the FEBS Journal

Translational incorporation of L-3,4-dihydroxyphenylalanine into proteins

Kiyoshi Ozawa[1], Madeleine J. Headlam[1], Dmitri Mouradov[2], Stephen J. Watt[3], Jennifer L. Beck[3], Kenneth J. Rodgers[4], Roger T. Dean[5], Thomas Huber[2], Gottfried Otting[1] and Nicholas E. Dixon[1]

1 Research School of Chemistry, Australian National University, Canberra, Australia
2 Departments of Biochemistry and Mathematics, University of Queensland, St. Lucia, Australia
3 Department of Chemistry, University of Wollongong, Australia
4 The Heart Research Institute, Sydney, Australia
5 The University of Canberra, Australia

Keywords
cell-free protein synthesis; DOPA; protein oxidation; protein NMR; protein misfolding

Correspondence
N. E. Dixon, Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia
Fax: +61 2 612 50750
Tel: +61 2 612 54391
E-mail: dixon@rsc.anu.edu.au

(Received 16 February 2005, revised 2 April 2005, accepted 25 April 2005)

doi:10.1111/j.1742-4658.2005.04735.x

An *Escherichia coli* cell-free transcription/translation system was used to explore the high-level incorporation of L-3,4-dihydroxyphenylalanine (DOPA) into proteins by replacing tyrosine with DOPA in the reaction mixtures. ESI-MS showed specific incorporation of DOPA in place of tyrosine. More than 90% DOPA incorporation at each tyrosine site was achieved, allowing the recording of clean ^{15}N-HSQC NMR spectra. A redox-staining method specific for DOPA was shown to provide a sensitive and generally applicable method for assessing the cell-free production of proteins. Of four proteins produced in soluble form in the presence of tyrosine, two resulted in insoluble aggregates in the presence of high levels of DOPA. DOPA has been found in human proteins, often in association with various disease states that implicate protein aggregation and/or misfolding. Our results suggest that misfolded and aggregated proteins may result, in principle, from ribosome-mediated misincorporation of intracellular DOPA accumulated due to oxidative stress. High-yield cell-free protein expression systems are uniquely suited to obtain rapid information on solubility and aggregation of nascent polypeptide chains.

The incorporation of non-natural amino acids opens up the possibility to endow proteins with properties that cannot be attained with the 20 natural amino acids encoded by DNA base triplets. The incorporation of non-natural amino acids can readily be achieved with the natural protein-translational machinery, if the structure of the modified amino acid is closely related to the natural amino acid, so that it can be loaded onto tRNA by one of the natural aminoacyl-tRNA synthetases.

A wide range of non-natural amino acids has been incorporated into proteins in this way [1]. In general, the efficiency of incorporation decreases with increasing K_M value of the aminoacyl-tRNA synthetase for the respective amino acid. This holds, in particular, for the *in vivo* incorporation of non-natural amino acids, where a pool of natural amino acids is always present. This problem can be circumvented by the use of auxotrophic strains [1] or cell-free protein production systems derived from nonauxotrophic strains combined with a suitably manipulated medium for protein synthesis [2,3].

Recently, high-yield, cell-free protein production systems have become available that allow the synthesis of proteins in quantities sufficient for structural genomics applications [4–7]. High-level incorporation of seleno-methionine (Se-Met) for X-ray crystallography and fluoro-tryptophan (F-Trp) for NMR has been

Abbreviations
DOPA, L-3,4-dihydroxyphenylalanine; GFP, cycle 3 mutant green fluorescent protein; *h*CypA, human cyclophilin A; His_6-PpiB, N-terminal His_6-tagged PpiB; HMP, *Escherichia coli* flavohaemoglobin; HSQC, heteronuclear single-quantum coherence; NBT, nitroblue tetrazolium; PpiB, *E. coli* peptidyl-prolyl *cis-trans* isomerase B; RNAP, RNA polymerase; TyrRS, tyrosyl-tRNA synthetase.

demonstrated [8,9], but limited dilution of isotope-labelled with unlabelled natural amino acids has also been reported [6].

This study investigated the high-level, high-yield incorporation of L-3,4-dihydroxyphenylalanine (DOPA) into proteins by replacing tyrosine with DOPA in the reaction mixture of an *Escherichia coli* cell-free transcription/translation system. The K_M value for *E. coli* tyrosyl-tRNA synthetase (TyrRS) was reported to be ≈200-fold higher for L-DOPA than for L-tyrosine (1.4 mM vs. 6 μM) [10,11], i.e. the natural enzyme discriminates against DOPA one order of magnitude more strongly than the respective aminoacyl tRNA synthetases incorporating Se-Met vs. Met [12] and F-Trp vs. Trp [13]. DOPA-enrichment is advantageous as it allows the facile assessment of protein production levels, because a highly specific staining method is available [14]. Finally, DOPA is produced naturally in humans by tyrosinase-catalysed oxidation of tyrosine in melanocytes for melanin production, and by tyrosine hydroxylase in the brain for biosynthesis of catecholamine neurotransmitters [15].

In addition, the accumulation of protein-bound DOPA in cells and tissues is a feature of a number of pathologies associated with ageing, such as atherosclerosis [15] and cataractogenesis [16], where it derives at least in part from oxygen-radical mediated post-translational oxidation of tyrosine side chains in proteins [15]. It has recently been shown that DOPA can be incorporated directly from the medium into proteins in cultured mouse [16] and human [17] cells, and that incorporation relies on protein synthesis [16]. If translational (ribosome-mediated) incorporation of DOPA is a distinct possibility, the structural and functional consequences of DOPA incorporation would be important to assess.

This study used a preparative *E. coli* cell-free transcription/translation system [5,18] to incorporate DOPA into four different *in vitro*-synthesized proteins. MS and NMR spectroscopy were used to verify whether DOPA incorporation occurred at positions normally occupied by tyrosine, and to assess the level of DOPA incorporation. The effects on folding of the four proteins (all of which have known structures) were assessed by examination of their solubility following their *in vitro* synthesis.

Results

Protein synthesis in the presence of DOPA

The effect of substitution of DOPA for tyrosine was investigated in a preparative *in vitro* protein synthesis system that employs an *E. coli* cell-free (S30) extract as the source of ribosomes, aminoacyl-tRNA synthetases and translation factors [5,18]. We chose to examine the synthesis of four different proteins whose three-dimensional structures are known from X-ray crystallographic studies: the peptidyl-prolyl *cis–trans* isomerases *E. coli* (PpiB; Protein Data Bank Accession no. 2NUL) [19] and human cyclophilin A (*h*CypA; Protein Data Bank Accession no. 2CPL) [20], the *E. coli* flavohaemoglobin (HMP; Protein Data Bank Accession no. 1GVH) [21] and the *Aequorea victoria* green fluorescent protein (GFP; Protein Data Bank Accession no. 1EMA) [22]. The first two proteins had previously been shown to be produced in good yield in the *in vitro* reaction [5,6,18,23].

All four proteins were found to be synthesized to similarly high levels in the presence of 1 mM tyrosine or DOPA (Fig. 1A; data not shown for PpiB). Analysis of the supernatant and pellet fractions by Coomassie Brilliant Blue staining of a SDS/PAGE gel indicated that all were soluble or mostly soluble when expressed using tyrosine (Fig. 1A). Whereas PpiB (18 kDa) and HMP (44 kDa) were still largely soluble when they were synthesized with DOPA, > 50% of *h*CypA (18 kDa) and GFP (27 kDa) were in the insoluble fraction. This implies that incorporation of DOPA can interfere with correct protein folding.

The yield of PpiB depended remarkably little on the concentration of DOPA or tyrosine in the reaction mixtures. High yields similar to those obtained with 1 mM tyrosine or DOPA were obtained with either amino acid at 50 μM (data not shown). With 10 and 5 μM DOPA, the yields were ≈20 and 50% lower, respectively, than with 10 μM tyrosine, and there was still discernible production of PpiB when both amino acids were omitted. This is presumably because of the presence of a trace of tyrosine as a contaminant in the S30 extract or its biosynthesis during the reaction.

DOPA is incorporated into proteins during cell-free synthesis

To show that DOPA was incorporated into the translated proteins, we first used a redox staining method employing nitroblue tetrazolium (NBT), which detects proteins containing *o*-catechols, like DOPA, after their separation by SDS/PAGE and western transfer to poly(vinylidene difluoride) membranes [14,17]. The staining method was verified using purified His_6-PpiB that had been produced by cell-free synthesis in the presence of 0.05 or 1.0 mM DOPA or 1.0 mM tyrosine. The protein was purified in similar yields from each reaction (≈2 mg per 2 mL of reaction mixture) by metal-ion

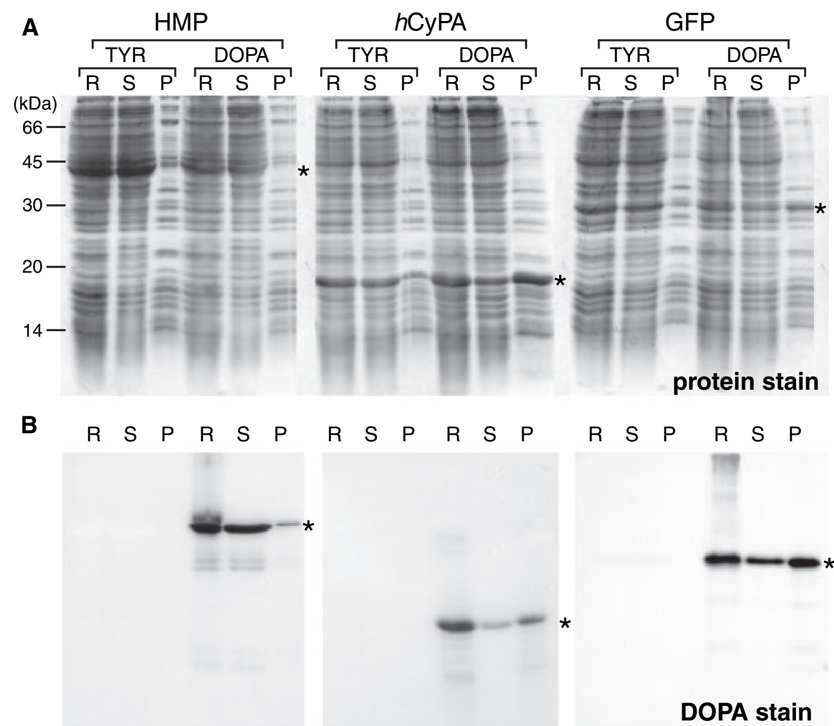

Fig. 1. Proteins are synthesized in good yields *in vitro* using DOPA in place of tyrosine. Duplicate *in vitro* synthesis reactions were carried out with 1 mM tyrosine (TYR) or with 1 mM DOPA. Proteins in equal portions of the complete reaction mixtures (R) or the fractionated soluble (S) and pelleted (P) fractions were separated by 15% SDS/PAGE. Duplicate gels were stained separately with Coomassie Brilliant Blue (A) and NBT (B) as described in Experimental procedures. The mobility of molecular mass marker proteins (sizes in kDa) were as indicated, and positions of newly synthesized full-length proteins are marked by asterisks.

affinity chromatography and analysed by duplicate SDS/PAGE gels that were stained either with Coomassie Brilliant Blue (Fig. 2A) or by redox staining

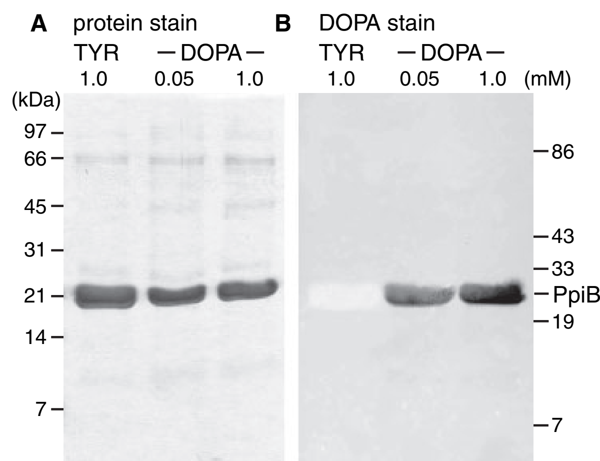

Fig. 2. DOPA incorporated into His_6-PpiB during its *in vitro* synthesis can be detected by redox staining. His_6-PpiB was isolated by Ni^{2+}-affinity chromatography from reaction mixtures containing 1 mM tyrosine (TYR) or 0.05 or 1 mM DOPA, as indicated. Approximately equal amounts (12.5 μg) of each of the purified protein samples were separated on duplicate 15% SDS/PAGE gels, which were stained with Colloidal Coomassie protein stain (A) or with NBT (B) as described in Experimental procedures. Mobilities of marker proteins (sizes in kDa; prestained in B) were as indicated.

(Fig. 2B). Only His_6-PpiB that had been produced in the presence of DOPA stained with NBT, and the staining intensity was somewhat higher for the sample produced with 1.0 mM in comparison with 0.05 mM DOPA.

This staining method could also be used to detect *de novo* synthesized proteins in the crude reaction mixtures. Proteins synthesized with tyrosine were not stained by NBT, but those made in the presence of DOPA were readily and specifically detected (Fig. 1B). These results show the incorporation of DOPA during cell-free protein synthesis. The sensitivity of this method is comparable with staining by Coomassie Brilliant Blue, and only newly synthesized proteins were detected, including some minor species presumed to have been produced by proteolysis and/or premature termination of translation (Fig. 1B). It confirmed that high-level DOPA incorporation results in mostly insoluble protein in the cases of GFP and *h*CypA and mostly soluble protein for HMP and PpiB. Given that the chromophore in GFP involves a tyrosine (Tyr66), and that its photophysical properties are particularly sensitive to substitution of this residue [22,24,25], it was of interest to examine fluorescence spectra of *in vitro* synthesized GFP. The excitation and emission spectra of crude mixtures containing the fraction of soluble GFP that had been produced with tyrosine or DOPA were found to be identical. Nevertheless, the yield of fluorescence was low (10–20%) in the soluble fraction from the

DOPA sample compared with that prepared with tyrosine (data not shown). This indicates that although some portion of GFP was capable of folding correctly into a soluble form when DOPA was incorporated in place of Tyr66, the chromophore either did not form or was not appreciably fluorescent. The insoluble fraction was not noticeably fluorescent.

Having shown that DOPA could be translationally incorporated into various proteins, we next established that this occurred specifically in place of tyrosine. This was done in three ways: (a) by showing that the mass of intact purified DOPA–His_6-PpiB, as determined by ESI-MS under native conditions, was increased by 16 mass units per tyrosine residue; (b) by showing that a relative increase in mass of tryptic fragments from DOPA–His_6-PpiB, as determined by ESI-MS after separation by RP-HPLC was observed only for peptides that would otherwise contain tyrosine and (c) that NMR chemical shift changes for samples of selectively ^{15}N-labelled DOPA–PpiB relative to native ^{15}N-labelled PpiB were consistent with the specific incorporation of DOPA in place of tyrosine.

ESI-MS of DOPA-labelled His_6-PpiB

The mass of purified His_6-PpiB produced in the presence of 0.05 mM DOPA was compared with that of the normal protein produced with 1.0 mM tyrosine. Two species were present in the tyrosine sample in almost equal proportions, with M_r values of 19 221.7 and 19 249.7 (Fig. 3A, peaks A and B, respectively). The larger component corresponds to a form of the protein that retains the N-formyl group on the N-terminal methionine residue (calculated M_r 19 249.6), and the smaller is the mature protein produced after deformylation (calculated M_r 19 221.6). This is in accord with the results of a previous NMR study, in which amide resonances could be observed for the N-terminal methionine as well as for the following residue in *h*CypA [18], indicating that our S30 extract is deficient in peptide deformylase activity [6].

His_6-PpiB contains three tyrosine residues. When produced with DOPA, the protein contained several species (Fig. 3B). The most abundant had masses of 19 297.9 (peak G), 19 281.9 (peak F) and 19 269.3 (peak E), in order of decreasing intensity. These species correspond to the N-formylated protein with three and two, and deformylated protein with three DOPA residues, respectively. Semiquantitative assessment of the incorporation level of DOPA was made by comparison of the sum of the peak heights of the 3-DOPA species (peaks E, G, H and I) vs. the sum of the peak heights of the 2-DOPA species (peaks D and F). This ratio was found to be 3 : 1. The same ratio was found by comparing the peak heights of peaks E (2-DOPA species) and D (1-DOPA species). This suggests that about three-quarters of all tyrosine residues had been replaced by DOPA, corresponding to an incorporation level of > 90% at each of the three tyrosine sites. The less than 100% efficiency was presumably due to traces of tyrosine (and/or tyrosyl-tRNA) remaining in, or synthesized by, the cell-free extract.

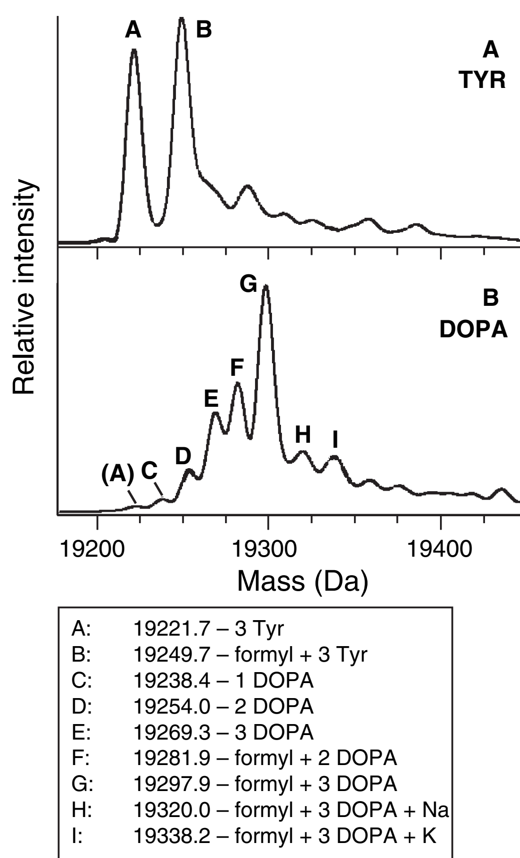

Fig. 3. Positive-ion ESI-MS analysis of *in vitro* synthesized His_6-PpiB. Samples of His_6-PpiB isolated by Ni^{2+}-affinity chromatography from reaction mixtures containing 1 mM tyrosine (A) or 0.05 mM DOPA (B) were analysed by ESI-MS in 10 mM ammonium acetate/formic acid (pH 3.0). Peaks in the transformed spectra are indicated, and their observed masses and compositions are indicated in the key shown below. Weak peaks in the spectra at higher masses are presumed to be adducts resulting from the presence of small amounts of sodium and potassium ions in the protein samples.

HPLC–ESI-MS of tryptic peptides from DOPA-labelled His_6-PpiB

Peptides resulting from partial tryptic digestion of His_6-PpiB that had been produced using 1 mM tyrosine or 1 mM DOPA were separated by HPLC and

analysed by in-line ESI-MS. Peptides of M_r > 500 (the threshold for ESI-MS), identified by correspondence between their M_r values and amino acid composition, spanned 165 (96%) of 172 residues of the amino acid sequence (Table 1). For both samples, a significant portion of the N-terminal peptides had a mass 28.0 units higher than expected; this confirms that the additional mass seen for much of the *in vitro* synthesized intact His$_6$-PpiB (Fig. 3B) is due to retention of the formyl group on the N-terminal methionine residue (see above) [6,18].

Three tryptic peptides, each containing one of the three tyrosine residues in His$_6$-PpiB, were observed at the expected mass in both samples. In the sample prepared with DOPA, three additional, more-abundant, peptides were observed, each with masses 16.0 units higher than these three (Table 1), as might be expected if each of the tyrosine residues was substantially replaced by DOPA. No peptides were observed that might correspond to significant replacement of another amino acid residue by DOPA.

NMR analysis of DOPA-labelled (^{15}N)PpiB

In a previous study [18], we showed that crude reaction mixtures containing *h*CypA produced *in vitro* in the presence of ^{15}N-labelled amino acids could be used directly, after dialysis into an NMR buffer (phosphate, pH 6.5), to record residue-specific ^{15}N-HSQC NMR spectra. The protein remained soluble during and following dialysis [18]. In contrast, in this study, the initially soluble portion of DOPA–*h*CypA precipitated quantitatively on dialysis into the NMR buffer, which indicates that incorporation of DOPA destabilizes the native structure of *h*CypA.

Therefore, the *E. coli* homologue of *h*CypA, PpiB, was used for NMR studies. We recorded ^{15}N-HSQC NMR spectra of crude mixtures containing PpiB that had been synthesized in the presence of tyrosine or DOPA using amino acid mixtures that contained both (^{15}N)cysteine and (^{15}N)phenylalanine in place of the corresponding unlabelled amino acids, and then dialysed into an NMR buffer (Fig. 4A). The buffer was identical to that used previously during assignment of the amide ^{15}N-^{1}H resonances in PpiB [23], and the availability of those data permitted the immediate assignment of most of the cysteine and phenylalanine amide resonances in the spectra of Fig. 4A. The protein samples were all completely soluble during their preparation, indicating that incorporation of DOPA did not irreversibly disrupt the structure of PpiB.

Table 1. ESI-MS tryptic peptide fingerprint of His$_6$-PpiB synthesized in the presence of tyrosine or DOPA. The His$_6$-PpiB used in these experiments contains, in addition to the wild-type PpiB sequence, an N-terminal Met-(His)$_6$ tag and an additional Asn residue at the C terminus. f-M, *N*-formyl-methionine. #, DOPA residue. Peptides containing DOPA are indicated in bold.

Observed mass (Da)		Theoretical	Amino	
Tyrosine	L-DOPA	mass (Da)	acids	Sequence
2592.5	2592.5	2593.2	1–21	f-MHHHHHHMVTFHTNHGDIVIK
2565.5	2565.6	2565.2	1–21	MHHHHHHMVTFHTNHGDIVIK
3825.9	3825.9	3824.8	1–32	f–MHHHHHHMVTFHTNHGDIVIKTFDDKAPETVK
3797.9	3797.9	3796.8	1–32	MHHHHHHMVTFHTNHGDIVIKTFDDKAPETVK
1250.7	1251.5	1250.6	22–32	TFDDKAPETVK
	2178.9	2178.0	22–39	**TFDDKAPETVKNFLD#CR**
2162.9	2162.9	2162.0	22–39	TFDDKAPETVKNFLDYCR
	946.4	946.4	33–39	**NFLD#CR**
930.4	930.4	930.4	33–39	NFLDYCR
	1413.7	1413.7	40–50	**EGF#NNTIFHR**
1397.5	1396.7	1397.7	40–50	EGFYNNTIFHR
1781.6	1781.6	1781.9	51–67	VINGFMIQGGGFEPGMK
1627.0	1627.8	1626.9	70–84	ATKEPIKNEANNGLK
1326.5	1326.6	1326.7	73–84	EPIKNEANNGLK
719.4	719.4	719.4	88–94	GTLAMAR
	5285.0	5283.4	95–142	**TQAPHSATAQFFINVVDNDFLNFS-GESLQGWG#CVFAEVVDGMDVVDK**
5269.0	5269.0	5267.4	95–142	TQAPHSATAQFFINVVDNDFLNFS-GESLQGWGYCVFAEVVDGMDVVDK
561.4	561.4	560.3	145–150	GVATGR
2413.7	2414.0	2413.2	151–172	SGMHQDVPKEDVIIESVTVSEN

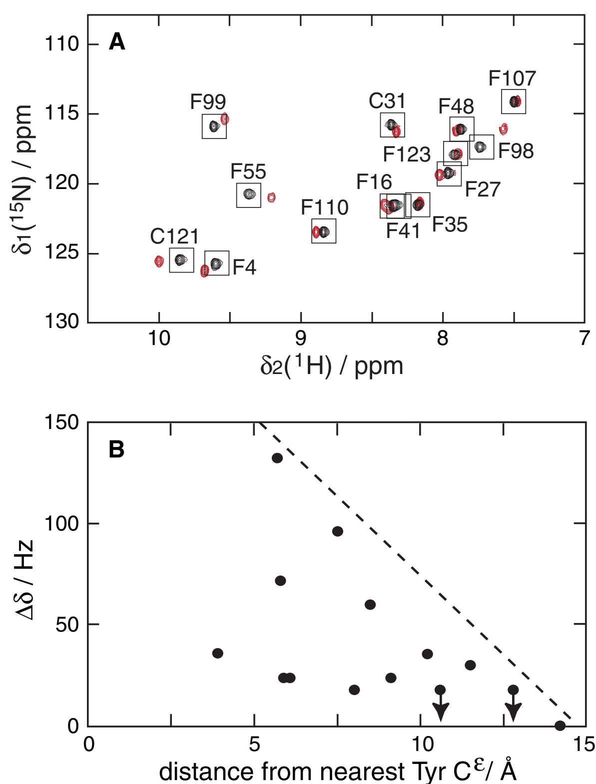

Fig. 4. Effect of incorporation of DOPA on the ^{15}N-HSQC NMR spectrum of crude reaction mixtures containing selectively ^{15}N-labelled PpiB. (A) PpiB was synthesized in reaction mixtures containing (^{15}N)cysteine and (^{15}N)phenylalanine (each at 0.35 mM) in place of unlabelled cysteine and phenylalanine, with either 1 mM tyrosine (peaks in black and highlighted with boxes) or 1 mM DOPA (in red). The amide resonances were assigned to the two cysteine and 12 phenylalanine residues of wild-type PpiB as shown [23]. (B) Correlation between amide chemical shift changes induced by DOPA and the distance from tyrosine side chains in the structure of PpiB. The chemical shift changes between DOPA-substituted and native PpiB are reported here as $\Delta\delta = ([\Delta\delta(^1\text{H})]^2 + [\Delta\delta(^{15}\text{N})]^2)^{1/2}$, calculated using values of $\Delta\delta(^1\text{H})$ and $\Delta\delta(^{15}\text{N})$ measured in Hz at 600 MHz ^{1}H NMR frequency. The distances were measured as the shortest distance between the respective amide protons to any of the $C^{\varepsilon1}$ or $C^{\varepsilon2}$ carbon atoms of a tyrosine residue in the crystal structure of PpiB (Protein Data Bank Accession no. 2NUL) [19] (Fig. 5). Arrows label the data points of the amide protons of Phe16 and Phe41 which overlap in the NMR spectra of wild-type and DOPA–PpiB. For one of these residues, the chemical shift change would be smaller than indicated. That all data points fall below the dashed line is consistent with expectations for substitution of DOPA at tyrosine residues in the protein.

PpiB contains 2 cysteine and 12 phenylalanine residues. The ^{15}N-HSQC spectrum of the DOPA–PpiB sample showed only the resonances expected for a single set of 14 ^{15}N-labelled residues, most of which were shifted relative to the signals of the tyrosine–PpiB sample. The remarkably homogeneous appearance of the ^{15}N-HSQC spectrum of the ^{15}N-labelled DOPA–PpiB sample could only be explained by uniform and > 90% incorporation of DOPA in place of any of the three tyrosines. The signal-to-noise ratio in the spectrum was insufficient to observe signals at the 10% level.

Figure 5 shows that the phenylalanine and cysteine residues are quite uniformly distributed in the three-dimensional structure of PpiB, generating a range of minimum distances with respect to the three tyrosine side chains. The introduction of additional hydroxyl groups resulting from replacement of tyrosine by DOPA changes the chemical environment and can thus, potentially, affect the chemical shifts of the protons in the immediate vicinity of the hydroxyl groups, whereas the chemical-shift changes of protons far from the DOPA side chains are always expected to be small. As expected, amide protons located far from the nearest C^{ε} atom of the tyrosine side chains in the three-dimensional structure of PpiB produced, at most, small changes in chemical shifts, and the largest changes in chemical shifts were observed for amide protons close to a DOPA side chain (Fig. 4B).

Discussion

At the outset of this study it was not clear that the cell-free synthetic reaction would produce good yields of proteins when DOPA was provided in place of tyrosine, nor whether they would be produced in a soluble and therefore native folded state. Purified *E. coli* TyrRS is capable of charging tRNA (presumably tRNA^{Tyr}) with DOPA *in vitro* [10,11] with a similar turnover number for L-tyrosine and DOPA, but a much higher value of K_M for DOPA [10]. Compared with the K_M values of the respective aminoacyl tRNA synthetases acting on Se-Met/Met or F-Trp/Trp pairs, the K_M value of TyrRS for the DOPA/Tyr pair is an order of magnitude less favourable for the non-natural amino acid [12,13]. It is thus remarkable [1] that highly efficient DOPA incorporation could be achieved using an S30 extract from a nonauxotrophic *E. coli* strain, and without significantly reduced protein yields (Fig. 1). The DOPA incorporation level of > 90% achieved at each tyrosine site compares favourably with a recent report, in which the cell-free incorporation of isotope-labelled tyrosine resulted in dilution with ≈ 3% unlabelled tyrosine [6]. Although the level of DOPA incorporation was sufficiently high to record a clean ^{15}N-HSQC spectrum, signals from incompletely DOPA-enriched protein molecules at the 10% level could interfere with the interpretation of weak NOESY cross-peaks. If desired, the incorporation level could

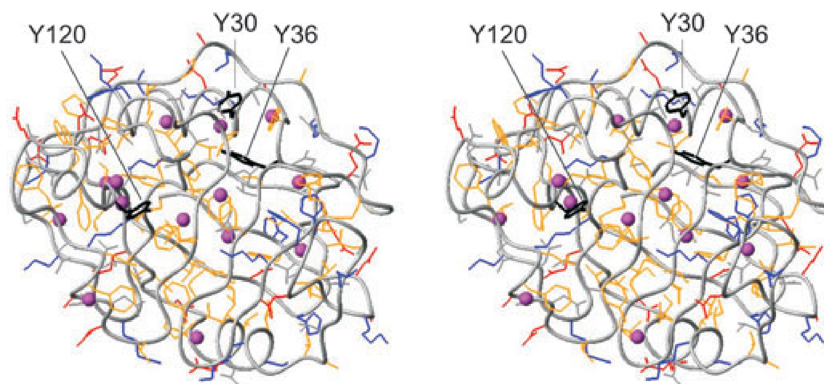

Fig. 5. Stereoview of the X-ray crystal structure of wild-type *E. coli* PpiB. Amide protons of the 12 phenylalanine and two cysteine residues are marked with spheres, and the side chains of the three tyrosine residues are highlighted in black. The following colour code is used for the other amino acid side chains: hydrophobic residues (Ala, Cys, Ile, Leu, Met, Phe, Pro, Trp, Val), yellow; positively (Arg, His, Lys) and negatively (Asp, Glu) charged residues, blue and red, respectively; hydrophilic residues (Asn, Gln, Ser, Thr), grey. The figure was prepared using MOLMOL [31].

probably be increased further by deriving the S30 extract from an *E. coli* strain that is auxotrophic for tyrosine. The N-terminal heterogeneity arising from incomplete deformylation did not result in any peak-doubling in the ^{15}N-HSQC spectrum, presumably because the nearest ^{15}N-labelled amide group was located > 9 Å from the N-terminus (Fig. 5).

This study further demonstrates that the presence of DOPA in newly synthesized proteins can easily be detected using a sensitive redox-staining procedure that detects catechols like DOPA [14,17]. Synthesis in the presence of DOPA and redox staining thus offers a convenient alternative to the use of antibodies or radiolabelled amino acids for the specific detection and quantification of *in vitro*-synthesized proteins. The residual fluorescence observed in the soluble fraction of DOPA-enriched GFP suggests that functional proteins can still be obtained in the presence of DOPA, provided that not all of the tyrosine residues are replaced.

MS (Fig. 3 and Table 1) and NMR (Fig. 4) data with PpiB produced by this route showed that DOPA is incorporated exclusively in place of tyrosine. This is consistent with our previous studies using mammalian cells in culture, which demonstrated that [^{14}C]DOPA competed with tyrosine for incorporation into cell proteins [16,17]. The current data strongly suggest that DOPA is loaded onto tRNA^{Tyr} by TyrRS and incorporated into proteins during their translational synthesis in *E. coli* in complete analogy to tyrosine.

Cell-free protein synthesis turns out to provide a fast route for the qualitative detection of the formation of protein aggregates, as insoluble proteins form visible precipitates already during protein synthesis. Two of the four proteins examined (*h*CypA and GFP) were seen to be largely insoluble when produced *in vitro* with DOPA, but were soluble when produced with tyrosine (Fig. 1). Because incorporation of DOPA was not 100% efficient, this suggests that proteins were incapable of correct folding when certain structurally important tyrosine residues were substituted. GFP, for example, contains nine tyrosines besides the one (Tyr66) in the fluorophore, two of which (Tyr92 and Tyr106) are completely buried in the structure.

*h*CypA and PpiB are highly homologous proteins that contain two and three tyrosine residues, respectively. One of these is at a conserved position and is buried completely, whereas the others are closer to the protein surface and are partially exposed to solvent. Whereas DOPA–PpiB remained soluble, DOPA–*h*CypA was found in the insoluble fraction. The greater tolerance of PpiB towards the incorporation of DOPA may be explained by the fact that, when comparing the natural proteins, *h*CypA is more prone to precipitation. Remarkably, DOPA was also incorporated into HMP without apparent effects on its solubility. The *ortho* positions of 8 of the 12 tyrosine rings in the *apo* form of HMP are solvent-exposed and can readily accommodate the extra hydroxyl group of DOPA.

We conclude that unless the DOPA side chains are solvent exposed, the translational incorporation of DOPA into proteins can affect their ability to fold in a native (soluble) structure, leading to misfolded/aggregated forms. This observation has additional importance in the context of oxidative damage and disease. We have recently shown that DOPA can be incorporated from the medium into mammalian cells in tissue

culture in a process that relies on *de novo* protein synthesis [16,17]. It is known that DOPA in proteins cannot be enzymatically repaired, yet it is redox active and potentially capable of inflicting further damage on biomolecules [15]. The presence of DOPA in misfolded proteins associated with various disease states in humans and animals is well established, and is assumed to arise largely by postranslational oxidation of tyrosine residues by oxygen free radicals [15]. Increased sensitivity to misfolding of newly translated proteins due to DOPA incorporation during translation could present an additional route for detrimental effects elicited by DOPA, aggravating the risks associated with DOPA accumulation. Determination of the ratio of K_M values of human TyrRS for DOPA and Tyr and measurement of intracellular DOPA concentrations will be required to substantiate the importance of this pathway.

Experimental procedures

In vitro cell-free protein synthesis

For production of the *E. coli* flavohaemoglobin HMP [26], transcription of the *hmp* gene was directed by tandem phage λ p_R and p_L promoters in plasmid pPL757 [27]. Plasmid derivatives of the T7-promoter vectors pETMCSI or pETMCSIII [28] were used to programme *in vitro* synthesis of *h*CypA (using pBH964) [18], PpiB (with and without an N-terminal His_6 tag, using pND1098 and pKO1154, respectively), and cycle 3 mutant GFP (using pMH1200). Plasmids pND1098 (His_6-PpiB) and pKO1154 (PpiB) were constructed by insertion of the appropriate genes as PCR-generated *Nde*I–*Eco*RI fragments between the corresponding sites in pETMCSIII or pETMCSI, respectively. To construct pMH1200 (GFP), a linker consisting of an equimolar mixture of oligonucleotides 683 (5′-TATGACTAGTAGCTAGGGATCCTAAG) and 684 (5′-AATTCTTAGGATCCCTAGCTACTAGTCA) was first inserted between the *Nde*I and *Eco*RI sites of pETMCSI to generate the new vector pETMCSIV (4670 bp) containing *Spe*I and *Bam*HI sites in its multiple cloning site (underlined in the sequences above). The *Spe*I–*Bgl*II fragment from plasmid pLEIGwt that encodes GFP was then inserted between these new sites in pETMCSIV to generate pMH1200; pLEIGwt [29] was a generous gift from Dr Peter Schultz (Scripps Research Institute, La Jolla, CA, USA). The orientations of inserted fragments and their integrity in the new plasmids were confirmed by nucleotide sequence determination (Biomolecular Resource Facility, Australian National University). Proteins were produced in the *E. coli* cell-free protein synthesis system described previously [5,18], with some modifications as described below. The concentrated S30 extract was prepared from *E. coli* A19 (*metB rna*) cells as described elsewhere [4]. Procedures for purification of phage T7 RNA polymerase (RNAP) [18] and *E. coli* RNAP holoenzyme [5] were as described.

The inner chamber reaction mixtures (usually 0.6–0.7 mL, in a dialysis sac) contained 55 mM Hepes/KOH pH 7.5, 1.7 mM dithiothreitol, 1.2 mM ATP, 0.8 mM each of CTP, UTP and GTP, 0.64 mM 3′,5′-cAMP, 68 μM folinic acid, 27.5 mM ammonium acetate, 208 mM potassium L-glutamate, 80 mM creatine phosphate, 250 $\mu g{\cdot}mL^{-1}$ creatine kinase, 1 mM each of the other 19 L-amino acids (unless specified otherwise), 15 mM magnesium acetate, 175 $\mu g{\cdot}mL^{-1}$ of total *E. coli* tRNA, 0.05% NaN_3, 210 $unit{\cdot}mL^{-1}$ RNase inhibitor, 16 $\mu g{\cdot}mL^{-1}$ of supercoiled plasmid DNA, 24% (v/v) of concentrated S30 extract (at 31 $mg{\cdot}mL^{-1}$ of total protein), and either additional *E. coli* (155 $\mu g{\cdot}mL^{-1}$) or T7 RNAP (93 $\mu g{\cdot}mL^{-1}$), as required for transcription from λ p_Rp_L or T7 promoters, respectively. The outer chamber dialysis buffer had the same composition as the inner chamber mixture, except that enzymes, tRNA and DNA were omitted, and magnesium acetate was present at 19.3 mM. To test whether DOPA was incorporated into newly synthesized proteins, it was provided in place of tyrosine in both reaction chambers. Protein synthesis was initiated by the addition of plasmid DNA to the inner reaction chamber, which was then immersed in the outer-chamber solution (12–14 mL) at 37 °C and shaken at 200 r.p.m. The outer-chamber solution was changed twice during the 8–9 h reaction [5,18]. Unless indicated otherwise, the products were separated into soluble and insoluble fractions by centrifugation at 30 000 ***g*** for 1 h at 4 °C. Soluble fractions from crude reaction mixtures containing *in vitro*-synthesized GFP were diluted 20-fold into 50 mM Hepes/KOH, pH 7.5 in a 3 mL cuvette, and fluorescence excitation and emission spectra were recorded at room temperature with a Cary Eclipse fluorescence spectrophotometer (Varian Inc., Palo Alto, CA, USA).

SDS/PAGE analysis

The soluble and insoluble fractions were analysed by duplicate 15% SDS/PAGE gels. One gel was stained with Coomassie Brilliant Blue for total protein detection. Proteins in the other were transferred electrophoretically to poly(vinylidene difluoride) membrane and stained with NBT using a redox cycling procedure for the detection of proteins that contain DOPA [14].

Purification of *in vitro* synthesized His_6-PpiB

Scaled-up *in vitro* cell-free reaction mixtures (2.0 mL) containing His_6-PpiB synthesized using either the complement of natural amino acids or with DOPA (0.05 or 1 mM) substituted for tyrosine were centrifuged at 100 000 ***g*** for 4 h at 4 °C to pellet ribosomes and ribosome-associated

proteins. The supernatant was mixed with an equal volume of Buffer P (20 mM sodium phosphate, pH 7.5, 0.5 M NaCl, 1 mM 2-mercaptoethanol) containing 10 mM imidazole and applied to a column of Ni-NTA agarose (Qiagen, Hilden, Germany; 1.5 × 1.8 cm) that had been equilibrated with the same buffer. After a washing with Buffer P + 10 mM imidazole (10 mL), bound proteins were eluted by sequential application of Buffer P containing 20, 100 and 500 mM imidazole (8 mL each). Fractions were analysed by 15% SDS/PAGE; those that contained highly purified His_6-PpiB were combined and concentrated to 1 mL by use of an Amicon (Billerica, MA, USA) Ultra-4 centrifugal concentrator (MWCO 10 000). The final protein concentration was estimated using the Bradford method [30] with bovine serum albumin as a standard.

Mass spectrometry

Samples of purified His_6-PpiB that had been synthesized *in vitro* with tyrosine or 0.05 or 1 mM DOPA were analysed by MS in several ways. ESI-MS of the native proteins after buffer exchange into 10 mM ammonium acetate, pH 6.8 and addition of formic acid to pH 3.0, were acquired with a Micromass (Wyntheshawe, UK) Q-TOF2 spectrometer operated in V-mode with a desolvation temperature of 180 °C, source temperature of 40 °C and cone voltage of 50 V. Ions from electrospray series were transformed to a mass scale using MASSLYNX™ software (Micromass).

To confirm that DOPA had been incorporated specifically in place of tyrosine, the masses of almost every tryptic peptide from samples of purified His_6-PpiB were determined. Protein samples (≈ 60 μg in 100 μL of 10 mM Hepes/KOH pH 7.5, 1 mM dithiothreitol) were digested with trypsin (0.2 μg; 24 h at 37 °C), and peptides were separated by RP-HPLC using a C_{18} capillary column (Agilent Technologies, Palo Alto, CA, USA), eluted with a gradient of 0–60% (v/v) acetonitrile in 0.1% aqueous acetic acid over 45 min at a flow rate of 0.1 $\mu L \cdot min^{-1}$. The capillary column was connected in-line to an Applied Biosystems QSTAR Pulsar mass spectrometer, which was used to record ESI-MS. Theoretical masses of peptides were calculated using the Expasy website (http://au.expasy.org/tools/peptide-mass.html). To determine the theoretical mass for proteins and peptides that contained DOPA in place of tyrosine, 16.0 mass units were added for each tyrosine residue.

NMR measurements

Samples of (untagged) PpiB containing either tyrosine or DOPA were prepared by *in vitro* synthesis in 0.7 mL reactions programmed with pKO1154, essentially as described above with tyrosine or DOPA at 1 mM, except that L-[^{15}N]phenylalanine and L-[^{15}N]cysteine (each at 0.35 mM, Cambridge Isotope Laboratories, Andover, MA, USA) were used instead of phenylalanine and cysteine [18]. The inner-chamber mixtures containing ^{15}N-labelled PpiB were dialysed overnight at 4 °C against 500 mL of an NMR buffer comprised of 50 mM sodium phosphate (pH 6.2) and 1 mM dithiothreitol, then clarified by centrifugation (100 000 ***g***, 4 h). The supernatants were concentrated to about 0.6 mL using Amicon Ultra-4 centrifugal filters (MWCO 10 000) and D_2O was added to a final concentration of 10% before NMR measurements.

NMR spectra were recorded at 36 °C using a Varian INOVA 600 MHz NMR spectrometer equipped with a quadruple resonance (^{1}H, ^{15}N, ^{13}C, ^{31}P) probe. ^{15}N-HSQC spectra were recorded with 5 mm sample tubes using $t_{1max} = 32$ ms, $t_{2max} = 102$ ms, and total recording times of 21 h.

Acknowledgements

We are grateful to Patrick Schaeffer for assistance with plasmid construction. This work was supported by grants from the National Health and Medical Research Council of Australia (to R.T.D., K.R.) and the Australian Research Council (to J.B., G.O., N.D.). K.O. and G.O. acknowledge an Australian Linkage (CSIRO) Postdoctoral Fellowship and a Federation Fellowship, respectively.

References

1 Hendrickson TL, de Crécy-Lagard V & Schimmel P (2004) Incorporation of non-natural amino acids into proteins. *Annu Rev Biochem* **73**, 147–176.

2 Budisa N (2003) Expression of ‘tailor-made’ proteins via incorporation of synthetic amino acids by using cell-free protein synthesis. In *Cell-Free Protein Expression* (Swartz JR, ed.), 2nd edn, pp. 89–98. Springer-Verlag, Berlin.

3 Sengupta K, Klammt C, Bernhard F & Rüterjans H (2003) Incorporation of fluorescent labels into cell-free produced proteins. *Cell-Free Protein Expression* (Swartz JR, ed.), 2nd edn, pp. 81–88. Springer-Verlag, Berlin.

4 Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T & Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett* **442**, 15–19.

5 Guignard L, Ozawa K, Pursglove SE, Otting G & Dixon NE (2002) NMR analysis of *in vitro*-synthesized proteins without purification: a high-throughput approach. *FEBS Lett* **524**, 159–162.

6 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.

7 Klammt C, Lohr F, Schafer B, Haase W, Dotsch V, Ruterjans H, Glaubitz C & Bernhard F (2004) High

level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271**, 568–580.

8 Kigawa T, Yamaguchi-Nunokawa E, Kodama K, Matsuda T, Yabuki T, Matsuda N, Ishitani R, Nureki O & Yokoyama S (2001) Selenomethionine incorporation into a protein by cell-free synthesis. *J Struct Funct Genomics* **2**, 29–35.

9 Neerathilingam M, Greene LH, Colebrooke SA, Campbell ID & Staunton D (2005) Quantitation of protein expression in a cell-free system: efficient detection of yields and ^{19}F NMR to identify folded protein. *J Biomol NMR* **31**, 11–19.

10 Calendar R & Berg P (1966) The catalytic properties of tyrosyl ribonucleic acid synthetases from *Escherichia coli* and *Bacillus subtilis*. *Biochemistry* **5**, 1690–1695.

11 Högenauer G, Kreil G & Bernheimer H (1978) Studies on the binding of DOPA (3,4-dihydroxyphenylalanine) to tRNA. *FEBS Lett* **88**, 101–104.

12 Hoffman JL, McConnell KP & Carpenter DR (1970) Aminoacylation of *Escherichia coli* methionine tRNA by selenomethionine. *Biochim Biophys Acta* **199**, 531–534.

13 Xu ZJ, Love ML, Ma LY, Blum M, Bronskill PM, Bernstein J, Grey AA, Hofmann T, Camerman N & Wong JT (1989) Tryptophanyl-tRNA synthetase from *Bacillus subtilis*. Characterization and role of hydrophobicity in substrate recognition. *J Biol Chem* **264**, 4304–4311.

14 Paz MA, Fluckiger R, Boak A, Kagan HM & Gallop PM (1991) Specific detection of quinoproteins by redox-cycling staining. *J Biol Chem* **266**, 689–692.

15 Rodgers KJ & Dean RT (2000) Metabolism of protein-bound DOPA in mammals. *Int J Biochem Cell Biol* **32**, 945–955.

16 Rodgers KJ, Wang H, Fu S & Dean RT (2002) Biosynthetic incorporation of oxidized amino acids into proteins and their cellular proteolysis. *Free Radical Biol Med* **32**, 766–775.

17 Rodgers KJ, Hume PM, Dunlop RA & Dean RT (2004) Biosynthesis and turnover of DOPA-containing proteins by human cells. *Free Radical Biol Med* **37**, 1756–1764.

18 Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE & Otting G (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively ^{15}N-labelled proteins for rapid analysis by NMR spectroscopy. *Eur J Biochem* **271**, 4084–4093.

19 Edwards KJ, Ollis DL & Dixon NE (1997) Crystal structure of cytoplasmic *Escherichia coli* peptidyl-prolyl isomerase: evidence for decreased mobility of loops upon complexation. *J Mol Biol* **271**, 258–265.

20 Ke H (1992) Similarities and differences between human cyclophilin A and other β-barrel structures. Structural refinement at 1.63 Å resolution. *J Mol Biol* **228**, 539–550.

21 Ilari A, Bonamore A, Farina A, Johnson KA & Boffi A (2002) The X-ray structure of ferric *Escherichia coli* flavohemoglobin reveals an unexpected geometry of the distal heme pocket. *J Biol Chem* **277**, 23725–23732.

22 Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY & Remington SJ (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395.

23 Kariya E, Ohki S-Y, Hayano T & Kainosho M (2000) Backbone ^{1}H, ^{13}C, and ^{15}N resonance assignments of an 18.2 kDa protein, *E. coli* peptidyl-prolyl *cis–trans* isomerase b (EPPIb). *J Biomol NMR* **18**, 75–76.

24 Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* **67**, 509–544.

25 Zimmer M (2002) Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem Rev* **102**, 759–781.

26 Vasudevan SG, Armarego WLF, Shaw DC, Lilley PE, Dixon NE & Poole RK (1991) Isolation and nucleotide sequence of the *hmp* gene that encodes a haemoglobin-like protein in *Escherichia coli* K-12. *Mol Gen Genet* **226**, 49–58.

27 Love CA, Lilley PE & Dixon NE (1996) Stable high-copy-number bacteriophage λ promoter vectors for overproduction of proteins in *Escherichia coli*. *Gene* **176**, 49–53.

28 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.

29 Wang L, Xie J, Deniz AA & Schultz PG (2003) Unnatural amino acid mutagenesis of green fluorescent protein. *J Org Chem* **68**, 174–176.

30 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal Biochem* **72**, 248–254.

31 Koradi R, Billeter M & Wüthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14**, 51–55.