodified poly(A) tails.

Results and discussion

We generated a structural model of EbPPK2 using AlphaFold2¹⁶ in order to examine its nucleotide-binding pocket. As other

PPK2-III enzymes are known to form tetrameric complexes,^{12c} we used four subunits to generate a multimeric model.¹⁷

The core of the structure was modelled with very high confidence, while the two lid helices of the nucleotide-binding pocket have a slightly lower confidence, suggesting a certain degree of conformational flexibility (Fig. S1†). Our model aligned closely to the crystal structure of a class III PPK2 enzyme from *Meiothermus ruber* bound to ATP (MrPPK2, PDB: 5LD1) (Fig. 1A).^{12c} Thus, we inserted the nucleotide from the MrPPK2 structure into our predicted EbPPK2 model and used the YASARA force field minimization server¹⁹ to improve the accuracy of the nucleotide-binding mode. Analyzing the nucleobase-binding pocket (Fig. 1A, inset), we observed that the C2- and C8-positions of the purine base are not involved in tight contacts with the protein, such that modifications might have sufficient space to be tolerated. Similarly, as the N⁶-position is coordinated by a single aspartate residue (D127), N⁶-modifications were also expected to have the potential to be tolerated.

Therefore, in addition to natural AMP (1a), we synthesized AMP analogues bearing modifications at the C2- and N⁶-positions (2a–5a) using the common procedure reported by Yoshikawa *et al.* in 1967¹⁰ (Fig. 1B). The compounds were purified using C₁₈-flash chromatography (Fig. S2†) and obtained as sodium salts after ion exchange with high yield (Fig. 1C). Identity was confirmed by mass spectrometry and NMR (Fig. S3–S18†). This panel was complemented with commercially available base-modified AMP analogues (6a–10a) to

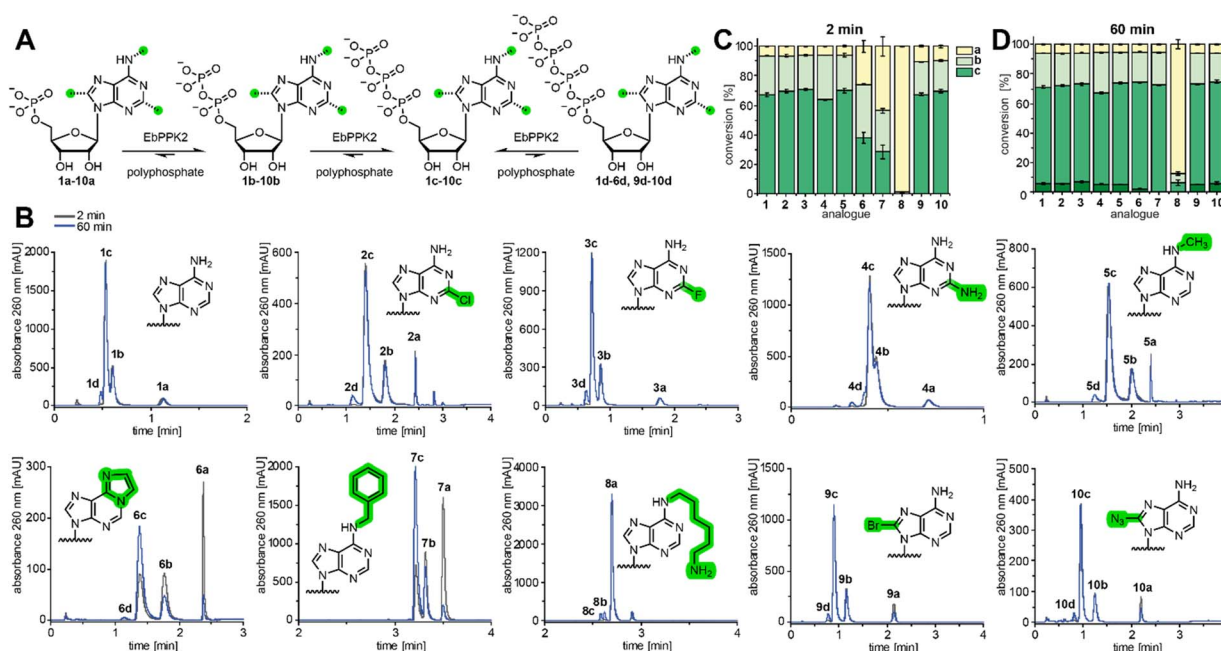


Fig. 2 Substrate scope of EbPPK2. (A) Scheme of the EbPPK2-catalysed reactions using polyphosphate. AMP analogues 1a–10a are converted to ADP analogues 1b–10b and ATP analogues 1c–10c until an equilibrium is approached. After 60 min the corresponding tetraphosphates (d) are found for some analogues. (B) Representative UV-chromatograms (260 nm) of the LC-MS analysis of the EbPPK2 reactions using 1a–10a after 2 min and 60 min. Conditions: 1 mM AMP analogue (1a–10a), 6.1 g L⁻¹ sodium polyphosphate, 4 μM EbPPK2 incubated for 2 min or 60 min at 30 °C in a total volume of 20 μL. Buffer: 20 mM Tris pH = 8, 20 mM MgCl₂. The polyphosphate used had an average chain length of 11 phosphate units, according to ³¹P NMR analysis (Fig. S24†), making the phosphate chain concentration approximately 5 mM. (C) and (D) Conversions of the EbPPK2 reactions after 2 min (C) and 60 min (D) as average of three independent experiments with standard deviation (SD).



include further modifications at the N⁶- and C8-positions to gain insight into the promiscuity of EbPPK2.

EbPPK2 was expressed in *E. coli* BL21(DE3) cells, yielding 125 mg of soluble protein per litre of culture after NiNTA purification (Fig. S19 and S20†). Gel filtration chromatography confirmed that EbPPK2 formed a tetramer in solution (Fig. S21†). Thermal stability of EbPPK2 was determined in different buffer conditions, and melting temperatures (T_m) of up to 54 °C were measured (Fig. S22 and S23†). The activity of EbPPK2 was first confirmed using the natural substrate **1a** (Fig. 2).

The reaction was performed under mild conditions, with incubation at 30 °C and pH 8. Aliquots were analysed by LC-MS after 2 min and 60 min incubation. The UV chromatogram at 260 nm (Fig. 2B) indicated that the equilibrium of 6% AMP (**1a**), 26% ADP (**1b**) and 67% ATP (**1c**) was approached in only 2 min, with minor conversion to the tetraphosphate (**1d**) after 1 hour.

Enzymatic reactions with AMP analogues **2a–10a** were then performed and analysed (Fig. 2B–D). All UV signals could be unambiguously assigned to the expected $[M + H]^+$ ion in ESI-positive mode (Fig. S25–S44†). To our delight, EbPPK2 accepted a broad range of C2-, N⁶-, and C8-modified analogues.

The C2-modified AMP analogues were tolerated well, as 2-chloro-AMP (**2a**), 2-fluoro-AMP (**3a**) and 2-amino-AMP (**4a**) were converted to 70% (**2c**), 71% (**3c**) and 64% (**4c**) of the corresponding ATP analogue in 2 min, respectively. N⁶-modifications led to more variable effects, particularly with respect to the kinetics of the reaction. N⁶-methyl-AMP (**5a**) gave 70% of the corresponding triphosphate (**5c**) in 2 min. Reactions using the larger N⁶-modifications etheno-AMP (**6a**) and N⁶-benzyl-AMP (**7a**) proceeded more slowly, with only 38% of **6c** and 28% of **7c** in 2 min, respectively. However, conversions increased to 72% of **6c** and 73% of **7c** after 60 min, respectively. N⁶-aminohexyl-AMP (**8a**) illustrates the limit of promiscuity of EbPPK2. In 2 min, only 1% conversion to the corresponding diphosphate (**8b**) was found, while **8c** was not detected. After 60 min, 6% conversion to **8b** and 6% to **8c** was observed. The C8-modified analogues were also tolerated well, with 8-bromo-AMP (**9a**) and 8-azido-AMP (**10a**) giving conversions of 67% (**9c**) and 70% (**10c**) in 2 min, respectively. Interestingly, the percent AMP remaining in the C8-modified reactions was higher than observed in other reactions after 2 minutes (~10% vs. ~6%). After 60 min, the percent AMP in these was reduced to 6%, with concomitant increase of the tetraphosphate product.

We then considered the utility of this process for producing sufficient quantities of modified ATP analogues to purify for further use. Therefore, we chose AMP (**1a**) and analogues **2a–5a** for preparative (10 μ mol) EbPPK2 reactions. Anion-exchange chromatography allowed excellent resolution of products, indicating 70–73% conversions to ATP analogues **1c–5c** (Fig. 3A and S45†). Precipitation of the ATP analogues with acetone resulted in isolated yields of 54–59% with excellent purity (Fig. 3B, C and S46–S50†). Isolated yields were determined *via* UV-absorbance, and the purity was examined

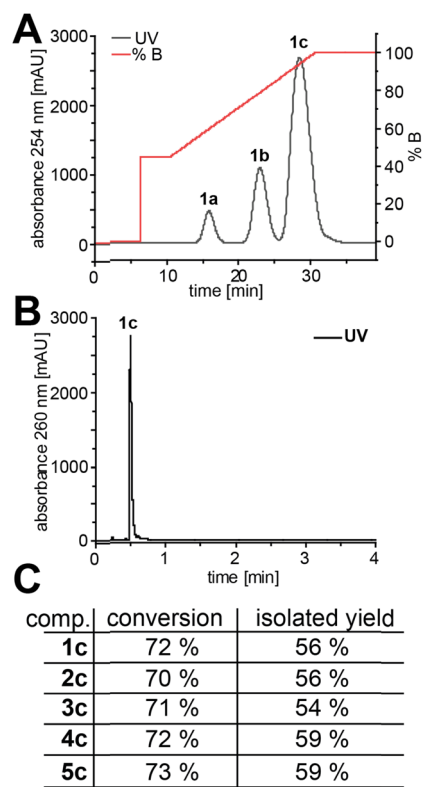


Fig. 3 Purification and purity of preparative EbPPK2 reactions (10 μ mol). (A) Chromatogram of the anion exchange chromatography exemplified for **1c** (for **2c–5c** see Fig. S45†). Performed on an ÄKTA purifier. Buffer A: ddH₂O; buffer B: 100 mM NaClO₄, pH 4.2. (B) Chromatogram of the LC-MS analysis of ATP (**1c**). The corresponding chromatograms and EIC for **1c–5c** can be found in Fig. S46–S50.† (C) Conversion and isolated yield for **1c–5c**.

via LC-MS. The separation observed suggested that this method would also allow purification of ADP analogues for other applications.

Poly(A) polymerase from *Saccharomyces cerevisiae* (ScPAP) can be used to incorporate ATP analogues at the 3' end of RNA in a template-independent fashion, but it does not appear to tolerate C8-modifications.²⁰ Therefore, we tested the incorporation of purified analogues **1c–5c** into 30-mer RNA using ScPAP (Fig. 4A). Under the conditions tested, ATP (**1c**) led to a poly(A) tail length distribution of 300–1000 nt (Fig. 4B), underscoring the quality of our ATP preparation. To our surprise, the C2-halogenated ATP analogues **2c** and **3c** were not only incorporated, but resulted in longer RNAs with distributions of 300 to >1000 nt and 400 to >1000 nt, respectively. The C2-amino-modified analogue **4c** led to poly(A) tails distributed around 90 nt and 120 nt. Finally, the N⁶-modified analogue **5c** was incorporated with a most common distribution around 125 nt. These results show that all the tested ATP analogues (**1c–5c**) can be incorporated by ScPAP, leading to hypermodified poly(A) tails of 90 to more than 1000 nt.

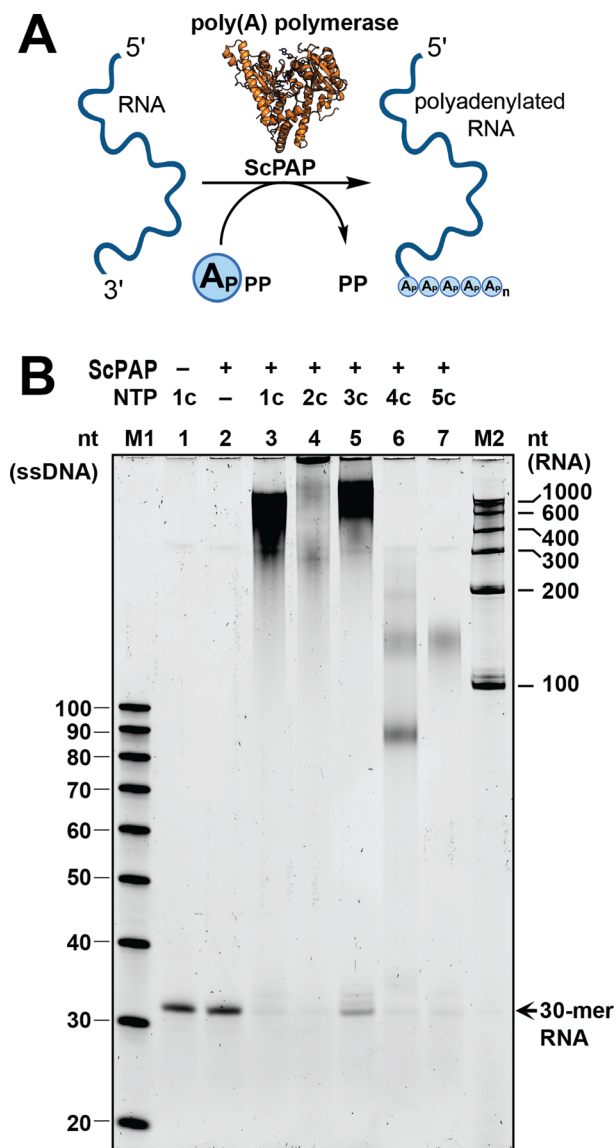


Fig. 4 Base-modified ATP analogues can be incorporated into RNA poly(A) tails using yeast poly(A) polymerase (ScPAP). (A) Scheme for uniform incorporation of ATP analogues. Created with <http://biorender.com/>. (B) 10% denaturing urea polyacrylamide gel (representative) of the polyadenylation reactions using ScPAP and compounds **1c–5c**. Replicates shown in Fig. S51.† RNA separation was performed for 1 h at 25 W, gel stained with SYBR Gold, and results visualised with a Typhoon scanner. M1: ssDNA Oligo Length standard 20/100 (Integrated DNA Technologies); M2: RiboRuler RNA Ladder Low Range (Thermo Fisher Scientific).

Conclusion

Here, we present a chemo-enzymatic strategy for the synthesis of base-modified ATP analogues. The 5'-monophosphorylation of adenosine analogues is performed by classical chemical synthesis, and the consecutive 5'-triphosphorylation is achieved by a single enzyme, EbPPK2, a polyphosphate kinase 2 class III enzyme. In comparison to other PPKs, the fast catalysis of EbPPK2 (5 s^{-1})^{13a} at mild conditions allow the reaction to

approach the equilibrium in 2 min at 30 °C in Tris buffer, pH 8 for most of the analogues tested. At analytical scale, conversions of 64–71% were achieved in 2 min for **1c–5c**, **9c**, and **10c**, while analogues **6c** and **7c** reached 72–73% conversion in 60 min. At preparative scale (10 μmol), 70–73% conversion to ATP analogues (**1c–5c**) was found, all of which could be purified and then incorporated into RNA using poly(A) polymerase. Next to ATP (**1c**), the C2-halogenated analogues **2c** and **3c** were incorporated best (>300 nt). Our chemo-enzymatic approach will help to further explore the chemical space of base-modified ATP analogues, currently limited by unreliable and inefficient 5'-phosphorylation strategies. This will benefit various areas of research, such as genetic code expansion, nucleotide-based inhibitors, RNA labelling and RNA therapeutics. Furthermore, incorporation into RNA by poly(A) polymerases paves the way for mRNAs with hypermodified poly(A) tails.

Data availability

The data supporting this article have been included as part of the ESI.†

Author contributions

R. M. M.-F., A. R and N. V. C. conceived the project. R. M. M.-F., J. E., H. S., R. R., M. E. and N. V. C. designed and performed experiments. R. R. and D. K. performed the structure analysis. All authors discussed the results. R. M. M.-F. and N. V. C. wrote the manuscript. All authors discussed and agreed on the final version.

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

This project has received funding from the German research council (DFG, KU 2531/6-1 to D. K. and RE 2796/10-1 to A. R.). R. M. M.-F. acknowledges support from the R.C. Good Faculty Fellowship (Denison University) and from the Denison University Research Foundation. M. E. is a member of CiM-IMPRS, the joint graduate school of the Cells-in-Motion Interfaculty Centre (CiM), University of Münster, Germany and the International Max Planck Research School – Molecular Biomedicine, Münster, Germany. The mass spectrometry and NMR facilities of the organic chemistry department of the University of Münster are gratefully acknowledged for analytical services. AlphaFold2 predictions were performed on the HPC cluster PALMA II of the University of Münster, subsidised by the DFG (INST 211/667-1).

References

- (a) K. Karikó, H. Muramatsu, F. A. Welsh, J. Ludwig, H. Kato, S. Akira and D. Weissman, *Mol. Ther.*, 2008, **16**, 1833–1840;



- (b) K. Karikó, M. Buckstein, H. Ni and D. Weissman, *Immunity*, 2005, **23**, 165–175.
- 2 (a) M. Brunderová, V. Havlíček, J. Matyašovský, R. Pohl, L. Poštová Slavětinská, M. Krömer and M. Hocek, *Nat. Commun.*, 2024, **15**, 3054; (b) M. Zhang, N. Singh, M. E. Ehmann, L. Zheng and H. Zhao, *iScience*, 2023, **26**, 107739.
- 3 (a) L. Anhäuser, S. Hüwel, T. Zobel and A. Rentmeister, *Nucleic Acids Res.*, 2019, **47**, e42; (b) K. J. Westerich, K. S. Chandrasekaran, T. Gross-Thebing, N. Kueck, E. Raz and A. Rentmeister, *Chem. Sci.*, 2020, **11**, 3089–3095; (c) A. Aditham, H. Shi, J. Guo, H. Zeng, Y. Zhou, S. D. Wade, J. Huang, J. Liu and X. Wang, *ACS Chem. Biol.*, 2022, **17**(12), 3352–3366; (d) H. Chen, D. Liu, J. Guo, A. Aditham, Y. Zhou, J. Tian, S. Luo, J. Ren, A. Hsu, J. Huang, F. Kostas, M. Wu, D. R. Liu and X. Wang, *Nat. Biotechnol.*, 2024, DOI: [10.1038/s41587-024-02174-7](https://doi.org/10.1038/s41587-024-02174-7).
- 4 (a) N. Klöcker, F. P. Weissenboeck and A. Rentmeister, *Chem. Soc. Rev.*, 2020, **49**, 8749–8773; (b) P. J. McCown, A. Ruszkowska, C. N. Kunkler, K. Breger, J. P. Hulewicz, M. C. Wang, N. A. Springer and J. A. Brown, *Wiley Interdiscip. Rev.: RNA*, 2020, **11**, e1595; (c) A. Cappannini, A. Ray, E. Purta, S. Mukherjee, P. Boccaletto, S. N. Moafinejad, A. Lechner, C. Barchet, B. P. Klaholz, F. Stefaniak and J. M. Bujnicki, *Nucleic Acids Res.*, 2024, **52**, D239–D244.
- 5 (a) S. M. Shepard, H. J. Jessen and C. C. Cummins, *J. Am. Chem. Soc.*, 2022, **144**, 7517–7530; (b) J. Y. Liao, S. Bala, A. K. Ngor, E. J. Yik and J. C. Chaput, *J. Am. Chem. Soc.*, 2019, **141**, 13286–13289.
- 6 (a) C. Domnick, F. Eggert and S. Kath-Schorr, *Chem. Commun.*, 2015, **51**, 8253–8256; (b) F. Eggert and S. Kath-Schorr, *Chem. Commun.*, 2016, **52**, 7284–7287; (c) M. B. Walunj, A. A. Tanpure and S. G. Srivatsan, *Nucleic Acids Res.*, 2018, **46**, e65; (d) S. G. Srivatsan and Y. Tor, *J. Am. Chem. Soc.*, 2007, **129**, 2044–2053; (e) P. Rosenqvist, J. J. Mäkinen, K. Palmu, J. Jokinen, R. K. Prajapati, H. J. Korhonen, P. Virta, G. A. Belogurov and M. Metsä-Ketelä, *Eur. J. Med. Chem.*, 2022, **237**, 114342; (f) C. P. M. Scheitl, M. Ghaem Maghami, A. K. Lenz and C. Höbartner, *Nature*, 2020, **587**, 663–667; (g) A. Hegelein, D. Müller, S. Gröbl, M. Göbel, M. Hengesbach and H. Schwalbe, *Chem.–Eur. J.*, 2020, **26**, 1800–1810.
- 7 M. Hocek and M. Fojta, *Org. Biomol. Chem.*, 2008, **6**, 2233.
- 8 (a) M. Brunderová, M. Krömer, M. Vlková and M. Hocek, *Angew. Chem., Int. Ed.*, 2023, **62**, e202213764; (b) C. Creech, M. Kanaujia and C. P. Causey, *Org. Biomol. Chem.*, 2015, **13**, 8550–8555; (c) N. Milisavljević, P. Perlíková, R. Pohl and M. Hocek, *Org. Biomol. Chem.*, 2018, **16**, 5800–5807.
- 9 I. Gillerman, J. Lecka, L. Simhaev, M. N. Munkonda, M. Fausther, M. Martín-Satué, H. Senderowitz, J. Sévigny and B. Fischer, *J. Med. Chem.*, 2014, **57**, 5919–5934.
- 10 M. Yoshikawa, T. Kato and T. Takenishi, *Tetrahedron Lett.*, 1967, **50**, 5065–5068.
- 11 (a) P. Benčić, M. Keppler, M. Kuge, D. Qiu, L. M. Schütte, M. Häner, K. Strack, H. J. Jessen, J. N. Andexer and C. Loenarz, *FEBS J.*, 2023, **290**, 4899–4920; (b) J. H. Overbeck, J. Vögele, F. Nussbaumer, E. Duchardt-Ferner, C. Kreutz, J. Wöhnert and R. Sprangers, *Angew. Chem., Int. Ed.*, 2023, **62**, e202218064; (c) J. Frisch, T. Marsic and C. Loderer, *Biomolecules*, 2021, **11**, 346; (d) J. A. McIntosh, Z. Liu, B. M. Andresen, N. S. Marzijarani, J. C. Moore, N. M. Marshall, M. Borrarske, J. V. Obligacion, P. S. Fier, F. Peng, J. H. Forstater, M. S. Winston, C. An, W. Chang, J. Lim, M. A. Huffman, S. P. Miller, F. R. Tsay, M. D. Altman, C. A. Lesburg, D. Steinhuebel, B. W. Trotter, J. N. Cumming, A. Northrup, X. Bu, B. F. Mann, M. Biba, K. Hiraga, G. S. Murphy, J. N. Koley, A. Makarewicz, W. Pan, I. Farasat, R. S. Bade, K. Stone, D. Duan, O. Alvizo, D. Adressa, E. Guetschow, E. Hoyt, E. L. Regalado, S. Castro, N. Rivera, J. P. Smith, F. Wang, A. Crespo, D. Verma, S. Axnanda, Z. E. X. Dance, P. N. Devine, D. Tschaen, K. A. Canada, P. G. Bulger, B. D. Sherry, M. D. Truppo, R. T. Ruck, L. C. Campeau, D. J. Bennett, G. R. Humphrey, K. R. Campos and M. L. Maddess, *Nature*, 2022, **603**, 439–444.
- 12 (a) B. Nocek, S. Kochinyan, M. Proudfoot, G. Brown, E. Evdokimova, J. Osipiuk, A. M. Edwards, A. Savchenko, A. Joachimiak and A. F. Yakunin, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 17730–17735; (b) B. P. Nocek, A. N. Khusnutdinova, M. Ruszkowski, R. Flick, M. Burda, K. Batyrova, G. Brown, A. Mucha, A. Joachimiak, Ł. Berlicki and A. F. Yakunin, *ACS Catal.*, 2018, **8**, 10746–10760; (c) A. E. Parnell, S. Mordhorst, F. Kemper, M. Giurrandino, J. P. Prince, N. J. Schwarzer, A. Hofer, D. Wohlwend, H. J. Jessen, S. Gerhardt, O. Einsle, P. C. F. Oyston, J. N. Andexer and P. L. Roach, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 3350–3355.
- 13 (a) M. Tavanti, J. Hosford, R. C. Lloyd and M. J. B. Brown, *Green Chem.*, 2021, **23**, 828–837; (b) M. Becker, I. Nowak, K. Hildebrand, S. Lütz and K. Rosenthal, *Catal. Sci. Technol.*, 2024, **14**, 3335–3345.
- 14 (a) K. Motomura, R. Hirota, M. Okada, T. Ikeda, T. Ishida and A. Kuroda, *Appl. Environ. Microbiol.*, 2014, **80**, 2602–2608; (b) T. Querengässer, J. Wenzlaff and C. Loderer, *ChemCatChem*, 2024, e202400181; (c) L. Eltoukhy and C. Loderer, *ChemBioChem*, 2022, **23**, e202100596.
- 15 M. Keppler, S. Moser, H. J. Jessen, C. Held and J. N. Andexer, *Beilstein J. Org. Chem.*, 2022, **18**, 1278–1288.
- 16 (a) J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583–589; (b) M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Židek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis and S. Velankar, *Nucleic Acids Res.*, 2021, **50**, D439–D444.



- 17 R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Židek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper and D. Hassabis, *bioRxiv*, 2022, preprint, DOI: [10.1101/2021.10.04.463034](https://doi.org/10.1101/2021.10.04.463034).
- 18 (a) T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris and T. E. Ferrin, *Protein Sci.*, 2018, **27**, 14–25; (b) E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris and T. E. Ferrin, *Protein Sci.*, 2021, **30**, 70–82.
- 19 E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker and K. Karplus, *Proteins*, 2009, **77**(Suppl 9), 114–122.
- 20 L. S. Chen and T. L. Sheppard, *J. Biol. Chem.*, 2004, **279**, 40405–40411.

