Biochimica et Biophysica Acta 1798 (2010) 1244–1249

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Signal sequence non-optimal codons are required for the correct folding of mature maltose binding protein

Yaramah M. Zalucki [a,c,1,2], Christopher E. Jones [a,b,1], Preston S.K. Ng [a,c], Benjamin L. Schulz [a], Michael P. Jennings [a,c,*]

[a] School of Chemistry & Molecular Biosciences, The University of Queensland, St. Lucia, 4072, Brisbane, QLD, Australia
[b] Queensland Brain Institute, The University of Queensland, St. Lucia, 4072, Brisbane, QLD, Australia
[c] Institute for Glycomics, Griffith University, Gold Coast, Queensland, 4222, Australia

ARTICLE INFO

Article history:
Received 13 December 2009
Received in revised form 7 March 2010
Accepted 9 March 2010
Available online 15 March 2010

Keywords:
Codon usage
Protein folding
Protein export
Maltose binding protein

ABSTRACT

Non-optimal codons are generally characterised by a low concentration of isoaccepting tRNA and a slower translation rate compared to optimal codons. In a previous study, we reported a 20-fold reduction in maltose binding protein (MBP) level when the non-optimal codons in the signal sequence were optimised. In this study, we report that the 20-fold reduction is rescued when MBP is expressed at 28 °C instead of 37 °C, suggesting that the signal sequence optimised MBP protein (MBP-opt) may be misfolded, and is being degraded at 37 °C. Consistent with this idea, transient induction of the heat shock proteases prior to MBP expression at 28 °C restores the 20-fold difference, demonstrating that the difference in production levels is due to post-translational degradation of MBP-opt by the heat-shock proteases. Analysis of the structure of purified MBP-wt and MBP-opt grown at 28 °C showed that although they have similar secondary structure content, MBP-opt is more resistant to thermal unfolding than is MBP-wt. The two proteins also exhibit different tryptic fragment profiles, further confirming that they are folded into conformationally different states. This is the first study to demonstrate that signal sequence non-optimal codons can influence the folding of the mature exported protein.

1. Introduction

A characteristic of all general secretory pathway proteins (sec) is that they have an N-terminal signal peptide. These are typically 20–30 amino acids long that help direct the pre-secretory protein to the cytoplasmic membrane [1]. The signal peptide is divided into three domains, a positively charged N-terminus, a hydrophobic core and a C-terminal cleavage signal [2]. The signal peptide is predicted to fold into an α-helix due to the hydrophobic core upon interaction with the lipid membrane [3–5], and is cleaved by leader peptidase I upon translocation of the pre-secretory protein into the periplasm [6].

Non-optimal codons occur at the highest frequency in the signal sequence of *Escherichia coli* secretory genes compared to any other region of the genome [7]. A similar finding has been reported in the Gram-positive bacterium *Streptomyces coelicolor* [8]. Non-optimal codons are translated at a slower rate than optimal codons [9–11], and have been shown to occur at higher frequencies at protein domain junctions [12,13]. One of the roles attributed to non-optimal codons is to allow the correct folding of the protein [14]. Previous studies have shown that replacing non-optimal codons with optimal codons from the same synonymous codon family causes the protein to misfold [15–17]. However, no study has directly investigated the effect of non-optimal codons in the signal sequence of the precursor form of secretory proteins on the folding of the mature form of the protein.

In two previous studies, codon usage was optimised in the signal sequence of maltose binding protein (MBP) and β-lactamase (Bla) (*malE*-opt, *bla*-opt). The protein produced from these optimised alleles (MBP-opt, Bla-opt) was expressed at 20-fold [18] and 4-fold [19] lower levels when compared to the expression of wild-type protein produced from *malE*-wt and *bla*-wt (MBP-wt, Bla-wt) respectively. For β-lactamase, this difference could be reduced to 2-fold by growing at a lower temperature (28 °C compared to 37 °C) [19]. In the case of MBP-opt, its production was partially recovered in strains deficient in multiple periplasmic (OmpT, DegP) or cytoplasmic proteases (Lon, ClpXP, HslVU) [18], indicating that the MBP-opt may be misfolded. In this study, we followed these indications to further examine the production and structure of MBP-opt and MBP-wt. In particular, we wanted to determine if MBP-opt had a different tertiary structure compared to MBP-wt, which would help explain the

* Corresponding author. Institute for Glycomics, Griffith University, Gold Coast, Queensland, 4222, Australia. Tel.: +617 55527050; fax: +617 55528908.
E-mail address: m.jennings@griffith.edu.au (M.P. Jennings).
[1] These authors contributed equally to the study.
[2] Current address: Department of Microbiology and Immunology, Emory University, Atlanta, GA 30322, USA.

0005-2736/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2010.03.010

protease sensitivity phenotype of MBP-opt. We investigated this through the biochemical assays on purified protein, the results of which are reported below.

2. Materials and methods

2.1. Bacterial growth conditions

All bacteria were grown in Luria-Bertani media supplemented with the appropriate antibiotic: ampicillin 100 μg/ml. The bacterial strains used in this study MC4100 (*F-araD139 A(argF-lac)U169 rpsL150 deoCI relAl thiA ptsF25 flbB5301*), TST1 (MC4100 *malE*, CGSC # 6137).

2.2. Expression of malE-wt and malE-opt after induction of heat-shock proteins

The *malE*-wt and *malE*-opt alleles are the same as described in [18]. A 3 ml LB-ampicillin starter culture of MC4100 was grown overnight at 28 °C, subbed 1:100 into 4 ml LB-ampicillin and grown until A_{600} 0.5. The heat shock proteins were induced by incubating the cultures for 5 min at 42 °C. MBP production was induced with 0.3 mM IPTG and the cultures grown at 28 °C. Expression of *malE*-wt and *malE*-opt without induction of the heat-shock response was carried out at the constant temperature of 28 °C for 1 h. Samples were taken at 15 min and 1 h, and prepared for Western analysis on 4–12% SDS-PAGE gel as described in [18]. Western blots were probed with anti-MBP (NEB, E8030S) at a 1:10,000 dilution and detected with anti-rabbit IgG conjugated with alkaline phosphatase.

2.3. Purification of MBP-wt and MBP-opt from strain TST1

The proteins MBP-wt and MBP-opt were purified from the *malE* deficient strain TST1, following the protocol in New England Biolabs. Briefly a starter culture of 3 ml LB-ampicillin was grown overnight and added to 500 ml LB-ampicillin. This was grown until A_{600} ~0.5, then MBP production was induced by addition of 0.3 mM IPTG and grown for 3 h. Cells were pelleted by centrifugation and left overnight at −20 °C. Cells were resuspended in column buffer (20 mM Tris–HCl pH 7.5, 200 mM NaCl, 1 mM EDTA), containing one tablet of EDTA-free protease inhibitor cocktail (Roche, Cat. No. 11873580001) and lysed by French-press by passing through five cycles at 1000 mPa. Both MBP-wt and MBP-opt were only present in the soluble fraction and not in the insoluble fraction. The cell lysate was centrifuged at 16,442×*g* and the supernatant passed by gravity flow over a 2 ml amylose column and washed with 20 ml column buffer. There was no MBP-opt or MBP-wt present in the unbound fraction. The MBP was eluted in 2 ml fractions with column buffer with 10 mM maltose. Fractions containing pure MBP were pooled and buffer exchanged on a 10 kDa column (Pall Life Science), into 20 mM sodium phosphate buffer, pH 7.4 for further analysis. Samples were spotted onto PVDF membrane, and N-terminally sequenced by Edman degradation on an Applied Biosystem's Procise 492cLC.

2.4. Circular Dichroism (CD) spectroscopy

CD spectra were obtained on a Jasco J-710 spectrometer (Jasco, Japan). Spectra were collected from 195 to 260 nm using a 1 mm path-length cuvette. For each sample five scans were collected and averaged. A baseline (20 mM sodium phosphate, pH 7.1) was subtracted from each sample. The direct CD measurements were converted to molar ellipticity, $\Delta\varepsilon$ (M^{-1} cm^{-1}) using the relationship $\Delta\varepsilon = \theta/33{,}000 \times c \times l$, where c is the concentration (M) and l is the path length (cm). For the thermal unfolding experiments, a Neslab RTE-111 microprocessor water bath was used and controlled via the Jasco software. The heating rate was 15 °C/h and the ellipticity at 222 nm was collected every 0.5 °C between 20 °C and 85 °C. The ellipticity was converted to % folded protein assuming the protein is 100% folded at 20 °C. The data is presented as the average of two independent scans and is representative of two independent purifications. CD wavelength scans were obtained at 20 °C before and after each thermal unfolding experiment. Data were fitted using a monomer-to-monomer unfolding equation using Sigma-plot. Reversibility was assessed after each run by cooling the sample from 85 °C and equilibrating at 20 °C then collecting the CD spectrum between 195 and 260 nm.

2.5. Maltose binding assay

Maltose binding was assessed by monitoring tryptophan fluorescence during a maltose titration. All fluorescence titrations were carried out on a Perkin Elmer LS50B luminescence spectrometer operating WinFl software (Perkin Elmer, Massachusetts, USA) and all spectra were collected at room temperature (22 °C). Aliquots of a maltose stock solution (1.4 mM in 20 mM phosphate buffer, pH 7.5) were titrated into protein solutions (2.0 ml). Titrations were conducted on protein concentrations of 1.0 μM (MBP $\varepsilon_{280\,nm} = 82{,}949\ M^{-1}\ cm^{-1}$). The fluorescence excitation wavelength was 290 nm (slit width = 2.5 nm) and the emission (slit width = 6.5 nm) was scanned from 300 to 400 nm. Five scans were collected and averaged for each maltose addition. The fluorescence intensity at 350 nm was recorded as a function of the maltose concentration. The extent of maltose binding after each addition was determined as $\Delta F_i/\Delta F_{max}$ where ΔF_i is the fluorescence change after each maltose increment and ΔF_{max} is the maximal fluorescence change at infinite maltose concentration. ΔF_{max} was determined from: $\Delta F_i = \Delta F_{max}$ ([maltose]/([maltose] + c)) where $\Delta F_i = F_0 - F_i$ (F_0 is the initial fluorescence, F_i is the fluorescence after maltose addition) and c is a constant. Data were fitted to using WinCurveFit (Kevin Rainer software). Plots of $\Delta F_i/\Delta F_{max}$ as a function of maltose concentration were prepared using Prism (Graphpad Software Inc, California, USA) and association constants (K_a) were determined using non-linear regression assuming a single, specific maltose binding site.

2.6. Digest of MBP with trypsin

Aliquots (20 μL) of MBP-wt and MBP-opt adjusted to an equal concentration of 0.191 mg/ml in 20 mM Tris–HCl pH 7.4, 200 mM NaCl and 1 mM EDTA were digested by addition of 1 μl of trypsin at 1, 0.1, 0.01 or 0 μg/μl, giving a trypsin/protein ratio of approximately 1:4, 1:40, 1:400 or 0, respectively, and incubation at 37 °C for 3 h. Samples were separated on a 4–12% SDS-PAGE and stained with Coomassie blue. For N-terminal sequencing, duplicate samples were semi-dry electroblotted to PVDF membrane and stained with Coomassie blue.

3. Results

3.1. Production of MBP-wt and MBP-opt equivalent at 28 °C

In our previous study, six non-optimal codons, defined by Burns and Beacham [20], in the plasmid encoded signal sequence of *malE*:: *lacZα* were changed to the most optimal codon from the same synonymous codon family. Protein production from the two alleles (*malE*-wt, *malE*-opt) saw a dramatic 20-fold drop in MBP-opt production compared to MBP-wt, which could not be explained by changes in plasmid copy number or RNA transcript levels [18]. However, production of MBP-opt was partially recovered by expression in strains deficient in multiple periplasmic or cytoplasmic proteases, implying that MBP-opt was misfolded. All of the previous experiments were carried out at 37 °C.

A common method used to achieve better production of unstable and misfolded proteins is to lower the temperature at which they are produced [21,22]. One reason for this is the lower degradation rate from ATP-dependent proteases at 30 °C compared to both 37 °C [23]

and 42 °C [24]. These ATP-dependent proteases include the heat-shock proteases Lon, ClpXP and HslVU [25]. To test if MBP-opt production levels could be improved by expressing at a lower temperature, strain MC4100 containing either the *malE*-wt or *malE*-opt alleles was grown at 28 °C. At this temperature, MBP-opt and MBP-wt was produced at equivalent amounts, which is in stark contrast to the 20-fold difference observed at 37 °C (Fig. 1).

The lower production of MBP-opt at 37 °C was hypothesised to be due to preferential degradation of a misfolded conformation of MBP-opt [18]. The equivalent production of MBP-wt and MBP-opt observed at 28 °C in Fig. 1 raises the possibility that MBP-opt produced at this temperature is correctly folded, and therefore not susceptible to proteolytic degradation by the heat-shock proteases. To test this hypothesis, production of MBP-wt and MBP-opt was compared in strain MC4100 grown at 28 °C that had the heat-shock proteases transiently induced by incubating the cultures 42 °C for 5 min prior to induction of MBP expression with IPTG at 28 °C. A temperature shift to 42 °C is a common way to induce expression of heat-shock genes [26], including the proteases Lon, ClpXP and HslVU. After induction, the cultures were shifted back to 28 °C and samples taken 1 h later for analysis by Western blot. Production of MBP-opt in the temperature shifted cultures was severely reduced compared to MBP-wt (Fig. 1), and the ratio between MBP-wt and MBP-opt was more comparable to that observed at 37 °C (Fig. 1). When no temperature shift was introduced, the production of MBP-wt and MBP-opt grown at 28 °C was equivalent. The change in production levels observed in the temperature shifted cultures was not due to differences in plasmid copy number or RNA transcript levels, which were equivalent throughout the experiment (data not shown). This suggests that MBP-opt folds into the same protease-sensitive conformation at both 28 °C and 37 °C, but requires induction of the heat-shock proteases for degradation to occur at 28 °C.

3.2. MBP-opt is more stable than MBP-wt

The previous experiment demonstrated that MBP-opt produced from cultures grown at either 28 °C or 37 °C is sensitive to proteolytic degradation, implying that MBP-opt is structurally different to protease-insensitive MBP-wt. To determine if a structural difference could underlie the protease-sensitivity, we purified MBP-wt and MBP-opt from cultures grown at 28 °C. Purified MBP-wt and MBP-opt was Coomassie stained (Fig. 3, Lanes 1 and 2) and sequenced by Edman degradation which showed that both had identical sequences, which confirmed that mature MBP was the major species present in the purified sample. There was very little unprocessed precursor (preMBP) in either of the purified samples.

Circular dichroism (CD) spectroscopy (Fig. 2A) revealed that the overall composition of α-helices, β-sheets and coil structures in the mature proteins is very similar, confirming no major differences in secondary structure between the two proteins. Using CD we next investigated the thermal unfolding features of both MBP-wt and MBP-opt over the temperature range 20–85 °C (Fig. 2B). The results show that MBP-opt ($T_m = 65.4$ °C) is more resistant to temperature-induced unfolding than MBP-wt ($T_m = 63.1$ °C), and did not unfold as extensively as MBP-wt, suggesting MBP-opt is more thermostable

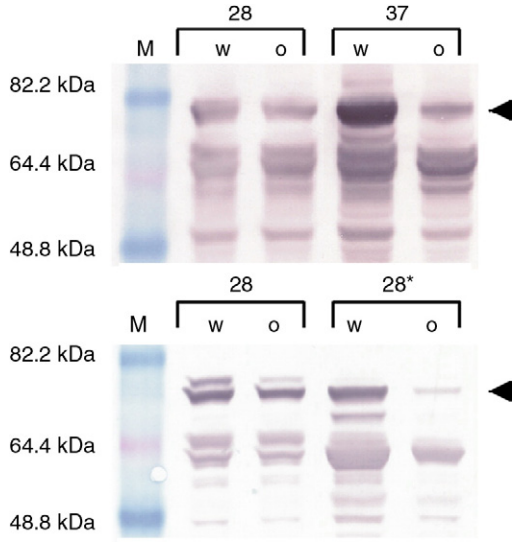

Fig. 1. Effect of temperature on MBP production in strain MC4100. Whole cell lysates were run on a 4–12% SDS-PAGE, and the Western and probed with anti-MBP (NEB E8032S). Cultures were grown at the temperature indicated above the lane. The top Western blot shows MBP production at 28° and 37°. The bottom Western blot compares MBP production at 28° with and without induction of the heat-shock proteins. The 28* indicates both *malE*-wt and *malE*-opt were heat-shocked at 42 °C for 5 min prior to induction of MBP with 0.3 mM IPTG, and then grown back at 28 °C. The arrow head indicates mature MBP. Key: M—prestained molecular weight marker (Invitrogen 10748-010), w—MBP-wt, o—MBP-opt.

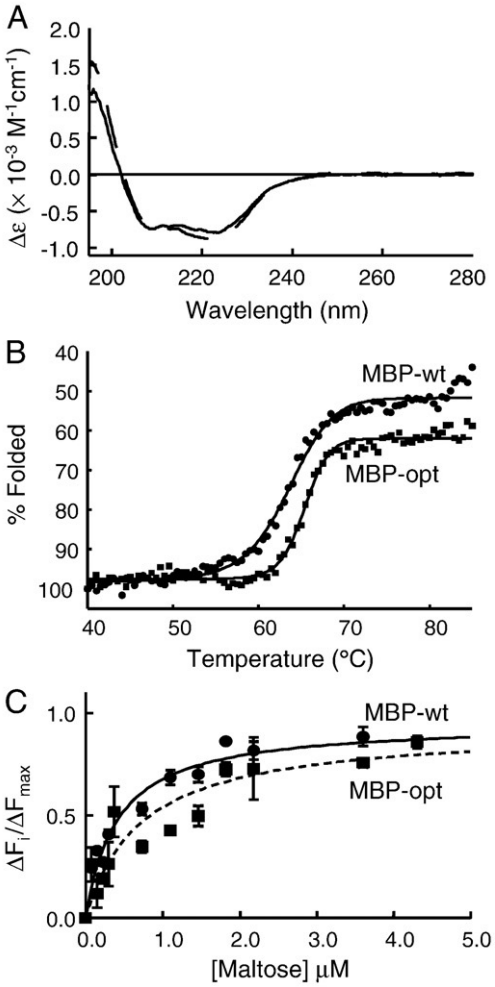

Fig. 2. A. CD spectra of the MBP-wt and MBP-opt. Samples were prepared in 20 mM sodium phosphate buffer (pH 7.4) and scans were taken in a range from 195 to 280 nm on a Jasco-710 spectrometer (Jasco). The *dashed line* is the MBP-opt spectrum and the *solid line* is the MBP-wt spectrum. B. Temperature induced unfolding of MBP-wt and MBP-opt. MBP-opt (5.8 μM, *squares*, ■) and MBP-wt (6.7 μM, *circles*, ●) in 20 mM sodium phosphate buffer (pH 7.4) were unfolded over the temperature range 20–85 °C (data shown from 40 °C). The solid lines represent the data fitted to a monomer-to-monomer unfolding equation. C. Maltose binding to MBP assessed by tryptophan fluorescence. The decrease in tryptophan fluorescence at 348 nm was monitored during the titration of maltose. The fractional saturation ($\Delta F_i/\Delta F_{max}$) was plotted as a function of maltose concentration. The *circles* (●) represent data obtained for MBP-wt and *squares* (■) for MBP-opt (both 1.0 μM). The *solid lines* represent the best-fits obtained from non-linear regression as described in Materials and methods.

than MBP-wt. Furthermore, Fig. 2B shows that MBP-wt (~52 °C) clearly begins to unfold before MBP-opt (~59 °C). The T_m obtained for MBP-wt is consistent with previous studies (63.0 °C at pH 7.4) [27]. The unfolding was not completely reversible for both MBP-wt and MBP-opt under these conditions and neither completely unfolded. Inspection of the diode current (measure of turbidity) obtained during the thermal unfolding experiment shows that neither MBP-wt nor MBP-opt aggregated during the experiment and the changes observed at 222 nm are a direct reflection of unfolding protein (data not shown).

3.3. The maltose binding properties of MBP-wt and MBP-opt

MBP-opt was capable of binding maltose as it was purified using a maltose affinity column and previous experiments demonstrated that MBP-opt can complement the *malE* deficient strain TL225, suggesting it is able to transport maltose into the cell [18]. Using fluorescence spectroscopy we followed the quenching of tryptophan fluorescence (at 350 nm) as maltose is titrated into a solution of MBP. Using the fluorescence data, we determined an affinity of MBP-opt for maltose of 0.98 ± 0.24 μM ($n = 2$) and MBP-wt of 0.44 ± 0.10 μM ($n = 2$) (Fig. 2C). Maltose affinities in the low μM range are consistent with previous experiments [28,29]. These data suggest that the structural difference that increases the thermal stability of MBP-opt and protease sensitivity does not extend to a major difference in maltose affinity. On this basis we can predict that the structural difference is likely to be remote from the maltose binding site.

3.4. MBP-wt and MBP-opt exhibit different tryptic digest profiles

To further confirm whether MBP-opt is structurally different to MBP-wt, both purified proteins were subject to digest by the protease trypsin. Four micrograms of MBP-wt and MBP-opt was digested with a range of trypsin concentrations for 30 min (see Materials and methods). Coomassie staining on a 4–12% SDS-PAGE gel revealed a band at about 13 kDa that appeared in the MBP-opt sample with greater intensity than in the corresponding MBP-wt sample (Fig. 3, see (*)). The greatest difference in intensity was observed in the 1:400 trypsin/protein ratio; at higher trypsin amounts the difference in intensity was not observed (Fig. 3). The N-terminal sequence of the first nine residues of this band in MBP-opt sample showed that it was the start of mature MBP (minus the signal peptide). This implies that the region at the N-terminus of MBP-opt is structurally distinct from the corresponding region in MBP-wt, allowing greater access to trypsin, hence the more intense band seen on the Coomassie gel.

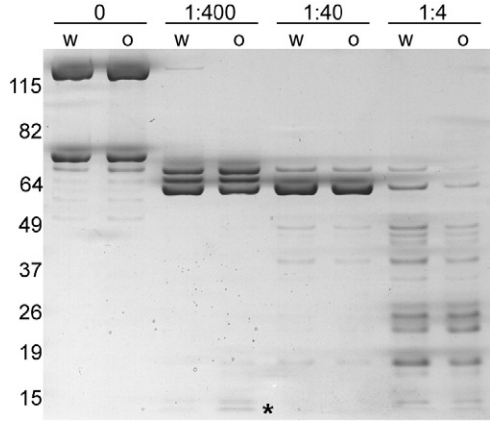

Fig. 3. Non-denatured MBP-wt (w) and MBP-opt (o) were digested with increasing ratios of trypsin/protein (no trypsin, 1:400, 1:40 and 1:4), separated by SDS-PAGE and stained with Coomassie Blue. N-terminal sequencing of a band present at higher intensity in MBP-opt (*) revealed the sequence KIEEGKLVI, corresponding to the N-terminus of the mature MBP protein.

4. Discussion

The results of this study strongly indicate that two proteins, MBP-wt and MBP-opt, identical at the amino acid level, are structurally different. By altering codon usage in the MBP signal sequence so that six non-optimal codons were changed to their most optimal codon, MBP produced at 37 °C was folded into a protease-sensitive conformation, and was degraded by heat-shock proteases. To our knowledge, this is the first study to report that optimising the codon usage in the cleaved secretory signal sequence of the precursor form of a secreted protein can generate a structurally distinct mature form of the same protein.

CD data suggest that MBP-wt and MBP-opt do not have major differences in their secondary structure content (Fig. 2A), yet whether the tertiary structures of these two proteins are the same is unclear from this data alone. However, the protease-sensitive conformation of MBP-opt results in an increase in thermal stability as evidenced by a change in both the melting temperature and the temperature at which unfolding begins (Fig. 2B) suggests that a structural difference is indeed behind the protease sensitivity. The structural difference does not appear to extend to a major difference in maltose binding affinity, which is consistent with previous experiments showing that MBP-opt can transport maltose efficiently.

The question remains as to how optimising the codon usage in the signal peptide of MBP, which is cleaved upon translocation to the periplasm, results in a conformation change in the mature protein. Optimising the codon usage in the signal sequence could affect a number of biological processes. These include mRNA secondary structures, ribosome spacing on the transcript, and the translation rate. The results of this study and our previous work [18] argues strongly against the possibility of mRNA secondary structures interfering with translation, since equal production of MBP-opt and MBP-wt is observed at 28 °C. For if such a structure did exist, it would be more stable at 28 °C than at 37 °C, and the difference in MBP-opt production compared to MBP-wt should still be observed. It is also unlikely that any effects on ribosome spacing, which have been modelled theoretically by Zhang et al. [30], would influence protein folding. Therefore, it is most likely that increasing the translation rate across the signal sequence of *malE*-opt alters a time-critical early event, which results in a change in the structure of the mature MBP-opt.

In our previous study we presented a hypothetical model in which the signal peptide possibly emerges misfolded from the ribosomal tunnel due to the increased translation rate across this region [18]. The folding of preMBP is influenced by interactions with the molecular chaperone SecB [31,32] and by the interaction of the signal peptide with preMBP [32–34]. It is assumed that the signal peptide interacts with hydrophobic regions of the mature protein to slow its folding, however the exact region of interaction is unknown. Therefore we hypothesised that a misfolded signal peptide alters the folding pathway of the mature protein resulting in a changed conformation that is still present after the signal peptide is cleaved [18]. The results of this study add further evidence to this hypothesis, as they confirm that mature MBP-opt has an altered tertiary structure, which results in a more thermostable, protease sensitive conformation compared to MBP-wt. One difference in tertiary structure could be located at the N-terminus of the mature protein, as the tryptic digests of MBP-opt revealed a 13 kDa band that occurred at lower concentrations of trypsin compared to MBP-wt. This could be the result of an altered structure at the N-terminus, or the result of a folding change that allows greater accessibility to trypsin at the N-terminus. Sequencing this band showed that it is the N-terminus of mature MBP. If the 13 kDa band is the product of a single digest, its size would imply that an arginine or lysine at approximately 120 residues downstream is more exposed in MBP-opt than MBP-wt. Given the proximity of this region to the signal peptide, the altered

tertiary structure of MBP-opt could be due to interactions between a misfolded signal peptide and the N-terminus of the mature protein prior to translocation to the periplasm.

Whilst it is unusual for higher ordered structures to pass through the SecYEG translocon, it is not without precedent. Several studies have shown that secondary and tertiary structures in ProOmpA do not prevent export [35–37], and a biotinated protein is capable of export *in vivo* [38]. In our previous study, MBP-opt was recovered in a strain deficient in both of the periplasmic proteases DegP and OmpT [18]. In this study, biochemical analysis of mature MBP-wt and MBP-opt strongly suggests a structural difference between the two proteins is the underlying reason behind the protease sensitivity of MBP-opt. The nature of protein export implies that the mature protein used in this study was from the periplasm. During export, the signal peptide is cleaved by signal peptidase I, which lies on the periplasmic face of the inner membrane [6], which then releases the protein from the inner membrane into the periplasm [39]. Taking together the results of this study and our previous one, they strongly suggest that folded mature MBP-opt is structurally different to MBP-wt, and that this structural difference is capable of passing through the SecYEG translocon.

It has well documented that the 5′ ends of genes have different codon usage patterns compared to the middle of genes [7,40–42]. In particular, signal sequences have highly biased codon usage compared to both the same region in other genes and downstream regions in secretory genes [7,8,43]. Various factors are thought to explain the observed bias, including a requirement for less mRNA secondary structure to allow efficient translation initiation [44], and lack of selection against non-optimal codons at this region [42]. A requirement for protein folding has not been thought to constrain codon usage at the 5′ end of genes. However, in this study we have shown that optimising the codon usage in the signal sequence of MBP causes a structural change in the mature protein. Given the propensity of signal peptides to interact with and slow the folding of the mature region [33,34,45], any alteration of this interaction through a misfolded signal peptide could induce a structural change in the mature protein. This could result in a conformation of the protein that is either less competent for export, or more sensitive to protease degradation. Hence, the folding requirements for signal peptides of exported proteins could be a selecting force that influences high frequency of non-optimal codons at the 5′ end of exported genes.

This study is the first to our knowledge to strongly suggest through direct biochemical data on pure protein that optimising the codon usage in the signal sequence alters the conformation of the mature exported protein. Further experiments are required to determine if a misfolded signal peptide is the cause of this change, or whether other mechanisms are involved. This will give insight into how proteins fold, especially early in translation of the nascent protein, and the role codon usage has in this dynamic process.

Acknowledgements

Thanks to Prof. Bostjan Kobe for his discussion and advice on protein structure and folding during the course of this study. Benjamin Schulz is funded by a University of Queensland Postdoctoral Research Fellowship; Yaramah Zalucki is funded by a CJ Martin Biomedical Fellowship (NHMRC #569913).

References

[1] A.P. Pugsley, The complete general secretory pathway in gram-negative bacteria, Microbiol. Rev. 57 (1993) 50–108.
[2] G. von Heijne, Signal sequences. The limits of variation, J. Mol. Biol. 184 (1985) 99–105.
[3] H. Bedouelle, M. Hofnung, Functional implications of secondary structure analysis of wild type and mutant bacterial signal peptides, Prog. Clin. Biol. Res. 63 (1981) 399–403.
[4] A.M. Batenburg, R. Brasseur, J.M. Ruysschaert, G.J. van Scharrenburg, A.J. Slotboom, R.A. Demel, B. de Kruijff, Characterization of the interfacial behavior and structure of the signal sequence of *Escherichia coli* outer membrane pore protein PhoE, J. Biol. Chem. 263 (1988) 4202–4207.
[5] M.S. Briggs, D.G. Cornell, R.A. Dluhy, L.M. Gierasch, Conformations of signal peptides induced by lipids suggest initial steps in protein export, Science 233 (1986) 206–208.
[6] R.E. Dalbey, Leader peptidase, Mol. Microbiol. 5 (1991) 2855–2860.
[7] P.M. Power, R.A. Jones, I.R. Beacham, C. Bucholtz, M.P. Jennings, Whole genome analysis reveals a high incidence of non-optimal codons in secretory signal sequences of *Escherichia coli*, Biochem. Biophys. Res. Commun. 322 (2004) 1038–1044.
[8] Y.D. Li, Y.Q. Li, J.S. Chen, H.J. Dong, W.J. Guan, H. Zhou, Whole genome analysis of non-optimal codon usage in secretory signal sequences of Streptomyces coelicolor, Biosystems 85 (2006) 225–230.
[9] S. Pedersen, Escherichia coli ribosomes translate in vivo with variable rate, EMBO J. 3 (1984) 2895–2898.
[10] M.A. Sorensen, C.G. Kurland, S. Pedersen, Codon usage determines translation rate in *Escherichia coli*, J. Mol. Biol. 207 (1989) 365–377.
[11] J.F. Curran, M. Yarus, Rates of aminoacyl-tRNA selection at 29 sense codons in vivo, J. Mol. Biol. 209 (1989) 65–77.
[12] T.A. Thanaraj, P. Argos, Protein secondary structural types are differentially coded on messenger RNA, Protein Sci. 5 (1996) 1973–1983.
[13] M. Oresic, D. Shalloway, Specific correlations between relative synonymous codon usage and protein secondary structure, J. Mol. Biol. 281 (1998) 31–48.
[14] I.J. Purvis, A.J. Bettany, T.C. Santiago, J.R. Coggins, K. Duncan, R. Eason, A.J. Brown, The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis, J. Mol. Biol. 193 (1987) 413–417.
[15] T. Crombie, J.C. Swaffield, A.J. Brown, Protein folding within the cell is influenced by controlled rates of polypeptide elongation, J. Mol. Biol. 228 (1992) 7–12.
[16] P. Cortazzo, C. Cervenansky, M. Marin, C. Reiss, R. Ehrlich, A. Deana, Silent mutations affect in vivo protein folding in *Escherichia coli*, Biochem. Biophys. Res. Commun. 293 (2002) 537–541.
[17] A.A. Komar, T. Lesnik, C. Reiss, Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation, FEBS Lett. 462 (1999) 387–391.
[18] Y.M. Zalucki, M.P. Jennings, Experimental confirmation of a key role for non-optimal codons in protein export, Biochem. Biophys. Res. Commun. 355 (2007) 143–148.
[19] Y.M. Zalucki, K.L. Gittins, M.P. Jennings, Secretory signal sequence non-optimal codons are required for expression and export of beta-lactamase, Biochem. Biophys. Res. Commun. 366 (2008) 135–141.
[20] D.M. Burns, I.R. Beacham, Rare codons in *E. coli* and *S. typhimurium* signal sequences, FEBS Lett. 189 (1985) 318–324.
[21] J.A. Chesshyre, A.R. Hipkiss, Low-temperatures stabilize Interferon-Alpha-2 against proteolysis in *Methylophilus methylotrophus* and *Escherichia coli*, Appl. Microbiol. Biotechnol. 31 (1989) 158–162.
[22] A.J. Carr, R.F. Rosenberger, A.R. Hipkiss, Microbial Physiology Symposium, European Federation of Biotechnology, Italy, 1986.
[23] M.J. Kosinski, U. Rinas, J.E. Bailey, Proteolytic response to the expression of an abnormal beta-galactosidase in *Escherichia coli*, Appl. Microbiol. Biotechnol. 37 (1992) 335–341.
[24] D.B. Straus, W.A. Walter, C.A. Gross, *Escherichia coli* heat shock gene mutants are defective in proteolysis, Genes Dev. 2 (1988) 1851–1858.
[25] S. Gottesman, Proteases and their targets in *Escherichia coli*, Annu. Rev. Genet. 30 (1996) 465–506.
[26] F.C. Neidhardt, R.A. VanBogelen, V. Vaughn, Genetics and regulation of heat shock proteins, Annu. Rev. Genet. 18 (1984) 295–329.
[27] C. Ganesh, A.N. Shah, C.P. Swaminathan, A. Surolia, R. Varadarajan, Thermodynamic characterization of the reversible, two-state unfolding of maltose binding protein, a large two-domain protein, Biochemistry 36 (1997) 5020–5028.
[28] P.G. Telmer, B.H. Shilton, Insights into the conformational equilibria of maltose-binding protein by analysis of high affinity mutants, J. Biol. Chem. 278 (2003) 34555–34567.
[29] F.A. Saul, M. Mourez, B. Vulliez-Le Normand, N. Sassoon, G.A. Bentley, J.M. Betton, Crystal structure of a defective folding protein, Protein Sci. 12 (2003) 577–585.
[30] S. Zhang, E. Goldman, G. Zubay, Clustering of low usage codons and ribosome movement, J. Theor. Biol. 170 (1994) 339–354.
[31] D.N. Collier, V.A. Bankaitis, J.B. Weiss, P.J. Bassford Jr., The antifolding activity of SecB promotes the export of the *E. coli* maltose-binding protein, Cell 53 (1988) 273–283.
[32] S.R. Kulothungan, M. Das, M. Johnson, C. Ganesh, R. Varadarajan, Effect of crowding agents, signal peptide, and chaperone SecB on the folding and aggregation of *E. coli* maltose binding protein, Langmuir 25 (2009) 6637–6648.
[33] S. Park, G. Liu, T.B. Topping, W.H. Cover, L.L. Randall, Modulation of folding pathways of exported proteins by the leader sequence, Science 239 (1988) 1033–1035.
[34] K. Beena, J.B. Udgaonkar, R. Varadarajan, Effect of signal peptide on the stability and folding kinetics of maltose binding protein, Biochemistry 43 (2004) 3608–3619.
[35] K. Tani, H. Tokuda, S. Mizushima, Translocation of ProOmpA possessing an intramolecular disulfide bridge into membrane vesicles of *Escherichia coli*. Effect of membrane energization, J. Biol. Chem. 265 (1990) 17341–17347.
[36] K. Tani, S. Mizushima, A chemically cross-linked nonlinear proOmpA molecule can be translocated into everted membrane vesicles of *Escherichia coli* in the presence of the proton motive force, FEBS Lett. 285 (1991) 127–131.
[37] S.H. Lecker, A.J. Driessen, W. Wickner, ProOmpA contains secondary and tertiary structure prior to translocation and is shielded from aggregation by association with SecB protein, EMBO J. 9 (1990) 2309–2314.

[38] K.E. Reed, J.E. Cronan Jr., *Escherichia coli* exports previously folded and biotinated protein domains, J. Biol. Chem. 266 (1991) 11425–11428.
[39] R.E. Dalbey, W. Wickner, Leader peptidase catalyzes the release of exported proteins from the outer surface of the *Escherichia coli* plasma membrane, J. Biol. Chem. 260 (1985) 15925–15931.
[40] G.D. Stormo, T.D. Schneider, L.M. Gold, Characterization of translational initiation sites in *E. coli*, Nucleic Acids Res. 10 (1982) 2971–2996.
[41] A. Eyre-Walker, M. Bulmer, Reduced synonymous substitution rate at the start of enterobacterial genes, Nucleic Acids Res. 21 (1993) 4599–4603.
[42] P.M. Sharp, W.H. Li, The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias, Mol. Biol. Evol. 4 (1987) 222–230.
[43] Y.M. Zalucki, P.M. Power, M.P. Jennings, Selection for efficient translation initiation biases codon usage at second amino acid position in secretory proteins, Nucleic Acids Res. 35 (2007) 5748–5754.
[44] C.M. Stenstrom, E. Holmgren, L.A. Isaksson, Cooperative effects by the initiation codon and its flanking regions on translation initiation, Gene 273 (2001) 259–265.
[45] A.A. Laminet, A. Pluckthun, The precursor of beta-lactamase: purification, properties and folding kinetics, EMBO J. 8 (1989) 1469–1477.