Combined biocatalytic reductive amination and deuteration to prepare isotopically labelled amino acids for NMR analysis of large proteins

Jack S. Rowbotham,[a,b,*] Jake H. Nicholson,[a] Miguel A. Ramirez,[a,c] Kouji Urata,[a] Peter M.T. Todd,[a] Gogulan Karunanithy,[d,e] Lars Lauterbach,[f,g] Holly A. Reeve,[a,c] Andrew J. Baldwin,[d,*] Kylie A. Vincent [a,*]

[(a)] Department of Chemistry, University of Oxford, Inorganic Chemistry Laboratory, South Parks Road, Oxford, UK. [(b)] Current address: Department of Chemistry, University of Manchester, Manchester Institute of Biotechnology, Manchester, UK. [(c)] Current address: HydRegen Limited, Centre for Innovation and Enterprise, Begbroke Science Park, Oxford, OX5 1PF, UK. [(d)] Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, Oxford, UK. [(e)] Current address: Department of Structural and Molecular Biology, Division of Biosciences, University College London, Darwin Building, Gower Street, WC1E 6BT, London, UK. [(f)] Technische Universität Berlin, Institute for Chemistry, Straße des 17. Juni 135, 10437 Berlin, Germany. [(g)] Current address: RWTH Aachen University iAMB – Institute of Applied Microbiology Worringer Weg 1, 52074 Aachen, Germany.

*Authors for correspondence: jack.rowbotham@manchester.ac.uk; andrew.baldwin@chem.ox.ac.uk; kylie.vincent@chem.ox.ac.uk

Biocatalysis offers many advantages for selective isotopic labelling of valuable small molecules, such as the deuterated amino acids utilised in protein NMR. Until recently, applications of biocatalytic deuteration systems have been restricted by their requirement for a supply of super-stoichiometric quantities of a specifically labelled ^{2}H-pre-cursor, which can be both costly to purchase and complex to prepare. Overcoming this hurdle, we have demonstrated a novel and easy to use H_2-driven biocatalytic platform for the incorporation of ^{2}H-atoms across a number of molecular functional groups. By combining the biocatalytic deuteration catalyst with enzymes capable of reductive amination, we synthesised a suite of multiply isotopically labelled amino acids from low-cost isotopic precursors, such as $^{2}H_2O$ and $^{15}NH_4^+$. Notably, this strategy enables the introduction of a ^{15}N-label, ^{2}H-label, and chiral centre all in a single-step, and gives rise to amino acid isotopologues on a half gram scale for use directly in the preparation of isotopically labelled proteins. To demonstrate the applicability of the approach in the workflow of protein NMR chemists, we prepared *L*-[*α*-^{2}H,^{15}N, *β*-^{13}C]-alanine and integrated it into a large (> 400 kDa) heat-shock protein, which was subsequently analysable by Methyl-TROSY techniques, revealing new structural information.

Introduction

Selective isotopic labelling is a commonly adopted strategy for probing chemical and biochemical systems.[1,2] Deuterium (^{2}H) is a particularly attractive substitution as it can often 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 mechanistic, dynamical, and structural studies.[4]

Labelled proteins can be obtained by feeding microorganisms such as *E.coli* with small molecule pre-cursors bearing the desired isotopic labelling patterns,[2,4,5] such as *α*-deuterated *L*-alanine. This approach has been particularly successful in the NMR spectroscopy of large proteins.[6] As molecules increase in size, they tumble more slowly, leading to broadening of resonances in their NMR spectra. Typically, when a protein exceeds around 30 kDa, resonances are unsuitable for study using conventional NMR methods.[7] By deuterating the protein and introducing specifically isotopically-enriched amino acids, the molecular weight of proteins that can be studied is vastly increased. Notably, the 'methyl-TROSY' (transverse relaxation optimised spectroscopy) methods developed by Kay *et al.*[8,9] allow structural and dynamical characterisation of proteins approaching a molecular weight of 1 MDa,[6,10] a limit which encompasses the majority of known proteins.[11]

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 auxillaries.[12–17] 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.[18–22] These methods benefit from the mild reaction conditions and inherent selectivity associated with enzyme reactivity, but the complex reaction mixtures hinder product isolation.

Two notable advances in the fields of chemo- and bio-catalysis have recently been achieved by Roche *et al.*[23] and Chun and Narayan.[24] Both groups performed hydrogen isotope exchange (HIE) on amino acids to yield *α*-deuterated isotopologues with full retention of stereochemistry (Figure 1A).

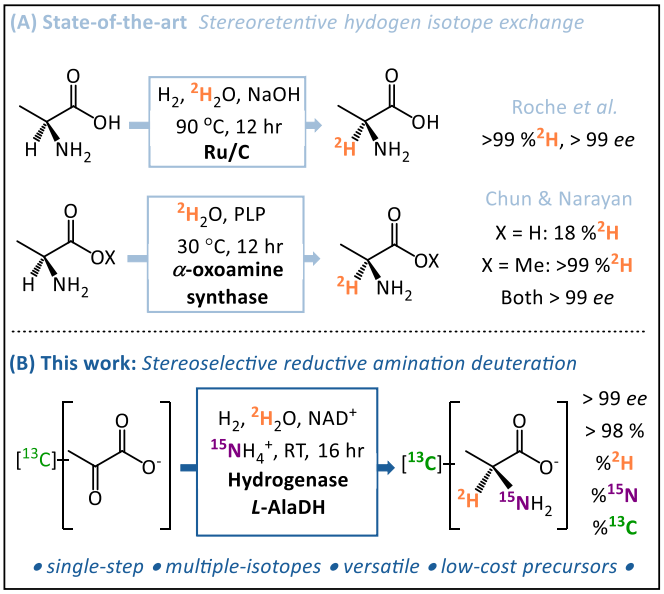

Figure 1 Routes to *α*-deuterated chiral amino acids (*L*-alanine in this case). (A) Stereoretentive hydrogen isotope exchange methodologies utilise chemo- or bio-catalysts to install a ^{2}H atom at a pre-formed asymmetric centre. (B) The stereoselective reductive amination approach used in this work installs the N-atom and ^{2}H-atom with simultaneous formation of the chiral centre, enabling many labelling patterns.

Whilst these HIE strategies represent a significant development in the preparation of [^{2}H]-amino acids, they still require a pre-formed chiral centre prior to the installation of the deuterium. In the case of amino acids for protein NMR, multiple isotope labels {^{2}H, ^{15}N, ^{13}C} are usually required,[25] which are ideally installed in as few steps as possible with low-cost isotope sources. Motivated by this problem, we sought to develop a direct and inexpensive route to *α*-deuterated amino acids with high stereoselectively, minimal work-up requirements, and in-built versatility regarding the pattern of isotopic labels.

In contrast to the HIE strategies employed elsewhere, we utilised a reductive amination deuteration approach, whereby the N-atom, ^{2}H-label, and chiral centre are all introduced into the target molecule in a single-step (Figure 1B). This transformation is achievable by the use of an amino acid dehydrogenase; these are robust, widely available enzymes with naturally high activity and selectivity.[26] The challenge in employing amino acid dehydrogenases for *α*-deuteration is the requirement to supply a suitably [^{2}H]-labelled cofactor: [4-^{2}H]-NADH. We have developed an atom-efficient strategy for generating [4-^{2}H]-NADH from NAD^{+} *in situ* using only H_2 as a clean reductant and $^{2}H_2O$ as a cheap source of deuterium.[27] The approach to obtain the deuterated cofactor uses an electronically coupled H_2-oxidising site and NAD^{+} reducing

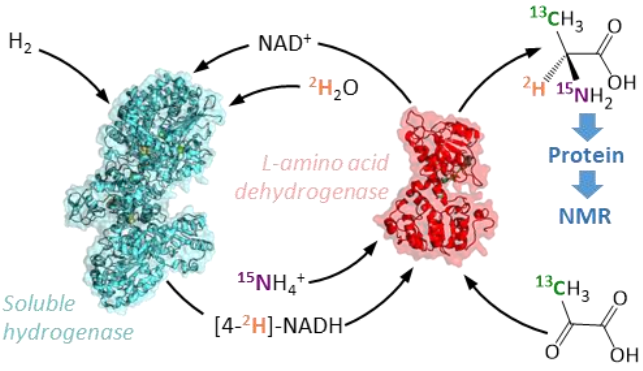

Figure 2 A H_2-driven route to selectively isotopically labelled forms of *L*-alanine bearing an asymmetric deuterium centre. The scheme utilises H_2 gas to drive the formation and recycling of [4-^{2}H]-NADH from NAD^{+} and $^{2}H_2O$ by the action of a soluble hydrogenase.

site, in this case as part of the same enzyme, the soluble hydrogenase from *Ralstonia eutropha*, (now known as *Cupriavidus necator), Re*SH.[28,29] The [4-^{2}H]-NADH can then be supplied to an amino acid dehydrogenase (such as *L*-alanine dehydrogenase, *L*-AlaDH) in the presence of $^{14}NH_4^{+}$ / $^{15}NH_4^{+}$ and substrate to give a chiral [^{2}H]-amino acid (Figure 2).

This method effectively uses two enzymes in one solution to produce *α*-deuterated amino acids in gram quantities, with H_2 in the local atmosphere providing the driving force. The approach is customisable, to allow for different $^{2}H/^{13}C/^{15}N$ labelling patterns with little variation of the method. We demonstrate the effectiveness of this procedure by incorporating it into a 400 kDa molecular chaperone, Hsp16.5, and acquiring a ^{1}H-^{13}C heteronuclear multiple quantum coherence (HMQC) spectrum suitable for high resolution structural and dynamical study. We anticipate that access to larger quantities of specifically labelled amino acids at substantially lower costs using our method will increase their use in biochemical studies.

Results and discussion

Verification of H_2-driven biocatalytic system for producing a library of isotopically labelled amino acids

We have previously demonstrated that *Re*SH is capable of simultaneous reduction and isotopic labelling of NAD^{+} in $^{2}H_2O$ under low pressures of H_2.[28,29] Here, the action of the *Re*SH was coupled to that of an *L*-AlaDH in order to drive the reductive 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-

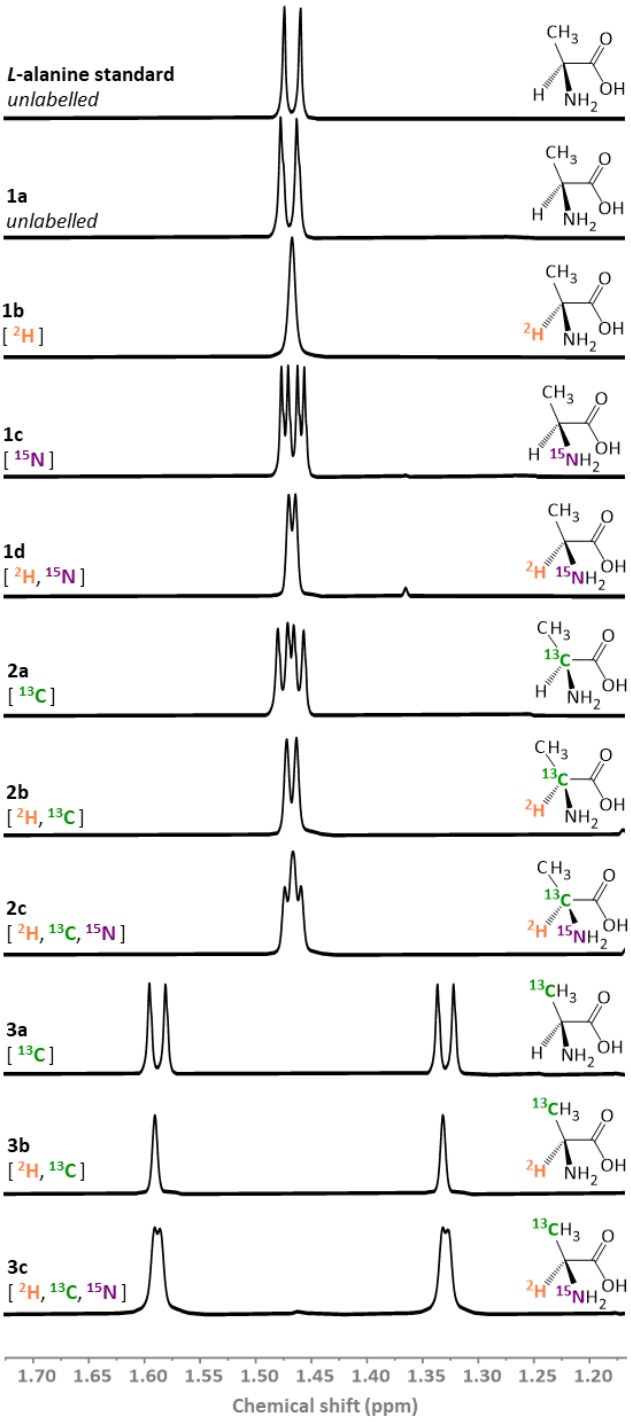

Figure 3 ^{1}H NMR spectra of variously isotopically labelled L-alanine molecules prepared biocatalytically (500 MHz, 293 K, $^{2}H_2O$, p^{2}H 8.0).

take of this amino acid by *E. coli* and incorporation into proteins for NMR structure determination.[2]

Using the combination of *Re*SH and *L*-AlaDH it was possible to generate a library of ten *L*-alanine isotopologues (compounds **1a** - **d**, **2a** - **c**, **3a** – **c** in Figure 3) simply by varying the solvent (H_2O/$^{2}H_2O$), ammonium salt ($^{14}NH_4^+$/$^{15}NH_4^+$), and the pyruvate substrate (^{12}C/^{13}C). In each case, the ^{1}H signals from the methyl peak of the alanine (-C<u>H</u>$_3$, δ = 1.46 ppm) in the corresponding NMR

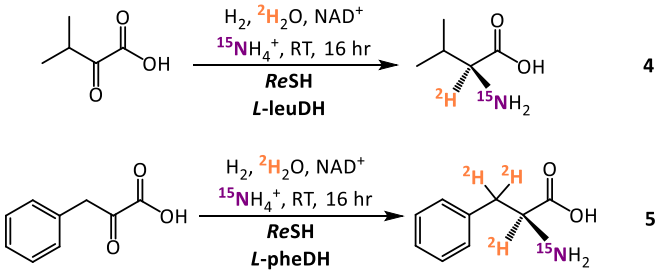

Figure 4 Preparation of ^{2}H and ^{15}N-labelled *L*-leucine and *L*-phenylalanine by H_2 driven biocatalytic reductive amination and deuteration.

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 H_2 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 SI 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 H_2-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-^{2}H,^{15}N]-leucine (**4**) and *L*-[α-^{2}H,^{15}N, β-$^{2}H_2$]-phenylalanine (**5**) with similarly high levels of conversion and isotopic and enantioselectivity (see Figure 4 and SI Section S.3.1).

Scale-up of H_2-driven synthesis of α-deuterated amino acids

Following initial demonstration of the H_2-driven biocatalytic system for preparing multiply labelled amino acid isotopologues, the reaction was tested on a preparative scale. Initially, 50 mg of **1d** was prepared by reductive amination of pyruvate (25 mM) in $^{2}H_2O$ (25 mL, p^{2}H 8.0) in a H_2 atmosphere (2 bar) at room temperature, in the presence of [^{15}N]H_4HCO_3 (50 mM), with *Re*SH, *L*-AlaDH, and NAD^+ (0.1 mM). The role of the [^{15}N]H_4HCO_3 was to serve as both an amine source and volatile buffer, enabling straightforward isolation of the product by rotary evaporation following the reaction. The reaction reached

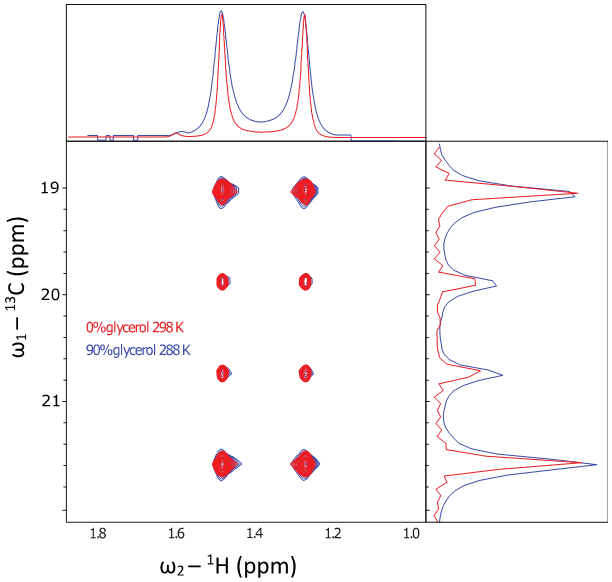

Figure 5 In order to simulate the rotational environment of the labelled amino acid inside a protein, *L*-[*α*-^{2}H,^{15}N, *β*-^{13}C]-alanine was dissolved in a viscous glycerol medium and studied by NMR. The figure shows the ^{13}C-^{1}H decoupled HSQC experiment of the sample in $^{2}H_2O$ at 298 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. By contrast at 90 vol.% glycerol, the ratio is approximately 2:1:1:2. This is due to the outer resonances having 9x faster relaxation than the inner lines, 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.

>99% conversion after 20 h, indicating a turnover number (ToN) for the *Re*SH 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 *Re*SH as a deuteration catalyst, a similar reaction to synthesise **3c** was set up aerobically in standard laboratory glassware. Here, [3-^{13}C]-pyruvate (35 mM) was reacted in a stirred solution of $^{2}H_2O$ (65 mL, p^{2}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 $^{2}H_2O$ with 150 mM NH_4HCO_3 (Figure 4b), 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 final products closely followed the isotopic purity of the starting reaction mixture and simple precautions (such as washing the enzymes in $^{2}H_2O$) were sufficient to keep the %^{2}H high (>95 %).

NMR analysis of triple labelled *L*-alanine (3c) under conditions representative of a high molecular weight protein

Having isolated sufficient quantities of *L*-[*α*-^{2}H,^{15}N, *β*-^{13}C]-alanine (**3c**), the sample was tested for its suitability for solution-state methyl-TROSY experiments prior to integration into a large molecular weight protein. Two decoupled ^{13}C-^{1}H HSQC spectra of the methyl group of the sample were recorded: the first at 298 K in $^{2}H_2O$ (Figure 5, red trace), and the second at 288 K in 90 vol.% glycerol (Figure 5, blue trace). The viscosity of the latter sample was calculated to be around 450-times that of the former, effectively simulating the local tumbling behaviour of the molecule as if it were part of a 50 kDa protein. In the decoupled HSQC experiment, a freely tumbling methyl group should yield an inverse quartet with intensity 3:1:1:3 on ω_1, which was observed for the sample in pure $^{2}H_2O$. In the macro-molecular limit, the relaxation rates of the outer lines should be R2=nKtc where n=9 for outer lines, and n=1 for the inner lines. Under slowly tumbling conditions, the intensity of the quartet was altered to 2:1:1:2, reflecting increased relaxation of these resonances. Nevertheless, the similarity of the two spectra confirmed the suitability of this reagent as a useful probe inside the slowly tumbling environment of a large protein.

Incorporation of triple-labelled amino acid into protein and subsequent analysis by NMR spectroscopy

To test the feasibility of using *L*-[*α*-^{2}H,^{15}N, *β*-^{13}C]-alanine (**3c**) in biomolecular studies, we incorporated it into the small heat shock protein Hsp16.5. This protein is a molecular chaperone from *Methanococcus jannaschii* that assembles into a 24mer of molecular weight 396 kDa (Figure 6a).[30] *In vitro*, this protein is a potent inhibitor of protein aggregation and amyloid formation. *L*-[*α*-^{2}H,^{15}N, *β*-^{13}C]-Alanine (**3c**, 100 mg) was added 1 hour prior to induction, and the incorporation was > 99%. A ^{13}C-^{1}H NMR spectrum was acquired. This is among the largest proteins for which an alanine spectrum has been recorded.[4,10] Interestingly, although 8 resonances were expected, we observed 12, which suggests that we are observing multiple conformations of the protein (Figure 6b). This spectrum validates the utility of this precursor in future biochemical studies.

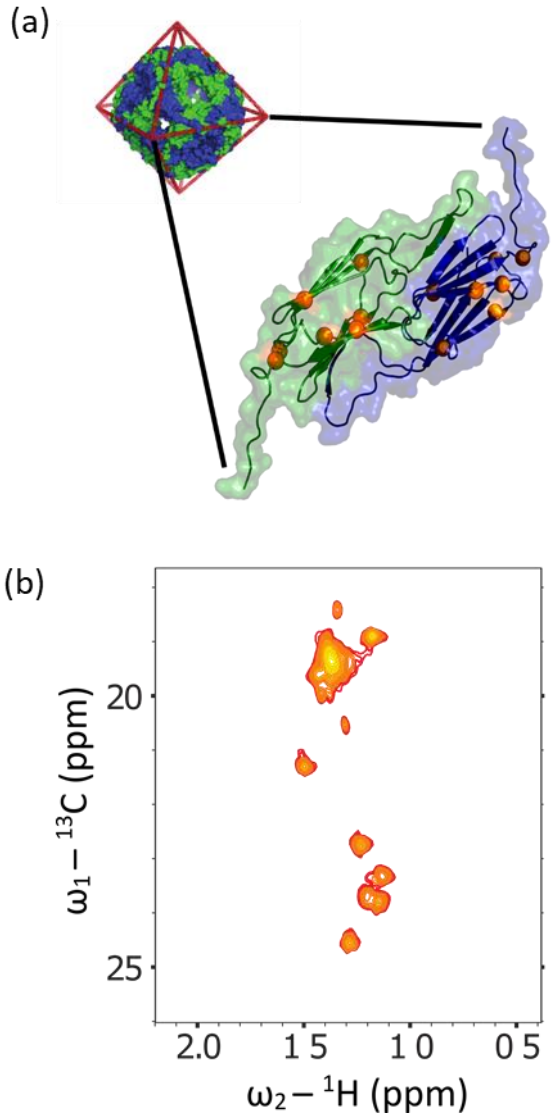

Figure 6 (a) Representation of the dimeric unit (16.5 kDa) from the small heat shock protein (sHSP) from *M. jannaschii*, showing the sites of the isotopically labelled alanine CH_3 groups. The homo 24-mer has a molecular weight of 396 kDa. (b) Alanine region of methyl spectrum revealing 12 clear resonances. 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.

Conclusions

We have developed a convenient means for the production of selectively deuterated amino acids. We demonstrated the success of the method in preparing variously ^{2}H, ^{15}N, and ^{13}C-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 ^{2}H and ^{15}N labels simultaneously with the formation of the chiral centre, and enables wide versatility in the choice of isotopes added. We anticipate this approach greatly facilitating mechanistic and structural studies of large biomolecules.

Conflicts of interest

A patent application detailing some of this research was filed through Oxford University Innovation (Feb 2018).

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.). In addition, J.S.R. was supported by an EPA Cephalosporin Junior Research Fellowship at Linacre College, Oxford.

References

1 Z. Kelman, *Isotope Labeling of Biomolecules - Applications*, Elsevier, New York, 2016, vol. 566.
2 G. W. Becker, *Briefings Funct. Genomics Proteomics*, 2008, **7**, 371–382.
3 J. Atzrodt, V. Derdau, W. J. Kerr and M. Reid, *Angew. Chemie Int. Ed.*, 2018, **57**, 1758–1784.
4 M. Kainosho, Y. Miyanoiri and M. Takeda, in *Experimental Approaches of NMR Spectroscopy*, Springer Singapore, Singapore, 2018, pp. 37–61.
5 M. Kainosho, Y. Miyanoiri, T. Terauchi and M. Takeda, *J. Biomol. NMR*, 2018, **71**, 119–127.
6 A. Velyvis, A. M. Ruschak and L. E. Kay, *PLoS One*, 2012, **7**, e43725.
7 V. Tugarinov, P. M. Hwang and L. E. Kay, *Annu. Rev. Biochem.*, 2004, **73**, 107–146.
8 J. E. Ollerenshaw, V. Tugarinov and L. E. Kay, *Magn. Reson. Chem.*, 2003, **41**, 843–852.
9 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.
10 T. Yuwen, R. Huang, P. Vallurupalli and L. E. Kay, *Angew. Chemie Int. Ed.*, 2019, 6250–6254.
11 L. Brocchieri and S. Karlin, *Nucleic Acids Res.*, 2005, **33**, 3390–3400.
12 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.
13 R. Sakuta, K. Takeda, K. Igarashi, H. Ohno and N.

Nakamura, *Energy Technol.*, 2018, **6**, 273–279.
14 S. Bhatia, G. Spahlinger, N. Boukhumseen, Q. Boll, Z. Li and J. E. Jackson, *European J. Org. Chem.*, 2016, **2016**, 4230–4235.
15 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. Chemie Int. Ed.*, 2015, **54**, 10474–10477.
16 K. Moozeh, S. M. So and J. Chin, *Angew. Chemie Int. Ed.*, 2015, **54**, 9381–9385.
17 H. Ohtsuki, M. Takashima, T. Furuta and T. Kawabata, *Tetrahedron Lett.*, 2018, **59**, 1188–1191.
18 U. M. Babu and R. B. Johnston, *Biochem. Biophys. Res. Commun.*, 1974, **58**, 460–466.
19 C. H. Wong and G. M. Whitesides, *J. Am. Chem. Soc.*, 1983, **105**, 5012–5014.
20 J. A. Chanatry, P. H. Schafer, M. S. Kim and D. M. Lemaster, *Anal. Biochem.*, 1993, **213**, 147–151.
21 J. Feeney, B. Birdsall, G. Ostler, M. D. Carr and M. Kairi, *FEBS Lett.*, 1990, **272**, 197–199.
22 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.
23 A. Michelotti, F. Rodrigues and M. Roche, *Org. Process Res. Dev.*, 2017, **21**, 1741–1744.
24 S. W. Chun and A. R. H. Narayan, *ACS Catal.*, 2020, **10**, 7413–7418.
25 M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono and P. Güntert, *Nature*, 2006, **440**, 52–57.
26 H. C. Lo and R. H. Fish, *Angew. Chemie Int. Ed.*, 2002, **41**, 478–481.
27 J. S. Rowbotham, M. A. Ramirez, O. Lenz, H. A. Reeve and K. A. Vincent, *Nat. Commun.*, 2020, **11**, 1454.
28 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.
29 A. Al-Shameri, M. Petrich, K. junge Puring, U. Apfel, B. M. Nestl and L. Lauterbach, *Angew. Chemie Int. Ed.*, 2020, **59**, 10929–10933.
30 K. K. Kim, R. Kim and S. H. Kim, *Nature*, 1998, **394**, 595–599.