Journal of Biomolecular NMR (2018) 71:287–293
https://doi.org/10.1007/s10858-017-0157-y

ARTICLE

Genetically encoded amino acids with *tert*-butyl and trimethylsilyl groups for site-selective studies of proteins by NMR spectroscopy

Choy Theng Loh[1,2] · Luke A. Adams[3] · Bim Graham[3] · Gottfried Otting[1]

Received: 19 October 2017 / Accepted: 24 November 2017 / Published online: 2 December 2017

Abstract

The amino acids 4-(*tert*-butyl)phenylalanine (Tbf) and 4-(trimethylsilyl)phenylalanine (TMSf), as well as a partially deuterated version of Tbf (dTbf), were chemically synthesized and site-specifically incorporated into different proteins, using an amber stop codon, suppressor tRNA and the broadband aminoacyl-tRNA synthetase originally evolved for the incorporation of *p*-cyano-phenylalanine. The ^{1}H-NMR signals of the *tert*-butyl and TMS groups were compared to the ^{1}H-NMR signal of *tert*-butyltyrosine (Tby) in protein systems with molecular weights ranging from 8 to 54 kDa. The ^{1}H-NMR resonance of the TMS group appeared near 0 ppm in a spectral region with few protein resonances, facilitating the observation of signal changes in response to ligand binding. In all proteins, the R_2 relaxation rate of the *tert*-butyl group of Tbf was only little greater than that of Tby (less than two-fold). Deuteration of the phenyl ring of Tbf made only a relatively small difference. The effective T_2 relaxation time of the TMS signal was longer than 140 ms even in the 54 kDa system.

Introduction

tert-Butyltyrosine (Tby, Fig. 1) is an unnatural amino acid that can be site-specifically incorporated into proteins in response to an amber stop codon. The *tert*-butyl group in *tert*-butyltyrosine appears in the 1D ^{1}H-NMR spectrum as an intense singlet. For solvent-exposed *tert*-butyl groups, this signal is also narrower than other resonances of the protein, due to its enhanced mobility arising from rapid rotations of the methyl groups and around the carbon–oxygen bonds linking the *tert*-butyl group to the aromatic ring of the amino acid. The narrow and intense singlet can thus readily be identified in the 1D ^{1}H-NMR spectrum even for protein systems of high molecular weight (> 300 kDa), producing a highly sensitive NMR probe for measuring submicromolar ligand binding affinities (Chen et al. 2015) and measuring intermolecular NOEs (Jabar et al. 2017). The intensity of a ^{1}H-NMR signal produced by nine ^{1}H spins is a clear advantage over ^{19}F-NMR of unnatural amino acids containing a CF_3 group (Chen et al. 2015; Hammill et al. 2007; Loscha et al. 2012), especially as ^{19}F-NMR resonances are much more susceptible to broadening by chemical shift anisotropy than the ^{1}H-NMR signals of CH_3 groups.

The chemical shift of a *tert*-butyl group is typically at about 1.3 ppm in aqueous solution, with small variations depending on the local chemical environment. This means that, despite its high intensity and narrow lineshape, the ^{1}H-NMR resonance of the *tert*-butyl group can be masked by the protein background. In principle, this problem can be addressed by using Tby with a uniformly ^{13}C-labelled *tert*-butyl group, allowing its selective detection by a simple ^{13}C-HSQC spectrum. We recently demonstrated this approach with a 95 kDa protein-DNA complex (Jabar et al. 2017), but it is expensive and therefore requires a highly efficient way of incorporation into proteins, such as cell-free protein synthesis (Loscha et al. 2012). A simpler approach would be to substitute Tby by an amino acid with a trimethylsilyl (TMS) group instead of the *tert*-butyl group. As the TMS group produces a singlet at about 0 ppm, it is in a spectral region where few protein resonances are found. We

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10858-017-0157-y) contains supplementary material, which is available to authorized users.

✉ Gottfried Otting
gottfried.otting@anu.edu.au

[1] Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

[2] Zhejiang Provincial Key Laboratory of Chemical and Biological Processing Technology of Farm Products, Zhejiang University of Science and Technology, Hangzhou 310023, People's Republic of China

[3] Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia

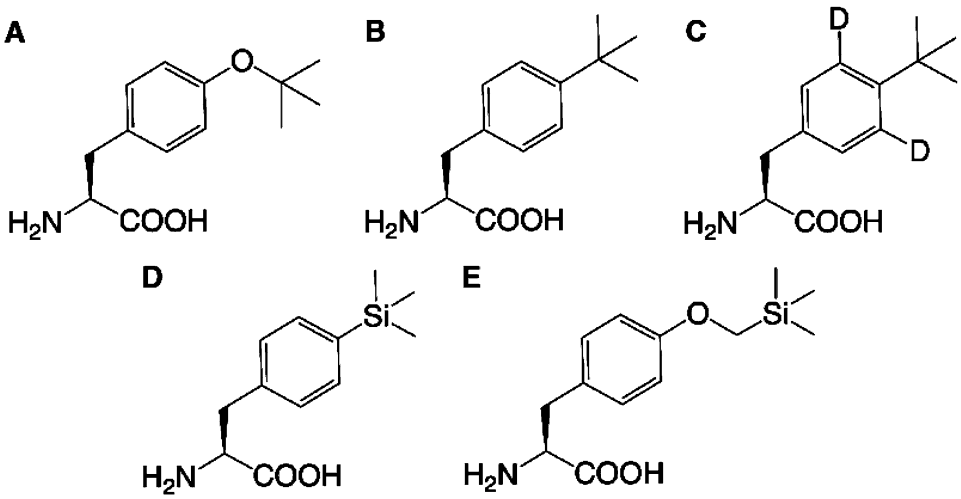

Fig. 1 Chemical structures of unnatural amino acids studied in the present work for NMR analysis following incorporation into different proteins. **a** *tert*-butyltyrosine (Tby). **b** 4-(*tert*-butyl)phenylalanine (Tbf). **c** 4-(*tert*-butyl)phenylalanine-(phenyl-3,5-d_2) (dTbf). **d** 4-(trimethylsilyl)phenylalanine (TMSf). **e** 4-[(trimethylsilyl)methoxy]phenylalanine (TMSmy)

recently confirmed that a TMS tag ligated to the thiol group of a solvent-exposed cysteine produces a ^{1}H-NMR signal at 0 ppm that is readily observable even in systems of high molecular weight (Jabar et al. 2017). In addition, it would be attractive if the location of the *tert*-butyl or TMS group could be predicted more accurately relative to the protein backbone by the design of a more rigid tether than in Tby.

Here we report the chemical synthesis of 4-(*tert*-butyl) phenylalanine-(phenyl-3,5-d_2) (dTbf), 4-(trimethylsilyl) phenylalanine (TMSf) and 4-((trimethylsilyl)methoxy)phenylalanine (TMSmy) for genetically encoded incorporation into proteins. Tbf, which is commercially available, dTbf and TMSf were successfully incorporated into different proteins using the polysubstrate-specific *p*-cyanophenylalanine aminoacyl-tRNA synthetase (*p*CNF-RS) reported by Peter G. Schultz and co-workers (Young et al. 2011), whereas the same system failed to incorporate TMSmy. To the best of our knowledge, the present work is the first describing the successful incorporation of Tbf and TMSf into proteins. For comparison, we also incorporated Tby. As the *tert*-butyl and TMS groups contain nine equivalent protons generating a singlet in the ^{1}H-NMR spectrum, we refer to them as 9PS groups (nine proton singlet). To assess their relative merits for NMR spectroscopy, we investigated the $R_2(^1H)$ relaxation rates of the 9PS amino acids in proteins of different molecular weight and tested TMSf as a probe for ligand binding.

Materials and methods

Materials

RF1-depleted S30 extract was prepared from the genomically modified *E. coli* strain BL21 Star (DE3)::RF1-CBD3 as described previously (Apponyi et al. 2008; Loscha et al. 2012). Tby and Tbf were purchased from Sigma-Aldrich and Santa Cruz Biotechnology, respectively. Detailed protocols for the synthesis of TMSmy, TMSf and deuterated Tbf are provided in the supporting information. Total tRNA including optimized suppressor-tRNA was prepared as described previously (Ozawa et al. 2012). *p*CNF-RS was expressed from the pEvol-*p*CNF plasmid (Young et al. 2011). The protease inhibitor 4-nitrophenyl-4-guanidinobenzoate hydrochloride was purchased from Sigma-Aldrich.

Expression and purification of *p*CNF-RS

The gene encoding *p*CNF-RS was amplified by PCR from pEvol-*p*CNF (Young et al. 2010, 2011), cloned into the T7 vector pETMCSI (Neylon et al. 2000) and transformed into *E. coli* BL21(DE3). The cells were grown in 1 L of LB medium containing 100 µg/mL ampicillin at 37 °C and induced with 1 mM IPTG at $OD_{600} = 0.7$. Cells were harvested after overnight expression (about 16 h) at 25 °C by centrifugation. Pellets were suspended in buffer A (50 mM Tris–HCl, pH 8.0, 1 mM EDTA, 1 mM DTT and 10% glycerol) and cells were lysed by French press at 12,000 psi. The cell lysates were centrifuged for 1 h at 30,000×*g* at 4 °C. The supernatant was subsequently loaded onto a DEAE Toyopearl 650 M anion exchange column. *p*CNF-RS was eluted with a linear gradient of 0–600 mM NaCl in buffer A. Fractions containing *p*CNF-RS were pooled and dialyzed against buffer B (20 mM Tris–HCl, pH 8, 1 mM EDTA and 1 mM DTT). The concentrated solution of *p*CNF-RS was stored at − 70 °C.

Preparation of total tRNA including optimized suppressor tRNA

The pEvol plasmid encoding optimized suppressor-tRNA (Young et al. 2010) was transformed into BL21(DE3) cells.

Total tRNA containing optimized suppressor-tRNA was prepared as described (Ozawa et al. 2012). About 20 mg of purified total tRNA was obtained from 8 L of LB medium. The concentration of total tRNA was measured by its absorbance at 260 nm.

Cell-free protein synthesis

Amber stop mutants of human ubiquitin (T66U), *S. aureus* sortase A (SrtA Y17U), rat ERp29 (S114U) and Zika virus NS2B-NS3 protease (ZiPro *S81U, where the star indicates that the unnatural amino acid is in NS2B) were designed with a C-terminal His_6-tag and cloned into pETMCSI (Neylon et al. 2000). The proteins were expressed by continuous exchange cell-free protein synthesis. The cell-free reactions were carried out at 30 °C for 16 h as described previously (Apponyi et al. 2008; Loscha et al. 2012; Ozawa et al. 2012). Additional components added into the cell-free reaction mixtures were: 0.2 mg/mL total tRNA including optimized suppressor-tRNA, 30 µM *p*CNF-RS in the inner reaction mixture and 1 mM unnatural amino acid in the inner and outer compartments. The volumes of the inner and outer mixtures were 2 and 20 mL, respectively. Proteins were purified using a 1 mL Ni–NTA gravity column equilibrated with buffer C (50 mM Tris–HCl, pH 8.0, 300 mM NaCl) and washed with buffer D (same as buffer C but with 20 mM imidazole). Proteins were then eluted with buffer E (same as buffer C but with 300 mM imidazole). Purified ubiquitin was buffer-exchanged and concentrated against NMR buffer F (50 mM HEPES-KOH, pH 7.0) by ultrafiltration using an Amicon centrifugal filter unit. The other proteins were buffer-exchanged and concentrated against NMR buffer G (20 mM MES-KOH, pH 6.5, 20 mM NaCl) in the same way. Filter cutoffs were 3 kDa for ubiquitin and 10 kDa for the other proteins.

Tby mutant of *Mt* LeuRS by in vivo protein expression

The gene encoding the Y115U mutant of C-terminal His_6-tagged *M. thermoautotrophicum* leucyl-tRNA synthetase (LeuRS Y115U) was cloned into pETMCSI (Neylon et al. 2000) and co-transformed with pEvol-*p*CNF (Young et al. 2011) into *E. coli* B–95.ΔA (Mukai et al. 2015). Cells were grown in 20 mL of LB medium containing 100 µg/mL ampicillin, 33 µg/mL chloramphenicol and 0.02% arabinose, induced with 1 mM IPTG at $OD_{600} = 0.5$ and supplemented with 1 mM Tby. The cells were harvested after overnight expression by centrifugation at 25 °C. The pellet was re-suspended in buffer C, lysed by ultra-sonication and the lysate was centrifuged for 1 h at 30,000×*g* at 4 °C. The supernatant was loaded onto a 1 mL Ni–NTA gravity column and the purification was carried out as described above. Purified protein was buffer-exchanged and concentrated against NMR buffer G by ultrafiltration.

NMR and R_2 relaxation measurement

NMR spectra were recorded at 25 °C on a 800 MHz Bruker NMR spectrometer equipped with a cryoprobe, using the double spin echo sequence for water suppression (Hwang and Shaka 1995). Protein concentrations ranged from 50 to 150 µM in NMR buffer with 10% D_2O.

$R_2(^1H)$ relaxation rates were assessed using the spin-echo sequence 90°–τ/2–180°–τ/2 followed by the double spin echo sequence for water suppression (Hwang and Shaka 1995). The peak intensities observed in two spectra recorded with $\tau = 1$ and 101 ms were converted into relaxation rates by assuming single exponential decays. The relaxation measurements were performed on a 600 MHz Bruker NMR spectrometer equipped with a cryoprobe.

Results

The unnatural amino acids of Fig. 1 are either commercially available (Tby and Tbf) or can be synthesized in a few steps (TMSmy, TMSf and dTbf, see the Supporting Information). Using cell-free synthesis reactions supplemented with suppressor-tRNA, *p*CNF-RS and unnatural amino acid, full-length protein samples were obtained in good yield for ubiquitin T66U, SrtA Y17U, ZiPro *S81U and ERp29 S114U, when the unnatural amino acid was Tby, Tbf or TMSf. In contrast, we failed to incorporate TMSmy using the same system. This result may in part have arisen from the relatively low solubility of TMSmy in aqueous solution, but it is also possible that the larger size of the TMS-methylene group compared with the *tert*-butyl group renders TMSmy incompatible with the amino-acid binding pocket of the *p*CNF-RS enzyme.

The best expression yields were obtained with Tby and Tbf, where the final yields of purified protein were about 0.8 mg per 2 mL cell-free inner reaction mixture for ubiquitin T66U, SrtA Y17U and ZiPro *S81U, and about 0.6 mg for ERp29 S114U. The yields were significantly lower with TMSf (between 0.3 and 0.5 mg per 2 mL inner reaction mixture). Expression of larger proteins with any of these unnatural amino acids proved difficult. No protein was obtained in attempts to produce T7 RNA polymerase (99 kDa) with amber stop codons in positions 312, 385, 623 or 846 by cell-free synthesis. Similarly, the Y115U and Y124U mutants of LeuRS (108 kDa) both failed to express under cell-free conditions with any of the unnatural amino acids and only LeuRS Y115Tby gave an acceptable yield in vivo (0.25 mg from 20 mL cell culture). In vivo expression of the amber stop mutants of ubiquitin

and SrtA produced protein in good yield with Tby, but the yields were low with Tbf and no protein was obtained with TMSf.

The chemical shifts of the 9PS signal of the free amino acids in water at 25 °C are 1.34 ppm for Tby, 1.30 ppm for Tbf and 0.28 ppm for TMSf. Incorporation into a protein changes these chemical shifts depending on the specific environment. In general, the spectra of ERp29 S114Tbf and SrtA Y17Tbf show that the 9PS signal of Tbf is about 0.1 ppm further upfield than the corresponding signal of Tby. The 9PS signal of TMSf was between 0.1 and 0.3 ppm (Fig. 2). In ubiquitin, the 9PS signals of all these amino acids are shifted upfield by about 0.4 ppm due to ring currents from Phe4, which is located near position 66 containing the unnatural amino acid residue.

The 9PS signals of TMSf tended to be broader than those of Tby and Tbf (Fig. 2), suggesting that the TMS group is less mobile than the *tert*-butyl group because of greater steric hindrance from the phenyl ring affecting the bond rotation rates. The methyl groups in Tbf and TMSf are in closer proximity to the aromatic ring protons than in Tby. To minimize dipole–dipole relaxation between the *tert*-butyl group in Tbf and the aromatic ring protons, we synthesized Tbf with deuterium substituting protons in the aromatic ring (dTbf). In all three proteins studied with Tbf and dTbf, the resulting 9PS signal was not much narrower than the signal from undeuterated Tbf (Fig. 2). The results showed that the relaxation rate of the 9PS signal of Tbf was generally greater than for Tby (Table 1), which may be attributed to the greater mobility of an *O-tert*-butyl group compared

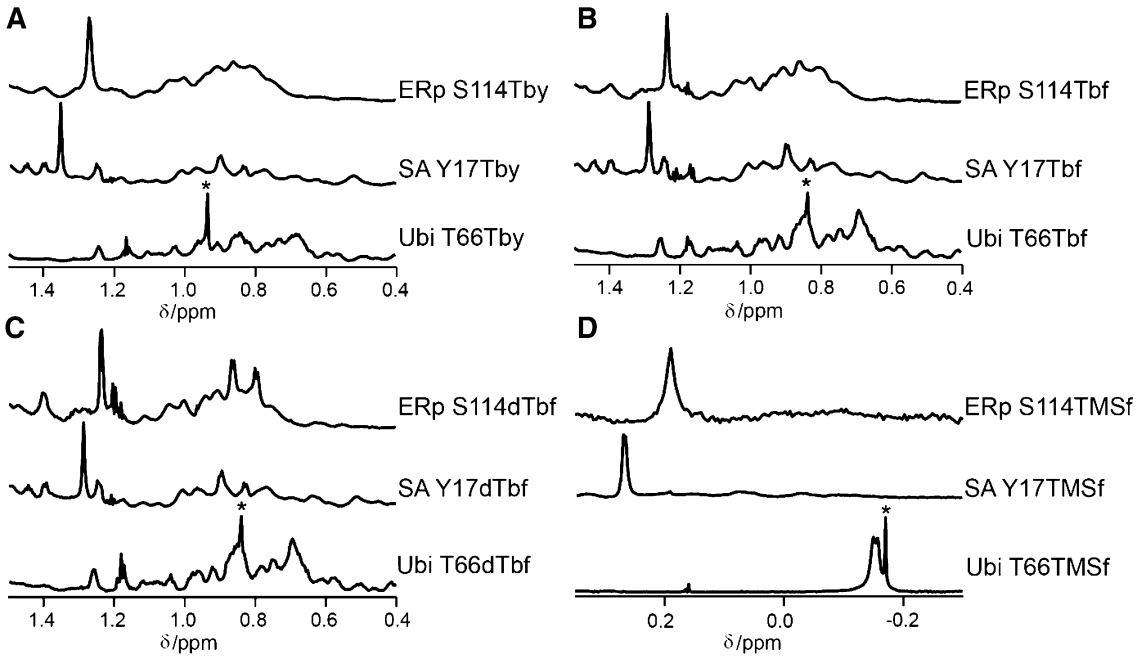

Fig. 2 1D ^{1}H-NMR spectra of 0.1 mM solutions of ubiquitin T66U, sortase A Y17U and ERp29 S114U with the unnatural amino acids of Fig. 1. All spectra were measured at 25 °C on a 800 MHz NMR spectrometer. The panels **a–d** show spectra with Tby, Tbf, dTbf and TMSf, respectively. The chemical shifts of the *tert*-butyl and TMS groups in the ubiquitin T66U mutant are identified by a star. They are shifted due to ring currents from the side chain of Phe4 nearby. In the case of ubiquitin T66TMSf, this leads to overlap with the resonance of Leu50 $^{\delta}CH_3$, which is also at about −0.16 ppm

Table 1 $R_2(^1H)$ relaxation rates of the ^{1}H NMR signals of the *tert*-butyl and TMS groups in ubiquitin T66U, Sortase A Y17U, Erp29 S114U and LeuRS Y115U

	MW (kDa)	Tby (s^{-1})	Tbf (s^{-1})	dTbf (s^{-1})	TMSf (s^{-1})
Ubiquitin T66U	8	4.2	5.3	4.2	2.8
Sortase A Y17U	19	4.0	7.5	7.1	4.9
ERp29 S114U	54	8.4	11.4	7.3	7.0
LeuRS Y115U	108	23.3	n.d	n.d	n.d

The transverse relaxation rates R_2 were determined at 600 MHz and 25 °C from two-time point measurements in a spin-echo experiment with relaxation delays differing by 100 ms

n.d. Not determined

with an oxygen-free *tert*-butyl group. Use of dTbf versus Tbf hardly reduced the relaxation rates for the smaller proteins, but made a difference for ERp29. Despite the broad line width observed for the TMS resonance in the 1D ^{1}H-NMR spectrum of ERp29 (Fig. 2d), the apparent relaxation rate of this 9PS signal measured in a spin-echo experiment performed with and without 100 ms relaxation delay was comparable to that of dTbf (Table 1). This counterintuitive result may be explained by multi-exponential relaxation, which is a well-known feature of methyl groups and most pronounced under non-extreme narrowing conditions (Hubbard 1958; Werbelow and Marshall 1973; Müller et al. 1987). Indeed, the resonance observed after a 301 ms delay of transverse relaxation was narrower than after a relaxation delay of 1 ms (6 vs. 11 Hz).

Quite generally, the relaxation rates increased with molecular mass, but the increase was not linear, as clearly demonstrated by the four proteins made with Tby. The width or relaxation rate of a 9PS signal thus is not a simple measure of the molecular weight. Unfortunately, LeuRS Y115U (like another protein in the molecular weight range > 100 kDa) could only be produced with Tby (Fig. S1) but not with Tbf or TMSf. Overall, our current set of data indicates that Tby and dTbf yield the narrowest 9PS signals, whereas the 9PS signal of TMSf, despite its broader width, comprises components that may relax equally or even more slowly.

As the 9PS signal of TMSf is in a spectral region with little signal overlap, it can more readily be tracked in ligand binding studies than the corresponding signal of Tby. This is illustrated by the Tby and TMSf samples of ZiPro *S81U, where the 9PS signal of Tby overlaps with other, similarly narrow resonances of the protein, whereas the 9PS signal of TMSf is much better resolved at about −0.1 ppm (Fig. 3a). Titration of ZiPro *S81TMSf with the generic serine protease inhibitor 4-nitrophenyl-4-guanidinobenzoate hydrochloride (Fig. S2) greatly broadened the 9PS signal, indicating that the exchange rate of this inhibitor is in the slow to intermediate time regime (Fig. 3b). It is interesting that the TMSf residue in position *81 senses the binding of the inhibitor, which is known to bind to NS3 whereas residue *81 resides in NS2B. Previous results obtained from ^{15}N-HSQC spectra of ^{15}N-labelled ZiPro indicated that the C-terminal β-hairpin of NS2B, which comprises *Ser81, is dissociated from NS3 and mostly unaffected by the presence or absence of the inhibitor (Mahawaththa et al. 2017).

Discussion

The present work demonstrates, for the first time, the incorporation of Tbf and TMSf into proteins as genetically encoded amino acids. We obtained these results by using the *p*CNF-RS enzyme, which has previously been shown to be poly-specific for a broad range of unnatural amino acids related to tyrosine (Young et al. 2011). Using cell-free synthesis, Tby, Tbf, dTbf and TMSf were successfully incorporated into human ubiquitin, *S. aureus* sortase A and rat ERp29. For proteins with molecular weights greater than

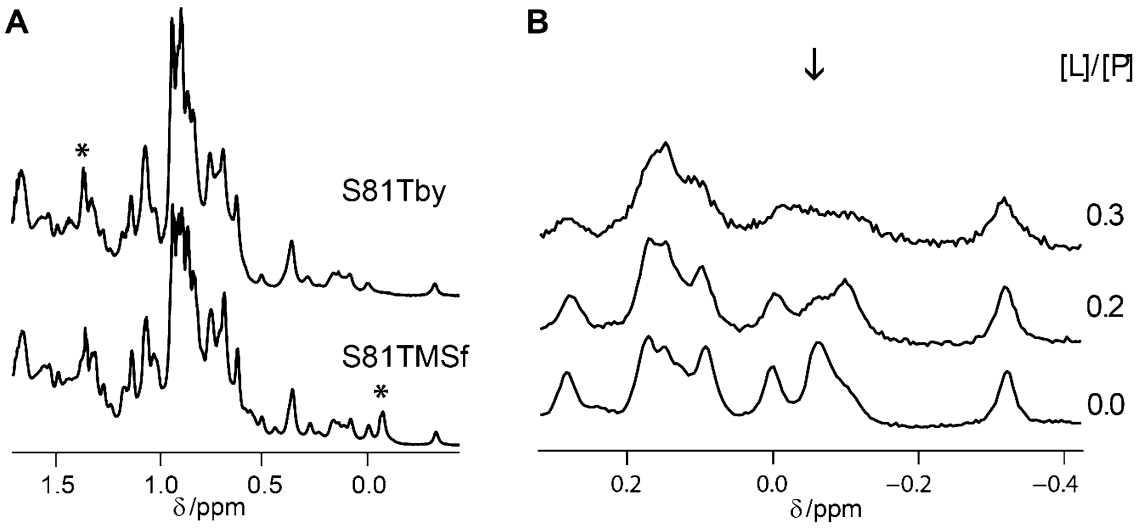

Fig. 3 1D ^{1}H-NMR spectra of the Zika virus NS2B-NS3 protease with the mutation *S81U in the NS2B co-factor. **a** Spectra of 0.1 mM solutions of the protease with either Tby (top panel) or TMSf (bottom panel) incorporated at position 81. Stars mark the 9PS signals of the *tert*-butyl and TMS groups, respectively. The spectra illustrate the improved resolution in the spectral region of the TMS group versus that of the *tert*-butyl group. **b** Spectra of a 0.05 mM solution of the Zika virus NS2B-NS3 protease mutant *Ser81 TMSf. The arrow marks the resonance of the TMS group. Spectra are shown for different ratios of ligand (L) and protein (P) concentrations, where the ligand is the generic protease inhibitor 4-nitrophenyl-4-guanidinobenzoate hydrochloride

100 kDa, expression yields by cell-free synthesis proved to be too low to produce NMR samples, even for the 108 kDa protein LeuRS Y115U, which gave good expression yields with Tby in vivo. It is increasingly recognised that *E. coli* cell-free extracts work best for proteins smaller than 50 kDa (Chong 2001). Attempts to produce LeuRS with Tbf and TMSf in vivo failed. In general, in vivo protein expression requires significantly larger amounts of amino acids than cell-free synthesis (Torizawa et al. 2004), discouraging sample preparation with costly unnatural amino acids. The problem is amplified in the case of Tbf and TMSf, where in vivo expression of ubiquitin T66U and SrtA Y17U yielded little protein with Tbf and none with TMSf, suggesting that these amino acids are not easily imported across the cell membrane.

While *p*CNF-RS readily incorporates Tby, the cell-free incorporation yields of Tbf and, in particular, TMSf were lower, which would be expected for a lesser activity of the *p*CNF-RS with these amino acids. For difficult unnatural amino acids, it is well known that *E. coli* strains deficient in the release factor 1 (such as the B–95.ΔA strain) can, with unexpectedly high yields, read through an amber stop codon by misincorporation of a natural amino acid, if the suppressor tRNA is not efficiently loaded with the unnatural amino acid (Mukai et al. 2010, 2015; Ohtake et al. 2012; Nilsson and Rydén-Aulin 2003; O'Donoghue et al. 2012; Gan and Fan 2017; George et al. 2016). The S30 extract used for our cell-free expressions was prepared from a cell-line, where the release factor 1 was present but had been tagged with chitin binding domains for selective removal prior to cell-free synthesis (Loscha et al. 2012). Nonetheless, we observed that also in this case full-length protein can be produced that does not contain the desired unnatural amino acid. Specifically, mass spectrometric analysis of tryptic digests of the ubiquitin T66TMSf sample revealed peptides containing either glutamine or lysine in place of the unnatural amino acid. Therefore, the observation that the 9PS signal in the spectrum of ubiquitin T66TMSf (Fig. 2d) was about ten times smaller than expected may be attributed to limited compatibility of the *p*CNF-RS enzyme with TMSf. The specific reasons for different incorporation yields in different proteins, however, are not understood.

In view of the greater rigidity of the *tert*-butyl group in Tbf than in Tby, the relaxation rate of the 9PS signal in Tbf was expected to be approximately proportional to the molecular mass of the protein. This was not observed. Potential explanations for this result include local mobility of the peptide segment harbouring the unnatural amino acid, different degrees of solvent exposure and multi-exponential relaxation effects.

It is interesting that the 9PS signal of TMSf can be broader than that of Tby, yet show slower $R_2(^1H)$ relaxation rates. The effect was most clearly manifested for the 54 kDa dimer of ERp29, suggesting that it is governed by multi-exponential dipole–dipole relaxation, which is expected to be more pronounced at high molecular weight. Evidence for multi-exponential relaxation is also presented by the pronounced mismatch between the linewidths and $R_2(^1H)$ relaxation rates measured. For example, none of the linewidths measured in the 1D ^{1}H-NMR spectra was below 5 Hz, which would correspond to a relaxation rate > 15 s^{-1}, but the relaxation rates reported by the peak heights after transverse relaxation during 100 ms were significantly lower (Table 1). The 9PS resonance of TMSf thus is advantageous not only for its chemical shift in a less crowded spectral region, but its slower transverse relaxation rates may also promote its detection in high-molecular weight systems using editing experiments, for example following ^{13}C labelling of the methyl groups.

Conclusion

In summary, we succeeded in genetic encoding of the amino acids Tbf and TMSf. The *tert*-butyl group of Tbf and the TMS group of TMSf produce ^{1}H-NMR resonances that can readily be identified without assigning any other NMR resonances. Their signals are sensitive reporters of the local environment, which can be used for ligand binding studies. Further optimization of the aminoacyl-tRNA synthetase may improve the incorporation yield of TMSf, which is attractive for its ^{1}H-NMR signal in a relatively well-resolved region of the protein NMR spectrum.

Acknowledgements We thank Mr Alireza Bahramzadeh for an expression construct and sample of wild-type Leu-RS, Dr Yao Wang and Mr Mithun Chamikara Mahawaththa for mass spectrometry measurements, and Professor Peter G. Schultz for the pEvol-*p*CNF-RS plasmid. Financial support by the Australian Research Council is gratefully acknowledged.

References

Apponyi MA, Ozawa K, Dixon NE, Otting G (2008) Cell-free protein synthesis for analysis by NMR spectroscopy. In: Kobe B, Guss M, Huber T (eds) Structural proteomics: high-throughput methods. Humana Press, Totowa, pp 257–268

Chen W-N, Kuppan KV, Lee MD, Jaudzems K, Huber T, Otting G (2015) O-*tert*-butyltyrosine, an NMR tag for high-molecular-weight systems and measurements of submicromolar ligand binding affinities. J Am Chem Soc 137:4581–4586

Chong S (2001) Overview of cell-free protein synthesis: historic landmarks, commercial systems, and expanding applications. In Current protocols in molecular biology. Wiley, New York

Gan Q, Fan C (2017) Increasing the fidelity of noncanonical amino acid incorporation in cell-free protein synthesis. Biochim Biophys Acta 1861:3047–3052

George S, Aguirre JD, Spratt DE, Bi Y, Jeffery M, Shaw GS, O'Donoghue P (2016) Generation of phospho-ubiquitin variants

by orthogonal translation reveals codon skipping. FEBS Lett 590:1530–1542

Hammill JT, Miyake-Stoner S, Hazen JL, Jackson JC, Mehl RA (2007) Preparation of site-specifically labeled fluorinated proteins for ^{19}F-NMR structural characterization. Nat Protoc 2:2601–2607

Hubbard PS (1958) Nuclear magnetic relaxation of three and four spin molecules in a liquid. Phys Rev 109:1153–1158

Hwang T-L, Shaka AJ (1995) Water suppression that works-excitation sculpting using arbitrary waveforms and pulsed field gradients. J Magn Reson A 112:275–279

Jabar S, Adams LA, Wang Y, Aurelio L, Graham B, Otting G (2017) Chemical tagging with *tert*-butyl and trimethylsilyl groups for measuring intermolecular nuclear Overhauser effects in a large protein–ligand complex. Chem Eur J 23:13033–13036

Loscha KV, Herlt AJ, Qi R, Huber T, Ozawa K, Otting G (2012) Multiple-site labeling of proteins with unnatural amino acids. Angew Chem Int Ed 51:2243–2246

Mahawaththa MC, Pearce BJG, Szabo M, Graham B, Klein CD, Nitsche C, Otting G (2017) Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. Antiviral Res 142:141–147

Mukai T, Hayashi A, Iraha F, Sato A, Ohtake K, Yokoyama S, Sakamoto K (2010) Codon reassignment in the *Escherichia coli* genetic code. Nucleic Acids Res 38:8188–8195

Mukai T, Hoshi H, Ohtake K, Takahashi M, Yamaguchi A, Hayashi A, Yokoyama S, Sakamoto K (2015) Highly reproductive *Escherichia coli* cells with no specific assignment to the UAG codon. Sci Rep 5:9699

Müller N, Bodenhausen G, Ernst RR (1987) Relaxation-induced violations of coherence transfer selection rules in nuclear magnetic resonance. J Magn Reson 75:297–334

Neylon C, Brown SE, Kralicek AV, Miles CS, Love CA, Dixon NE (2000) Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. Biochemistry 39:11989–11999

Nilsson M, Rydén-Aulin M (2003) Glutamine is incorporated at the nonsense codons UAG and UAA in a suppressor-free *Escherichia coli* strain. Biochim Biophys Acta 1627:1–6

O'Donoghue P, Prat L, Heinemann IU, Ling J, Odoi K, Liu WR, Söll D (2012) Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNAPyl for genetic code expansion. FEBS Lett 586:3931–3937

Ohtake K, Sato A, Mukai T, Hino N, Yokoyama S, Sakamoto K (2012) Efficient decoding of the UAG triplet as a full-fledged sense codon enhances the growth of a prfA-deficient strain of *Escherichia coli*. J Bacteriol 194:2606–2613

Ozawa K, Loscha KV, Kuppan KV, Loh CT, Dixon NE, Otting G (2012) High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites. Biochem Biophys Res Commun 418:652–656

Torizawa T, Shimizu M, Taoka M, Miyano H, Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. J Biomol NMR 30:311–325

Werbelow LG, Marshall AG (1973) Internal rotation and nonexponential methyl nuclear relaxation for macromolecules. J Magn Reson 11:299–313

Young TS, Ahmad I, Yin JA, Schultz PG (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. J Mol Biol 395:361–374

Young DD, Young TS, Jahnz M, Ahmad I, Spraggon G, Schultz PG (2011) An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. Biochemistry 50:1894–1900