

Cite this: *Chem. Sci.*, 2023, 14, 12160

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

Biocatalytic reductive amination as a route to isotopically labelled amino acids suitable for analysis of large proteins by NMR†

Jack S. Rowbotham,^{‡*} Jake H. Nicholson,[‡] Miguel A. Ramirez,[‡] Kouji Urata,^a Peter M. T. Todd,^a Gogulan Karunanithy,[‡] Lars Lauterbach,[‡] Holly A. Reeve,[‡] Andrew J. Baldwin,[‡] and Kylie A. Vincent^{‡*}

We demonstrate an atom-efficient and easy to use H₂-driven biocatalytic platform for the enantioselective incorporation of ²H-atoms into amino acids. By combining the biocatalytic deuteration catalyst with amino acid dehydrogenase enzymes capable of reductive amination, we synthesised a library of multiply isotopically labelled amino acids from low-cost isotopic precursors, such as ²H₂O and ¹⁵NH₄⁺. The chosen approach avoids the use of pre-labelled ²H-reducing agents, and therefore vastly simplifies product cleanup. Notably, this strategy enables ²H, ¹⁵N, and an asymmetric centre to be introduced at a molecular site in a single step, with full selectivity, under benign conditions, and with near 100% atom economy. The method facilitates the preparation of amino acid isotopologues on a half-gram scale. These amino acids have wide applicability in the analytical life sciences, and in particular for NMR spectroscopic analysis of proteins. To demonstrate the benefits of the approach for enabling the workflow of protein NMR chemists, we prepared L-[α-²H,¹⁵N, β-¹³C]-alanine and integrated it into a large (>400 kDa) heat-shock protein oligomer, which was subsequently analysable by methyl-TROSY techniques, revealing new structural information.

Received 3rd April 2023

Accepted 20th September 2023

DOI: 10.1039/d3sc01718d

rsc.li/chemical-science

Introduction

Selective isotopic labelling is widely used for probing chemical and biochemical systems.^{1,2} Deuterium (²H) is a particularly attractive substitution as the heavy isotope of hydrogen can be installed into a wide range of molecular sites to facilitate analysis by a suite of techniques.³ Notably, the development of

isotopically labelled amino acids has been transformative for the analysis of proteins, enabling atomic resolution mechanistic, dynamical, and structural studies *via* a range of experimental techniques.^{3–7} Solution NMR spectroscopy particularly benefits selective isotopic enrichment, where the combination of carefully matched labelling schemes and experiments can greatly enable characterisation of complexes of molecular weight approaching *ca.* 1 MDa,⁸ greatly exceeding what is conventionally thought to be possible. There is therefore a need for straightforward synthetic methods to routinely prepare the necessary isotopically enriched precursors for such studies. Similarly, the emergence of deuterium-stabilised drug compounds⁹ and therapeutic peptides¹⁰ also places demand on new synthetic routes to isotopically labelled amino acids.

From a synthetic standpoint, chiral-deuteration of target precursors is challenging. Chemical strategies to prepare α-deuterated amino acids often rely on precious metals and expensive chiral ligands or auxiliaries.^{11–16} Such approaches can suffer from imperfect selectivity, diminished isotopic purities, and complex work-up procedures, which can make the product cost prohibitively high for structural biology research. As an alternative to chemo-routes, several biocatalytic strategies have also been demonstrated for preparing α-deuterated amino acids.^{17–21} These methods benefit from the mild reaction conditions and inherent selectivity associated with enzyme

^aDepartment of Chemistry, University of Oxford, Inorganic Chemistry Laboratory, South Parks Road, Oxford, UK. E-mail: jack.rowbotham@manchester.ac.uk; kylie.vincent@chem.ox.ac.uk

^bDepartment of Chemistry, University of Oxford, Physical and Theoretical Chemistry Laboratory, Oxford, UK

^cTechnische Universität Berlin, Institute for Chemistry, Straße des 17. Juni 135, 10437 Berlin, Germany

^dKavli Institute for Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, UK. E-mail: andrew.baldwin@chem.ox.ac.uk

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3sc01718d>

‡ Current address: Department of Chemistry, University of Manchester, Manchester Institute of Biotechnology, Manchester, UK.

¶ Current address: Department of Structural and Molecular Biology, Division of Biosciences, University College London, Darwin Building, Gower Street, WC1E 6BT, London, UK.

|| Current address: RWTH Aachen University iAMB – Institute of Applied Microbiology Worringer Weg 1, 52074 Aachen, Germany.

§ Current address: HydRegen Limited, Centre for Innovation and Enterprise, Begbroke Science Park, Oxford, OX5 1PF, UK.



Results and discussion

Verification of H₂-driven biocatalytic system for producing a library of isotopically labelled amino acids

We have previously demonstrated that *ReSH* is capable of simultaneous reduction and isotopic labelling of NAD⁺ in ²H₂O under low pressures of H₂.^{37,38} Here, the action of the *ReSH* was coupled to that of an *L*-AlaDH in order to drive the reductive

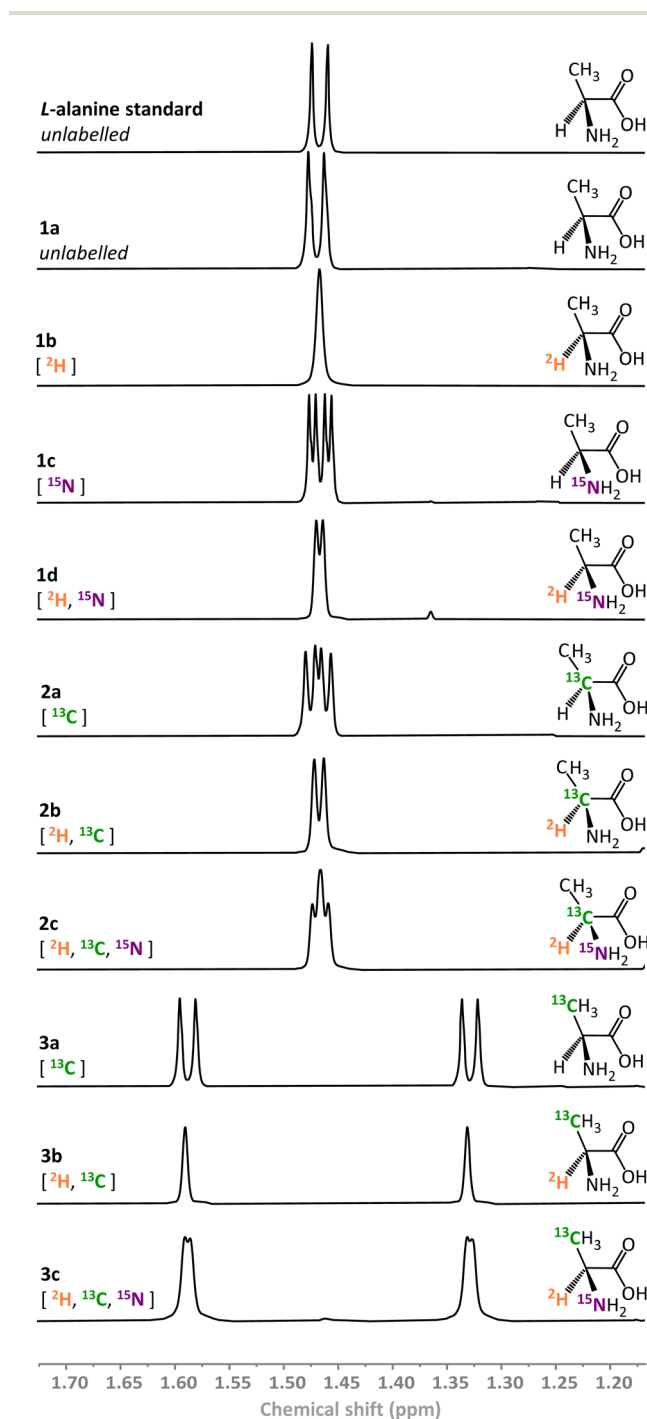


Fig. 3 ¹H NMR spectra of variously isotopically labelled *L*-alanine molecules prepared biocatalytically (500 MHz, 293 K, ²H₂O, p²H 8.0).

amination of pyruvate in the presence of ammonium bicarbonate (acting as both buffer and amine source). We selected *L*-alanine as the target compound, owing to the well-established strategies for up-take of this amino acid by *E. coli* and incorporation into proteins for NMR structure determination.²

Using the combination of *ReSH* and *L*-AlaDH it was possible to generate a library of ten *L*-alanine isotopologues (compounds **1a-d**, **2a-c**, **3a-c** in Fig. 3) simply by varying the solvent (H₂O/²H₂O), ammonium salt (¹⁴NH₄⁺/¹⁵NH₄⁺), and the pyruvate substrate (¹²C/¹³C). In each case, the ¹H signals from the methyl peak of the alanine (–CH₃, δ = 1.46 ppm) in the corresponding NMR spectra could be used to verify the product (full characterisation given in SI Section S.3.1†). Within the limits of detection, only a single isotopologue was formed in each case, verifying the high selectivity of the biocatalytic system, even in the presence of unlabelled H₂ gas. It was found that the commercial *L*-AlaDH used to generate this library led to an enantiomeric excess (ee) of 90%, owing to the presence of a contaminating racemase. As such, a purified *L*-AlaDH (produced in house, see ESI Section S.2.2†) was used for all subsequent experiments, and delivered the expected >99% ee.

In addition to the experiments with *L*-AlaDH, the H₂-driven system was also coupled to commercial *L*-leucine dehydrogenase (*L*-LeuDH) and *L*-phenylalanine dehydrogenase (*L*-PheDH) enzymes. Using this system, it was straightforward to prepare *L*-[α-²H, ¹⁵N]-leucine (**4**) and *L*-[α-²H, ¹⁵N, β-²H₂]-phenylalanine (**5**) with similarly high levels of conversion and isotopic and enantioselectivity (see Fig. 4 and ESI Section S.3.1†).

Scale-up of H₂-driven synthesis of α-deuterated amino acids

Following initial demonstration of the H₂-driven biocatalytic system for preparing multiply labelled amino acid isotopomers, the reaction was tested on a preparative scale. Initially, 50 mg of **1d** was prepared by reductive amination of pyruvate (25 mM) in ²H₂O (25 mL, p²H 8.0) in a H₂ atmosphere (2 bar) at room temperature, in the presence of [¹⁵N]H₄HCO₃ (50 mM), with *ReSH*, *L*-AlaDH, and NAD⁺ (0.1 mM). The role of the [¹⁵N]H₄HCO₃ was to serve as both an amine source and volatile

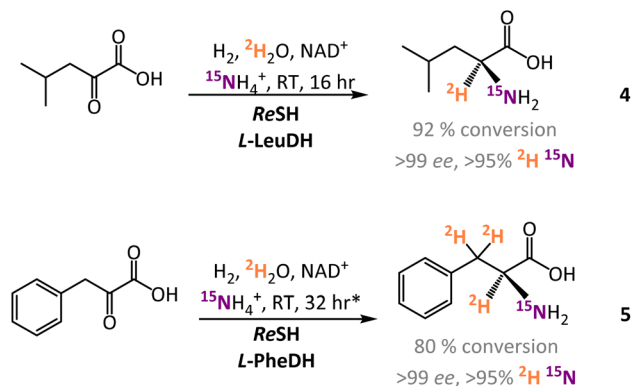


Fig. 4 Preparation of ²H and ¹⁵N-labelled *L*-leucine and *L*-phenylalanine by H₂ driven biocatalytic reductive amination and deuteration. *For **5**, β-²H atoms were incorporated by pre-incubation of the substrate in ²H₂O overnight before the enzymes were added.



buffer, enabling straightforward isolation of the product by rotary evaporation following the reaction. The reaction reached >99% conversion after 20 h, indicating a turnover number (ToN) for the *ReSH* of over 130 000.

This initial scaled-up reaction was set up anaerobically in a pressure-vessel, which may not be available in all laboratories. Therefore, in order to probe the wider applicability of the *ReSH* as a deuteration catalyst, a similar reaction to synthesise **3c** was set up on the bench in standard laboratory glassware. Here, [$3\text{-}^{13}\text{C}$]-pyruvate (35 mM) was reacted in a stirred solution of $^2\text{H}_2\text{O}$ (65 mL, $p^2\text{H}$ 8.0) in a round-bottom flask with a balloon of H_2 , under otherwise identical conditions. Again, this reaction proceeded to full conversion, and enabled the isolation of 198 mg of product.

Finally, the reaction conditions were further intensified to increase the amount of isolated product. Here, the synthesis of **1b** was performed using 100 mM pyruvate in 60 mL $^2\text{H}_2\text{O}$ with 150 mM NH_4HCO_3 , again in a round-bottom flask. Whilst full conversion was observed after 19 hours, giving rise to 520 mg of product (95% isolated yield), a slight decrease in ee was observed (97%) under the higher loading conditions. In all cases however, the isotopic purity of the isolated products (**1d**, **1b**, **3c**) closely followed the isotopic purity of the starting reaction mixture and simple precautions (such as washing the enzymes in $^2\text{H}_2\text{O}$) were sufficient to keep the $\% ^2\text{H}$ high (>95%).

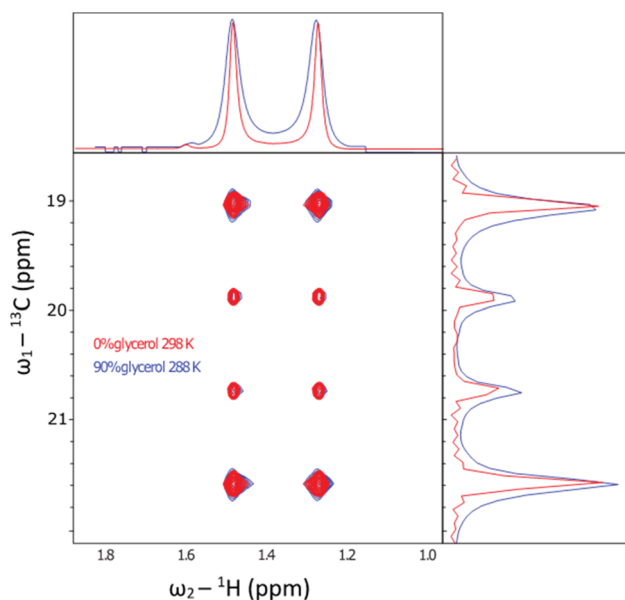


Fig. 5 In order to simulate the rotational environment of the labelled amino acid inside a protein, L-[$\alpha\text{-}^2\text{H}$, ^{15}N , $\beta\text{-}^{13}\text{C}$]-alanine was dissolved in a viscous glycerol medium and studied by NMR. The figure shows the ^{13}C - ^1H decoupled HSQC experiment of the sample in $^2\text{H}_2\text{O}$ at 298 K (red) and 90 vol% glycerol at 288 K (blue trace). Under 0 vol% glycerol, the ratio of intensities is 3 : 1 : 1 : 3, as expected for a small molecule. At 90 vol% glycerol and low temperature, the resonances remained intense. The intensity ratio changed to approximately 2 : 1 : 1 : 2. Due to the outer resonances having $9\times$ faster relaxation than the inner lines, in the macromolecular limit when the molecule tumbles slowly. Although the effects of slow tumbling are clear, the resonances remain sharp indicating the utility of this reagent inside high molecular weight proteins.

Engineering isotopically enriched proteins with triply labelled L-alanine (**3c**) for methyl-TROSY

A specific application that stands to benefit tremendously from these methods is biomolecular NMR spectroscopy. Dipolar interactions between hydrogens dominate relaxation in proteins, where the slower they tumble, the broader the detected resonances in spectra. Typically, when a protein exceeds around 30 kDa, resonances are then unsuitable for study using conventional solution-state NMR methods.³⁹ To overcome this limit, a combination of deuteration, and transverse relaxation optimised spectroscopy (TROSY) methods can be applied. Notably, the 'methyl-TROSY' methods developed by Kay *et al.*^{40,41} allow structural and



Fig. 6 (a) The X-ray crystallographic structure of the small heat shock protein (HSP) from *M. jannaschii*, (PDB id 1SHS).³⁴ The structure is built from dimeric sub-units, and the expected alanine $^{13}\text{CH}_3$ groups are shown. The homo 24-mer has a molecular weight of 396 kDa. (b) Alanine region of methyl spectrum revealing 12 clear resonances with a range of linewidths. This contrasts with the X-ray structure that anticipates 8 unique environments. In solution, the complex adopts multiple conformations, and a range of environments with different motional regimes. This spectrum demonstrates that L-alanine produced using our synthesis scheme can be readily incorporated into high molecular weight biomolecules for mechanistic and structural studies.



dynamical characterisation of proteins approaching a molecular weight of 1 MDa,^{8,42} a limit which encompasses the majority of known proteins (excluding oligomers).⁴³ This requires the selective incorporation of amino acids into proteins that contain ¹³C¹H₃ methyl groups, which are otherwise deuterated. Having isolated gram quantities of L-[α -²H, ¹⁵N, β -¹³C]-alanine (**3c**), we sought to ascertain its suitability for solution-state methyl-TROSY experiments.

Initially, a decoupled ¹³C-¹H HSQC spectra of the methyl group of the sample were recorded at 298 K in ²H₂O (Fig. 5, red trace), yielding a pair of quartets with signal intensity 3 : 1 : 1 : 3 in the ¹³C dimension, as expected for a small molecule where we detect an anti-phase coherence (which would be in the ratio 1 : 3 : 3 : 1 in a ¹³C 1D spectrum detecting in-phase coherence). In a spectrum acquired at 288 K in 90 vol% glycerol (Fig. 5, blue trace), the spectra were highly similar. Under these conditions, the viscosity was increased by a factor of approximately 450, effectively simulating the local tumbling behaviour of the molecule as if it were part of a 50 kDa protein. This demonstrates that we expect sharp signals from this amino acid, even when incorporated in high molecular weight complexes. We therefore sought to incorporate this in a suitably chosen protein. We incorporated L-[α -²H, ¹⁵N, β -¹³C]-alanine (**3c**) into the small heat shock protein Hsp16.5. This protein is a molecular chaperone from *Methanococcus jannaschii* that assembles into a 24-mer of molecular weight 396 kDa (Fig. 6a).⁴⁴ *In vitro*, this protein is a potent inhibitor of protein aggregation and amyloid formation. L-[α -²H, ¹⁵N, β -¹³C]-Alanine (**3c**, 100 mg) was added 1 hour prior to induction, and the incorporation was >99%. A ¹³C-¹H HMQC NMR spectrum was acquired. 12 resonances were clearly observed, with a range of linewidths, indicating differential mobility within different parts of the complex. As 8 unique resonances were expected from the symmetric structure characterised by X-ray crystallography, our precursor reveals multiple conformations of the protein (Fig. 6b). This spectrum validates the utility of this precursor in future biochemical studies. This is among the largest proteins for which an alanine methyl-TROSY spectrum has been recorded.^{4,42}

Conclusions and outlook

We have developed a convenient means for the production of selectively deuterated amino acids. We demonstrated the success of the method in preparing variously ²H, ¹⁵N, and ¹³C-labelled L-alanines, and further showed that the principle behind the technique is applicable to a wider range of amino acids. Our scheme exploits enzymes that can be produced with simple molecular biology strategies in most laboratories, uses easily implemented reaction conditions and workup procedures, and cheap and readily available isotopic precursors. The technique installs the ²H and ¹⁵N labels simultaneously with the formation of the chiral centre, and enables wide versatility in the choice of isotopes added. We anticipate this approach facilitating rapid access to arrays of variously labelled canonical and non-canonical amino acids. This in turn will enable mechanistic, structural, and biophysical studies of large

biomolecules such as proteins and peptides in biologically relevant environments.

Data availability

The datasets supporting this article have been uploaded as part of the ESI.†

Author contributions

J. S. R., K. A. V., H. A. R., and A. J. B. jointly conceptualised and supervised the work. J. S. R., J. H. N., K. U., and P. M. T. T. synthesised amino acids. M.A.R. performed molecular biology. G. K. and A. J. B. carried out protein NMR. All authors contributed to data analysis and assisted with the drafting of the manuscript.

Conflicts of interest

A patent application by J. S. R., H. A. R. and K. A. V. detailing some of this research was filed through Oxford University Innovation (Feb 2018). The remaining authors declare no competing interests.

Acknowledgements

The research of K. A. V., H. A. R., J. S. R., M. A. R., K. U. and J. H. N. was supported by Engineering and Physical Sciences Research Council (EPSRC) IB Catalyst award EP/N013514/1 (to K. A. V. and H. A. R.) and ERC-2018-CoG BiocatSusChem 819580 to K. A. V. In addition, J. S. R. was supported by an EPA Cephalosporin Junior Research Fellowship at Linacre College, Oxford and BBSRC Discovery Fellowship (BB/V005464/1). L. L. acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy – EXC 2186 – 390919832 – “The Fuel Science Center”, and Germany's Excellence Strategy – EXC 2008 – 390540038 – “UniSysCat”. A. J. B. was supported by ERC Grant No. (101002859). We thank Prof O. Lenz (TU Berlin) for his gift of the *Ralstonia eutropha* strain HF904. The authors are grateful to A. Tam, Dr K. Ebrahimi and Dr J. Walsby-Tickle (U. Oxford) for assistance with high-resolution mass spectrometry (HRMS), and Dr L. Pirvu (U. Manchester) for GC.

References

- Z. Kelman, *Isotope Labeling of Biomolecules – Applications*, Elsevier, New York, 2016, vol. 566.
- G. W. Becker, *Briefings Funct. Genomics Proteomics*, 2008, 7, 371–382.
- J. Atzrodt, V. Derdau, W. J. Kerr and M. Reid, *Angew. Chem., Int. Ed.*, 2018, 57, 1758–1784.
- M. Kainosho, Y. Miyanoiri and M. Takeda, in *Experimental Approaches of NMR Spectroscopy*, Springer Singapore, Singapore, 2018, pp. 37–61.
- R. Ashkar, H. Z. Bilheux, H. Bordallo, R. Briber, D. J. E. Callaway, X. Cheng, X.-Q. Chu, J. E. Curtis,



- M. Dadmun, P. Fenimore, D. Fushman, F. Gabel, K. Gupta, F. Herberle, F. Heinrich, L. Hong, J. Katsaras, Z. Kelman, E. Kharlampieva, G. R. Kneller, A. Kovalevsky, S. Krueger, P. Langan, R. Lieberman, Y. Liu, M. Losche, E. Lyman, Y. Mao, J. Marino, C. Mattos, F. Meilleur, P. Moody, J. D. Nickels, W. B. O'Dell, H. O'Neill, U. Perez-Salas, J. Peters, L. Petridis, A. P. Sokolov, C. Stanley, N. Wagner, M. Weinrich, K. Weiss, T. Wymore, Y. Zhang and J. C. Smith, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2018, **74**, 1129–1168.
- 6 M. Haertlein, M. Moulin, J. M. Devos, V. Laux, O. Dunne and V. Trevor Forsyth, *Methods Enzymol.*, 2016, **566**, 113–157.
- 7 H. El Mkami and D. G. Norman, *Methods Enzymol.*, 2015, **564**, 125–152.
- 8 A. Velyvis, A. M. Ruschak and L. E. Kay, *PLoS One*, 2012, **7**, e43725.
- 9 T. Pirali, M. Serafini, S. Cargnin and A. A. Genazzani, *J. Med. Chem.*, 2019, **62**, 5276–5297.
- 10 L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang and C. Fu, *Signal Transduction Targeted Ther.*, 2022, **71**(7), 1–27.
- 11 C. Taglang, D. E. Korenchan, C. Von Morze, J. Yu, C. Najac, S. Wang, J. E. Blecha, S. Subramaniam, R. Bok, H. F. Vanbrocklin, D. B. Vigneron, S. M. Ronen, R. Sriram, J. Kurhanewicz, D. M. Wilson and R. R. Flavell, *Chem. Commun.*, 2018, **54**, 5233–5236.
- 12 R. Sakuta, K. Takeda, K. Igarashi, H. Ohno and N. Nakamura, *Energy Technol.*, 2018, **6**, 273–279.
- 13 S. Bhatia, G. Spahlinger, N. Boukhumseen, Q. Boll, Z. Li and J. E. Jackson, *Eur. J. Org. Chem.*, 2016, **2016**, 4230–4235.
- 14 C. Taglang, L. M. Martínez-Prieto, I. del Rosal, L. Maron, R. Poteau, K. Philippot, B. Chaudret, S. Perato, A. Sam Lone, C. Puente, C. Dugave, B. Rousseau and G. Pieters, *Angew. Chem., Int. Ed.*, 2015, **54**, 10474–10477.
- 15 K. Moozeh, S. M. So and J. Chin, *Angew. Chem., Int. Ed.*, 2015, **54**, 9381–9385.
- 16 H. Ohtsuki, M. Takashima, T. Furuta and T. Kawabata, *Tetrahedron Lett.*, 2018, **59**, 1188–1191.
- 17 U. M. Babu and R. B. Johnston, *Biochem. Biophys. Res. Commun.*, 1974, **58**, 460–466.
- 18 C. H. Wong and G. M. Whitesides, *J. Am. Chem. Soc.*, 1983, **105**, 5012–5014.
- 19 J. A. Chanatry, P. H. Schafer, M. S. Kim and D. M. Lemaster, *Anal. Biochem.*, 1993, **213**, 147–151.
- 20 J. Feeney, B. Birdsall, G. Ostler, M. D. Carr and M. Kairi, *FEBS Lett.*, 1990, **272**, 197–199.
- 21 N. G. Faleev, S. B. Ruvinov, M. B. Saporovskaya, V. M. Belikov, L. N. Zakomyrdina, I. S. Sakharova and Y. M. Torchinskii, *Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1989, **38**, 2154–2156.
- 22 J. Atzrodt, V. Derdau, W. J. Kerr and M. Reid, *Angew. Chem., Int. Ed.*, 2018, **57**, 3022–3047.
- 23 M. Lepron, M. Daniel-Bertrand, G. Mencia, B. Chaudret, S. Feuillastre and G. Pieters, *Acc. Chem. Res.*, 2021, **54**, 1465–1480.
- 24 S. Kopf, F. Bourriquen, W. Li, H. Neumann, K. Junge and M. Beller, *Chem. Rev.*, 2022, **122**, 6634–6718.
- 25 S. W. Chun and A. R. H. Narayan, *ACS Catal.*, 2020, **10**, 7413–7418.
- 26 T. J. Doyon and A. R. Buller, *J. Am. Chem. Soc.*, 2022, **144**, 7327–7336.
- 27 A. Michelotti, F. Rodrigues and M. Roche, *Org. Process Res. Dev.*, 2017, **21**, 1741–1744.
- 28 M. Valero, R. Weck, S. Güssregen, J. Atzrodt and V. Derdau, *Angew. Chem., Int. Ed.*, 2018, **57**, 8159–8163.
- 29 M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono and P. Güntert, *Nature*, 2006, **440**, 52–57.
- 30 H. C. Lo and R. H. Fish, *Angew. Chem., Int. Ed.*, 2002, **41**, 478–481.
- 31 N. M. Kelly, B. C. O'Neill, J. Probert, G. Reid, R. Stephen, T. Wang, C. L. Willis and P. Winton, *Tetrahedron Lett.*, 1994, **35**, 6533–6536.
- 32 N. M. Kelly, R. G. Reid, C. L. Willis and P. L. Winton, *Tetrahedron Lett.*, 1996, **37**, 1517–1520.
- 33 J. H. Schrittwieser, S. Velikogne and W. Kroutil, *Catal*, 2018, **8**, 205.
- 34 A. Galkin, L. Kulakova, T. Yoshimura, K. Soda and N. Esaki, *Appl. Environ. Microbiol.*, 1997, **63**, 4651–4656.
- 35 S. Velasco-Lozano, E. S. da Silva, J. Llop and F. López-Gallego, *ChemBioChem*, 2018, **19**, 395–403.
- 36 J. S. Rowbotham, M. A. Ramirez, O. Lenz, H. A. Reeve and K. A. Vincent, *Nat. Commun.*, 2020, **11**, 1454.
- 37 L. A. Thompson, J. S. Rowbotham, J. H. Nicholson, M. A. Ramirez, C. Zor, H. A. Reeve, N. Grobert and K. A. Vincent, *ChemCatChem*, 2020, **12**, 3913–3918.
- 38 A. Al-Shameri, M. Petrich, K. Junge Puring, U. Apfel, B. M. Nestl and L. Lauterbach, *Angew. Chem., Int. Ed.*, 2020, **59**, 10929–10933.
- 39 V. Tugarinov, P. M. Hwang and L. E. Kay, *Annu. Rev. Biochem.*, 2004, **73**, 107–146.
- 40 J. E. Ollerenshaw, V. Tugarinov and L. E. Kay, *Magn. Reson. Chem.*, 2003, **41**, 843–852.
- 41 R. L. Isaacson, P. J. Simpson, M. Liu, E. Cota, X. Zhang, P. Freemont and S. Matthews, *J. Am. Chem. Soc.*, 2007, **129**, 15428–15429.
- 42 T. Yuwen, R. Huang, P. Vallurupalli and L. E. Kay, *Angew. Chem., Int. Ed.*, 2019, 6250–6254.
- 43 L. Brocchieri and S. Karlin, *Nucleic Acids Res.*, 2005, **33**, 3390–3400.
- 44 K. K. Kim, R. Kim and S. H. Kim, *Nature*, 1998, **394**, 595–599.

