Chemical Science

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2019, **10**, 4994

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

Received 20th November 2018
Accepted 15th April 2019

DOI: 10.1039/c8sc05167d

rsc.li/chemical-science

Synthesis and characterization of an unnatural boron and nitrogen-containing tryptophan analogue and its incorporation into proteins†

Katherine Boknevitz, James S. Italia, Bo Li, Abhishek Chatterjee* and Shih-Yuan Liu*

A boron and nitrogen containing unnatural analogue of tryptophan is synthesized through the functionalization of BN-indole. The spectroscopic properties of BN-tryptophan are reported with respect to the natural tryptophan, and the incorporation of BN-tryptophan into proteins expressed in *E. coli* using selective pressure incorporation is described. This work shows that a cellular system can recognize the unnatural, BN-containing tryptophan. More importantly, it presents the first example of an azaborine containing amino acid being incorporated into proteins.

Introduction

The strategy of BN/CC isosterism, or the substitution of a carbon–carbon (CC) double bond with a boron-nitrogen (BN) bond, offers a unique way to alter the physical and chemical properties of arenes without significantly altering the steric profile.[1] This concept can be utilized to expand the chemical space of arene-containing, organic molecules that are prominent in medicinal chemistry by replacing a CC bond with a BN bond in benzene (1,2-dihydro-1,2-azaborine) or other arenes.[2] The inclusion of the BN bond results in a modification of certain electronic characteristics,[3] such as dipole moment,[4] reactivity,[5] and hydrogen bonding capabilities,[6] while maintaining the same number of atoms and π-electrons. These features could potentially be utilized in biomedical studies to alter the properties of known, biologically active, arene-containing compounds with minimal disturbance to the size of the molecule.

Little research has been done on the interactions of this class of arenes with biological systems. So far, it has been shown that 1,2-azaborines can bind to the non-polar binding site in biological macromolecules.[7] More recently, through isothermal titration calorimetry and protein crystal structure analysis, we have demonstrated the increased binding strength of benzene and ethylbenzene BN-analogues in the polar binding pocket of T4 lysozyme mutants through a hydrogen bonding interaction unavailable in the carbonaceous compounds.[6] As a pharmacophore in medicinal chemistry, BN analogues of naphthalene-containing biologically active compounds have been studied through ADMET (absorption, distribution, metabolism, excretion, toxicity) and biological activity profiling, both *in vivo* and *in vitro*.[8,9] We have shown, also through ADMET profiling, that the 1,2-azaborine analogues of biologically active compounds can increase the bioavailability of the drug without significantly altering the activity of the all-carbon substrate.[10] To further expand our research into biological applications, we pursued the synthesis of a boron and nitrogen-containing analogue of tryptophan (Fig. 1).

Unnatural amino acids (UAAs) are tools used to explore and engineer the function, properties, and stability of proteins.[11] Some UAAs contain main group elements, such as boron,[12] phosphorous,[13] selenium,[14] and tellurium,[14c] which are elements not typically found in biological studies. These UAAs have been successfully used as handles for bioconjugation,[14a,d,16] fluorescent tags,[13,15] and biosensors.[15]

Numerous other UAAs have been developed as endogenous amino acid analogues to alter the fluorescent character of proteins with minimal perturbation to the native structure of the protein,[17] while others have been developed to probe the cellular mechanisms of protein synthesis.[18] In many cases, tryptophan residues serve as the target of these studies.[19] Tryptophan is the main source of intrinsic fluorescence in proteins,[20] participates in various π-system interactions,[21] and

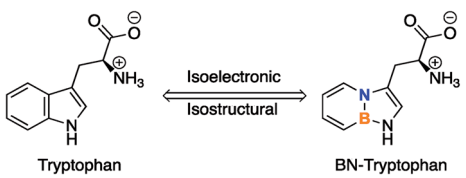


Fig. 1 BN/CC isosterism and BN-tryptophan.

Department of Chemistry, Boston College, Chestnut Hill, MA 02467, USA. E-mail: shihyuan.liu@bc.edu; abhishek.chatterjee@bc.edu

† Electronic supplementary information (ESI) available. CCDC 1879929. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c8sc05167d

accounts for about 1% of amino acids found in proteins.[22] Tryptophan's relative rarity, role in protein stability, and utility as a spectroscopic handle mark it as a useful target for protein studies.

The addition of a BN-analogue of tryptophan to the UAA toolbox could offer an alternative approach to probing protein function and fluorescent properties while also demonstrating the utility of azaborine chemistry in a biological context. It has been shown that the pK_a of the N–H proton in the internal BN-indole is higher than that of the natural indole (~30 and 21, respectively) and has altered spectroscopic characteristics.[23,24] These distinctive properties encouraged us to pursue the synthesis and application of a BN-analogue of tryptophan. In this article, we report the synthesis, characterization, and application of the first azaborine-containing canonical amino acid analogue.

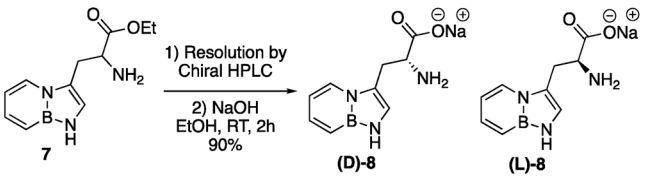
Results and discussion

The synthesis of BN-tryptophan (Scheme 1) begins with the regioselective electrophilic aromatic substitution (EAS) of TBS-BN-indole **1** with dimethyliminium chloride, as demonstrated previously in our group.[25] The resulting EAS product **2** is methylated with iodomethane and subsequently displaced with silyl-ketene-acetal **4**.[26] In aqueous acidic conditions, the Schiff base protecting group of **5** can be removed. This is followed by deprotection of the silyl protecting group to yield BN-tryptophan ethyl ester **7**.

By chiral HPLC, the enantiomers were resolved to give enantiomerically enriched D- and L-BN-tryptophan esters (Scheme 2). Lastly, through a basic hydrolysis in ethanol, the parent resolved amino acids **8** were obtained. A crystal structure was also acquired for the triflate salt of ester **6** (Fig. 2).

We also determined the absorption and emission properties of the resulting amino acid sodium salts in comparison with natural tryptophan L-**9** (Fig. 3). Tryptophan's absorbance maximum was found at $\lambda = 279$ nm in water, whereas BN-tryptophan's absorbance maximum is bathochromically shifted to $\lambda = 294$ nm. The fluorescence spectrum of L-**9** shows an emission maximum of $\lambda = 356$ nm while L-**8** shows a maximum at $\lambda = 394$ nm. Consequently, BN-tryptophan shows a larger Stokes shift of 8633 cm^{-1} compared to 7752 cm^{-1} of tryptophan L-**9**. Both display a similar quantum yield of ~0.3. The solvatochromic properties were explored as well and were consistent with previous studies (see ESI Fig. S6 for details†).[28]

With BN-tryptophan in hand, we evaluated the possibility of its incorporation into proteins as a surrogate of L-tryptophan using the selective pressure incorporation method with a tryptophan auxotrophic strain of *E. coli*.[19b] In Fig. 4, we demonstrated that BN-tryptophan does not inhibit cell growth in the



Scheme 2 Synthesis of (L/D)-BN-tryptophan **8**.

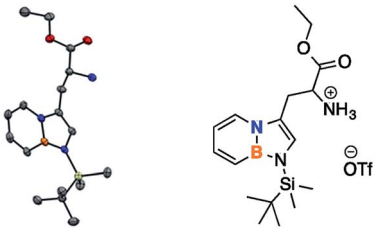
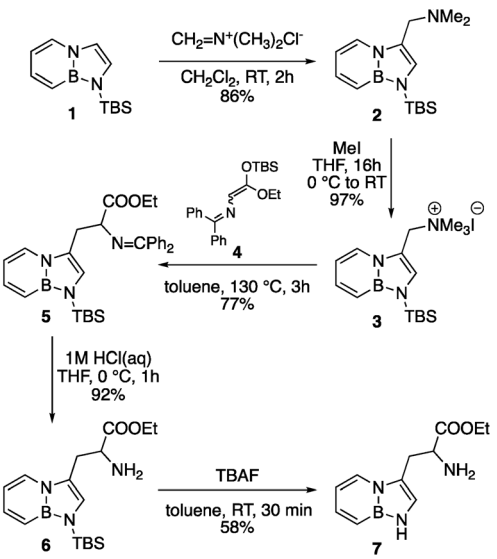
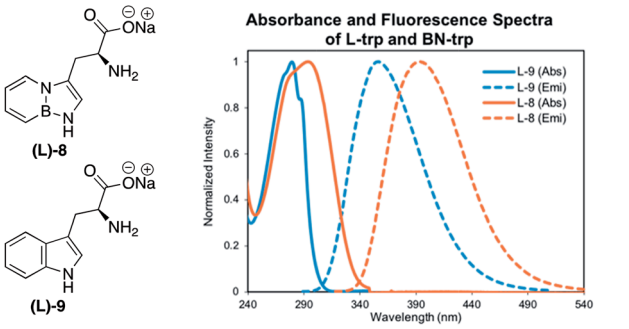


Fig. 2 Crystal structure of BN-tryptophan ester **6**. Hydrogen atoms are omitted for clarity.



Scheme 1 Synthesis of BN-tryptophan ester **7**.



	λ_{abs} (nm)	λ_{em} (nm)	ε (M^{-1}cm^{-1})	Stokes Shift (cm^{-1})	ϕ
L-trp **9**	279	356	6154	7752	0.25 (0.31)[a]
BN-trp **L8**	295	394	6410	8633	0.29

Measured in water, pH 10. [a] Literature value.[27]

Fig. 3 Absorbance and emission properties of BN-trp *vs.* L-trp.

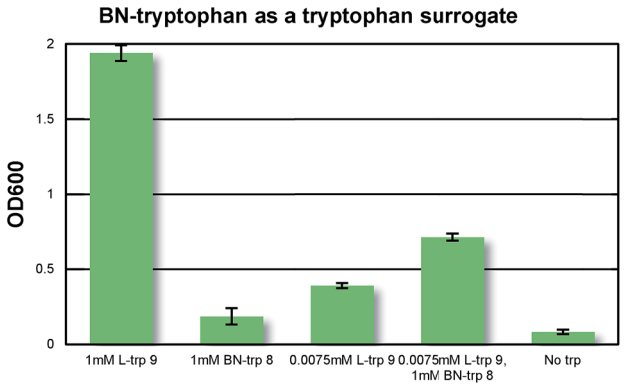


Fig. 4 Evaluation of BN-trp **8** as a tryptophan surrogate in *E. coli*. Cells were grown in defined minimal medium with the indicated concentration of tryptophan/BN-tryptophan and optical densities were measured after 16 hours.

presence of natural tryptophan, confirming its lack of general toxicity. In a defined growth medium lacking tryptophan, addition of BN-trp led to weak *E. coli* growth relative to a culture that did not receive any tryptophan analogue. However, addition of L-tryptophan to the same culture resulted in significantly higher growth. This indicates that BN-trp is accepted as a substrate for the endogenous tryptophanyl-tRNA synthetase, but proteome-wide replacement of L-tryptophan with BN-trp is not well-tolerated by *E. coli*.

To further our studies, we incorporated our UAA, along with three selected tryptophan analogues, which are known substrates for the endogenous tryptophanyl-tRNA synthetase, into superfolder green fluorescent protein (sfGFP).[29] In this protein, there is a single tryptophan residue, which precedes the chromophore, as seen in Fig. 5a. Expression of the full-length fluorescent sfGFP is contingent upon successful decoding of the TGG codon. Indeed, this is what we see (Fig. 5c); the similar levels of fluorescence normalized with respect to optical density supports the notion that BN-tryptophan was incorporated into the sfGFP. We also see similar or improved incorporation efficiency of BN-trp relative to 7-aza, 5-F, and 5-Me-tryptophans, which are established surrogates of L-tryptophan.[30]

We also incorporated BN-trp into a sfGFP mutant which harbors an additional tryptophan codon (sfGFP-151-TGG)[31] at a surface exposed site and isolated the protein by immobilized metal-ion chromatography using a C-terminal poly-histidine tag. Analysis by LC/MS shows the expected mass (Fig. S11†). To further confirm the presence of the BN-trp in this protein, which has essentially the same mass as natural tryptophan, we took advantage of its unique sensitivity to the mild oxidant hydrogen peroxide. sfGFP incorporating either tryptophan or BN-trp was briefly subjected to 1 mM hydrogen peroxide followed by full-protein MS analysis. It was found that, as shown in Fig. S11,† the protein containing tryptophan underwent partial oxidation, which is expected. In stark contrast, the identical treatment led to near-complete loss of the original peak for sfGFP-BN-trp. Since these two otherwise identical proteins only differ by the presence of BN-trp, we can conclude that BN-tryptophan displays different reactivity from

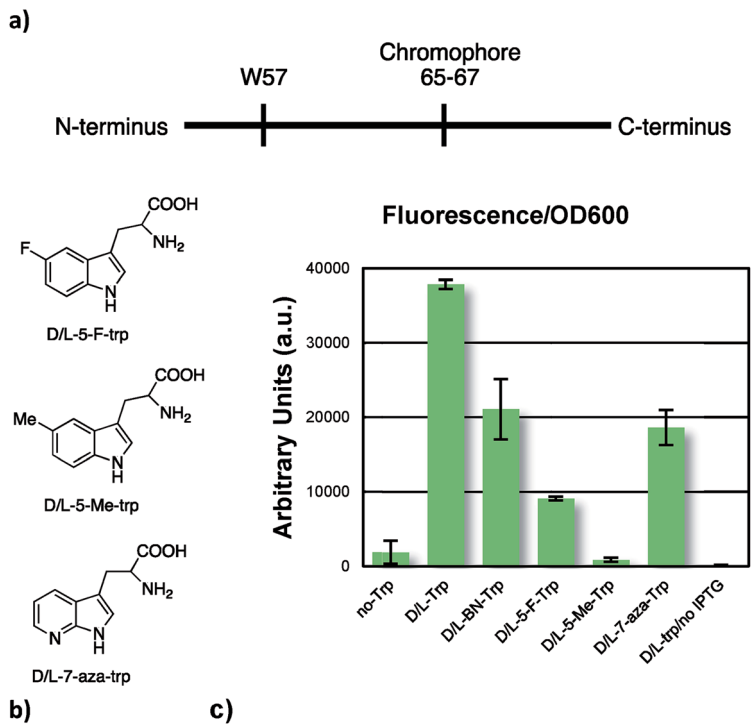


Fig. 5 (a) Representation of sfGFP sequence. (b) Tryptophan analogues used for comparison. (c) Fluorescence/OD plot. Cells transformed with sfGFP plasmid were grown with a limiting concentration of L-trp (0.0075 mM) to 0.5 OD_{600}. Tryptophan substrate and IPTG were added and cells were allowed to grow to confluence. Fluorescence and OD_{600} were measured.

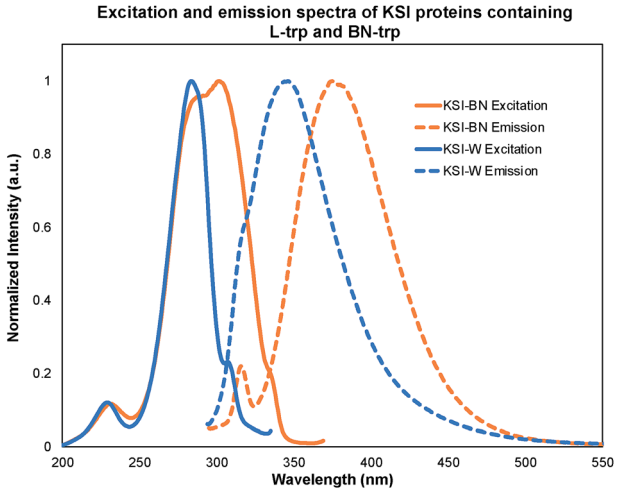


Fig. 6 Excitation and emission spectra of BN-trp and L-trp containing KSI. Conditions: PBS buffer (phosphate buffered saline: 20 mM Na_2HPO_4, 300 mM NaCl, pH 7.4), imidazole: ~0.2 mM. Excitation wavelength: 284 nm for KSI-W and 285 nm for KSI-BN.

tryptophan, a characteristic that can potentially be utilized in future studies.

Since tryptophan serves as the primary source of protein fluorescence, the incorporation of the BN-trp **8** should result in a bathochromic shift in the excitation and emission spectra when compared to the wild-type protein. To probe this possibility, we incorporated BN-trp **8** into the non-fluorescent protein, ketosteroid isomerase (KSI), which is a homo-dimeric protein that houses two tryptophan residues in each monomer. As expected, the fluorescence of the BN-containing protein (KSI-BN) was bathochromically shifted with respect to the L-trp containing protein (KSI-W) (Fig. 6). The KSI-BN and KSI-W fluorescence maxima were found to be 372 and 342 nm, respectively. Remarkably, the tyrosine fluorescence peak[32] (λ_{max} = 316 nm) can be clearly distinguished from the BN-trp peak in KSI-BN. On the other hand, the tyrosine emission signal cannot be resolved from the signal associated with the natural tryptophan in KSI-W under otherwise identical conditions. The distinct fluorescence of BN-Trp and its nearly indistinguishable geometric structure relative to tryptophan makes it a useful probe to study protein functions.

Conclusions

In summary, we synthesized a BN-analogue of the canonical amino acid, tryptophan, and characterized its spectroscopic properties with respect to its natural tryptophan counterpart. We demonstrated through selective pressure incorporation that BN-tryptophan can be incorporated into proteins *in vivo* and that two different (GFP and KSI) proteins containing BN-tryptophan can be isolated. This work shows that the natural tryptophanyl-tRNA synthetase can recognize an azaborine containing amino acid, further solidifying the potential use of BN/CC isosterism of arenes in a biological context. We hope to utilize this UAA in further protein studies and further classify its properties with respect to the natural substrate.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the National Institutes of Health NIGMS (R01-GM094541) to S.-Y. L. and by NIGMS (R01-GM124319) to A. C.

Notes and references

1 For recent reviews, see: (*a*) Z. X. Giustra and S.-Y. Liu, *J. Am. Chem. Soc.*, 2018, **140**, 1184–1194; (*b*) G. Belanger-Chabot, H. Braunschweig and D. K. Roy, *Eur. J. Inorg. Chem.*, 2017, 4353–4368; (*c*) M. M. Morgan and W. E. Piers, *Dalton Trans.*, 2016, **45**, 5920–5924.

2 For early work on BN/CC isosterism in biomedical research, see: (*a*) H.-B. Zhou, K. W. Nettles, J. B. Bruning, Y. Kim, A. Joachimiak, S. Sharma, K. E. Carlson, F. Stossi, B. S. Katzenellenbogen, G. L. Greene and J. A. Katzenellenbogen, *Chem. Biol.*, 2007, **14**, 659–669; (*b*) H. Ito, K. Yumura and K. Saigo, *Org. Lett.*, 2010, **12**, 3386–3389.

3 For examples, see: (*a*) C. R. McConnell, P. G. Campbell, C. R. Fristoe, P. Memmel, L. N. Zakharov, B. Li, C. Darrigan, A. Chrostowska and S.-Y. Liu, *Eur. J. Inorg. Chem.*, 2017, 2207–2210; (*b*) Z. Liu, J. S. A. Ishibashi, C. Darrigan, A. Dargelos, A. Chrostowska, B. Li, M. Vasiliu, D. A. Dixon and S.-Y. Liu, *J. Am. Chem. Soc.*, 2017, **139**, 6082–6085; (*c*) S. Xu, T. C. Mikulas, L. N. Zakharov, D. A. Dixon and S.-Y. Liu, *Angew. Chem., Int. Ed.*, 2013, **52**, 7527–7531; (*d*) P. G. Campbell, E. R. Abbey, D. Neiner, D. J. Grant, D. A. Dixon and S.-Y. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 18048–18050.

4 A. Chrostowska, S. Xu, A. N. Lamm, A. Mazière, C. D. Weber, A. Dargelos, P. Baylère, A. Graciaa and S.-Y. Liu, *J. Am. Chem. Soc.*, 2012, **134**, 10279–10285.

5 (*a*) K. Edel, X. Yang, J. S. A. Ishibashi, A. N. Lamm, C. Maichle-Mossmer, Z. X. Giustra, S.-Y. Liu and H. F. Bettinger, *Angew. Chem., Int. Ed.*, 2018, **57**, 5296–5300; (*b*) R. J. Burford, B. Li, M. Vasiliu, D. A. Dixon and S.-Y. Liu, *Angew. Chem., Int. Ed.*, 2015, **54**, 7823–7827; (*c*) A. W. Baggett, M. Vasiliu, B. Li, D. A. Dixon and S.-Y. Liu, *J. Am. Chem. Soc.*, 2015, **137**, 5536–5541; (*d*) A. N. Brown, B. Li and S.-Y. Liu, *J. Am. Chem. Soc.*, 2015, **137**, 8932–8935; (*e*) A. N. Lamm, E. B. Garner, D. A. Dixon and S.-Y. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 8157–8160; (*f*) J. Pan, J. W. Kampf and A. J. Ashe, *Org. Lett.*, 2007, **9**, 679–681.

6 H. Lee, M. Fischer, B. K. Shoichet and S.-Y. Liu, *J. Am. Chem. Soc.*, 2016, **138**, 12021–12024.

7 L. Liu, A. J. V. Marwitz, B. W. Matthews and S.-Y. Liu, *Angew. Chem., Int. Ed.*, 2009, **48**, 6817–6819.

8 (*a*) F. J. R. Rombouts, F. Tovar, N. Austin, G. Tresadern and A. A. Trabanco, *J. Med. Chem.*, 2015, **58**, 9287–9295; (*b*) A. Vlasceanu, M. Jessing and J. P. Kilburn, *Bioorg. Med. Chem.*, 2015, **23**, 4453–4461.

9 For leading references on the general method for their preparation, see: (*a*) G. H. Davies, Z. Z. Zhou, M. Jouffroy and G. A. Molander, *J. Org. Chem.*, 2017, **82**, 549–555; (*b*) G. A. Molander and S. R. Wisniewski, *J. Org. Chem.*, 2014, **79**, 6663–6678; (*c*) G. A. Molander, S. R. Wisniewski and J. Amani, *Org. Lett.*, 2014, **16**, 5636–5639.

10 P. Zhao, D. O. Nettleton, R. Karki, F. J. Zecri and S.-Y. Liu, *ChemMedChem*, 2017, **12**, 358–361.

11 For an overview, see: (*a*) D. D. Young and P. G. Schultz, *ACS Chem. Biol.*, 2018, **13**, 854–870; (*b*) J. W. Chin, *Nature*, 2017, **550**, 53–60.

12 D. B. Diaz and A. K. Yudin, *Nat. Chem.*, 2017, **9**, 731–742.

13 M. Arribat, E. Rémond, S. Clément, A. V. D. Lee and F. Cavelier, *J. Am. Chem. Soc.*, 2018, **140**, 1028–1034.

14 (*a*) A. Dantas de Araujo, S. R. Perry and D. P. Fairlie, *Org. Lett.*, 2018, **20**, 1453–1456; (*b*) T. Mukai, A. Sevostyanova, T. Suzuki, X. Fu and D. Söll, *Angew. Chem., Int. Ed.*, 2018, **57**, 7215–7219; (*c*) K. Satheeshkumar, S. Raju, H. B. Singh and R. J. Butcher, *Chem.–Eur. J.*, 2018, **24**, 17513–17522; (*d*) D. T. Cohen, C. Zhang, C. M. Fadzen, A. J. Mijalis, L. Hie, K. D. Johnson, Z. Shriver, O. Plante, S. J. Miller, S. L. Buchwald and B. L. Pentelute, *Nat. Chem.*, 2019, **11**, 78–85.

15 For an example of the use of the main group element arsenic as a fluorogenic probe, see: G.-M. Fang and O. Seitz, *ChemBioChem*, 2017, **18**, 189–194.

16 C. Zhang, E. V. Vinogradova, A. M. Spokoyny, S. L. Buchwald and B. L. Pentelute, *Angew. Chem., Int. Ed.*, 2019, **58**, 4810–4839.

17 (*a*) N. Budisa, M. Rubini, J. H. Bae, E. Weyher, W. Wenger, R. Golbik, R. Huber and L. Moroder, *Angew. Chem., Int. Ed.*, 2002, **41**, 4066–4069; (*b*) J. B. Alexander Ross, A. G. Szabo and C. W. V. Hogue, *Methods Enzymol.*, 1997, **278**, 151–190.

18 J. T. Ngo and D. A. Tirrell, *Acc. Chem. Res.*, 2011, **44**, 677–685.

19 (*a*) C. Minks, R. Huber, L. Moroder and N. Budisa, *Biochemistry*, 1999, **38**, 10649–10659; (*b*) N. Budisa, S. Alefelder, J. H. Bae, R. Golbik, C. Minks, R. Huber and L. Moroder, *Protein Sci.*, 2001, **10**, 1281–1292; (*c*) S. Lepthien, B. Wiltschi, B. Bolic and N. Budisa, *Appl. Microbiol. Biotechnol.*, 2006, **73**, 740–754; (*d*) I. Kwon and D. A. Tirrell, *J. Am. Chem. Soc.*, 2007, **129**, 10431–10437; (*e*) Z. Zhang, L. Alfonta, F. Tian, B. Bursulaya, S. Uryu, D. S. King and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 8882–8887; (*f*) S. Lepthien, M. G. Hoesl, L. Merkel and N. Budisa, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 16095–16100; (*g*) J. H. Bae, S. Alefelder, J. T. Kaiser, R. Friedrich, L. Moroder, R. Huber and N. Budisa, *J. Mol. Biol.*, 2001, **309**, 925–936.

20 P. R. Callis, *Methods Enzymol.*, 1997, **278**, 113–150.

21 (*a*) C. Ruan and M. T. Rodgers, *J. Am. Chem. Soc.*, 2004, **126**, 14600–14610; (*b*) D. A. Dougherty, *J. Nutr.*, 2007, **137**, 1504S–1508S; (*c*) D. A. Dougherty, *J. Org. Chem.*, 2008, **73**, 3667–3673; (*d*) S. K. Burley and G. A. Petsko, *Science*, 1985, **229**, 23–28.

22 C. P. McCaul and R. D. Ludescher, *Photochem. Photobiol.*, 1999, **70**, 166–171.

23 For an overview, see: E. R. Abbey and S.-Y. Liu, *Org. Biomol. Chem.*, 2013, **11**, 2060–2069.

24 E. R. Abbey, L. N. Zakharov and S.-Y. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 11508–11511.

25 (*a*) E. R. Abbey, L. N. Zakharov and S.-Y. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 16340–16342; (*b*) A. Chrostowska, S. Xu, A. Mazière, K. Boknevitz, B. Li, E. R. Abbey, A. Dargelos, A. Graciaa and S.-Y. Liu, *J. Am. Chem. Soc.*, 2014, **136**, 11813–11820.

26 T. Wartmann and T. Lindel, *Eur. J. Org. Chem.*, 2013, **9**, 1649–1652.

27 A. G. Szabo and D. M. Rayner, *J. Am. Chem. Soc.*, 1980, **102**, 554–563.

28 M. Saif, J. R. Widom, S. Xu, E. R. Abbey, S.-Y. Liu and A. H. Marcus, *J. Phys. Chem. B*, 2015, **119**, 7985–7993.

29 J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo, *Nat. Biotechnol.*, 2006, **24**, 79–88.

30 (*a*) A. B. Pardee, V. G. Shore and L. S. Prestidge, *Biochim. Biophys. Acta*, 1956, **21**, 406–407; (*b*) S. Barlati and O. Ciferri, *J. Bacteriol.*, 1970, **101**, 166–172; (*c*) E. A. Pratt and C. Ho, *Biochemistry*, 1975, **14**, 3035–3040.

31 Refer to: J. S. Italia, P. S. Addy, C. J. J. Wrobel, L. A. Crawford, M. J. Lajoie, Y. Zheng and A. Chatterjee, *Nat. Chem. Biol.*, 2017, **13**, 446–450. See ESI for details.†

32 D. E Duggan and S. Udenfriend, *J. Biol. Chem.*, 1956, **223**, 313–319.