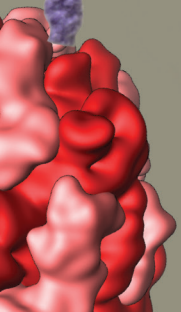
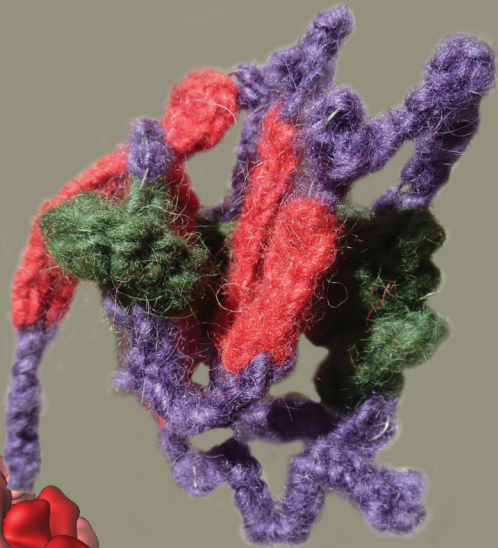
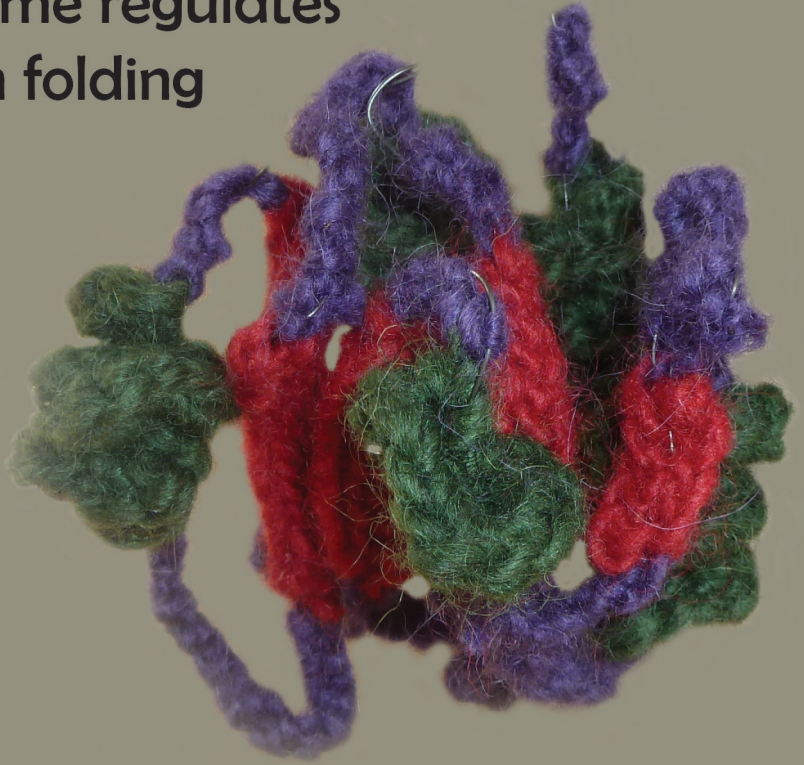


The ribosome regulates flavodoxin folding



Joseline A. Houwman

Propositions

1. In vivo there are more states between a folded and misfolded protein than can be explained by in vitro studies.
(this thesis)
2. Cofactor binding kinetics is a useful tool to study cotranslational folding.
(this thesis)
3. Expressing protein concentration in mg/mL is nonsensical for multi-domain and/or heteromeric proteins.
4. Finding new antibiotics is not the problem, convincing people not to take them for the flu is.
5. To be a great scientist one should not only be a specialist but also a generalist.
6. The ability of foreigners to learn Dutch is negatively correlated with the ability of Dutch people to speak English.
7. Knitting during conferences is a great way to keep up your attention span.

Propositions belonging to the thesis, entitled

‘The ribosome regulates flavodoxin folding’

Joseline A. Houwman
Wageningen, 12 May 2017

The ribosome regulates flavodoxin folding

Joseline A. Houwman

Thesis committee

Promotor

Prof. Dr. W.J.H. van Berkel
Personal chair at the Laboratory of Biochemistry
Wageningen University & Research

Co-promotor

Dr. C.P.M. van Mierlo
Assistant professor, Laboratory of Biochemistry
Wageningen University & Research

Other members

Prof. Dr. J. van der Oost, Wageningen University & Research
Dr. S.G.D. Rüdiger, Utrecht University
Dr. A. Matagne, University of Liège, Belgium
Dr. W. Holtkamp, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

The ribosome regulates flavodoxin folding

Joseline A. Houwman

Thesis

Submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University & Research
by the authority of the Academic Board,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 12 May 2017
at 1.30 p.m in the Aula

Joseline A. Houwman
The ribosome regulates flavodoxin folding
178 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2017)
With references, with summaries in English and Dutch

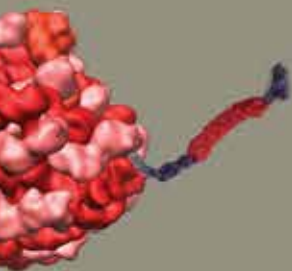
ISBN 978-94-6343-145-3
DOI <http://dx.doi.org/10.18174/410493>

Contents

Chapter 1	p.7
<i>Introduction</i>	
Chapter 2	p. 27
<i>Concurrent presence of on- and off-pathway molten globules under physiological conditions</i>	
Chapter 3	p. 59
<i>Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome</i>	
Chapter 4	p. 85
<i>The ribosome restrains molten globule formation in stalled nascent flavodoxin</i>	
Chapter 5	p. 109
<i>Folding of proteins with a flavodoxin-like architecture</i>	
Chapter 6	p. 151
<i>General discussion</i>	
Chapter 7	p. 163
<i>English summary Nederlandse samenvatting</i>	
About the athour	p. 171
<i>Acknowledgements Publications Curriculum Vitae Overview of completed training activities</i>	

1

Introduction



THE PROTEIN-FOLDING PROBLEM

Proteins have evolved to become some of the most important components in living cells. Not only do proteins confer stability through the cytoskeleton, provide immune responses, and control gene expression, they are also essential for cellular homeostasis through catalysis of metabolic reactions and transport of the resulting products. The ability of proteins to perform such a wide range of functions and reactions stems from folding their amino acid sequences into three-dimensional (3D) structures. Every protein consists of a linear chain of amino acid residues with a specific sequence. This peptide sequence is referred to as the protein's primary structure, which folds into secondary structure elements such as α -helices, β -strands and loops. In most proteins these elements fold into a specific, often tightly packed and globular, 3D conformation, which is referred to as the tertiary structure or the native structure. Some proteins, so called multi-domain proteins, are composed of multiple chains of amino acid residues and the resulting tertiary structures must come together in a quaternary structure for correct functionality.

How proteins are able to form such a wide variety of complex 3D-structures from differing amino acid sequences and how similar sequences can fold to dissimilar structures have been long-standing questions in the field of protein folding (1-5). Two principles of protein folding have been essential for trying to understand these difficult questions. The first principle is Anfinsen's thermodynamic hypothesis, which states that, at least for small globular proteins, the native structure is only determined by the amino acid sequence (1). The native structure is a unique, stable conformation that has the lowest free energy in the protein's relevant physiological environment. All possible conformations arising from a single amino acid chain can be visualized in a folding funnel (2,6). This funnel depicts the free energy of folding (z-axis) as a function of conformational entropy (x,y-plane). Starting from the unfolded state of a protein (i.e., high entropy and free energy), the chain of amino acid residues randomly searches for the native state, which is the lowest point in the energy landscape for protein folding (Fig. 1A, (1)). This search highlights the second principle that governs protein folding, namely Levinthal's paradox. Calculating the time it would take for a 100-residue protein to randomly search all possible conformations, Levinthal concluded that it would take longer than the lifetime of the universe (7,8). However, most small proteins fold to their native structure within milliseconds or even microseconds. The paradox can be resolved if the conformational search is not random, but involves intermediates in which local interactions form that determine the subsequent folding of the polypeptide chain (7,8). Due to formation of intra-chain contacts, the conformational freedom and the peptide's free energy is lowered. As a result, a pathway emerges in the folding landscape that directs and thus accelerates the folding reaction (Fig. 1B).

The starting point for folding (i.e., the unfolded state) is not just one defined state. Instead, it comprises a conformational ensemble of rapidly interconverting unfolded structures. Thus, there is not just one "correct" pathway to the native state, but folding can start from any conformation within the unfolded state. This phenomenon can be represented in an idealized funnel, as depicted in Fig. 1C, where folding routes from any unfolded conformation to the native structure are direct and kinetically identical. Though such landscapes promote unrestricted folding, most proteins have to pass through local energetic minima on their way to their native fold.

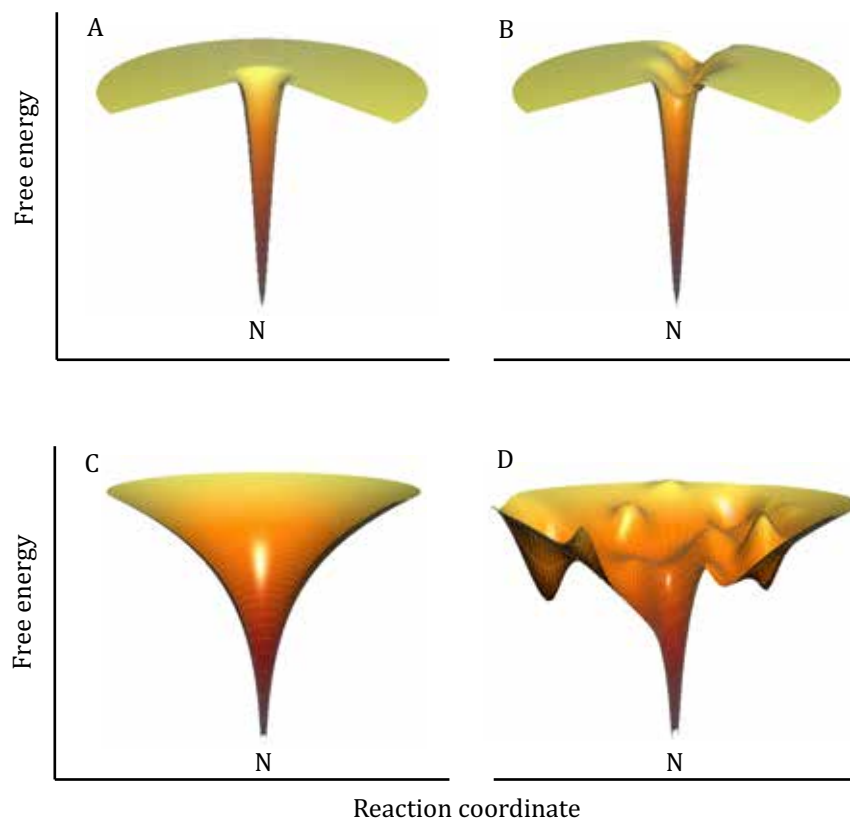


Figure 1 Visualisation of various types of folding landscapes, also called folding funnels. *A*, the classic Levinthal "golf course", wherein the search for the native state is random. *B*, because of formation of folding intermediates, a "pathway" emerges. *C*, an idealized funnel, in which all possible starting conformations lead to the native state. *D*, kinetic traps and energy barriers lead to a "rugged" landscape.

This causes the folding funnel to become rugged (Fig. 1D), producing kinetic traps and barriers that folding proteins must overcome to reach the native state. Such traps and barriers are associated with the formation of intermediate folding species, which can be either on- or off-pathway to the native structure. While on-pathway intermediates reside on a productive folding route, off-pathway intermediates have non-native contacts between amino acid residues and need to (partially) unfold before they can resume a productive folding route. Due to the high-energy barrier between the off-pathway intermediate and the native state, the folding protein can become trapped in the off-pathway state, which can subsequently lead to aggregation with similarly trapped molecules. By studying the formation and behaviour of intermediates, we may comprehend the principles underlying protein folding.

A particular type of intermediate is the molten globule (MG), which can be either on- or off-pathway to the native state. MGs were first described for the protein α -lactalbumin (9-11). They are characterized by a substantial amount of secondary structure, yet without the tight side-chain packing and burial of hydrophobic surfaces associated with tertiary protein structure. Furthermore, while MGs possess a

loosely packed core, they are relatively compact compared to native protein (typical radius increase of around 10 to 30 %) (11,12). Under mildly denaturing conditions MGs can be found for various proteins, such as α -lactalbumin, apomyoglobin and cytochrome c (13-18). Detailed structural characterization of molten globular proteins is challenging, since MGs often consist of transient structures that possess inherent flexibility and these states are hardly populated compared to native or unfolded protein.

THE PROCESS OF PROTEIN PRODUCTION AND FOLDING *IN VIVO*

In vitro folding experiments most often use the full-length protein sequence to probe the folding landscape. This does not reflect the situation *in vivo*, as proteins may already fold during their translation from mRNA by the ribosome, so-called cotranslational folding.

Translation of proteins (described in detail in (19)) starts with the positioning of two ribosomal subunits (i.e., the 30S and the 50S subunit in case of bacteria) along mRNA with help from initiation factor proteins (Fig. 2A). On the mRNA, the start codon for methionine recognises the anti-codon of fMet-tRNA (tRNA charged with a formyl-methionine). Within the peptidyl transferase center (PTC) of the ribosome, where amino acids become covalently attached to one another, there are three sites for tRNAs: the amino-acyl (A), the peptidyl (P) and the exit (E) site. Elongation factor Tu (EF-Tu) brings a charged tRNA to the A-site (Fig. 2B), where the anti-codon of this tRNA binds to the corresponding codon of mRNA. The P-site already holds tRNA attached to the growing nascent chain (i.e., fMet-tRNA when translation starts). Upon C-terminal formation of a peptide bond between the amino acid attached to the tRNA in the A-site and the nascent chain, this chain is transferred to the tRNA in the A-site. Subsequently, translocase EF-G binds to the ribosome and shifts tRNAs from A- and P-sites to P- and E-sites, respectively. In addition, mRNA shifts downstream in relation to the ribosome, and provides the next mRNA codon in the now unoccupied A-site (Fig. 2C and D). Finally, tRNA in the E-site exits the ribosomal complex and a new round of translation starts. This repetitive process continues until a stop codon is reached.

Upon emergence of the nascent chain from the ribosomal exit tunnel, cellular proteins can start to interact with it. Some of the first proteins that interact with the nascent chain are processing proteins such as peptide deformylase (PDF), which removes the formyl group of N-terminal formyl-methionine. Another processing protein is methionine aminopeptidase (MAP), which hydrolyses the N-terminal methionine in case of more than 50 % of nascent chains (Fig. 2E) (20). Chaperones such as Trigger Factor (Fig. 2F) and the DnaK/DnaJ (Fig. 2G) complex may already bind to the nascent chain, either preventing aggregation by holding the nascent chain in an unfolded state, or stimulating folding towards the native state. These chaperones can bind to the nascent chain on their own or they can bind sequentially, thereby passing the nascent chain on from one chaperone to the other. Often chaperones interact with numerous co-chaperones, many of which have not yet been studied extensively. Chaperones can affect proteins during or after translation and may be confined to specific environments such as the ER, periplasmic space or the Golgi apparatus.

Other processes that can happen during translation are further processing of the

nascent chain through glycosylation, phosphorylation, etcetera, and translocation of the nascent chain into membranes, such as the cell membrane or in eukaryotes the ER. Upon reaching the stop codon, which cannot be recognized or bound by tRNA, release factor protein (RF) binds and disassembles the ribosomal complex (Fig. 2G and H). Subsequently, the nascent chain releases and diffuses out of the ribosome through the exit tunnel (Fig. 2I). Once in the cytoplasm a protein may either fold by itself or with help from chaperones such as the DnaK/DnaJ or GroEL/GroES complexes (Fig. 2J).

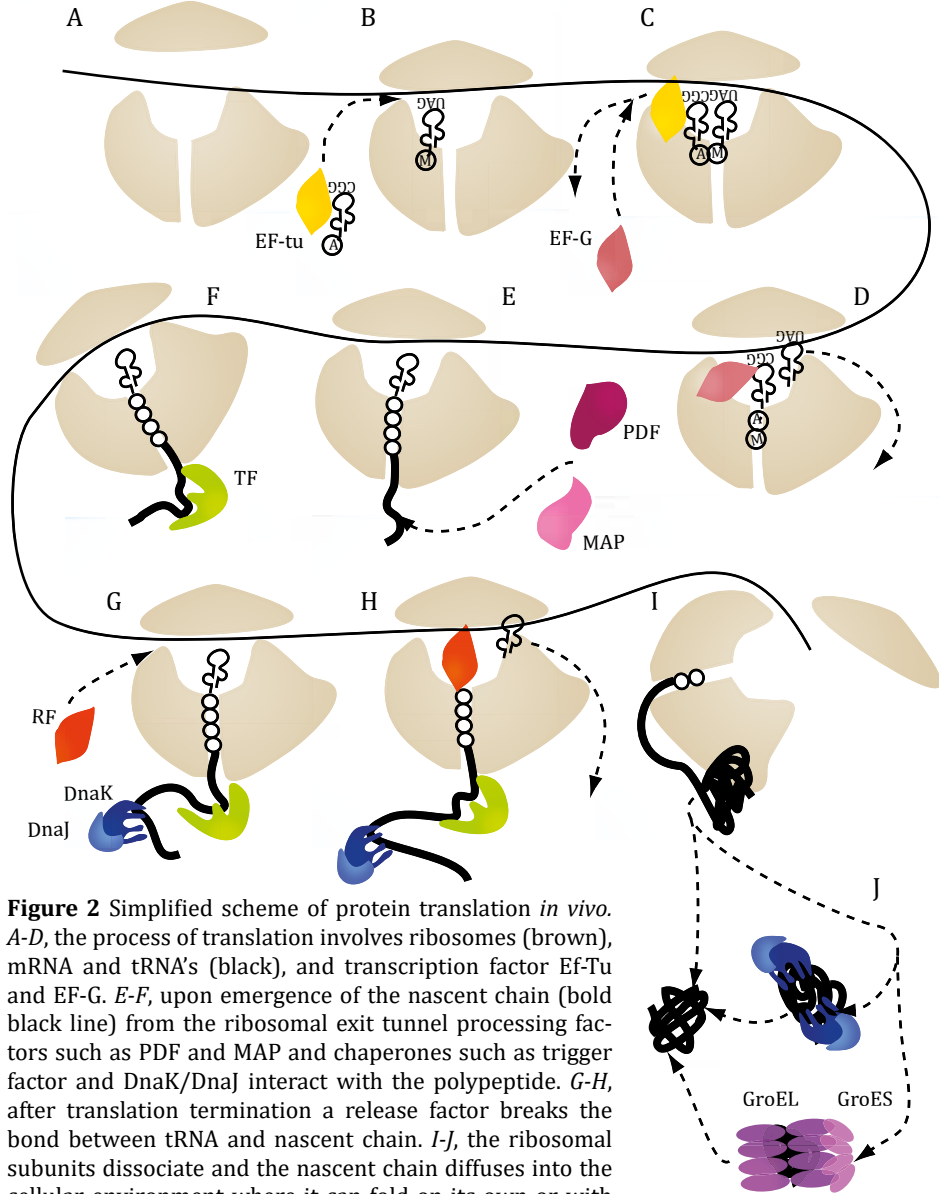


Figure 2 Simplified scheme of protein translation *in vivo*. A-D, the process of translation involves ribosomes (brown), mRNA and tRNA's (black), and transcription factor Ef-Tu and EF-G. E-F, upon emergence of the nascent chain (bold black line) from the ribosomal exit tunnel processing factors such as PDF and MAP and chaperones such as trigger factor and DnaK/DnaJ interact with the polypeptide. G-H, after translation termination a release factor breaks the bond between tRNA and nascent chain. I-J, the ribosomal subunits dissociate and the nascent chain diffuses into the cellular environment where it can fold on its own or with the help of chaperones like GroES/GroEL.

DIFFERENCES BETWEEN FOLDING *IN VIVO* AND *IN VITRO*

While *in vitro* folding studies have contributed much to the understanding of protein folding, their results cannot be transplanted indiscriminately to the folding of proteins *in vivo*. Four major differences between folding *in vivo* and *in vitro* are: presence of chaperones; macromolecular crowding; the use of denaturants such as GuHCl in *in vitro* experiments and cotranslational folding. As explained above, the circumstances under which proteins are introduced in *in vivo* or *in vitro* experiments (i.e., full-length sequence versus growing polypeptide) differ considerably. Though in some cases chemically or sequentially truncated versions of proteins have been studied for their folding behaviour (21-24), these experiments do not take into account the other major differences between *in vivo* and *in vitro* circumstances. As the lifetimes of proteins in cells can range from minutes to years (25) there is ample time for multiple folding rounds for fully synthesized protein.

During these folding and unfolding cycles *in vivo*, chaperone proteins may aid the unfolded or misfolded proteins to achieve their correct fold. Irreversibly misfolded proteins can be targeted for degradation by the proteasome or may be sequestered in inclusion bodies. Chaperones can be purified and their effect on folding can be studied *in vitro*. However, the effect of the entirety of all proteins in the chaperone network is difficult to test adequately *in vitro*, as it numbers in the hundredfold for eukaryotes (26).

Another major difference is the amount of crowding in the cellular environment versus a diluted buffer. Estimated concentrations of proteins and other macromolecules range between 300 to 400 g/L (27). These macromolecules are extremely varied with regard to their biophysical properties, spanning the entire spectrum of size, charge and hydrophobicity. In studies on macromolecular crowding effects *in vitro*, often only one type of crowder is used (28).

Finally, the use of GuHCl and urea to denature proteins *in vitro* and subsequently study their folding is something that is hardly relevant in a cellular context. To transplant findings from such experiments to understand protein folding *in vivo* is difficult. Of more relevance for this are *in vitro* experiments that unfold proteins through changes in temperature or pH. However, also with these techniques it is important to not stray too much from the physiologically relevant conditions, as these would not occur in the cell.

TECHNIQUES TO STUDY PROTEIN FOLDING *IN VIVO*

One of the greatest challenges to study protein folding *in vivo* is differentiating between the signals of the protein of interest and those of other proteins in the cellular environment. By using sophisticated and very specific labelling techniques it is possible to distinguish certain proteins from those of the cellular bulk. These techniques can be used for in-cell NMR (29-31) or fluorescence (32-34) studies of already translated and folded proteins. Studying cotranslational folding *in vivo* is even more difficult, as the ribosome itself contains more than 50 proteins in *E. coli*, a number that increases up to 80 in eukaryotes (35). In addition, chaperones such as Trigger factor and DnaK, but also enzymes such as PDF, MAP and other modifying enzymes interact (transiently) with the nascent chain (20). All these factors complicate characterization of folding pathways of nascent chains *in vivo*. Therefore, most high-resolution cotranslational folding studies employ purified ribosomal nascent chain complexes

(RNCs) or purified transcription/translation systems. These systems consist of the minimal constituents required for successful translation, such as for example tRNAs, aminoacyl-tRNA synthetases, and transcription factors. Sometimes transcription is performed *in vitro* and purified mRNAs are added to the translation system, though transcription/translation can also be achieved in a one-pot reaction series.

Investigation of most RNCs involves complete ribosomes, containing both a small and a large subunit (i.e., 30S and 50S in case of *E. coli*), to which a nascent chain is tethered through a stalling mechanism. Purification of these RNCs is relatively easy, either by ultracentrifugation or through attachment of an affinity tag to the N-terminus of the nascent chain followed by an affinity chromatography step (36,37). Depending on the stalling mechanism used, RNCs can have lifetimes of up to 1 or 2 days (38). These lifetimes attest to the relative strength of the stalling mechanism and makes purified RNCs suitable for biophysical and biochemical research. Various ways exist to stall translation, ranging from truncated mRNAs lacking a stop codon to naturally occurring stalling sequences such as TnaC or SecM (39). Stalling sequences can work in a variety of ways, using effector molecules (TnaC) or by altering the geometry of the ribosome at the PTC (SecM) (40). One of the most widely used stalling sequences is derived from the SecM protein, which is essential for correct incorporation of the SecA protein into membranes (39).

Folding of RNCs can be probed either at equilibrium or, as described recently, in real time (i.e., during the process of translation) (41). Studies of cotranslational folding in real time do not employ stalling mechanisms, as this would halt the translation reaction. Techniques to probe folding states of RNCs include NMR (38,42,43), Förster Resonance Energy Transfer (FRET) (41,44,45), force-induced unfolding/refolding by optical tweezers (46) and proteolytic digestion (47). To be able to distinguish the nascent chain from ribosomal proteins by NMR spectroscopy, NMR-sensitive isotopes needed for labelling the nascent chain are added upon induction of translation. This procedure ensures that ribosomes themselves are not labelled and thus do not drown out the signal of the nascent chain (38,48). Similarly, to selectively probe nascent chain folding by FRET requires the necessary fluorescent labels to be added cotranslationally. In contrast, conventional labelling reactions, such as maleimide labelling of cysteines, would also label the 30 cysteines present in *E. coli* ribosomal proteins (41,44). In summary, the methods used to follow nascent chain folding are quite challenging.

Up to now, cotranslational folding studies have elucidated the timing of several steps during protein production, though there is much that still needs clarification. Folding of α -helices in the ribosomal exit tunnel has been observed for several proteins. A 29-residue zinc finger domain protein was shown to bind its metal cofactor Zn^{2+} while it still resides in the ribosomal exit tunnel (49). Upon emerging from this tunnel various proteins interact with the nascent chain in processes that are precisely coordinated. The proteins PDF and MAP were shown to interact with the nascent chain as soon as it emerges from the exit tunnel (20). These proteins are superseded by Trigger factor after approximately 100 amino acids have been translated (50). When the entire amino acid sequence is available outside the exit tunnel, various proteins can achieve their native fold. These proteins range from all α -proteins, such as globins (44,51), calmodulin (52) and HemK (41), all β -proteins like GFP (53), Dom5 (48), tailspike protein (37,54) and SH3 (38,47), to α/β proteins such as luciferase (55), CFTR NBD1 (56) and barnase (57). The ribosome itself also affects

protein folding, as shown for T4 lysozyme (46), an immunoglobulin-like domain (58), and for the N-terminal domain of HemK (41). Presence of non-native intermediates during nascent chain folding has been detected for a few proteins, including GFP and the N-terminal domain of HemK (41,53). The conformations of these intermediates could not be determined, except that for most proteins probed they seem to be rather compact. In case of SH3, addition of Trigger factor actually unfolded the folding intermediate observed, possibly to prevent misfolding of the nascent chain and subsequent aggregation (47). The real-time folding events of the N-terminal domain of HemK showed that it forms a non-native state in the exit tunnel and rearranges itself into native-like structure upon emerging from the ribosome (41). While important steps have been made to probe cotranslational folding, the folding of nascent polypeptides with complex native structure remains largely unexplored.

ELUCIDATING COTRANSLATIONAL FOLDING OF FLAVODOXIN

One such complex structure is the α - β parallel native-state topology. We aim to clarify how nascent proteins with such a topology fold cotranslationally. We choose to follow nascent chain folding of the α - β parallel protein flavodoxin, which is well known for producing an off-pathway MG during its folding *in vitro*. This choice enables us to probe the influence the ribosome and the intracellular environment have on (mis)folding of nascent chains.

Flavodoxins are small monomeric flavoproteins that contain a non-covalently bound FMN cofactor. Flavodoxins act as low-potential one-electron carriers and adopt a α - β parallel fold, with a central β -sheet surrounded by α -helices (Fig. 3). This fold is often called the flavodoxin-like fold, but it is not constricted to flavodoxins only, as proteins unrelated in sequence or function, such as chemotactic proteins also adopt this fold (59). The flavodoxin-like topology is considered to be one the most ancient protein folds (60). While the flavodoxin protein itself is absent from higher kingdoms of life such as plants and animals (61), it is present in bacteria and many algal species (62). Furthermore, flavodoxin-like folds are found as domains

in many crucial multi-domain eukaryotic proteins such as cytochrome P450 reductase (63) and nitrogen oxide synthase (64).

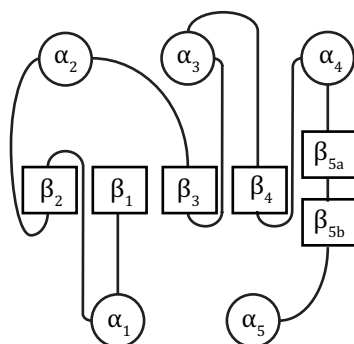
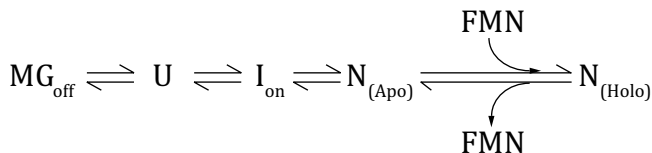


Figure 3 Schematic picture of the flavodoxin-like topology, which consists of a central β -sheet (squares) involving 5 β -strands and is surrounded by 5 α -helices (circles).

The obligate aerobic bacterium *Azotobacter vinelandii* contains three flavodoxins, of which flavodoxin II is most abundant under nitrogen fixing conditions (65). The folding of flavodoxin II from *A. vinelandii* (65,66) has been studied extensively. Un- and re-folding of both the apo- and the holo-form of this protein have been probed, using GuHCl as denaturant (67-74). It has been shown that apoflavodoxin folds autonomously to its native state, after which it binds FMN as the last step in flavodoxin formation (Fig. 4). The 3D structures of natively folded apoflavodoxin and flavodoxin are almost identical, except for considerable flexibility in the flavin-binding region of apoflavodoxin (75,76).

The folding landscape of kinetic apoflavodoxin folding involves two intermediates and is described by the following scheme:



Intermediate Ion is an obligatory, high-energy on-pathway species, which is non-molten globular and rapidly converts to native apoflavodoxin (67). Due to its instability it is not observed during denaturant-dependent equilibrium folding. In contrast, the off-pathway intermediate MG_{off} is a molten globule that is relatively stable and needs to unfold extensively before productive folding to the native state can resume (71,77). Of all folding apoflavodoxin molecules approximately 90 % first misfolds to this MG (67). Its off-pathway nature is due to docking of non-native α -helices, thereby preventing formation of the central β -sheet of native flavodoxin (Fig. 4) (70,71,78). Formation of off-pathway folding intermediates seems to be typical for proteins with a flavodoxin-like fold, as they have also been observed for *Anabaena* apoflavodoxin and chemotactic protein (59,79-81). MG_{off} aggregates severely at elevated protein concentrations or in the presence of macromolecular crowders (77). While formation of native apoflavodoxin is highly cooperative, MG_{off} folds non-cooperatively (71). *In vitro*, the folding of the C-terminal part of MG_{off} precedes the folding of the N-terminal part and the MG gradually compacts due to progressive extension of its ordered core (74).

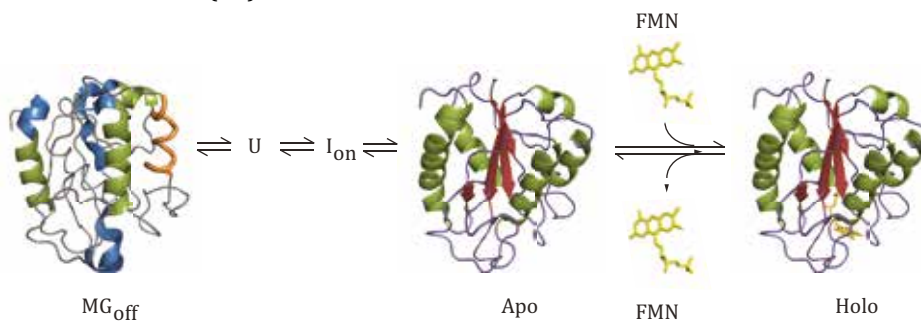


Figure 4 Folding of flavodoxin proceeds from the unfolded state to either native apoflavodoxin (Apo) or to an off-pathway MG (MG_{off}). Binding of FMN to native apoflavodoxin forms flavodoxin (Holo). Structures of Apo and Holo are shown as cartoon renderings of PDB entry 1YOB (83), with α -helices in green and β -strands in red. FMN is shown in yellow. The cartoon rendering of MG_{off} is a model for the known secondary structure elements within the off-pathway MG. α -helices in MG_{off} that are also found in native apoflavodoxin are coloured green, non-native helices in MG_{off} are coloured blue and the structured part of MG_{off} that is neither α -helix nor β -strand is coloured orange (82). This model does not represent the relative positioning of these secondary structure elements within MG_{off} because this positioning is currently unknown.

In flavodoxin variant F44Y, in which phenylalanine at position 44 is substituted for a tyrosine, MG_{off} can be induced in the absence of denaturant by lowering the ionic strength (82). This mutation de facto introduces an oxygen atom into a hydrophobic pocket of apoflavodoxin and destabilizes its native state (78,82). F44Y apoflavodoxin is thus an interesting candidate to study formation of off-pathway MGs *in vivo*, as substantial population of MG_{off} occurs at physiologically relevant conditions.

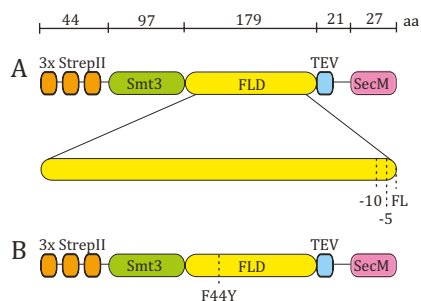


Figure 5 The constructs used for production of RNCs of flavodoxin. All constructs contain a triple N-terminal StrepII-tag (orange), an Smt3-domain (green) fused to flavodoxin (yellow), a recognition site for TEV protease (blue) and a linker that spans the ribosomal exit tunnel and concludes with the SecM stalling sequence (magenta). All constructs contain the C69A mutation in the flavodoxin gene. A, construct used to produce C-terminally shortened RNCs. Through deletion of 5 or 10 residues at the C-terminus of flavodoxin, respectively, 2 constructs are obtained that produce stalled nascent proteins of specified length. B, construct used to produce F44Y RNCs.

To study cotranslational folding of flavodoxin, we prepared RNCs of flavodoxin variants C69A (in which cysteine at position 69 is substituted for an alanine) and F44Y. The native states of both protein variants are similar to native WT flavodoxin (75,82). Besides examining the folding of nascent chain in which the entire apoflavodoxin domain is exposed outside the ribosome, we also prepared C-terminally shortened RNCs of C69A apoflavodoxin. The latter enables us to assess distinct stages during cotranslational folding of flavodoxin. The construct used for RNC production consists of several components (Fig. 5) (38,57): a triple Strep-tag for purification purpose; a Smt3-domain with a C-terminal Ulp1-protease cleavage site to produce authentic N-termini for RNCs; the flavodoxin sequence; a TEV-protease site located C-terminally of the flavodoxin sequence to enable release of the nascent chain from the ribosome; a linker that spans the ribosomal exit tunnel; and the SecM sequence, which stalls the nascent chain to the ribosome.

OUTLINE OF THIS THESIS

The main focus of this thesis is the investigation of how *A. vinelandii* flavodoxin folds during its translation by the ribosome.

This flavodoxin is notorious for forming an off-pathway MG during folding and this intermediate can be populated extensively in flavodoxin variant F44Y. Lowering ionic strength facilitates this population, thus the F44Y variant is an excellent candidate for exploring cotranslational folding. Various methods exist to reveal the presence of MGs, including the use of extrinsic dyes, circular dichroism and intrinsic tryptophan fluorescence. However, these methodologies are unsuitable for studying cotranslational MG formation, as RNCs not only contain the emerging polypeptide, but also more than 50 ribosomal proteins. Thus, detection of MG formation on the ribosome is particularly difficult.

Chapter 2 describes detection of MG formation *in vitro* through a combination of cofactor binding kinetics and polarized time-resolved tryptophan fluorescence spectroscopy. We exploit the fact that while the cofactor FMN is fluorescent in solution, upon binding to apoflavodoxin its fluorescence becomes severely quenched. The off-pathway MG needs to unfold before productive folding and FMN only binds to natively folded apoflavodoxin. Thus, whenever MG_{off} is present in an apoflavodoxin sample the corresponding FMN binding rate ought to be delayed. We show that FMN binding kinetics indeed allows us to follow the presence of MG_{off} under physiological conditions. Interestingly, in combination with polarized time-resolved fluorescence spectroscopy, the binding kinetics reveals the presence of another, simultaneously occurring MG. We present proof that the two concurrent MGs of F44Y apoflavodoxin are kinetically distinct, with one species being on-pathway and the other being off-pathway. Presence of this on-pathway MG would explain why approximately 10 % of folding apoflavodoxin molecules directly follows the productive folding route, as opposed to the 90 % that misfold to MG_{off}.

In **Chapter 3** we present the first insights into cotranslational folding of flavodoxin. We address the binding of flavin mononucleotide (FMN) to nascent flavodoxin, by generating ribosome-arrested nascent chains that expose either the entire protein or C-terminally truncated segments outside the exit tunnel. We used constructs for flavodoxin that lack zero, five or ten amino acid residues at the C-terminus of stalled polypeptide (Fig. 5A). This procedure allows us to mimic the late stages of protein translation and enables the elucidation of the phase at which nascent apoflavodoxin is capable of binding FMN. **Chapter 3** demonstrates that sequestering of only five C-terminal amino acids of flavodoxin within the ribosome already forces the polypeptide chain in a non-native conformation, which is incapable of binding FMN. In contrast, released protein that lacks the five C-terminal residues of flavodoxin can bind FMN. We show that because the ribosome affects protein folding, apoflavodoxin cannot bind FMN during its translation. **Chapter 3** also proofs that in *E. coli* FMN is limiting for saturation of this flavoprotein on time-scales vastly exceeding those of flavodoxin synthesis (hours versus seconds). We show in **Chapter 3** that binding of cofactor to released protein is the last step in production of this flavoprotein in the cell. Finally, **Chapter 3** reveals that once apoflavodoxin is entirely synthesized and exposed outside the ribosome to which it is stalled by an artificial linker containing the SecM sequence, the protein is natively folded and capable of binding FMN. For multi-domain proteins that contain a flavodoxin-like fold as subdomain, our work suggests that incorporation of cofactor probably occurs cotranslationally once this fold has fully emerged from the ribosome.

Chapter 4 reports on the existence of MGs during translation, which is a phenomenon that has not been investigated to date. We utilize apoflavodoxin variant F44Y to probe this phenomenon, as extensive population of MG_{off} occurs with this variant at physiological ionic strength. **Chapter 4** shows that ascertaining the binding rate of FMN as a function of ionic strength can be used as a tool to determine the presence of the off-pathway MG on the ribosome. Application of this methodology to F44Y apoflavodoxin RNCs reveals that at physiological ionic strength the ribosome restrains molten globule formation in stalled nascent flavodoxin and forces the nascent chain towards the native state. Confinement of MG formation during translation is an important observation that emphasizes differences between folding *in vivo* and *in vitro*.

The folding properties of various bacterial flavodoxins and other proteins with flavodoxin-like topologies such as CheY have been extensively studied *in vitro*. This research of the past two decades is reviewed in **Chapter 5**. The common theme of the folding of these proteins is the occurrence of a misfolded intermediate, which for several proteins is molten globular. These intermediates arise due to topological frustration in the folding pathway, as an α -helix forms much quicker than a parallel β -sheet. While the misfolded intermediate seems to be a conserved feature of the flavodoxin-like architecture, the productive transition states of the various proteins differ significantly from one another. The final part of the review bridges the knowledge on *in vitro* folding of proteins with a flavodoxin-like topology with information obtained on the cotranslational folding of *A. vinelandii* flavodoxin.

The main insights obtained during the research that lead to this thesis are discussed in **Chapter 6**, as well as directions for future research and possible applications in biotechnology or pharmaceuticals.

REFERENCES

1. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230
2. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19
3. Jackson, S. E. (1998) How do small single-domain proteins fold? *Fold. Des.* **3**, R81-91
4. Dill, K. A., Ozkan, S. B., Weikl, T. R., Chodera, J. D., and Voelz, V. A. (2007) The protein folding problem: when will it be solved? *Curr. Opin. Struct. Biol.* **17**, 342-346
5. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. *Science* **338**, 1042-1046
6. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195
7. Levinthal, C. (1968) Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44-45
8. Levinthal, C. (1969) How to fold gracefully. in *Mössbauer Spectroscopy in Biological Systems* (Munck, J. T. P. D. a. E. ed., University of Illinois Press, Allerton House, Monticello, Illinois
9. Kuwajima, K., Nitta, K., Yoneyama, M., and Sugai, S. (1976) Three-state denaturation of alpha-lactalbumin by guanidine hydrochloride. *J. Mol. Biol.* **106**, 359-373
10. Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S., and Ptitsyn, O. B. (1981) Alpha-Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett.* **136**, 311-315
11. Ohgushi, M., and Wada, A. (1983) 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21-24
12. Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* **6**, 87-103
13. Kuroda, Y., Kidokoro, S., and Wada, A. (1992) Thermodynamic characterization of cytochrome c at low pH. Observation of the molten globule state and of the cold denaturation process. *J. Mol. Biol.* **223**, 1139-1153
14. Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995) Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5446-5450
15. Colón, W., and Roder, H. (1996) Kinetic intermediates in the formation of the cytochrome c molten globule. *Nat. Struct. Biol.* **3**, 1019-1025
16. Jamin, M., and Baldwin, R. L. (1998) Two forms of the pH 4 folding intermediate of apomyoglobin. *J. Mol. Biol.* **276**, 491-504
17. Arai, M., and Kuwajima, K. (2000) Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209-282
18. Mok, K. H., Nagashima, T., Day, I. J., Hore, P. J., and Dobson, C. M. (2005) Multiple subsets of side-chain packing in partially folded states of α -lactalbumins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8899-8904

19. Nierhaus, K. H., and Wilson, D. (2004) *Protein Synthesis and Ribosome Structure*, John Wiley & Sons
20. Jha, S., and Komar, A. A. (2011) Birth, life and death of nascent polypeptide chains. *Biotechnol. J.* **6**, 623-640
21. Hornemann, S., and Glockshuber, R. (1996) Autonomous and reversible folding of a soluble amino-terminally truncated segment of the mouse prion protein. *J. Mol. Biol.* **261**, 614-619
22. Maldonado, S., Jimenez, M. a. Å., Langdon, G. M., and Sancho, J. (1998) Cooperative stabilization of a molten globule apoflavodoxin fragment. *Biochemistry* **37**, 10589-10596
23. Nguyen, H., Jäger, M., Moretto, A., Gruebele, M., and Kelly, J. W. (2003) Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3948-3953
24. Lopez-Llano, J., Campos, L., Bueno, M., and Sancho, J. (2006) Equilibrium Φ -analysis of a molten globule: The 1-149 apoflavodoxin fragment. *J. Mol. Biol.* **356**, 354-366
25. Toyama, B. H., and Hetzer, M. W. (2013) Protein homeostasis: live long, won't prosper. *Nat. Rev. Mol. Cell Biol.* **14**, 55-61
26. Brehme, M., Voisine, C., Rolland, T., Wachi, S., Soper, J. H., Zhu, Y., Orton, K., Vilella, A., Garza, D., and Vidal, M. (2014) A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. *Cell. Rep.* **9**, 1135-1150
27. Zimmerman, S. B., and Trach, S. O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol.* **222**, 599-620
28. Rivas, G. N., and Minton, A. P. (2016) Macromolecular crowding *in vitro*, *in vivo*, and in between. *Trends Biochem. Sci.* **41**, 970-981
29. Selenko, P., and Wagner, G. (2007) Looking into live cells with in-cell NMR spectroscopy. *J. Struct. Biol.* **158**, 244-253
30. Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., Mikawa, T., Walchli, M., Smith, B. O., Shirakawa, M., Gunttert, P., and Ito, Y. (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102-105
31. Pielak, G. J., and Tian, F. (2012) Membrane proteins, magic-angle spinning, and in-cell NMR. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4715-4716
32. Ignatova, Z., and Gierasch, L. M. (2004) Monitoring protein stability and aggregation *in vivo* by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 523-528
33. Festy, F., Ameer-Beg, S. M., Ng, T., and Suhling, K. (2007) Imaging proteins *in vivo* using fluorescence lifetime microscopy. *Mol. Biosyst.* **3**, 381-391
34. Gershenson, A., and Gierasch, L. M. (2011) Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol.* **21**, 32-41
35. Anger, A. M., Armache, J. P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D. N., and Beckmann, R. (2013) Structures of the human and *Drosophila* 80S

ribosome. *Nature* **497**, 80-85

36. Cabrita, L. D., Dobson, C. M., and Christodoulou, J. (2010) Protein folding on the ribosome. *Curr. Opin. Struct. Biol.* **20**, 33-45

37. Evans, M. S., Sander, I. M., and Clark, P. L. (2008) Cotranslational folding promotes β -helix formation and avoids aggregation *in vivo*. *J. Mol. Biol.* **383**, 683-692

38. Eichmann, C., Preissler, S., Riek, R., and Deuerling, E. (2010) Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9111-9116

39. Cruz-Vera, L. R., Sachs, M. S., Squires, C. L., and Yanofsky, C. (2011) Nascent polypeptide sequences that influence ribosome function. *Curr. Opin. Microbiol.* **14**, 160-166

40. Bhushan, S., Hoffmann, T., Seidelt, B., Frauenfeld, J., Mielke, T., Berninghausen, O., Wilson, D. N., and Beckmann, R. (2011) SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. *PLoS Biol.* **9**, e1000581

41. Holtkamp, W., Kokic, G., Jäger, M., Mittelstaet, J., Komar, A. A., and Rodnina, M. V. (2015) Cotranslational protein folding on the ribosome monitored in real time. *Science* **350**, 1104-1107

42. Cabrita, L. D., Hsu, S. T., Launay, H., Dobson, C. M., and Christodoulou, J. (2009) Probing ribosome-nascent chain complexes produced *in vivo* by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22239-22244

43. Hsu, S. T., Cabrita, L. D., Fucini, P., Christodoulou, J., and Dobson, C. M. (2009) Probing side-chain dynamics of a ribosome-bound nascent chain using methyl NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 8366-8367

44. Ellis, J. P., Bakke, C. K., Kirchdoerfer, R. N., Jungbauer, L. M., and Cavagnero, S. (2008) Chain dynamics of nascent polypeptides emerging from the ribosome. *ACS Chem. Biol.* **3**, 555-566

45. Knight, A. M., Culviner, P. H., Kurt-Yilmaz, N., Zou, T., Ozkan, S. B., and Cavagnero, S. (2013) Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem. Biol.* **8**, 1195-1204

46. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I., Jr., and Bustamante, C. (2011) The ribosome modulates nascent protein folding. *Science* **334**, 1723-1727

47. Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012) Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* **48**, 63-74

48. Cabrita, L. D., Hsu, S.-T. D., Launay, H., Dobson, C. M., and Christodoulou, J. (2009) Probing ribosome-nascent chain complexes produced *in vivo* by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22239-22244

49. Nilsson, O. B., Hedman, R., Marino, J., Wickles, S., Bischoff, L., Johansson, M., Muller-Lucks, A., Trovato, F., Puglisi, J. D., O'Brien, E. P., Beckmann, R., and von Heijne, G. (2015) Cotranslational protein folding inside the ribosome exit tunnel. *Cell Rep.* **12**, 1533-1540

50. Oh, E., Becker, A. H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R. J., Typas, A., Gross, C. A., Kramer, G., Weissman, J. S., and Bukau, B. (2011) Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor *in*

vivo. *Cell* **147**, 1295-1308

51. Komar, A. A., Kommer, A., Krashennnikov, I. A., and Spirin, A. S. (1997) Cotranslational folding of globin. *J. Biol. Chem.* **272**, 10646-10651

52. Lamprou, P., Kempe, D., Katranidis, A., Büldt, G., and Fitter, J. (2014) Nano-second dynamics of calmodulin and ribosome-bound nascent chains studied by time-resolved fluorescence anisotropy. *ChemBioChem* **15**, 977-985

53. Kelkar, D. A., Khushoo, A., Yang, Z., and Skach, W. R. (2012) Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. *J. Biol. Chem.* **287**, 2568-2578

54. Clark, P. L., and King, J. (2001) A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early *in vitro* refolding intermediates. *J. Biol. Chem.* **276**, 25411-25420

55. Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U. (1999) Co-translational domain folding as the structural basis for the rapid *de novo* folding of firefly luciferase. *Nat. Struct. Biol.* **6**, 697-705

56. Khushoo, A., Yang, Z., Johnson, A. E., and Skach, W. R. (2011) Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1. *Mol. Cell* **41**, 682-692

57. Rutkowska, A., Beerbaum, M., Rajagopalan, N., Fiaux, J., Schmieder, P., Kramer, G., Oschkinat, H., and Bukau, B. (2009) Large-scale purification of ribosome-nascent chain complexes for biochemical and structural studies. *FEBS Lett.* **583**, 2407-2413

58. Cabrita, L. D., Cassaignau, A. M., Launay, H. M., Waudby, C. A., Wlodarski, T., Camilloni, C., Karyadi, M. E., Robertson, A. L., Wang, X., Wentink, A. S., Goodsell, L. S., Woolhead, C. A., Vendruscolo, M., Dobson, C. M., and Christodoulou, J. (2016) A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. *Nat. Struct. Mol. Biol.* **23**, 278-285

59. Bollen, Y. J. M., and van Mierlo, C. P. M. (2005) Protein topology affects the appearance of intermediates during the folding of proteins with a flavodoxin-like fold. *Biophys. Chem.* **114**, 181-189

60. Caetano-Anollés, G., Kim, H. S., and Mittenthal, J. E. (2007) The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9358-9363

61. Pierella Karlusich, J. J., Ceccoli, R. D., Grana, M., Romero, H., and Carrillo, N. (2015) Environmental selection pressures related to iron utilization are involved in the loss of the flavodoxin gene from the plant genome. *Genome Biol. Evol.* **7**, 750-767

62. Pierella Karlusich, J. J., Lodeyro, A. F., and Carrillo, N. (2014) The long goodbye: the rise and fall of flavodoxin during plant evolution. *J. Exp. Bot.* **65**, 5161-5178

63. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8411-8416

64. Garcin, E. D., Bruns, C. M., Lloyd, S. J., Hosfield, D. J., Tiso, M., Gachhui, R., Stuehr, D. J., Tainer, J. A., and Getzoff, E. D. (2004) Structural basis for isozyme-specific regulation of electron transfer in nitric-oxide synthase. *J. Biol. Chem.* **279**, 37918-

37927

65. Klugkist, J., Voorberg, J., Haaker, H., and Veeger, C. (1986) Characterization of three different flavodoxins from *Azotobacter vinelandii*. *Eur. J. Biochem.* **155**, 33-40
66. Taylor, M. F., Boylan, M. H., and Edmondson, D. E. (1990) *Azotobacter vinelandii* flavodoxin: purification and properties of the recombinant, dephospho form expressed in *Escherichia coli*. *Biochemistry* **29**, 6911-6918
67. Bollen, Y. J. M., Sanchéz, I. E., and van Mierlo, C. P. M. (2004) Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin. *Biochemistry* **43**, 10475-10489
68. Bollen, Y. J. M., Nabuurs, S. M., van Berkel, W. J. H., and van Mierlo, C. P. M. (2005) Last in, first out. The role of cofactor binding in flavodoxin folding. *J. Biol. Chem.* **280**, 7836-7844
69. Bollen, Y. J. M., Kamphuis, M. B., and van Mierlo, C. P. M. (2006) The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4095-4100
70. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2008) Extensive formation of off-pathway species during folding of an α - β parallel protein is due to docking of (non)native structure elements in unfolded molecules. *J. Am. Chem. Soc.* **130**, 16914-16920
71. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2009) Noncooperative formation of the off-pathway molten globule during folding of the α - β parallel protein apoflavodoxin. *J. Am. Chem. Soc.* **131**, 2739-2746
72. Lindhoud, S., Westphal, A. H., Borst, J. W., and van Mierlo, C. P. M. (2012) Illuminating the off-pathway nature of the molten globule folding intermediate of an α - β parallel protein. *PLoS ONE* 10.1371/journal.pone.0045746
73. Lindhoud, S., Westphal, A. H., Visser, A. J. W. G., Borst, J. W., and van Mierlo, C. P. M. (2012) Fluorescence of Alexa fluor dye tracks protein folding. *PLoS ONE* 10.1371/journal.pone.0046838
74. Lindhoud, S., Pirchi, M., Westphal, A. H., Haran, G., and van Mierlo, C. P. M. (2015) Gradual folding of an off-pathway molten globule detected at the single-molecule level. *J. Mol. Biol.* **427**, 3148-3157
75. Steensma, E., Heering, H. A., Hagen, W. R., and Van Mierlo, C. P. M. (1996) Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur. J. Biochem.* **235**, 167-172
76. Steensma, E., and van Mierlo, C. P. M. (1998) Structural characterisation of apoflavodoxin shows that the location of the stable nucleus differs among proteins with a flavodoxin-like topology. *J. Mol. Biol.* **282**, 653-666
77. Engel, R., Westphal, A. H., Huberts, D. H., Nabuurs, S. M., Lindhoud, S., Visser, A. J., and van Mierlo, C. P. (2008) Macromolecular crowding compacts unfolded apoflavodoxin and causes severe aggregation of the off-pathway intermediate during apoflavodoxin folding. *J. Biol. Chem.* **283**, 27383-27394
78. Nabuurs, S. M., and van Mierlo, C. P. M. (2010) Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of

α - β parallel protein flavodoxin. *J. Biol. Chem.* **285**, 4165-4172

79. Otzen, D. E., Giehm, L., Baptista, R. P., Kristensen, S. R., Melo, E. P., and Pedersen, S. (2007) Aggregation as the basis for complex behaviour of cutinase in different denaturants. *Biochim. Biophys. Acta* **1774**, 323-333

80. Kathuria, S. V., Day, I. J., Wallace, L. A., and Matthews, C. R. (2008) Kinetic traps in the folding of beta alpha-repeat proteins: CheY initially misfolds before accessing the native conformation. *J. Mol. Biol.* **382**, 467-484

81. Lorenz, T., and Reinstein, J. (2008) The influence of proline isomerization and off-pathway intermediates on the folding mechanism of eukaryotic UMP/CMP Kinase. *J. Mol. Biol.* **381**, 443-455

82. Nabuurs, S. M., Westphal, A. H., aan den Toorn, M., Lindhoud, S., and van Mierlo, C. P. M. (2009) Topological switching between an α - β parallel protein and a remarkably helical molten globule. *J. Am. Chem. Soc.* **131**, 8290-8295

83. Alagaratnam, S., van Pouderoyen, G., Pijning, T., Dijkstra, B. W., Cavazzini, D., Rossi, G. L., van Dongen, W. M. A. M., van Mierlo, C. P. M., van Berkel, W. J. H., and Canters, G. W. (2005) A crystallographic study of Cys69Ala flavodoxin II from *Azotobacter vinelandii*: Structural determinants of redox potential. *Protein Sci.* **14**, 2284-2295

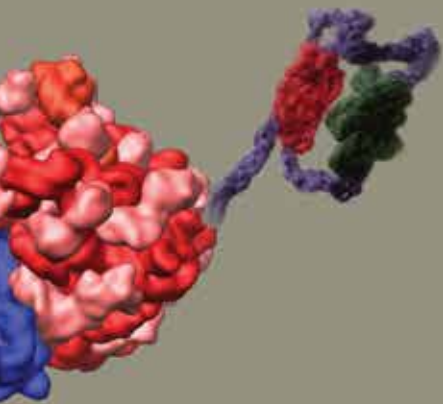
Concurrent presence of on- and off-pathway molten globules under physiological conditions

Joseline A. Houwman[§], Adrie H. Westphal[§], Antonie J.W.G. Visser^{§†},

Jan Willem Borst^{§†} and Carlo P.M. van Mierlo[§]

[§]Laboratory of Biochemistry, Wageningen University, the Netherlands

[†]Microspectroscopy Centre, Wageningen University, The Netherlands



Proteins fold along routes on which intermediates form that can be on- or off-pathway to the native structure. An important type of intermediate is the molten globule (MG), which is characterized by a substantial amount of secondary structure, yet without the tertiary side-chain packing of natively folded protein. The protein flavodoxin has an architecture that can be traced back to the universal ancestor of the three kingdoms of life. Proteins with a flavodoxin-like architecture tend to temporarily misfold during unassisted folding to their native state and as a result they form intermediates. Several of these intermediate species are MGs, with the MG of flavodoxin from *Azotobacter vinelandii* as a particular example. Without the use of denaturants or lowering of pH, this MG is induced in the F44Y variant of *A. vinelandii* apoflavodoxin at physiological ionic strength. Here, using a combination of cofactor binding kinetics and polarized time-resolved tryptophan fluorescence spectroscopy, we show that two MG species of F44Y apoflavodoxin co-exist. Besides the misfolded, off-pathway MG described previously, we discovered an on-pathway MG. Thus, distinct, simultaneously present MGs that reside on folding routes of decidedly different nature (i.e., on- and off-pathway) exist during apoflavodoxin folding. Detection of the presence of concurrent MGs enables future exploration of how the cellular environment, like involvement of chaperones, influences their formation.

INTRODUCTION

Proteins that fold from their unfolded to native structure often pass through intermediate folding states (i.e., local minima in free energy), which are partially structured species (1-4). These intermediates can either reside on productive folding pathways (on-pathway species) or form dead-ends within the folding funnel (off-pathway species) (5). On-pathway intermediates pre-dominantly contain native-like interactions in their structured parts. In contrast, off-pathway species have significant non-native, misfolded structure and are trapped in the corresponding energetic minima. To return to the productive folding ensemble, these trapped species have to overcome significant energy barriers.

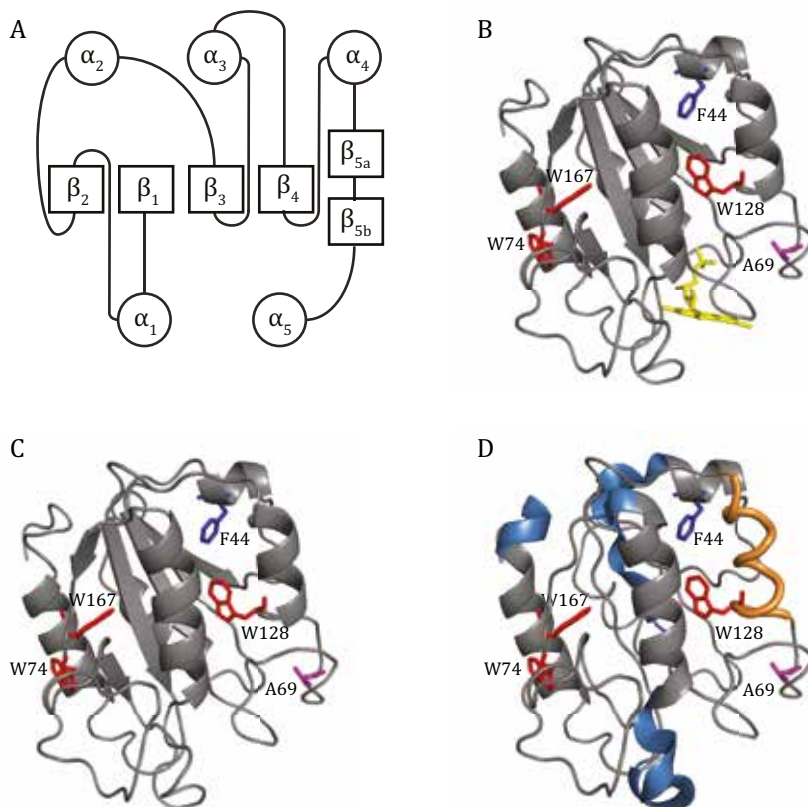
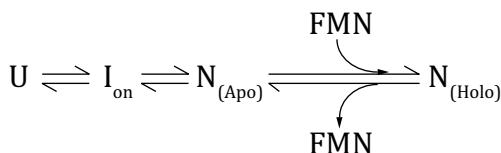


Figure 1 Cartoon drawings of folding species of *A. vinelandii* flavodoxin (PDB entry 1YOB (28)). *A*, Topology of the flavodoxin-like fold. α -Helices are represented as circles and β -strands as squares. *B*, flavodoxin, with the FMN cofactor coloured yellow. *C*, native apoflavodoxin (i.e., flavodoxin without FMN), which is structurally identical to flavodoxin except for disorder (not shown) in the flavin-binding region (74). *D*, Model for MG_{off} which highlights the four transiently structured. Helical parts also present in native protein are shown and regions that adopt α -helical structure not found in native protein are coloured blue; the orange element is structured, yet is neither α -helix nor β -strand (38). These regions dock non-natively and form the core of apoflavodoxin's MG_{off} (38-41). The cartoon shows the structural elements of MG_{off} but their relative positioning is unknown. Residues F44, A69, W74, W128 and W167 are represented as sticks and are labelled. Residue F44 is shown in blue, A69 is coloured magenta and the tryptophans are highlighted in red.

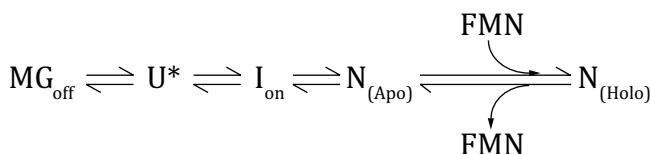
A specific type of folding intermediate is the molten globule (MG). MGs contain secondary structure but lack the tight tertiary packing of natively folded proteins. Furthermore, MGs are relatively compact (typical radius increase compared to native protein of about 10 to 30 %), possess a loosely packed hydrophobic core and expose hydrophobic patches to the solvent (6,7). MGs are prone to aggregation and consequently are implicated in various diseases (8). Several proteins form MGs *in vitro*, including α -lactalbumin, apomyoglobin and cytochrome c (9-16). The MGs of the latter proteins contain native-like secondary structure and packing, causing them to be on-pathway to the native state. In contrast, off-pathway MGs contain non-native secondary structure and/or packing. Nowadays, many proteins are thought to fold via MGs (17-23) and these intermediates are for example necessary during insertion of proteins into membranes or during translocation (24-26). Because of their importance in protein folding, considerable effort has been put into determining the structural properties of MGs. Investigating the characteristics of MGs is complicated by their generally low presence at equilibrium, their conformational heterogeneity, their often transient nature, and their tendency to aggregate (27). To research their folding properties, MGs are induced by for example lowering pH or through addition of denaturant. However, it remains a challenge to extrapolate the insights gained in such studies to physiologically relevant conditions.

Flavodoxins are monomeric, single domain flavoproteins that are involved in electron transfer. To shuttle electrons, flavodoxins contain a non-covalently bound flavin mononucleotide (FMN). The native protein has an α/β parallel topology, with five α -helices surrounding a central parallel β -sheet (Fig. 1A) and is the archetype for the $\alpha\beta\alpha$ sandwich class (28,29). Flavodoxins from *Anabaena*, *A. vinelandii*, (Fig. 1B) and *Desulfovibrio desulfuricans* tend to form off-pathway intermediates during their unassisted folding *in vitro* to their functionally active forms (20,30-32). Several of these species are MGs. Off-pathway intermediate formation is also characteristic for the folding of CheY-like proteins, which share the flavodoxin-like fold (33,34). A MG that has been extensively studied is the off-pathway MG (MG_{off}) of the 179-residue flavodoxin from *A. vinelandii* (32,35-43).

Characterisation of how *A. vinelandii* flavodoxin kinetically folds *in vitro* shows that two parallel folding routes are accessible (32,44). These routes, of which one is on- and the other off-pathway, are visualized in the energy landscape of Fig. 2. This landscape depicts a funnel to the native state and a trough in which a misfolded intermediate temporarily resides. Folding in the funnel and trough are described by schemes 1 and 2, respectively:



Scheme 1



Scheme 2

Scheme 1 represents the approximately 10 % of unfolded molecules (U; rim of the funnel towards native protein) that directly follow the productive folding route to native apoflavodoxin ($N_{(Apo)}$; bottom of funnel). On this path, intermediate I_{on} is an obligatory, high-energy on-pathway species that rapidly converts to native apoflavodoxin (i.e., flavodoxin without FMN; Fig. 1C) (32). I_{on} is native-like and resides in a moat in the funnel towards native protein. Due to its instability, I_{on} is not observed during denaturant-dependent equilibrium folding of the protein. This productive folding of 10 % of unfolded molecules happens on the microsecond timescale. Subsequently, flavodoxin ($N_{(Holo)}$; not depicted in Fig. 2) forms upon binding of FMN to native apoflavodoxin.

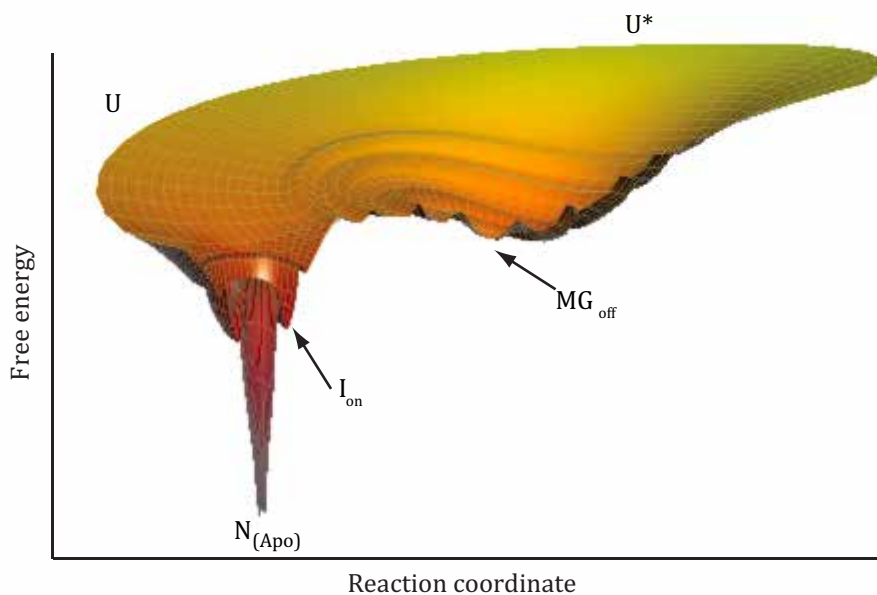


Figure 2 Schematic folding energy landscape of *A. vinelandii* apoflavodoxin. Free energy is shown vertically and the reaction coordinate is horizontally. Unfolded protein structures reside at the rim of each folding funnel. The left funnel shows the on-pathway route to native protein ($N_{(Apo)}$), on which an on-pathway intermediate (I_{on}) resides. The trough on the right displays formation of apoflavodoxin's off-pathway MG (MG_{off}). This molten globular species folds gradually, as exemplified by the presence of shallow moats. Only native apoflavodoxin binds FMN, which leads to considerable stabilisation of the protein and deepening of the corresponding funnel (not shown). In reality, the ratio of the circumferences of the rims of the on- and off-pathway funnels is about 1 to 9. This ratio reflects that 10 % of unfolded molecules directly follow the productive folding route to native apoflavodoxin, whereas 90 % of unfolded molecules misfold and temporarily form MG_{off} (32). Apoflavodoxin is in 100 mM potassium pyrophosphate, pH 6.0, at 25 °C.

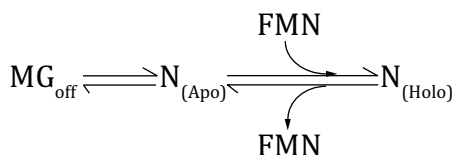
Around 90 % of unfolded molecules (U^* ; the outer edge of the off-pathway trough) temporarily misfold within milliseconds and form the off-pathway intermediate MG_{off} (scheme 2). This relatively stable species is also detected during denaturant-dependent equilibrium folding and is molten globular: MG_{off} is positioned at the bottom of the off-pathway trough (Fig. 2) (32,35,42). MG_{off} needs to unfold sub-

stantially (i.e., to U^*) before folding to the native state takes place. This productive folding occurs within microseconds and again involves I_{on} . Thus, all folding apoflavodoxin molecules, whether they reside on productive or non-productive pathways, ultimately pass through I_{on} before reaching the native state.

Unfolded apoflavodoxin is not a featureless statistical coil, but instead contains four transiently ordered regions. Three of these regions are α -helices, whereas the fourth adopts non-native structure that is neither α -helix nor β -strand (Fig. 1D) (38). During folding, these structured elements interact and can subsequently form the ordered core of MG_{off} (38,40). This propensity to interact is visualized in Fig. 2 and in scheme 2 by U^* . Non-native docking of the α -helices in MG_{off} prevents formation of the parallel β -sheet of native apoflavodoxin (38,40,41,45). Thus, the source for MG_{off} formation is situated in the unfolded state. The off-pathway intermediate has a drastically different architecture compared with native protein: it is largely α -helical and contains no β -sheet (39) and is slightly expanded compared to native apoflavodoxin (36). This α -helical MG acts as a trap and needs to unfold significantly in order to embark on a route to native α - β parallel protein.

While formation of native apoflavodoxin is highly cooperative (32,43,46,47), conversion of unfolded protein into MG_{off} is a non-cooperative, gradual process that simultaneously involves separate regions within apoflavodoxin (40,43). This suggests an energy landscape of MG folding with many barriers (visualised by shallow moats in the off-pathway trough of Fig. 2). Ultimately, after folding, the helical off-pathway MG of apoflavodoxin is almost entirely structured (40,43).

Unfolding of MG_{off} happens within seconds (32) and is symbolized, for use in the analysis of the data obtained in this study, in the following scheme:



Scheme 3

Note that the productive on-pathway intermediate I_{on} has not been included in this and following schemes, because its rapid transformation to native protein cannot be detected on the timescale of the cofactor binding kinetics and polarized time-resolved tryptophan fluorescence experiments described in this paper.

Most of our knowledge of apoflavodoxin's MG_{off} stems from studies that use denaturant. In case of apoflavodoxin variant F44Y, in which phenylalanine at position 44 is substituted for a tyrosine, the protein switches from natively folded to MG_{off} upon decreasing ionic strength to physiological values (i.e., from 100 to 10 mM potassium pyrophosphate (KPPi), which is equivalent to 75 to 345 mM NaCl) (39). The tyrosine introduces an extra oxygen atom into a hydrophobic pocket of native apoflavodoxin, causing considerable destabilization (39,41). F44Y protein allows characterization of apoflavodoxin's MG_{off} under physiologically relevant conditions. In this study we characterize the behaviour of MG_{off} by exploiting the quenching of FMN fluorescence upon binding to apoflavodoxin (48,49). FMN binding to apoflavodoxin involves a very specific combination and geometry of hydrogen bonds and aromat-

ic interactions (50). Therefore, binding of FMN only occurs to native apoflavodoxin and is the last step in folding of flavodoxin (44). This binding is diffusion-limited. As MG_{off} needs to unfold before native apoflavodoxin can form, in presence of this intermediate the rate of cofactor binding and accompanied FMN fluorescence quenching slows down compared to the situation where only natively folded apoflavodoxin is present. Thus, the rate of FMN fluorescence quenching ought to track the unfolding of MG_{off} .

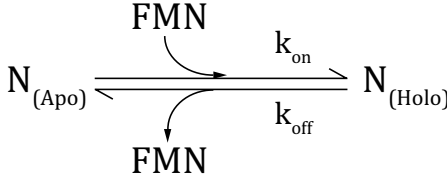
Here we show that FMN binding kinetics, in combination with polarized time-resolved fluorescence spectroscopy, reveals, besides MG_{off} , the presence of another, simultaneously occurring MG. This concurrent MG of F44Y apoflavodoxin is kinetically distinct from MG_{off} and is on-pathway to native protein.

MATERIAL AND METHODS

Protein expression and purification. To avoid covalent dimerization of purified *A. vinelandii* (apo)flavodoxin, the single cysteine at position 69 was substituted by an alanine (46). The properties of variant C69A are largely similar to that of wild-type (apo)flavodoxin (46,51). An additional F44Y replacement was introduced (39), and this protein variant is referred to as F44Y (apo)flavodoxin. Each construct (i.e., C69A or F44Y) was transformed in *Escherichia coli* (strain TG2) and purified as described previously (46). Apoprotein was prepared according to established protocols (52). Holo- and apoproteins were flash-frozen in liquid nitrogen and stored in KPPi buffer, pH 6.0, at -80 °C.

FMN binding. To assess the binding rates of cofactor to apoflavodoxin, FMN fluorescence was followed in time. Flavin fluorescence was recorded on a Cary Eclipse spectrofluorimeter (Varian) at 25 °C. Upon freezing and subsequent thawing of F44Y apoprotein a fraction of the protein deteriorates into soluble molecules that are incapable of binding FMN. As these molecules cannot be easily separated from FMN binding competent apoflavodoxin, determination of the concentration of cofactor binding competent apoprotein was achieved by titrating the protein solution with FMN (Sigma). Protein samples of 50 to 140 nM cofactor-binding competent molecules were prepared in KPPi at pH 6.0. Final concentration of FMN was 25 nM. FMN fluorescence of the apoflavodoxin samples was measured in stirred fluorescence cuvettes during five minutes, before and after manual addition of FMN at the same KPPi concentration as the corresponding protein solution. C69A apoflavodoxin samples were made in 100, 70, 40 and 10 mM KPPi and F44Y apoflavodoxin samples were prepared in 100, 90, 80, 70, 60, 50, 40, 30, 20, and 10 mM KPPi, respectively. Excitation was at 450 nm and emission was monitored at 540 nm. Excitation and emission slits were set to bandwidths of 10 and 20 nm, respectively, and PMT voltage was 970 V. The averaging time of the measurement was 0.1 s. The time between adding FMN and the start of fluorescence detection was approximately 2 s. Background correction was performed by subtracting the fluorescence at 540 nm before addition of FMN from the fluorescence data obtained after addition of FMN.

Analysis of FMN binding rates. Analysis of cofactor binding data was done in IGOR Pro (version 6.37). Fluorescence traces (Flu) resulting from predominantly native apoflavodoxin ($N_{(Apo)}$; i.e., all C69A traces and traces of F44Y in 100, 90 and 80 mM KPPi, respectively) were analysed according to the following scheme and equations:



Scheme 4

$$\frac{dC_{N(Apo)}}{dt} = -k_{on} \cdot C_{N(Apo)} \cdot C_{FMN} + k_{off} \cdot C_{N(Holo)}$$

Eq. 1

$$\frac{dC_{FMN}}{dt} = -k_{on} \cdot C_{N(Apo)} \cdot C_{FMN} + k_{off} \cdot C_{N(Holo)}$$

Eq. 2

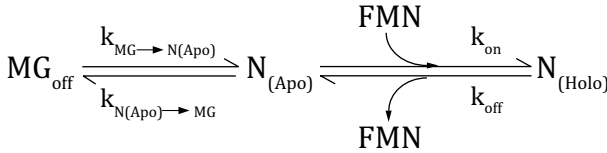
$$\frac{dC_{N(Holo)}}{dt} = k_{on} \cdot C_{N(Apo)} \cdot C_{FMN} - k_{off} \cdot C_{N(Holo)}$$

Eq. 3

$$\frac{dFlu}{dt} = F_{FMN} \cdot \frac{dC_{FMN}}{dt} + F_{N(Holo)} \cdot \frac{C_{N(Holo)}}{dt}$$

Eq. 4

Traces resulting from MG (i.e., when F44Y apoflavodoxin is in 10 to 70 mM KPPI) were analysed with the following scheme and equations, including equations 2 to 4:



Scheme 5

$$\frac{dC_{MG}}{dt} = -k_{MG \rightarrow N(Apo)} \cdot C_{MG} + k_{N(Apo) \rightarrow MG} \cdot C_{N(Apo)}$$

Eq. 5

$$\begin{aligned} \frac{dC_{N(Apo)}}{dt} = & k_{MG \rightarrow N(Apo)} \cdot C_{MG} - k_{N(Apo) \rightarrow MG} \cdot C_{N(Apo)} \\ & - k_{on} \cdot C_{N(Apo)} \cdot C_{FMN} + k_{off} \cdot C_{N(Holo)} \end{aligned}$$

Eq. 6

C_{MG} , $C_{N(Apo)}$, C_{FMN} , $C_{N(Holo)}$ are the initial concentrations of MG, native apoflavodoxin, FMN and flavodoxin, respectively. C_{MG} and $C_{N(Apo)}$ are determined by $\frac{C_{Total}}{1 + \frac{C_{N(Apo)}}{C_{MG}}}$ and

$\frac{C_{Total}}{1 + \frac{C_{MG}}{C_{N(Apo)}}}$, respectively, with C_{Total} the initial concentration of apoprotein (i.e., both native and MG) and with $\frac{C_{N(Apo)}}{C_{MG}}$ the initial ratio between native apoflavodoxin and MG. $k_{MG \rightarrow N(Apo)}$ and $k_{N(Apo) \rightarrow MG}$ are the rate constants for conversion of MG

to native apoflavodoxin and conversion of native apoflavodoxin to MG, respectively and $k_{\text{MG} \rightarrow \text{N(Apo)}}$ equals $\frac{C_{\text{N(Apo)}}}{C_{\text{MG}}} \cdot k_{\text{N(Apo)} \rightarrow \text{MG}}$. k_{on} is the second-order rate constant for

association of FMN to native apoflavodoxin ($\text{M}^{-1}\text{s}^{-1}$) and k_{off} is the first-order rate constant for dissociation of FMN (s^{-1}). k_{off} and K_{D} have been determined previously (50) and k_{on} can be determined because K_{D} equals $k_{\text{off}}/k_{\text{on}}$. F_{FMN} and F_{Holo} are the conversion factors from amount of molecules to fluorescence of FMN and flavodoxin, respectively. Before analysis, data were normalized.

Time-resolved fluorescence anisotropy. C69A and F44Y apoflavodoxin concentration was 2 μM . F44Y apoprotein was freshly prepared to avoid the cold denaturation described above. The KPPi concentrations were as used for the FMN binding experiments. Time-resolved fluorescence anisotropy was determined using time-correlated single photon counting, as described (37,53). Excitation was at 300 nm. Repetition rate of the excitation pulses was 3.8 MHz with a pulse duration of less than 0.2 ps and pulse energies at the pJ level, and samples were in 1 ml 10 mm path length quartz cuvettes at 25 °C. Use was made of a Schott UV-DIL 348.8 nm ($\Delta\lambda = 5.4$ nm) interference filter to detect the fluorescence photons. Decay curves were collected in 4096 channels of a multi-channel analyser using a channel time spacing of 5.0 ps. Ten repeated sequences of 10 s duration of parallel ($I_{\parallel}(t)$) and of perpendicularly ($I_{\perp}(t)$) polarized fluorescence emission were acquired. Background fluorescence emission was determined under identical conditions. For deconvolution purposes, the dynamic instrumental response function was recorded using a freshly made solution of p-terphenyl in a mixture of 50/50 (v/v) cyclohexane and CCl_4 (54,55).

Total fluorescence decay

$$I(t) = I_{\parallel}(t) + 2 I_{\perp}(t) \quad \text{Eq. 7}$$

was analysed using a sum of discrete exponentials with lifetimes τ_i and amplitudes α_i :

$$I(t) = E(t) \otimes \sum_{i=1}^N \alpha_i e^{-t/\tau_i} \quad \text{Eq. 8}$$

$E(t)$ is the instrumental response function and \otimes denotes a convolution product. Time-dependent fluorescence anisotropy

$$r(t) = (I_{\parallel}(t) - I_{\perp}(t)) / I(t) \quad \text{Eq. 9}$$

was analysed from parallel and perpendicular intensity components (37).

Analysis of time-resolved fluorescence anisotropy. Decay of fluorescence anisotropy of apoflavodoxin is caused by exchange of excited-state energy between two tryptophan residues by the Förster mechanism and by overall protein rotation and is described by (37,56,57):

$$r(t) = \{\beta_T \exp(-t/\phi_T) + \beta_r\} \exp(-t/\phi_r) \quad \text{Eq. 10}$$

with ϕ_T the correlation time of resonance energy transfer, ϕ_r the rotational correlation time of the protein. Both β_T and β_r depend on the intra- and intermolecular angles between absorption and emission transition moments.

Reversible energy transfer happens between a pair of identical, isoenergetic chromophores (i.e., tryptophans). In this case ϕ_T is given by:

$$\phi_T = \frac{1}{2k_T} \quad \text{Eq. 11}$$

with k_T the rate constant for energy transfer, which is given by the Förster equation (58):

$$k_T = 8.71 \times 10^{-17} \kappa^2 n^{-4} k_r J R^{-6} \quad \text{Eq. 12}$$

with κ^2 the orientation factor for the transition dipole moments of the tryptophans involved, n the refractive index of the medium between donor and acceptor (59) (for native apoflavodoxin $n=1.6$) (60), k_r the radiative rate constant (ns^{-1}) (61), J the integrated spectral overlap of acceptor (tryptophan) absorbance and donor (tryptophan) fluorescence emission (M^{-1}cm^3), and R the distance (nm) between donor and acceptor.

Unidirectional energy transfer happens if excited-state energy of one tryptophan is lower than that of the other and thereby the lowest excited-state acts as acceptor of resonance energy transfer. In this situation ϕ_T is given by:

$$\phi_T = \frac{1}{k_T} \quad \text{Eq. 13}$$

Steady-state fluorescence anisotropy. Polarized time-resolved fluorescence data are related to steady-state fluorescence anisotropy according to the following relation (62):

$$\langle r \rangle = \frac{\int_0^\infty r(t) \cdot I(t) dt}{\int_0^\infty I(t) dt} \quad \text{Eq. 14}$$

in which $I(t)$ is the total fluorescence decay and $r(t)$ the time-dependent fluorescence anisotropy.

This equation can be rewritten into the well-known Perrin equation:

$$\langle r \rangle = \frac{r(0)}{1 + \frac{\langle \tau \rangle}{\langle \phi \rangle}}$$

Eq. 15

in which $r(0)$ is fluorescence anisotropy at $t=0$, $\langle \tau \rangle$ is the average fluorescence lifetime, and $\langle \phi \rangle$ is a harmonic mean correlation time characteristic for the particular fluorescence depolarization process.

RESULTS

FMN binding follows formation of apoflavodoxin's off-pathway MG. As discussed, FMN binding kinetics ought to track the presence of apoflavodoxin's off-pathway MG under physiologically relevant conditions. To proof the latter, we mixed FMN with either native apoflavodoxin or MG and analysed cofactor binding using Schemes 4 and 5, respectively. As the pseudo-first-order rate constant for FMN binding is on the order of $\mu\text{M}^{-1}\text{s}^{-1}$ (44,50) low concentrations of both apoflavodoxin and FMN are required to enable tracking of cofactor binding on the seconds timescale after manual addition of FMN. The major fluorescence contribution to the binding curves reflects the decrease in free FMN due to FMN binding to apoflavodoxin. The minor fluorescence component to the binding curves tracks the increasing number of flavodoxin molecules, as freshly generated flavodoxin is slightly fluorescent (50). Thus, the FMN fluorescence trace obtained after titrating native apoflavodoxin with FMN is complex, is not simply biphasic, and needs to be analysed by using equations 1 to 4 (Fig. 3A-C, Tables 1 and 2). The FMN fluorescence trace resulting from molten globular apoflavodoxin is also multi-faceted and requires analysis with equations 2 to 6 (Fig. 3D).

In case of C69A apoflavodoxin, the protein is native in 10 and 100 mM KPPI (39). Indeed, FMN fluorescence shows that binding of the cofactor to C69A flavodoxin is independent of salt in this concentration range (Fig. 3A,B). In contrast, the FMN fluorescence traces of F44Y apoflavodoxin at both salt concentrations differ considerably (Fig. 3C,D). F44Y apoflavodoxin in 100 mM KPPI is predominantly natively folded (39) and hence Fig. 3C shows that it has a similar FMN binding curve as native C69A apoflavodoxin (Fig. 3A,B). In 10 mM KPPI however, F44Y apoflavodoxin predominantly adopts the off-pathway MG conformation, which first needs to unfold before native apoflavodoxin can form (Scheme 5) (39). FMN only binds to native apoflavodoxin and under the conditions used this binding happens much faster than the time it takes to unfold the off-pathway MG. As a result, under the latter condition, FMN fluorescence decreases much slower (Fig. 3D) than observed for this protein in 100 mM KPPI (Fig. 3C). This change in FMN binding rate allows us to track the unfolding of MG_{off} .

Previously, it was shown that the rate constants for FMN binding to native apoflavodoxin and FMN release from flavodoxin (i.e., k_{on} and k_{off} , respectively) are affected by ionic strength (63,64). We checked whether this phenomenon occurs in the rather narrow KPPI concentration range of 10 to 100 mM. A superposition of the FMN binding traces obtained for native C69A apoflavodoxin at 100 and 10 mM KPPI, respectively, shows no ionic strength dependence (Fig. 4A). Fitting of these FMN binding curves with equations 1 to 4 resulted in k_{off} and k_{on} values of $9.65 \cdot 10^{-5} \text{ s}^{-1}$ and

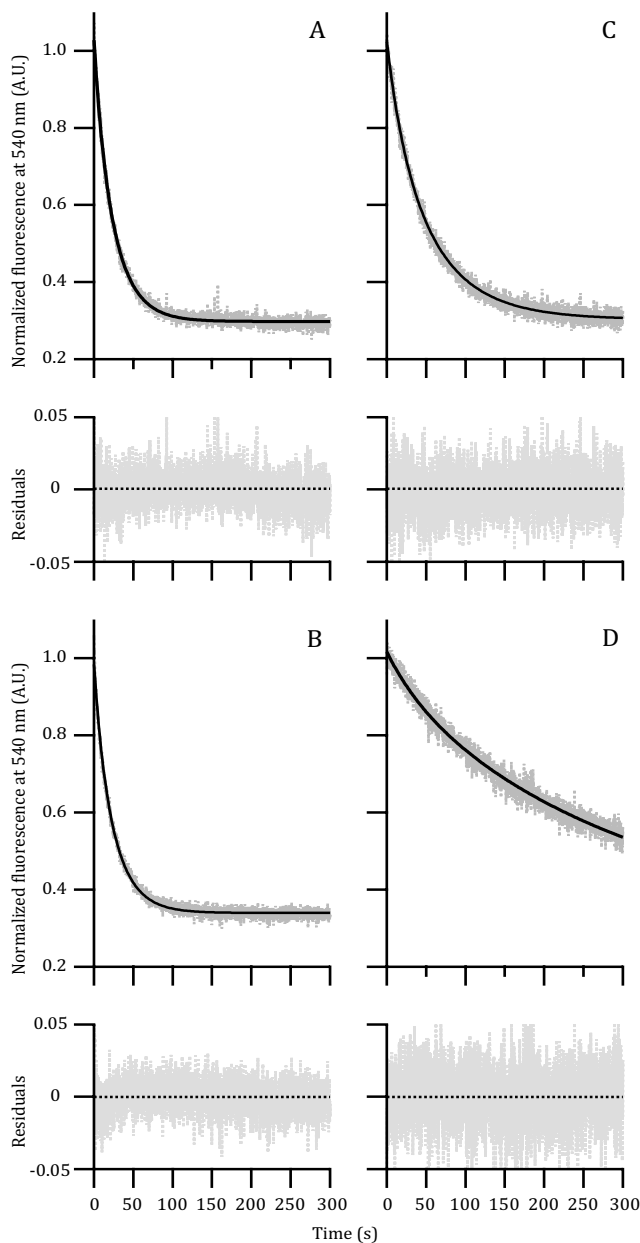


Figure 3 Quenching of FMN fluorescence upon binding of the cofactor to C69A and F44Y apoflavodoxin in 100 or 10 mM KPPI, respectively. Shown are the normalized FMN fluorescence (grey), the fits of the data (black) and the corresponding residuals. Table 1 and 2 list the parameters of the fits. *A*, C69A apoflavodoxin in 100 mM KPPI. *B*, C69A apoflavodoxin in 10 mM KPPI. Protein and FMN concentrations are 140 and 25 nM, respectively. *C*, F44Y apoflavodoxin in 100 mM KPPI. *D*, F44Y apoflavodoxin in 10 mM KPPI. Protein and FMN concentrations are 71 and 25 nM, respectively. (A-C) are analysed with differential equations 1 to 4. (D) is analysed with differential equations 2 to 6. The pH is 6.0 and temperature is 25 °C.

	C69A	F44Y#	F44Y\$
k_{on} ($\text{M}^{-1}\text{s}^{-1}$)	$3.45 \cdot 10^5^*$	$3.45 \cdot 10^5^*$	$3.45 \cdot 10^5^*$
k_{off} (s^{-1})	$9.65 \cdot 10^{-5}^*$	$9.65 \cdot 10^{-5}^*$	$9.65 \cdot 10^{-5}^*$
F_{FMN}	$4.08 \cdot 10^7$	$4.05 \cdot 10^7$	$4.05 \cdot 10^7$
F_{Holo}	$1.15 \cdot 10^7$	$1.18 \cdot 10^7$	$1.18 \cdot 10^7$
$C_{\text{N(Apo)}} \text{ (nM)}^a$	140	71	
$C_{\text{Total}} \text{ (nM)}^a$			71*
$C_{\text{FMN}} \text{ (nM)}^a$	25*	25*	25*
$C_{\text{Holo}} \text{ (nM)}^a$	0*	0*	0*
$\text{Flu}_0^{a,b}$	1.00	1.00	1.00

Table 1 Parameters that describe the binding of FMN to C69A and F44Y apoflavodoxin. For explanation of parameters see main text. C69A apoflavodoxin is in 100 mM KPPI and is native.

^a Values at $t = 0$ s.

^b Initial fluorescence of normalized FMN binding trace

* Parameter fixed

Parameters for traces of F44Y apoflavodoxin containing predominantly native protein (i.e., the protein is in 100, 90 or 80 mM KPPI), as analysed with equations 1 to 4

\$ Parameters for traces of F44Y apoflavodoxin containing predominantly MG (i.e., the protein is in 10 to 70 mM KPPI), as analysed with equations 2 to 6.

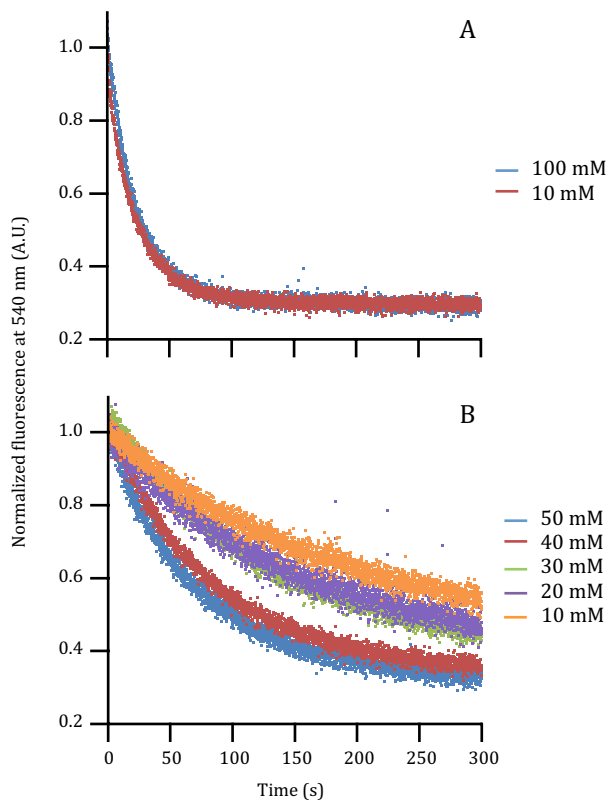


Figure 4 Dependence of the rate of FMN binding to C69A and F44Y apoflavodoxin on KPPI concentration. Shown are normalized FMN fluorescence traces acquired at various salt concentrations. *A*, superposition of FMN binding curves of C69A apoflavodoxin in 100 and 10 mM KPPI, respectively. *B*, superposition of FMN binding curves of F44Y apoflavodoxin at various KPPI concentrations. Concentrations of C69A apoflavodoxin, F44Y apoflavodoxin and FMN are 140, 71 and 25 nM, respectively.

	KPPi (mM)	χ^2	$\frac{C_{N(Apo)}}{C_{MG}}$	$k_{N(Apo) \rightarrow MG}$	$k_{MG \rightarrow N(Apo)}$ #	% Apo	% MG
C69A	100	0.43	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
	70	0.56	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
	40	0.47	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
	10	0.39	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
F44Y	100	0.63	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
	90	0.89	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
	80	1.42	n.a.	n.a.	n.a.	100 ^{\$}	0 ^{\$}
	70	0.67	6.36 ± 0.84	$1.70 \cdot 10^{-2} \pm 6.18 \cdot 10^{-2}$	$1.08 \cdot 10^{-1}$	84	16
	60	0.87	3.81 ± 0.08	$0.70 \cdot 10^{-2*}$	$2.67 \cdot 10^{-2}$	79	21
	50	0.74	2.11 ± 0.03	$0.23 \cdot 10^{-2} \pm 3.22 \cdot 10^{-4}$	$4.85 \cdot 10^{-3}$	66	34
	40	0.58	$1.53 \pm 1.78 \cdot 10^{-2}$	$0.18 \cdot 10^{-2} \pm 1.59 \cdot 10^{-4}$	$2.75 \cdot 10^{-3}$	57	43
	30	0.84	$0.59 \pm 7.38 \cdot 10^{-3}$	$0.69 \cdot 10^{-2} \pm 3.64 \cdot 10^{-4}$	$4.07 \cdot 10^{-3}$	40	60
	20	1.57	$0.44 \pm 6.85 \cdot 10^{-3}$	$0.57 \cdot 10^{-2} \pm 3.46 \cdot 10^{-4}$	$2.51 \cdot 10^{-3}$	29	71
	10	1.03	$0.31 \pm 4.26 \cdot 10^{-4}$	$0.77 \cdot 10^{-2} \pm 4.06 \cdot 10^{-4}$	$2.39 \cdot 10^{-3}$	20	80

Table 2 Ratio of MG to native apoflavodoxin.

For explanation of parameters see main text. Chi-square: χ^2 , n.a.: not applicable

* Fixed

Calculated using $\frac{C_{N(Apo)}}{C_{MG}} \cdot k_{N(Apo) \rightarrow MG}$

^{\$} The corresponding FMN fluorescence traces were analysed with equations 1 to 4, as the samples do not contain MG

$3.45 \cdot 10^5 \text{ M}^{-1}\text{s}^{-1}$, respectively (Table 1). We fixed these rates during further analysis of FMN binding data. Tables 1 and 2 summarize the parameters that describe the binding of FMN to C69A and F44Y apoflavodoxin at various KPPi concentrations.

Whereas FMN binding to native apoflavodoxin is independent of KPPi concentration in the range we use, Fig. 4B shows that the rate of FMN binding to F44Y apoprotein changes most noticeably upon lowering KPPi concentration from 50 to 10 mM KPPi. The FMN binding rate thus depends on the ratio of the initial concentrations of MG and native protein in the sample. This ratio relates to the rates for folding (i.e.) and unfolding (i.e.) of the off-pathway MG. These latter rates must therefore depend on ionic strength and can be derived using equations 2 to 6. To be able to fit the FMN binding data of F44Y apoflavodoxin at low KPPi concentration with these equations, one needs to know the ratio of MG to native protein. To obtain this ratio,

we use the average fluorescence lifetime as derived from time-resolved tryptophan fluorescence anisotropy, as presented in the next sections.

Time-resolved tryptophan fluorescence anisotropy. This methodology can be used to establish relative distances and orientations between pairs of tryptophans. It measures depolarization of fluorescence emission of tryptophans, which is caused by molecular rotation due to Brownian motion and/or by energy transfer to a molecule/residue with a different orientation (see Eq. 10). The time-resolved fluorescence anisotropy decays of native and molten globular apoflavodoxin are shown in Fig. 5. The traces of C69A apoflavodoxin in 10 and 100 mM KPPI and of F44Y apoflavodoxin in 100 mM KPPI are similar (Fig. 5A-C), as at these salt concentrations the samples contain natively folded apoprotein (37,39). In contrast, F44Y apoflavodoxin

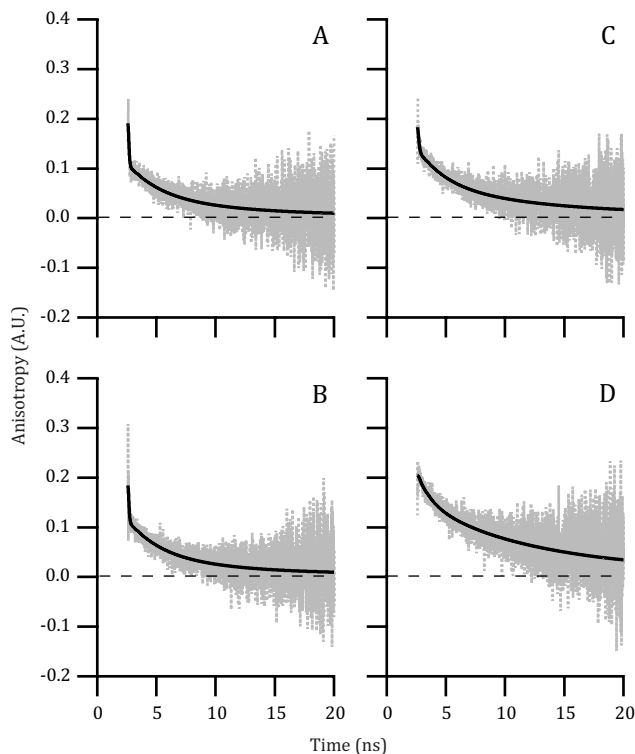


Figure 5 Time-resolved tryptophan fluorescence anisotropy (grey) and associated fits (black) for C69A and F44Y apoflavodoxin in 100 or 10 mM KPPI, respectively. Decays are fitted to $\{\beta_{T1} \exp(-t/\phi_{T1}) + \beta_{T2} \exp(-t/\phi_{T2}) + \beta_r \exp(-t/\phi_r)\}$. The data of C69A apoflavodoxin in 100 and 10 mM KPPI are globally fitted, with linked parameters for ϕ_{T1} , ϕ_{T2} and ϕ_r , and fixed ϕ_{T1} . A, C69A apoflavodoxin in 100 mM KPPI with $\phi_{T1} = 0.04$ ns (fixed) ($\beta_{T1} = 0.09$), $\phi_{T2} = 2.89$ ns ($\beta_{T2} = 0.07$) and $\phi_r = 12.02$ ns ($\beta_r = 0.04$), $\chi^2 = 1.34$. B, C69A apoflavodoxin in 10 mM KPPI with $\phi_{T1} = 0.04$ ns (fixed) ($\beta_{T1} = 0.08$), $\phi_{T2} = 2.89$ ns ($\beta_{T2} = 0.07$) and $\phi_r = 12.02$ ns ($\beta_r = 0.04$), $\chi^2 = 1.34$. The data of F44Y apoflavodoxin in 100 and 10 mM KPPI are globally fitted, with linked parameters for ϕ_{T2} and ϕ_r , and fixed ϕ_{T1} for F44Y apoflavodoxin in 100 mM KPPI. C, F44Y apoflavodoxin in 100 mM KPPI with $\phi_{T1} = 0.07$ ns (fixed) ($\beta_{T1} = 0.05$), $\phi_{T2} = 2.61$ ns ($\beta_{T2} = 0.07$) and $\phi_r = 12.65$ ns ($\beta_r = 0.06$), $\chi^2 = 1.30$. D, F44Y apoflavodoxin in 10 mM KPPI with $\phi_{T1} = 0.84$ ns, ($\beta_{T1} = 0.04$), $\phi_{T2} = 2.61$ ns ($\beta_{T2} = 0.04$) and $\phi_r = 12.65$ ns ($\beta_r = 0.13$), $\chi^2 = 1.30$.

in 10 mM KPPI gives rise to different behaviour (Fig. 5D), because now the protein molecules are predominantly MG. Examples of the analysis of fluorescence and anisotropy decays are shown in Fig. S1 and S2. In addition, we verified that the steady-state anisotropies derived from time-resolved fluorescence anisotropy data overlap with previously published ones (Fig. S3) (39).

In native protein, the three tryptophans of apoflavodoxin (i.e., W74, W128 and W167) are positioned and oriented in such a manner that they yield a surprisingly low steady-state anisotropy of 0.04 (32), instead of a commonly expected value of 0.1 to 0.2 (37). This low value is due to Förster resonance energy transfer (FRET) between tryptophans. In native apoflavodoxin two major energy transfers can be detected, i.e., from W167 to W128 and from W74 to W167, with W128 acting as a sink for FRET (37). As a result of the sink, uni-directional FRET with a 0.04 ns transfer correlation time ϕ_{τ_1} occurs from W167 to W128 (Fig. 5A,B).

Upon going from native F44Y apoflavodoxin to off-pathway MG, the steady-state fluorescence anisotropy increases from 0.04 to 0.08 (39). This increase is due to two phenomena. Firstly, a change in solvent accessibility of tryptophans shortens the average fluorescence lifetime thereby increasing steady-state fluorescence anisotropy (Eq. 15). Secondly, the fast rotation correlation time increases by a factor of about ten (i.e., from 0.07 to 0.84 ns, Fig. 5C,D), which is caused by a change in the positioning and/or orientation of the W167/W128 FRET pair, which decreases FRET and increases fluorescence anisotropy (eq. 10).

Average fluorescence lifetime reveals the fraction of natively folded F44Y apoflavodoxin. To determine the fraction of natively folded F44Y apoflavodoxin at various KPPI concentrations, and thereby establish the ratio of MG to native protein, we use the change in average fluorescence lifetime upon formation of molten globular apoflavodoxin molecules. This lifetime can be derived from the time-resolved fluorescence anisotropy data by using the following equation:

$$\langle \tau \rangle = \sum \alpha_i \cdot \tau_i \quad \text{Eq. 16}$$

in which τ_i are the fluorescence lifetimes and α_i are the corresponding amplitudes (Eq. 8), with the sum of α 's equaling one. The average fluorescence lifetime decreases as the MG state becomes populated, because in the MG of apoflavodoxin the tryptophans are solvent exposed, causing quenching of tryptophan fluorescence.

To correlate the average fluorescence lifetime with the relative fraction of native apoflavodoxin molecules, one requires values for the presence of 100 and 0 % native apoprotein, respectively. The for C69A apoflavodoxin (i.e., 4.13 ± 0.01 ns) is used as reference for the presence of 100 % native apoflavodoxin molecules. To obtain the reference value for the presence of 0 % native apoflavodoxin molecules (i.e., 100 % MG) we extrapolated the average fluorescence lifetime of F44Y apoflavodoxin as a function of KPPI concentration to 0 mM KPPI (i.e., 3.02 ± 0.04 ns). For extrapolation we used a sigmoidal fit, because the equilibrium (un)folding of the off-pathway MG of apoflavodoxin has a sigmoidal transition (43).

Table 3 and Fig. 6 (red squares) show the calculated fraction of natively folded F44Y apoflavodoxin molecules as a function of KPPI concentration. For protein in 100 mM KPPI the fraction of natively folded molecules is calculated to be about

61 %, whereas at 10 mM KPPi this fraction decreases to approximately 5 %. These numbers are consistent with previous estimations of about 88 and 2 % of native F44Y apoflavodoxin in 100 and 10 mM KPPi, respectively (39).

	C69A		F44Y						
mM KPPi	100	10	100	50	40	30	20	10	0
$\langle \tau \rangle$ (ns)	4.12	4.14	3.70	3.35	3.24	3.16	3.07	3.07	3.02*
% Apo	100	100	61	30	20	13	5	5	0
% MG	0	0	39	70	80	87	95	95	100

Table 3 Average fluorescence lifetime ($\langle \tau \rangle$) values of C69A and F44Y apoflavodoxin and derived percentages of native and molten globular apoprotein.

* Extrapolated value from a sigmoidal fit of the dependence of $\langle \tau \rangle$ values of F44Y apoflavodoxin on KPPi concentration.

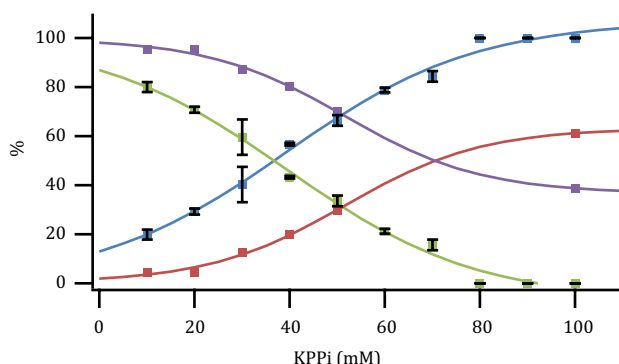


Figure 6 Comparison of the percentage of F44Y apoprotein molecules that bind FMN swiftly with the fraction of natively folded apoflavodoxin molecules. The average fluorescence lifetime has been used to derive the fraction natively folded apoflavodoxin molecules (red), see main text. The corresponding fraction molten globular apoprotein is shown in purple. The percentage of molecules that bind FMN rapidly (blue) is derived from fitting equations 1 to 6 to FMN binding traces (i.e., based on the determination of the ratio between MG and native apoflavodoxin). The corresponding percentage of molecules that bind FMN slowly (i.e., related to the presence of the off-pathway MG of apoflavodoxin) is shown in green. To calculate the percentage of apoprotein molecules that rapidly bind FMN the average of two independent data sets is used, and error bars show the standard deviations.

A folding intermediate concurrent with the off-pathway MG exists at physiological ionic strength. FMN only binds to native apoflavodoxin and under the conditions used this binding happens much faster than the time it takes to unfold the off-pathway MG. Using the ratio of MG to native apoflavodoxin derived from the average fluorescence lifetime (Table 3; Fig. 6), we fitted the FMN binding curves of F44Y apoflavodoxin in 10 to 70 mM KPPi (Fig. 4) to equations 2 to 6. Fixing the ratio between MG and native apo in these equations allows one to determine $k_{N(Apo) \rightarrow MG}$

and $k_{MG \rightarrow N(Apo)}$, i.e., the rates for folding and unfolding of the MG, respectively. Initial

guesses for these rates were obtained from a COPASI 4.16 (build 104) model, based on equations 2 to 6 and values for k_{off} and k_{on} of $9.65 \cdot 10^{-5} \text{ s}^{-1}$ and $3.45 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$, respectively (Table 1). As stated, we derived that the fraction of native apoprotein mol-

ecules in case of F44Y apoflavodoxin in 10 mM KPPI is approximately 5 %. When constraining the fit of the corresponding FMN binding curve to a ratio of native apoflavodoxin to MG of 0.047 ratio, no adequate fit of the data could be obtained (Fig. 7A). Fitting of this FMN binding curve with no constraints to the fraction of native apoflavodoxin shows that approximately 20 % of protein molecules bind FMN rapidly (Fig. 7B).

For KPPI concentrations between 10 and 70 mM, the percentage of molecules that binds FMN swiftly is consistently higher than anticipated from the amount of native apoflavodoxin molecules derived from the average fluorescence lifetime (Fig. 6; compare blue and red squares respectively). For KPPI concentrations of 80 mM and above, it is not possible to fit the FMN binding data taking an off-pathway MG fraction into account, and hence these data sets are fitted considering only native apoflavodoxin (equations 1 to 4). The FMN binding data show that the fraction of molecules that binds FMN swiftly is on average 30 % higher than anticipated from the amount of native apoflavodoxin molecules derived from the values (Fig. 6).

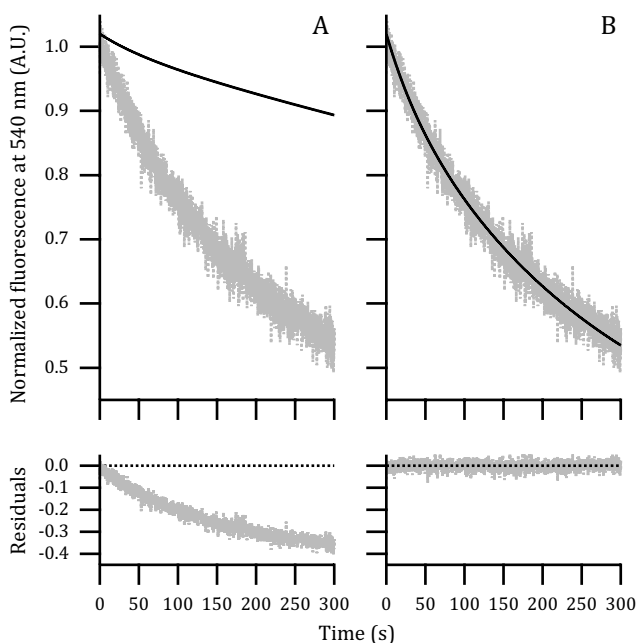


Figure 7 The FMN binding trace of F44Y apoflavodoxin in 10 mM KPPI cannot be properly fitted upon taking into account the percentage of native apoflavodoxin molecules as derived from tryptophan fluorescence anisotropy. Depicted are normalized FMN fluorescence (grey), fits of the data (black) and corresponding residuals. Shown is the FMN binding curve of 25 nM FMN that is added to 71 nM F44Y apoflavodoxin in 10 mM KPPI. *A*, data fitted with a ratio of native apoflavodoxin to MG of 0.047, which corresponds to 5 % natively folded protein. *B*, data fitted with a ratio of native apoflavodoxin to MG of 0.306, which corresponds to 20 % natively folded protein. All other fitting parameters are fixed (see Table 1 and 2).

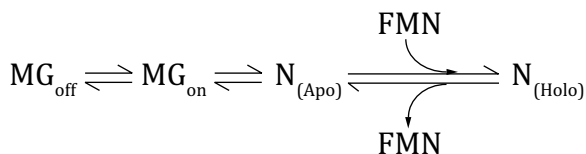
To explain the discrepancy, we propose that besides MG_{off} another folding intermediate of F44Y apoflavodoxin co-exists at equilibrium. The latter intermediate is

present at KPPi concentrations ranging from 10 to 100 mM, which is equivalent to 75 to 345 mM NaCl (i.e., similar to physiological ionic strengths). This intermediate interconverts rapidly with native apoflavodoxin, as opposed to the non-native, misfolded off-pathway MG. This newly discovered intermediate is most probably on-pathway during productive folding of apoflavodoxin.

DISCUSSION

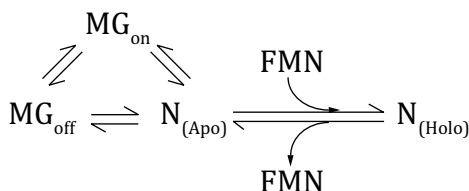
The discovered on-pathway intermediate of apoflavodoxin is molten globular. Virtually all F44Y apoflavodoxin molecules in 10 mM KPPi are molten globular as Fig. 6 (purple squares) implies. About 20 % of the molecules is on-pathway intermediate (blue squares) and the remaining 80 % of protein molecules form the well-known off-pathway MG of apoflavodoxin (green squares). The structural features of F44Y apoflavodoxin under these conditions have been probed by thermal denaturation, tryptophan fluorescence, far-UV circular dichroism and steady-state tryptophan fluorescence anisotropy (39) and hydrogen/deuterium exchange (41). These techniques show that the protein is flexible with solvent-exposed tryptophans, and is molten globular, which thus also applies for the on-pathway intermediate discovered here. The aforementioned methodologies are unsuitable to further structurally characterize the on-pathway MG (MG_{on}), because these techniques cannot distinguish the structural features of two simultaneously occurring MGs. As MG_{on} rapidly interconverts with native apoflavodoxin, it most likely has structural features of native apoprotein and possibly possesses its anti-parallel β -sheet.

Hence, to accurately fit the FMN binding data we would need to introduce additional terms in the differential equations used to describe the folding behaviour of MG_{on} . Furthermore, it is unknown whether MG_{on} is on route between MG_{off} and native apoflavodoxin, as shown in Scheme 6:



Scheme 6

or whether these three folding species are interconnected like in Scheme 7:



Scheme 7

The first scenario would require taking into account two additional rate constants, whereas the second scenario would require four extra ones. As these rate constants are unknown, it would be impossible to verify their accuracy when fitted.

MG on-pathway to native apoflavodoxin. During folding of apoflavodoxin approximately 90 % of molecules first form MG_{off} which subsequently unfolds and

then embarks on the productive folding route to native protein. The other 10 % of folding apoflavodoxin molecules directly follow the productive folding route (Fig. 2) (32). Using single-molecule FRET we demonstrated that all folding molecules convert into MG protein before native apoflavodoxin is formed (43). Our current findings show that molten globular apoflavodoxin contains a minor fraction of molecules that interconvert rapidly to native apoflavodoxin. This fraction is comparable to the 10 % of molecules in kinetic refolding experiments that directly follow the productive folding route. Therefore, we hypothesize that these molecules are the on-pathway MG discovered in this study.

Observations of distinct MGs at equilibrium under physiological conditions. Investigation of MGs is important to improve our understanding of protein folding and aggregation and hence considerable effort has been put into the study of MGs of various proteins. A handful of these studies demonstrated existence of distinct MGs under (semi-)physiological conditions.

In case of α -lactalbumins, MGs can be formed either by acid denaturation or by removal of Ca^{2+} ions. Each method results in MGs that share many secondary structure elements, yet they do exhibit differences in their respective overall structure. These MG species seem to have native-like topology (15,16). Other proteins that form multiple MGs are cytochrome c and apomyoglobin. Interestingly, each of these proteins forms two distinct MGs at equilibrium in acidic conditions, with the ratio between both MGs dictated by salt concentration. For cytochrome c the two MGs form at pH 2 and have a highly native α -helical content (11,13). Apomyoglobin forms two MGs at pH 4. These MGs have near-native α -helical content (12,14). The MGs of all aforementioned proteins are on-pathway and are situated sequentially on one route towards native protein. Barstar however, forms two on-pathway MGs that are not on the same folding pathway to native protein (65). To our knowledge, the study we present here is the first description of the existence of distinct, simultaneously present MGs that reside on folding routes of decidedly different nature (i.e., on- and off-pathway).

Formation of MGs in a cellular context. Several chaperones, such as α -crystallin and GroEL, have been shown to interact with MGs. α -Crystallin distinguishes between MG states of α -lactalbumin, and preferentially interacts with MGs on route to aggregation and precipitation (66). GroEL, sometimes in combination with GroES, recognizes MGs and may stimulate their productive folding, though the exact mechanisms by which it does this are still not completely understood (67-69). In addition, the ribosome seems to have a protein folding activity through Domain V rRNA (70). This folding activity has been shown to promote productive folding from denatured and MG states (71-73). As both MGs of F44Y apoflavodoxin arise under physiological conditions, these MGs are promising candidates for the investigation of how a cell deals with their potential intracellular formation.

ACKNOWLEDGMENTS

We thank Rob Koehorst for assistance in acquiring steady-state fluorescence anisotropy data. We thank Simon Lindhoud and Willem van Berkel for critical reading of the manuscript. This work was supported by an ECHO (711.011.007) grant of NWO to C.v.M.

REFERENCES

1. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnel, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195
2. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19
3. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. *Science* **338**, 1042-1046
4. Wolynes, P. (2016) BIOMOLECULAR FOLDING. Moments of excitement. *Science* **352**, 150-151
5. Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: polypeptide conformations on competing pathways. *Arch. Biochem. Biophys.* **469**, 100-117
6. Ohgushi, M., and Wada, A. (1983) 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21-24
7. Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* **6**, 87-103
8. Dobson, C. M. (2003) Protein folding and misfolding. *Nature* **426**, 884-890
9. Kuwajima, K., Nitta, K., Yoneyama, M., and Sugai, S. (1976) Three-state denaturation of α -lactalbumin by guanidine hydrochloride. *J. Mol. Biol.* **106**, 359-373
10. Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S., and Ptitsyn, O. B. (1981) α -Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett.* **136**, 311-315
11. Kuroda, Y., Kidokoro, S., and Wada, A. (1992) Thermodynamic characterization of cytochrome c at low pH. Observation of the molten globule state and of the cold denaturation process. *J. Mol. Biol.* **223**, 1139-1153
12. Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995) Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5446-5450
13. Colón, W., and Roder, H. (1996) Kinetic intermediates in the formation of the cytochrome c molten globule. *Nat. Struct. Biol.* **3**, 1019-1025
14. Jamin, M., and Baldwin, R. L. (1998) Two forms of the pH 4 folding intermediate of apomyoglobin. *J. Mol. Biol.* **276**, 491-504
15. Arai, M., and Kuwajima, K. (2000) Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209-282
16. Mok, K. H., Nagashima, T., Day, I. J., Hore, P. J., and Dobson, C. M. (2005) Multiple subsets of side-chain packing in partially folded states of α -lactalbumins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8899-8904
17. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., and Razgulyaev, O. I. (1990) Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **262**, 20-24
18. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544-1548

19. Lakshmikanth, G. S., Sridevi, K., Krishnamoorthy, G., and Udgaonkar, J. B. (2001) Structure is lost incrementally during the unfolding of barstar. *Nat. Struct. Biol.* **8**, 799-804
20. Bollen, Y. J. M., and van Mierlo, C. P. M. (2005) Protein topology affects the appearance of intermediates during the folding of proteins with a flavodoxin-like fold. *Biophys. Chem.* **114**, 181-189
21. Baldwin, R. L., and Rose, G. D. (2013) Molten globules, entropy-driven conformational change and protein folding. *Curr. Opin. Struct. Biol.* **23**, 4-10
22. Bhattacharyya, S., and Varadarajan, R. (2013) Packing in molten globules and native states. *Curr. Opin. Struct. Biol.* **23**, 11-21
23. Elms, P. J., Chodera, J. D., Bustamante, C., and Marqusee, S. (2012) The molten globule state is unusually deformable under mechanical force. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3796-3801
24. van der Goot, F. G., Gonzalez-Manas, J. M., Lakey, J. H., and Pattus, F. (1991) A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature* **354**, 408-410
25. Ren, J. H., Kachel, K., Kim, H., Malenbaum, S. E., Collier, R. J., and London, E. (1999) Interaction of diphtheria toxin T domain with molten globule-like proteins and its implications for translocation. *Science* **284**, 955-957
26. Benke, S., Roderer, D., Wunderlich, B., Nettels, D., Glockshuber, R., and Schuler, B. (2015) The assembly dynamics of the cytolytic pore toxin ClyA. *Nat. Comm.* **6**, 6198
27. Jaenicke, R., and Seckler, R. (1997) Protein misassembly *in vitro*. *Adv. Protein Chem.* **50**, 1-59
28. Alagaratnam, S., van Pouderoyen, G., Pijning, T., Dijkstra, B. W., Cavazzini, D., Rossi, G. L., van Dongen, W. M. A. M., van Mierlo, C. P. M., van Berkel, W. J. H., and Canters, G. W. (2005) A crystallographic study of Cys69Ala flavodoxin II from *Azotobacter vinelandii*: Structural determinants of redox potential. *Protein Sci.* **14**, 2284-2295
29. Caetano-Anolles, G., Kim, H. S., and Mittenthal, J. E. (2007) The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9358-9363
30. Fernández-Recio, J., Genzor, C. G., and Sancho, J. (2001) Apoflavodoxin folding mechanism: an α/β protein with an essentially off-pathway intermediate. *Biochemistry* **40**, 15234-15245
31. Apiyo, D., and Wittung-Stafshede, P. (2002) Presence of the cofactor speeds up folding of *Desulfovibrio desulfuricans* flavodoxin. *Protein Sci.* **11**, 1129-1135
32. Bollen, Y. J. M., Sanchéz, I. E., and van Mierlo, C. P. M. (2004) Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin. *Biochemistry* **43**, 10475-10489
33. Kathuria, S. V., Day, I. J., Wallace, L. A., and Matthews, C. R. (2008) Kinetic traps in the folding of beta alpha-repeat proteins: CheY initially misfolds before accessing the native conformation. *J. Mol. Biol.* **382**, 467-484
34. Hills, R. D., Jr., Kathuria, S. V., Wallace, L. A., Day, I. J., Brooks, C. L., 3rd, and

Matthews, C. R. (2010) Topological frustration in beta alpha-repeat proteins: Sequence diversity modulates the conserved folding mechanisms of alpha/beta/alpha sandwich proteins. *J. Mol. Biol.* **398**, 332-350

35. Bollen, Y. J. M., Kamphuis, M. B., and van Mierlo, C. P. M. (2006) The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4095-4100

36. Engel, R., Westphal, A. H., Huberts, D. H., Nabuurs, S. M., Lindhoud, S., Visser, A. J., and van Mierlo, C. P. (2008) Macromolecular crowding compacts unfolded apoflavodoxin and causes severe aggregation of the off-pathway intermediate during apoflavodoxin folding. *J. Biol. Chem.* **283**, 27383-27394

37. Visser, N. V., Westphal, A. H., van Hoek, A., van Mierlo, C. P. M., Visser, A. J. W. G., and van Amerongen, H. (2008) Tryptophan-tryptophan energy migration as a tool to follow apoflavodoxin folding. *Biophys. J.* **95**, 2462-2469

38. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2008) Extensive formation of off-pathway species during folding of an α - β parallel protein is due to docking of (non)native structure elements in unfolded molecules. *J. Am. Chem. Soc.* **130**, 16914-16920

39. Nabuurs, S. M., Westphal, A. H., aan den Toorn, M., Lindhoud, S., and van Mierlo, C. P. M. (2009) Topological switching between an α - β parallel protein and a remarkably helical molten globule. *J. Am. Chem. Soc.* **131**, 8290-8295

40. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2009) Noncooperative formation of the off-pathway molten globule during folding of the α - β parallel protein apoflavodoxin. *J. Am. Chem. Soc.* **131**, 2739-2746

41. Nabuurs, S. M., and van Mierlo, C. P. M. (2010) Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of α - β parallel protein flavodoxin. *J. Biol. Chem.* **285**, 4165-4172

42. Lindhoud, S., Westphal, A. H., Borst, J. W., and van Mierlo, C. P. M. (2012) Illuminating the off-pathway nature of the molten globule folding intermediate of an α - β parallel protein. *PLoS ONE* 10.1371/journal.pone.0045746

43. Lindhoud, S., Pirchi, M., Westphal, A. H., Haran, G., and van Mierlo, C. P. M. (2015) Gradual folding of an off-pathway molten globule detected at the single-molecule level. *J. Mol. Biol.* **427**, 3148-3157

44. Bollen, Y. J. M., Nabuurs, S. M., van Berkel, W. J. H., and van Mierlo, C. P. M. (2005) Last in, first out. The role of cofactor binding in flavodoxin folding. *J. Biol. Chem.* **280**, 7836-7844

45. Nabuurs, S. M., de Kort, B. J., Westphal, A. H., and van Mierlo, C. P. M. (2010) Non-native hydrophobic interactions detected in unfolded apoflavodoxin by paramagnetic relaxation enhancement. *Europ. Biophys. J.* **39**, 689-698

46. van Mierlo, C. P. M., van Dongen, W. M. A. M., Vergeldt, F., van Berkel, W. J. H., and Steensma, E. (1998) The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate. *Protein Sci.* **7**, 2331-2344

47. van Mierlo, C. P. M., van den Oever, J. M. P., and Steensma, E. (2000) Apoflavodoxin (un)folding followed at the residue level by NMR. *Protein Sci.* **9**, 145-157

48. Edmondson, D. E., and Tollin, G. (1971) Flavin-protein interactions and the redox properties of the Shethna flavoprotein. *Biochemistry* **10**, 133-145
49. Taylor, M. F., Boylan, M. H., and Edmondson, D. E. (1990) *Azotobacter vinelandii* flavodoxin: purification and properties of the recombinant, dephospho form expressed in *Escherichia coli*. *Biochemistry* **29**, 6911-6918
50. Bollen, Y. J. M., Westphal, A. H., Lindhoud, S., van Berkel, W. J. H., and van Mierlo, C. P. M. (2012) Distant residues mediate picomolar binding affinity of a protein cofactor. *Nature Comm.* 10.1038/ncomms2010
51. Steensma, E., Heering, H. A., Hagen, W. R., and Van Mierlo, C. P. M. (1996) Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur. J. Biochem.* **235**, 167-172
52. Houwman, J. A., Westphal, A. H., van Berkel, W. J., and van Mierlo, C. P. (2015) Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome. *BBA-Proteins Proteom.* **1854**, 1317-1324
53. Borst, J. W., Hink, M. A., van Hoek, A., and Visser, A. J. (2005) Effects of refractive index and viscosity on fluorescence and anisotropy decays of enhanced cyan and yellow fluorescent proteins. *J. Fluoresc.* **15**, 153-160
54. Visser, N. V., Visser, A. J. W. G., Konc, T., Kroh, P., and van Hoek, A. (1994) New reference compound with single, ultrashort lifetime for time-resolved tryptophan experiments. in *Time-resolved laser spectroscopy in biochemistry IV* (Lakowicz, J. R. ed.), Proc. SPIE, Los Angeles, USA. pp 618-626
55. Laptinok, S. P., Visser, N. V., Engel, R., Westphal, A. H., van Hoek, A., van Mierlo, C. P., van Stokkum, I. H., van Amerongen, H., and Visser, A. J. (2011) A general approach for detecting folding intermediates from steady-state and time-resolved fluorescence of single-tryptophan-containing proteins. *Biochemistry* **50**, 3441-3450
56. Tanaka, F., and Mataga, N. (1979) Theory of time-dependent photoselection in interacting fixed systems. *Photochem. Photobiol.* **29**, 1091-1097
57. Nunthaboot, N., Tanaka, F., Kokpol, S., Visser, N. V., van Amerongen, H., and Visser, A. J. W. G. (2014) Molecular dynamics simulation of energy migration between tryptophan residues in apoflavodoxin. *RSC Advances* **4**, 31443-31451
58. Förster, T. (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55-75
59. Knox, R. S., and van Amerongen, H. (2002) Refractive index dependence of the Förster resonance excitation transfer rate. *J. Phys. Chem. B* **106**, 5289-5293
60. Toptygin, D., Savtchenko, R. S., Meadow, N. D., Roseman, S., and Brand, L. (2002) Effect of the solvent refractive index on the excited-state lifetime of a single tryptophan residue in a protein. *J. Phys. Chem. B* **106**, 3724-3734
61. Engelborghs, Y. (2003) Correlating protein structure and protein fluorescence. *J. Fluoresc.* **13**, 9-16
62. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, 3rd ed., Springer US, New York
63. Shiga, K., and Tollin, G. (1974) Flavine-protein interactions in flavoenzymes. Effects of aggregation of the apoprotein of *Azotobacter* flavodoxin on coenzyme

binding. *Biochemistry* **13**, 3268-3273

64. Gast, R., Valk, B. E., Muller, F., Mayhew, S. G., and Veeger, C. (1976) Studies on the binding of FMN by apoflavodoxin from *Peptostreptococcus elsdenii*, pH and NaCl concentration dependence. *Biochim. Biophys. Acta* **446**, 463-471

65. Shastry, M. C., and Udgaonkar, J. B. (1995) The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates. *J. Mol. Biol.* **247**, 1013-1027

66. Lindner, R. A., Treweek, T. M., and Carver, J. A. (2001) The molecular chaperone α -crystallin is in kinetic competition with aggregation to stabilize a monomeric molten-globule form of α -lactalbumin. *Biochem. J.* **354**, 79-87

67. Moparthy, S. B., Sjölander, D., Villebeck, L., Jonsson, B. H., Hammarström, P., and Carlsson, U. (2013) Transient conformational remodeling of folding proteins by GroES-individually and in concert with GroEL. *J. Chem. Biol.* **7**, 1-15

68. Haldar, S., Gupta, A. J., Yan, X., Miličić, G., Hartl, F. U., and Hayer-Hartl, M. (2015) Chaperonin-assisted protein folding: Relative population of asymmetric and symmetric GroEL:GroES complexes. *J. Mol. Biol.* **427**, 2244-2255

69. Hayer-Hartl, M., Bracher, A., and Hartl, F. U. (2016) The GroEL-GroES chaperonin machine: A nano-cage for protein folding. *Trends Biochem. Sci.* **41**, 62-76

70. Voisset, C., Saupe, S. J., and Blondel, M. (2011) The various facets of the protein-folding activity of the ribosome. *Biotechnol. J.* **6**, 668-673

71. Pathak, B. K., Mondal, S., Ghosh, A. N., and Barat, C. (2014) The ribosome can prevent aggregation of partially folded protein intermediates: studies using the *Escherichia coli* ribosome. *PLoS ONE* 10.1371/journal.pone.0096425

72. Chakraborty, B., Bhakta, S., and Sengupta, J. (2016) Mechanistic insight into the reactivation of BCAII enzyme from denatured and molten globule states by eukaryotic ribosomes and Domain V rRNAs. *PLoS ONE* 10.1371/journal.pone.0153928

73. Chakraborty, B., Bhakta, S., and Sengupta, J. (2016) Disassembly of yeast 80S ribosomes into subunits is a concerted action of ribosome-assisted folding of denatured protein. *Biochem. Biophys. Res. Commun.* **469**, 923-929

74. Steensma, E., and van Mierlo, C. P. M. (1998) Structural characterisation of apoflavodoxin shows that the location of the stable nucleus differs among proteins with a flavodoxin-like topology. *J. Mol. Biol.* **282**, 653-666

SUPPLEMENTARY INFORMATION

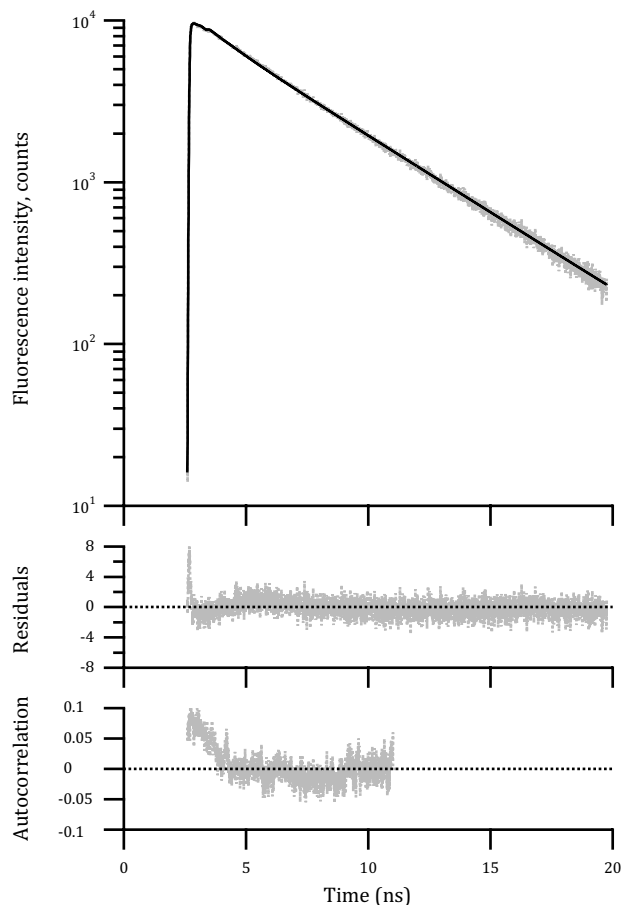


Figure S1. Total fluorescence decay (grey) and associated fit (black) obtained for 2 μM F44Y apoflavodoxin in 100 mM KPPI, pH 6.0, 25 $^{\circ}\text{C}$. The excitation wavelength is 300 nm and the emission is measured at 348.8 nm. To achieve an optimal fit to the experimental data a sum of three exponential terms ($N=3$) is required. The optimized lifetimes and relative amplitudes (percentage in parentheses) obtained are: $\tau_1 = 0.114$ ns (5 %), $\tau_2 = 1.496$ ns (21 %) and $\tau_3 = 4.545$ ns (74 %), the goodness-of-fit parameter $\chi^2 = 1$. The average fluorescence lifetime is determined to be 3.69 ns. To illustrate the quality of the fit the weighted residuals and autocorrelation function of the residuals are shown as well.

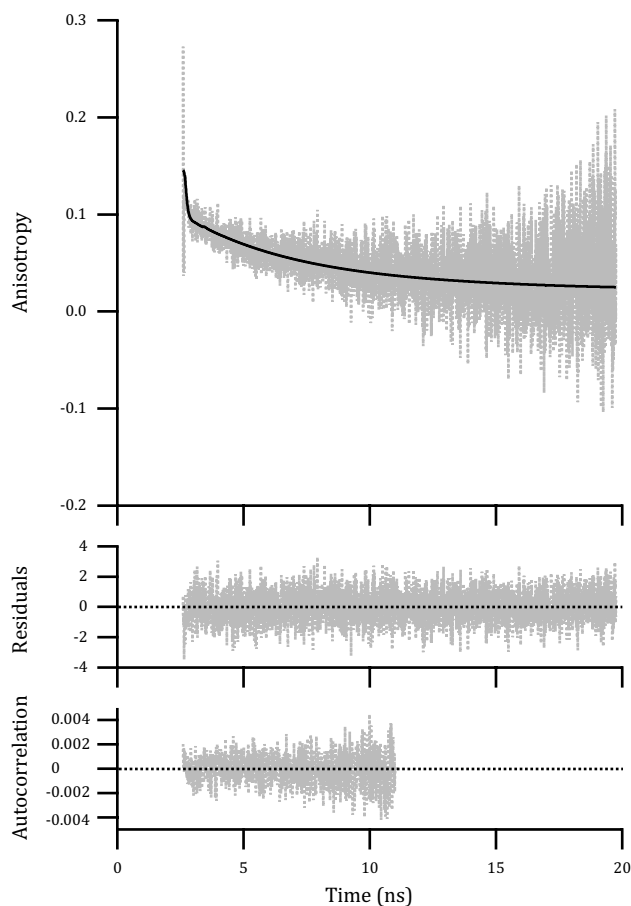


Figure S2. Fluorescence anisotropy decay (grey) and associated fit (black) obtained for 2 μM F44Y apoflavodoxin in 100 mM KPPI, pH 6.0, 25 $^{\circ}\text{C}$. The excitation wavelength is 300 nm and the emission is measured at 348.8 nm. The optimized parameters are: 0.07 ns (fixed) (0.05), 2.61 ns (0.07) and 12.65 ns (0.06), the goodness-of-fit parameter $\chi^2 = 1.30$. To illustrate the quality of the fit the weighted residuals and autocorrelation function of the residuals are shown as well.

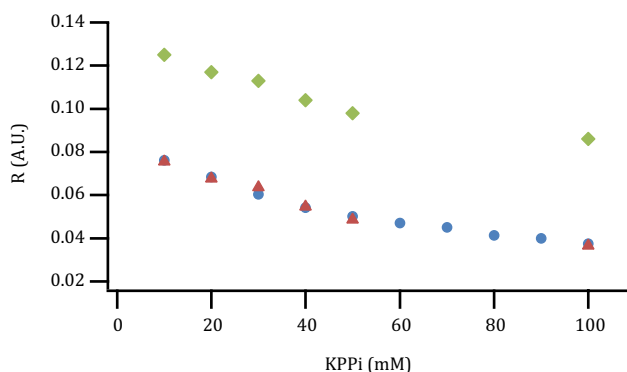


Figure S3. Steady-state tryptophan fluorescence anisotropy (R) shows that when KPPi concentration is decreased, F44Y apoflavodoxin forms a MG. Comparison between steady-state fluorescence anisotropy as measured by Nabuurs *et al.* on a Fluorolog 3.2.2 fluorometer (blue circles (39)) and values derived in this study from time-resolved anisotropy measurements using a time-correlated single photon counting device (green diamonds). As the anisotropies are recorded on set-ups with differences in excitation/emission wavelengths, the absolute anisotropy values differ. Upon correcting time-resolved data for this difference in offset (red triangles), both data sets overlap.

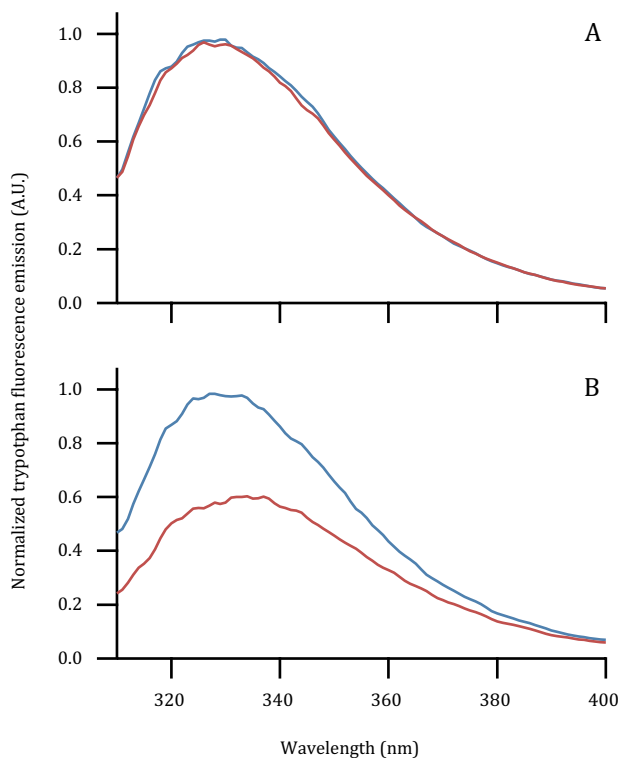
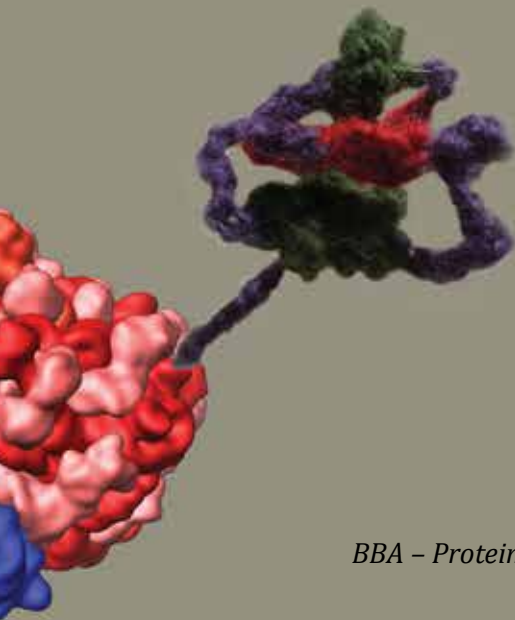


Figure S4. Tryptophan fluorescence emission of F44Y apoflavodoxin decreases with decreasing salt concentration, whereas it remains the same for C69A apoflavodoxin. The emission spectra obtained at 100 mM KPPi (blue) and 10 mM KPPi (red) are normalized to the emission obtained at 100 mM KPPi. *A*, the emission spectrum of C69A apoflavodoxin at 10 mM KPPi is identical to the corresponding spectrum at 100 mM KPPi. The tryptophans remain in a similar micro-environment and are not solvent exposed. *B*, A similar decrease in salt concentration in case of F44Y apoflavodoxin leads to a reduction in fluorescence emission. This observation is due to formation of MGs at low salt concentration. The tryptophans in these species are exposed to the solvent and have increased flexibility as compared to native apoflavodoxin. Concentrations of C69A and F44Y apoflavodoxin are 11.2 μM and 5.7 μM , respectively. Spectra are recorded at 25 $^{\circ}\text{C}$.

Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome

Joseline A. Houwman, Adrie H. Westphal, Willem J.H. van Berkel and Carlo
P.M. van Mierlo

Laboratory of Biochemistry, Wageningen University, the Netherlands



Correct folding of proteins is crucial for cellular homeostasis. More than thirty percent of proteins contain one or more cofactors, but the impact of these cofactors on cotranslational folding remains largely unknown. Here, we address the binding of flavin mononucleotide (FMN) to nascent flavodoxin, by generating ribosome-arrested nascent chains that expose either the entire protein or C-terminally truncated segments thereof. The native α/β parallel fold of flavodoxin is among the most ancestral and widely distributed folds in nature and exploring its cotranslational folding is thus highly relevant. In *Escherichia coli* (strain BL21(DE3) $\Delta tig::kan$) FMN turns out to be limiting for saturation of this flavoprotein on time-scales vastly exceeding those of flavodoxin synthesis. Because the ribosome affects protein folding, apoflavodoxin cannot bind FMN during its translation. As a result, binding of cofactor to released protein is the last step in production of this flavoprotein in the cell. We show that once apoflavodoxin is entirely synthesized and exposed outside the ribosome to which it is stalled by an artificial linker containing the SecM sequence, the protein is natively folded and capable of binding FMN.

INTRODUCTION

Proteins belong to the most important elements of a living cell. For example, they confer stability in the form of the cytoskeleton, act as biological catalysts and transporters, provide immune protection, and are essential in cellular homeostasis (1). An important aspect of homeostasis is proteostasis, which not only encompasses the amount and distribution of proteins, but their correct folding as well (2). Protein folding has been extensively studied *in vitro* and for some proteins folding energy landscapes have been derived. Nowadays, focus has shifted to understanding how proteins fold *in vivo*, which includes folding while a protein is synthesized by the ribosome. Whereas ribosomal structures have been elucidated in atomic detail (see e.g. (3-5)), relatively little is known about the conformational events polypeptide chains undergo while they are produced.

Several differences between *in vitro* and *in vivo* conditions may affect protein folding, such as the dissimilarity between the dilute environment of the test tube and the highly crowded condition in cells (6). In addition, a large fraction of cellular proteins comprise chaperones, which prevent aggregation of newly synthesized proteins (1). Moreover, the vectorial nature of polypeptide synthesis by the ribosome can open up entirely new folding pathways.

Folding can already begin cotranslationally, with only part of the polypeptide chain synthesized, and has been investigated for protein folds like all α (7-9), all β (10-13) and α/β (14-17). Some proteins seem to fold only when they are entirely exposed outside the ribosome (12) or when they are released (18). During synthesis, other proteins can form folding intermediates (8,9,11,13,17,19), and protein subdomains may properly fold while the remainder of the nascent sequence does not (14,16). In addition, the ribosome itself can in principle also modulate nascent chain folding (20).

Few nascent proteins investigated contain a cofactor in their native state. This under-representation is undesirable, considering that at least thirty percent of proteins have such a moiety (21). Cofactors range from metal ions to complexes like hemes, and organic moieties such as flavins. Heme has been suggested to bind to nascent globin (7), which may imply that the protein folds cotranslationally. Flavin adenine dinucleotide (FAD) was proposed to covalently attach to nascent 6-hydroxy-D-nicotine oxidase (22), but its cotranslational flavinylation is now judged unlikely (23).

Most cofactors bind non-covalently, and in case of flavoproteins ninety percent bind flavin in this way (24). Flavoproteins make up to three and a half percent of genes in certain genomes and fulfil many important roles (24). To reveal whether flavin can bind to nascent apo-protein, in this study we utilize a 179-residue flavodoxin from *Azotobacter vinelandii*. Flavodoxin is involved in electron transport in the nitrate reduction cycle. For this purpose it uses a non-covalently, but tightly bound flavin mononucleotide (FMN) (25). The protein consists of five α -helices sandwiching a central parallel β -sheet, which is an α/β parallel topology (26). Approximately twenty-five percent of proteins share this topology, according to SCOP classification in the Protein Data Bank (PDB). The flavodoxin-like fold is among the most ancestral folds (27) and is considered to be an archetype for the entire class of $\alpha\beta\alpha$ sandwiches. Though flavodoxin is absent from plants and animals (28), it is found in most algal lineages (29) and the flavodoxin-like fold is part of many multi-domain eukaryotic

proteins such as cytochrome P450 reductase (30) and nitric oxide synthase (31).

In vitro folding of full-length, isolated flavodoxin involves two intermediates: an on-pathway one, which is quickly converted into natively folded apo-protein, and a molten globule species (32-35). This species is off the productive folding pathway and needs to unfold to embark on a route to native protein (36). The molten globule consists of a loosely packed core and solvent exposed hydrophobic side chains (37-39). For many proteins with an α/β -parallel topology, formation of an off-pathway intermediate seems to be a characteristic feature (40). FMN binds to the flavin-binding site of native apoflavodoxin. Besides affecting this pocket, cofactor incorporation also influences thermodynamic stabilities of residues faraway from the binding site (25).

Here, we report properties of nascent flavodoxin. We prepare ribosome-nascent chain complexes (RNCs) that expose either the entire protein or C-terminally truncated segments of flavodoxin outside the exit tunnel. Use is made of constructs for flavodoxin that lack zero, five or ten amino acid residues at the C-terminus of the stalled polypeptide. This procedure allows the mimicking of late periods during protein translation and enables the elucidation of the stage at which nascent apoflavodoxin is capable of binding FMN. Translational stalling is made possible by attaching a sequence derived from *E. coli* SecM to the C-terminus of a nascent protein (41,42). The RNCs are isolated from *E. coli* strain BL21(DE3) $\Delta tig::kan$, which lacks chaperone Trigger factor. We address FMN binding to nascent flavodoxin by titrating RNCs with FMN or by treating them with trichloroacetic acid (TCA). It turns out that because the ribosome influences protein folding, FMN cannot bind to apoflavodoxin during translation.

MATERIALS AND METHODS

Engineering RNCs. The construct for arrested, nascent protein, containing a triple N-terminal StrepII-tag linked to a Smt3 domain, a multiple cloning site (MCS), a TEV-site, a linker and a SecM stalling motif (Fig. 1A), was a kind gift of Professor Elke Deuerling (University of Konstanz, Germany) (12). Ulp1 protease specifically recognises Smt3 and cleaves the polypeptide down-stream of Smt3, thereby producing nascent flavodoxin chains with Ala-1 as authentic N-terminus, if required. The recognition site for TEV protease enables future release on demand of nascent polypeptide. After the gene for flavodoxin was cloned into the MCS, two deletion variants of nascent protein were produced by using the method described in (43). These variants are shortened at the C-terminus by five or ten amino acids, respectively. To avoid covalent protein dimerization, the single cysteine at position 69 of flavodoxin was replaced by an alanine in all constructs (44). Plasmids were transformed in *E. coli* strain BL21(DE3) $\Delta tig::kan$ for expression of RNCs.

Expression and purification of RNCs. Based on the protocol described in (12), we developed the following procedure. Cells were grown O/N in 50 ml Terrific Broth (TB) with 100 $\mu\text{g}/\text{ml}$ ampicillin at 37 °C while shaking at 140 rpm. Cells were then pelleted at 5,000 g for 10 min at 4 °C, re-suspended in 3 l of fresh TB containing 100 $\mu\text{g}/\text{ml}$ ampicillin and grown until an OD_{600} of 0.8. For full-length construct, cells were pelleted at 11,000 g for 15 min at 4 °C. The pellet was re-suspended in 1 l of minimal (M9) medium with 100 $\mu\text{g}/\text{ml}$ ampicillin and shaken for 1 h at 37 °C at 140 rpm. Induction by 1 mM IPTG lasted 2.5 hours. After reaching an OD_{600} of 0.8, shortened constructs were induced by 1 mM IPTG for 1h in TB. Cells were subsequently pel-

leted at 11,000 g for 15 min at 4 °C and frozen in liquid nitrogen.

Frozen cells were thawed and re-suspended in buffer I (50 mM HEPES-KOH, 100 mM potassium acetate, 15 mM magnesium acetate, 1 mM dithiothreitol (DTT), pH 7.4) with added protease inhibitor cocktail (Roche) and lysed by French Press. After centrifugation at 40,000 g for 30 min at 4 °C, supernatant was filtered using a 0.22 µm filter (EMD Millipore) and loaded onto a Strep-Tactin sepharose (IBA GmbH) column, pre-equilibrated with buffer I. RNCs and released protein were eluted with buffer I containing 2.5 mM desthiobiotin (IBA GmbH). Eluate was concentrated with a 10-kDa spin filter (EMD Millipore) at 5,500 g and loaded onto a Superdex75 column to separate RNCs from released constructs. Both fractions were concentrated, frozen in liquid nitrogen and stored at -80 °C.

Samples of each purification step were analysed by Coomassie-stained SDS-PAGE or Western blotting using either flavodoxin antibody raised in rabbits (Eurogentec) followed by anti-rabbit HRP-IG (Eurogentec) or StrepMAB classic (IBA GmbH) against the N-terminal StrepII-tag.

To determine whether binding of FMN releases nascent chains and thus results in a different Superdex75 profile, the above purification protocol was used with one modification. The filtered cell extract was split in two and 50 µM FMN (Sigma) was added to one half. Both samples were incubated on ice for 1 hour.

FMN titration and TCA precipitation. To determine the extent of FMN occupation of RNCs and released constructs, samples were made after Superdex75 purification that contain 0.5 to 1 µM of RNC or 2 to 3 µM of released protein in buffer I. Samples were then split in two, of which one half was titrated with FMN to determine its FMN-binding capacity using flavin fluorescence. The other half was precipitated by adding 3 % (w/v) TCA to establish the total amount of FMN bound to protein. Precipitate was spun down for 10 min at 4 °C at 21,000 g and flavin fluorescence of the supernatant was subsequently measured. A 3 µM solution of FMN in 3 % (w/v) TCA in buffer I was used as reference.

Flavin fluorescence was measured on a Cary Eclipse spectrophotometer (Varian). Excitation was at 450 nm, with emission recorded between 460 and 625 nm. Each sample was scanned 5 times and the average was used for calculations. Excitation and emission slits were set to a bandwidth of 5 nm. For RNCs the PMT voltage was 900 V, whereas released protein was measured at 800 V.

Preparations of (truncated) apoflavodoxin variants. For production of apo-proteins, the constructs were loaded onto a Superdex75 column containing 6 M guanidine hydrochloride (GuHCl; Fluka) in its upper 20 % volume. As a result, holo-protein unfolds and releases FMN. Upon elution with buffer, apoflavodoxin and FMN separate. Due to migration into buffer without denaturant, apoflavodoxin autonomously folds to native protein on-column, as its unfolding is reversible (45).

RESULTS

Production and purification of RNCs. To assess cofactor binding during translation, we generated stalled nascent chains of flavodoxin. The corresponding constructs (Fig. 1A) enable the tracking of nascent flavodoxin folding outside the ribosome at sequential stages of elongation. Nascent protein lacking none, five or ten C-terminal amino acids of flavodoxin, termed FL, -5 and -10 respectively, was

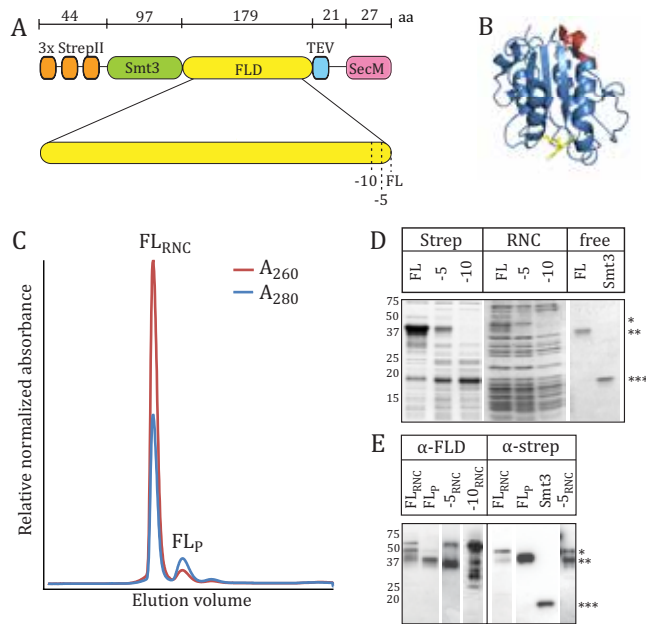


Figure 1 Schematic representation of the construct for production of arrested, nascent flavodoxin and assessment of full-length and truncated RNCs of flavodoxin. **A**, The gene for production of RNCs consists of a triple N-terminal StrepII-tag (orange), an Smt3 domain (green) fused to flavodoxin (yellow), a recognition site for TEV protease (blue) and a flexible hinge (i.e., GSGASGGAS) connected to a linker, which contains the SecM stalling sequence (magenta) and spans the ribosomal exit tunnel. Through deletion of 5 or 10 residues at the C-terminus of flavodoxin, respectively, 2 constructs are obtained that produce stalled nascent proteins of specified length. **B**, Cartoon model of native flavodoxin (PDB ID 1YOB (26)) with FMN coloured yellow. Flavodoxin's C-terminal 10 residues are highlighted in red. **C**, After induction in *E. coli* in M9 minimal medium, cell lysis, and elution from a Strep-Tactin column, samples were loaded onto a Superdex75 column. The Superdex75 elution profile of FL_{RNC} and FL_p, shows absorbance at 260 (red line) and 280 nm (blue line). RNCs elute in the void volume, whereas released protein is retarded, thus achieving separation of both species. **D**, Coomassie stained SDS-PAGE. Lanes marked "Strep" show the mixture of RNCs and released proteins eluted from a Strep-Tactin sepharose column. Lanes marked "RNC" display nascent proteins eluted from a Superdex75 column and illustrate the pattern typical for ribosomal proteins. Lanes marked "free" show released protein (FL) and Smt3-domain with StrepII-tag attached (Smt3). This domain is a degradation product of released protein. In (D) and (E), the position of nascent chain attached to tRNA is labelled (*) and the position of nascent chain without tRNA is marked (**). The position of StrepII-tag attached to Smt3-domain is labelled (***). **E**, Western blots of RNCs and released proteins, with antibodies against flavodoxin (α-FLD) or the N-terminal StrepII-tag (α-strep). The 60-kDa band observed in the RNC samples blotted with α-FLD is due to ribosomal protein that reacts with α-FLD.

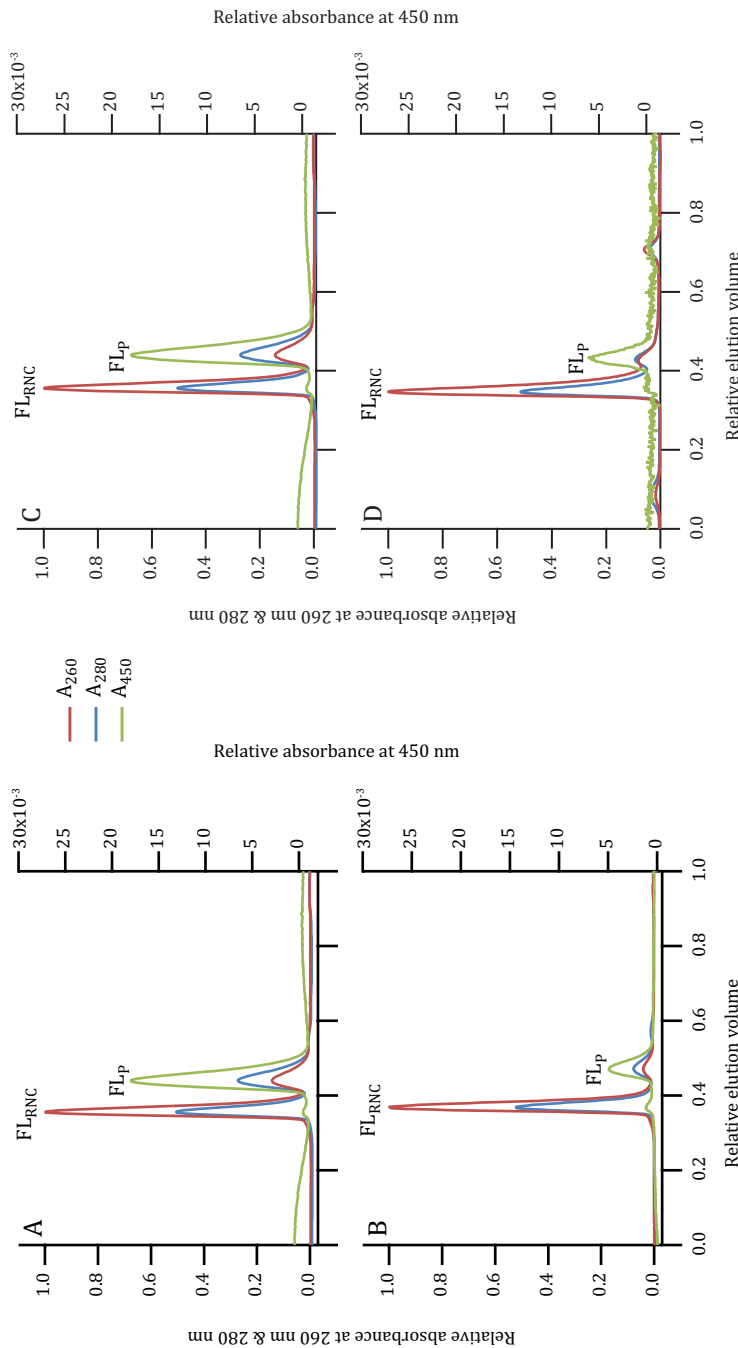


Figure 2 A switch from nutrient rich to minimal medium or from 37° C to 20° C during induction affects cellular release of stalled nascent protein. Shown are Superdex75 elution profiles, with all absorbance traces normalized relative to the maximum absorbance at 260 nm of FL_{RNC}. Absorbance at 450 nm tracks FMN incorporation. Elution volumes are normalized to the total column volume V_t. A, FL_p produced in *E. coli* in Terrific Broth after an induction period of 1 hour at 37 °C. B, FL_p produced in *E. coli* in M9 minimal medium after an induction period of 2.5 hours at 37 °C. Compared to (A) the number of RNCs is unaltered, but release of stalled nascent chains has decreased considerably. C, FL_p produced in *E. coli* in Terrific Broth after an induction period of 1 hour at 37 °C. D, FL_p produced in *E. coli* in Terrific Broth after an induction period of 2.5 hours at 20 °C. Compared to (C), the number of RNCs is unaltered, but release of stalled nascent chains has decreased considerably.

produced. The removed residues are part of the C-terminal α -helix of flavodoxin, which packs onto the central β -sheet of native protein (Fig. 1B). Expression of these constructs in *E. coli* and subsequent purification (see Materials and Methods) succeeded.

The SecM sequence allows for relatively tight stalling of nascent chains (41). However, some release of SecM-stalled nascent chains has been documented both during production (17,46) and from purified RNCs (12,47). The mechanism that overcomes stalling in a cell is not yet known. It has been suggested that physical force plays a role (41,48,49). We observe for all flavodoxin constructs that release of stalled nascent protein happens during production in *E. coli*. As a result, Strep-Tactin column eluates contain RNCs (i.e., FL_{RNC}, -5_{RNC}, -10_{RNC}, respectively) and released protein (i.e., FL_p, -5_p, -10_p, respectively). To separate RNCs from released protein, we employ size-exclusion chromatography (Fig. 1C). The absorbance ratio A_{260}/A_{280} enables distinction between RNCs and released flavodoxin, as this ratio is ~ 2 for RNCs, while A_{260}/A_{280} is <1 for released protein. Fig. 1D shows assessment of the obtained fractions after Strep-Tactin and Superdex75 column chromatography by Coomassie stained SDS-PAGE. The lanes marked "Strep" display ribosomal protein bands arising from RNCs and surplus released protein, i.e., FL_p, -5_p or a 18-kDa band containing the StrepII-tag attached to Smt3, labelled (***). This 18-kDa band is an *in vivo* degradation product of released protein, and its identity is confirmed by blots probed with antibodies against the StrepII-tag (Fig. 1E). As the quantity of released protein is many times more than the number of ribosomes in a cell, release of nascent chains likely occurs during their production and not during their purification. Co-elution of released nascent chains with RNCs was excluded (see Fig. S1 and S2). The lanes marked "RNC" show the typical pattern for ribosomal proteins and the presence of nascent protein, some of which is attached to tRNA. Fig. 1D confirms successful separation of RNCs from released protein for all three constructs. Fig. 1E displays Western blot bands detected with antibodies against StrepII-tag or flavodoxin. The lanes marked " α -FLD" show the presence of flavodoxin in the fractions containing RNCs or released proteins. As mentioned, some nascent chains are coupled to tRNA, which results in observation of two bands approximately 15 kDa apart in the lanes containing RNCs. Because a polyclonal antibody is used to detect flavodoxin, we observe differences in affinity for the constructs used. Highest affinity is observed for FL, and affinity decreases in the order of -5 to -10. The lanes marked " α -Strep" display a similar pattern for RNCs, showing nascent chain with or without tRNA.

Intracellular release of stalled nascent protein depends on growth condition and chain length. Production of protein in *E. coli* at 37 °C in nutrient rich (i.e., TB) or in minimal medium (i.e., M9) leads to similar amounts of FL_{RNC}. However, release of stalled nascent chains is much lower for cells grown in minimal medium (Fig. 2A and B). A similar decrease in release of stalled nascent chains is observed when RNC production in *E. coli* takes place in nutrient rich medium at 20 °C (Fig. 2C and D). These observations suggest that the more favourable the growth conditions are for *E. coli*, the less effective SecM stalling becomes. To compensate for lower synthesis rates during less favourable growth conditions, induction times were increased from 1 hour (for TB at 37 °C) to 2.5 hours (for M9 at 37 °C or TB at 20 °C).

Another factor that influences the efficiency of ribosomal stalling could be the length of the nascent chain. Fig. 3A shows observations for FL construct. It is estimated that about 350 times more FL_p than FL_{RNC} is present, based on absorbance at

280 and 260 nm, respectively. Absorbance at 450 nm reveals that part of FL_p contains FMN. In case of -5 construct we observe less released nascent chains as compared to released FL construct (Fig. 3B). The ratio of -5_p to -5_{RNC} is about 90. In case of -10 construct, we detect no released protein (i.e., -10_p) (Fig. 3C). The -5 and -10 purifications all show separate elution of StreptII-tag attached to Smt3 domain, as verified by Western blotting with antibodies against StreptII-tag. Thus, both nascent chains are rather unstable, because intracellular proteases can cleave off the tagged N-terminal parts of these constructs. This degradation might happen while the protein chain is attached to the ribosome and/or after its release. As we obtain similar amounts of FL_{RNC} , -5_{RNC} and -10_{RNC} , which must have a StreptII-tag coupled to the nascent chain during the first step of purification, degradation likely happens mostly after release of the stalled nascent chain. However, as can be seen in Fig. 1D, the

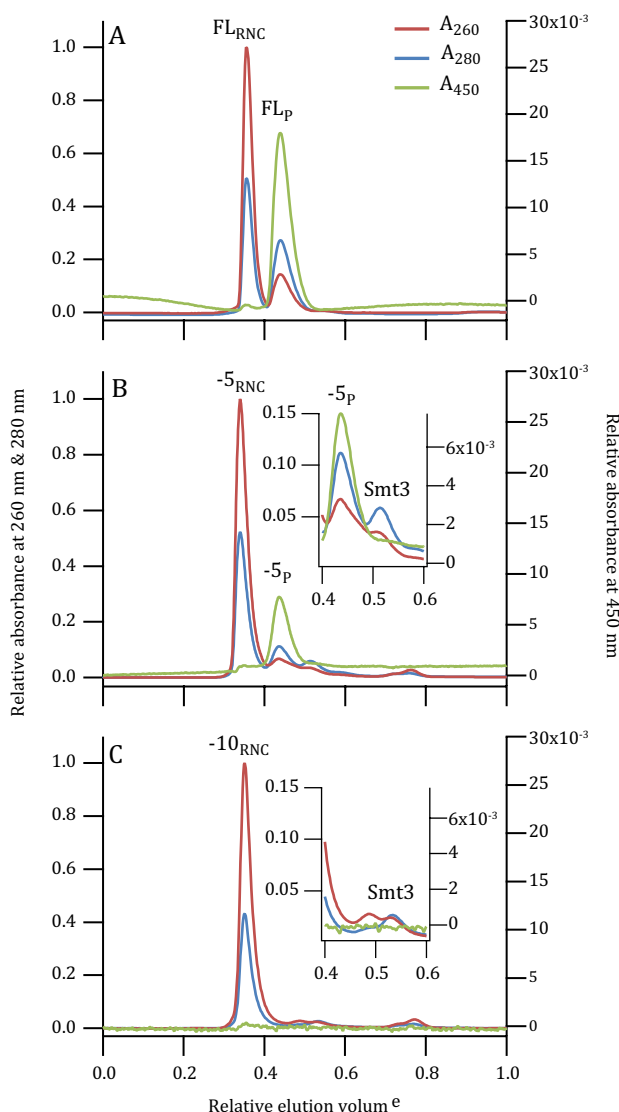


Figure 3 Comparison of release of shortened, stalled nascent chains. Shown are Superdex75 elution profiles with all absorbance traces normalized relative to the maximum absorbance at 260 nm of the respective RNCs. Absorbance at 450 nm tracks FMN incorporation. Elution volumes are normalized to the total column volume V_t . All constructs are induced for 1 hour in v in TB medium. StreptII-tag attached to the Smt3 domain is labelled (Smt3). A, FL construct produces FL_{RNC} and FL_p . B, -5 construct produces -5_{RNC} , -5_p and Smt3 domain. C, -10 construct produces -10_{RNC} and Smt3 domain.

intensity of the nascent chain in the -10_{RNC} lane (at approximately 40 kDa) is lower than the intensities corresponding to ribosomal proteins, indicating that some degradation of -10 nascent chain happens during subsequent Superdex75 purification. After purification we observe no further degradation of nascent chains and released proteins during our experiments, which last only several hours at room temperature. Clearly, the shorter the chain length of the C-terminus of nascent flavodoxin is, the more destabilized the protein becomes.

Released FL_p and -5_p contain FMN and are natively folded. Because we separate RNCs from released protein, a comparison of their cofactor binding capacities can be made. First, we focus on released flavodoxin constructs. FL_p and -5_p both show absorbance at 450 nm (Fig. 3A and B), indicating that FMN is bound. Binding of FMN to native apoflavodoxin occurs primarily through a very specific combination and geometry of hydrogen bonds and aromatic interactions (25,50). Thus, FMN-containing FL_p and -5_p must be natively folded.

To assess remaining FMN binding capacity, we titrated released protein from several purification batches of *E. coli* with FMN (Fig. 4A). Due to quenching of the fluorescence of FMN upon binding to protein, the slope of fluorescence arising from bound FMN is less steep than the slope from free FMN. This characteristic demonstrates FMN binding. Both FL_p and -5_p were evaluated, whereas -10_p was not, because of degradation by proteases in *E. coli*. In case of FL_p, FMN binding varies strongly between batches. Quantification of the amount of bound FMN, derived by precipitating FL_p with TCA and measuring fluorescence of released FMN, also shows considerable variation between batches. FMN content of FL_p ranges between 40 to 100 % of protein population. FMN synthesis is strictly regulated in *E. coli* and is only activated when needed. To make FMN, the enzyme riboflavin kinase converts riboflavin into FMN. Because most of the ribosomes in cells producing stalled nascent chains are occupied, we assume that they synthesize less riboflavin kinase and thus less FMN is made. Small variations during growth may lead to alterations in the amount of riboflavin kinase and thus FMN. As the amount of protein we produce is relatively small, in terms of percentage, variation between batches can be large. Apparently, a significant fraction of apo-FL_p resists intracellular proteolysis. Of the -5_p molecules 90 – 100 % contain FMN according to TCA precipitation. As a result, FMN titration of -5_p shows no additional binding of flavin (Fig. 4A).

The question arises why a higher proportion of -5_p is observed to be in the holo form than is the case for FL_p. We note that the apo form of -5_p is more susceptible to degradation than apo-FL_p, which is in agreement with our observation that site-directed mutagenesis of apoflavodoxin decreases the stabilities of the corresponding native apo-proteins (38,45,51,52). In addition, flexibility of the flavin-binding site in flavodoxin is much lower than in apoflavodoxin and holo-protein is much more stable than apo-protein (25,53,54). As a result, contrary to apoflavodoxin, flavodoxin is hardly degraded by proteases, as trypsin digestion corroborates (Fig. 4C). For FL_p it is possible to prepare apo-protein, using GuHCl unfolding of the construct and its subsequent refolding on a Superdex75 column (see Materials and Methods). Apo-FL_p shows the same degradation pattern as apoflavodoxin (Fig. 4C). The apo form of -5_p could not be generated due to aggregation of the sample during refolding, an indication of -5_p's instability. A -5_p mimic was created by N-terminally moving the original stop codon of flavodoxin, generating a -5_p without StrepII-tag, Smt3-domain and SecM sequence. Upon overexpression, this -5_p mimic is sequestered into the in-

soluble part of cell lysate or degraded by proteases (see Fig. S3), whereas flavodoxin is soluble and stable. This observation is another indication of the instability of -5_p . During production of stalled nascent chains the accompanying apo -5_p is degraded by proteases, whereas holo -5_p is protected. Observation of the elution of the isolated StrepII-tag attached to Smt3 supports this scenario (Fig. 3B). As a result, in case of -5_p 90 – 100 % of purified, released nascent chains contain FMN and are properly folded.

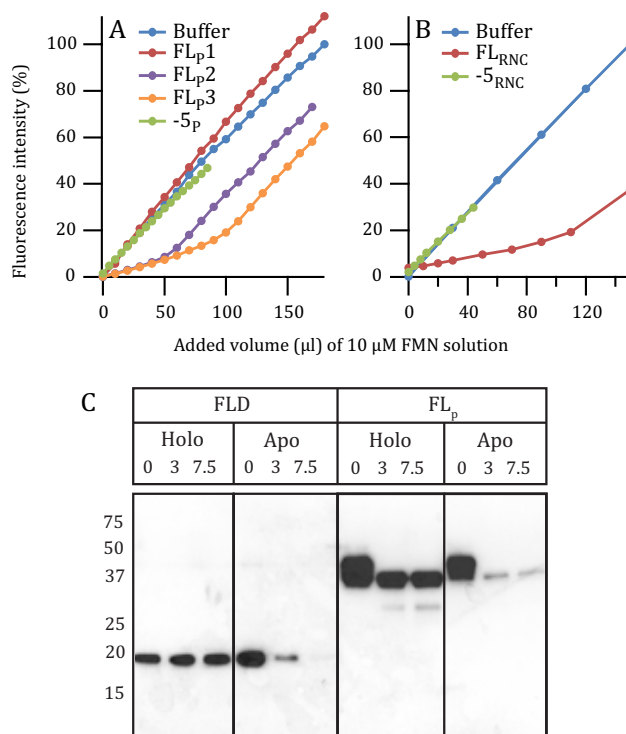


Figure 4 Effect of FMN binding capacity on protection against proteolysis and release of nascent chains. For (A) and (B), observation of an alternating slope in the titration data reveals FMN binding. All fluorescence is normalized to the fluorescence of the end-point of the titrated buffer sample. A, Titration of buffer, FL_p and -5_p with FMN. FL_p preparations of three *E. coli* batches have been used, labeled FL_p1, FL_p2 and FL_p3, respectively. FL_p1 and FL_p3 were induced in M9 minimal medium, whereas FL_p2 was induced in TB medium. FL_p1 and -5_p do not bind FMN, as they are already saturated. Concentrations of FL_p and -5_p were between 2 to 3 μM. B, Titration of buffer, FL_{RNC} and -5_{RNC} with FMN. FL_{RNC} was induced in M9 minimal medium. -5_{RNC} does not bind FMN and no fluorescence arises after TCA precipitation, therefore it cannot bind FMN. Concentrations of FL_{RNC} and -5_{RNC} were between 0.5 and 1 μM. C, Purified holo and apo samples of flavodoxin and FL_p (1 pmol/μl) are digested with trypsin (10 μg/ml). Samples are taken at 0, 3 and 7.5 minutes after trypsin addition and analysed by Western blot probed with anti-flavodoxin. While the respective apo-proteins are degraded within 3 minutes, holo protein resists proteolysis up to 7.5 minutes. The approximately 5-kDa downward shift of holo FL_p observed between 0 and 3 minutes is due to proteolytic removal of triple StrepII-tag. The faint band in the apo FL_p samples of 3 and 7.5 minutes is a small remainder of holo FL_p after preparation of apo protein.

In *E. coli* (strain BL21(DE3) Δ tig::kan) nascent flavodoxin does not bind FMN.

In contrast to FL_p and -5_p, FL_{RNC} and -5_{RNC} show no fluorescence of free FMN after TCA precipitation. As none of the ribosomal proteins contain FMN or FAD as cofactor, we conclude that during an induction period of up to 2.5 hours, and under the growth conditions used, nascent flavodoxin does not bind FMN during production in *E. coli*.

To reveal whether RNCs of flavodoxin have intrinsic FMN-binding capacity, we titrated them with FMN. Upon binding, FMN fluorescence is quenched, resulting in a difference in fluorescence intensity between free and bound FMN. The altering slope in the fluorescence titration data of FL_{RNC} (Fig. 4B) shows that FMN binds to FL_{RNC}. Non-native folding states, like for example unfolded and molten globular apo-flavodoxin, do not show a change in slope of the titration data and thus they do not bind FMN (50). The titration data of Fig. 4B thus demonstrate that flavodoxin can obtain its native fold while fully exposed outside the ribosome. Folding to native apoflavodoxin can occur *in vivo* as well as *in vitro*, and consequently FMN binding also happens in both conditions. It is unlikely that FMN titration induces structural rearrangements that permit FMN binding in stalled or released apoflavodoxin and its C-terminally truncated variants. Only natively folded apoflavodoxin is capable of binding FMN, because it provides the required combination and geometry of hydrogen bonds and aromatic interactions (25). As FMN titration of -5_{RNC} shows, incomplete flavodoxin polypeptide that emerges from the ribosomal exit tunnel does not incorporate FMN (Fig. 4B). The corresponding released product (i.e., -5_p) does bind FMN (Fig. 4A). Docking of the C-terminal α -helix of flavodoxin onto the core of the protein is required for binding of FMN, as several residues of this α -helix are involved (25). Thus, upon dissociation from the ribosome, the C-terminal α -helix forms within -5_p, with possibly some residues of the TEV linker involved (Fig. 1A), enabling subsequent binding of FMN. FMN incorporation is also seen in the -5_p mimic (which has no StrepII-tag, Smt3 and SecM sequence) though to a lesser degree (see Fig. S4). The difference in FMN binding capacities of -5_{RNC} and -5_p reveals that the ribosome affects nascent flavodoxin folding and cofactor binding.

The C-terminally fused SecM sequence, which is used as the stalling sequence, does not constrain native folding of apoflavodoxin, because it is well inside the ribosomal exit tunnel and relatively distant from the apoflavodoxin sequence (Fig. 1A). Using similar constructs, it has been shown that the presence of the SecM sequence does not affect nascent SH3 domain and barnase, because they fold properly when fully exposed outside the ribosome (12,15). Indeed, full-length ribosome-attached apoflavodoxin folds natively and binds FMN, as demonstrated above.

Binding of FMN does not release FL protein from the ribosome. As FL_{RNC} can bind FMN the question arises whether FMN binding overcomes stalling. Incorporation of FMN stabilizes flavodoxin by about 7 kcal/mol (50) and this drop in free energy might suffice to dissociate SecM from the ribosome.

To test this hypothesis, 50 μ M of FMN has been added to cell extract obtained through lysis of *E. coli* cells expressing FL construct. This FMN concentration is about 20 times the total concentration of released protein and RNCs. Subsequent purification of FL_{RNC} and FL_p happens according to protocol. If FMN binding would release stalled nascent chains, the ratio of released protein versus RNC should increase. However, no such increase is observed within experimental error (Fig. 5A and B). Incubation of purified FL_{RNC} with excess FMN for 24 hours shows some release of stalled nascent chain, but no more than without FMN. Clearly, binding of

FMN does not dissociate FL protein from the ribosome.

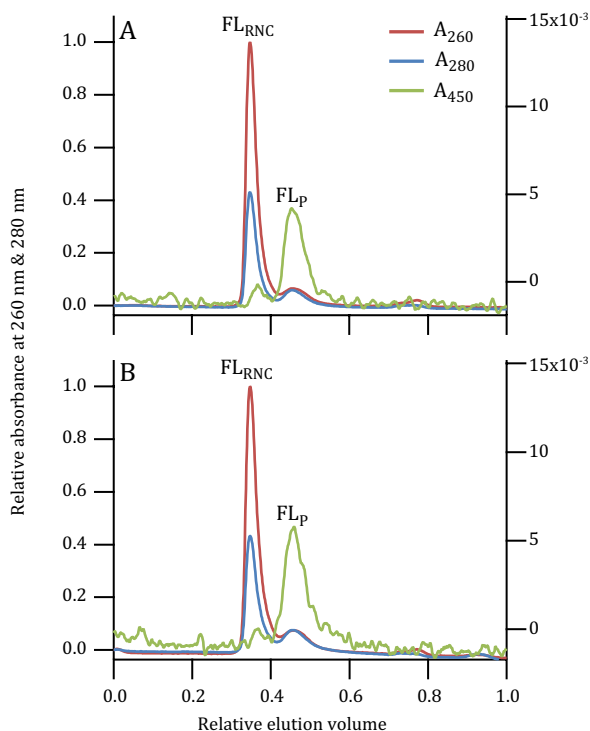


Figure 5 Binding of FMN does not release stalled nascent FL_P from ribosomes. Shown are Superdex75 elution profiles with all absorbance traces normalized relative to the maximum absorbance at 260 nm of FL_{RNC}. Absorbance at 450 nm tracks the incorporation of FMN. Elution volumes are normalized to the total column volume V_t . *A*, FL_{RNC} and FL_P obtained from cell extract of *E. coli* to which no excess FMN has been added. *B*, FL_{RNC} and FL_P obtained from cell extract of *E. coli* with an excess of 50 μ M FMN.

DISCUSSION

Whereas the understanding of folding of isolated proteins in a test tube is considerable, knowledge about how polypeptides fold while they are synthesized by the ribosome is largely lacking. As the α/β parallel architecture of flavodoxin is a common structural design in nature, and the flavodoxin-like fold is a subdomain in various multi-domain proteins, elucidating its nascent chain folding is highly relevant.

Synthesis of the 179-residue flavodoxin lasts about 9 seconds in *E. coli* under optimal conditions, because the translation rate is about 20 amino acid residues per second (55). Our data imply that, during an induction period of up to 2.5 hours, FMN in *E. coli* (strain BL21(DE3) Δ *tig::kan*) is limiting for saturating apoflavodoxin, because FL_P and FL_{RNC} can still bind additional FMN (Fig. 4A and B). However, we cannot exclude that the presence of chaperones may affect folding of full-length ribosome-attached apoflavodoxin in *E. coli* (strain BL21(DE3) Δ *tig::kan*), which could adversely influence FMN binding. We note minute, sub-stoichiometric amounts of chaperones DnaK and GroEL/ES with mass spectrometry in purified samples of RNCs and released proteins. Apoflavodoxin molecules that are unable to incorporate FMN explain why we previously observed less over-production in *E. coli* (strain TG2) of protein variants of flavodoxin (like e.g. F44Y, W128F, W167F and W128F/W167F (38,56)) compared to wild-type protein. During the 12 to 20 hours lasting over-production of these protein variants, cells have to produce a large amount of FMN to satisfy the requirements for cofactor. As FMN synthesis is relatively slow, it takes much longer to assemble all the protein in holo form than the time in which

the maximum of apo-protein is synthesized (57). Due to amino acid residue replacement, F44Y, W128F, W167F or W128F/W167F apoflavodoxin variants are destabilized compared to wild-type protein (38,45,51,52). Thus, these apo-proteins are more efficiently degraded by intracellular proteases before they can take advantage of the stabilization conferred through binding of FMN. For larger flavoproteins the timely incorporation of cofactor and thus stabilization of the protein may be even more important, as larger (multi-domain) proteins often do not refold efficiently once they have been unfolded (58).

Our study demonstrates that the ribosome affects cofactor binding during birth of flavodoxin, because translation emulators like -5_{RNC} cannot bind FMN (Fig. 6). The corresponding released protein construct (i.e., -5_{p}) is shown to contain FMN. As a result, non-covalent binding of cofactor to released apoflavodoxin is the last step in production of this flavoprotein in the cell. Full-length, isolated chains of proteins with an α - β parallel native-state topology tend to temporarily misfold during unassisted folding *in vitro* to their functionally active forms. This folding defect causes the presence of molten globules, which are prone to aggregation (32,37,40). Whether the ribosome modulates formation of molten globules of nascent flavodoxin awaits further investigation.

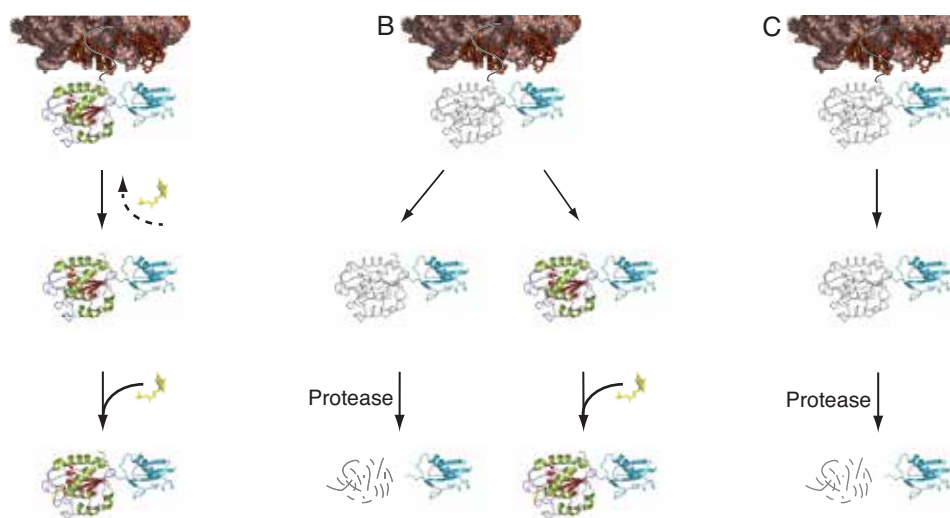


Figure 6 Proposed model for cotranslational flavodoxin folding and cofactor binding. Flavodoxin and Smt3-domain (PDB ID 1L2N) are coloured red/green and blue, respectively. FMN is shown in yellow. *A*, Full-length RNCs can fold while stalled to the ribosome (coloured brown) and are capable of binding FMN. However, most nascent chains release as apoFL_p, which subsequently bind FMN. *B*, -5_{RNC} does not acquire flavodoxin's native fold while stalled to the ribosome. After release, either -5_{p} does not fold properly and is degraded by proteases, resulting in purification of only the Smt3 domain. Or, -5_{p} folds properly and binds FMN, which protects it against proteolytic degradation. *C*, -10_{RNC} does not fold properly while stalled to the ribosome. Released -10_{p} is degraded by proteases, and as a result only the Smt3 domain is purified.

Using a linker with a SecM stalling sequence, we have shown that flavodoxin is natively folded while fully exposed outside the ribosome and that it can bind FMN (Fig. 6). Separate domains of multi-domain proteins are expected to fold independently of each other during translation (14). For multi-domain proteins that contain a flavodoxin-like fold as subdomain, our work suggests that incorporation of cofactor probably occurs cotranslationally once this fold has fully emerged from the ribosome. This concept contributes to understanding proteostasis during translation of cofactor containing, multi-domain proteins. Incorporation of cofactors would hamper proteolytic degradation of nascent chains, due to increased stability conferred by cofactor binding. Especially in eukaryotes, where protein synthesis is much slower than in prokaryotes (i.e., 3 instead of 20 amino acids per second (59)), such impediment of proteolysis could increase protein production efficiency. Indeed, for several flavoproteins involved in diseases, a beneficial effect has been observed when the respective flavin was supplemented (60).

ACKNOWLEDGEMENTS

We thank Prof. Elke Deuerling (University of Konstanz) for the BL21(DE3) *Δtig::kan E. coli* strain and the plasmid containing the Strep-Smt3-SecM construct. We acknowledge Sebastian Müller, who cloned the shortened constructs. We also thank Martina Radli, Tom Ewing, dr. Simon Lindhoud and dr. Evert Houwman for helpful comments on this manuscript. This work was supported by an ECHO grant of NWO to C.v.M.

REFERENCES

1. Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324-332
2. Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.* **78**, 959-991
3. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905-920
4. Harms, J., Schlutzenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) High resolution structure of the large ribosomal subunit from a mesophilic Eubacterium. *Cell* **107**, 679-688
5. Selmer, M., Dunham, C. M., Murphy, F. V., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942
6. Ellis, R. J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.* **11**, 114-119
7. Komar, A. A., Kommer, A., Krashennnikov, I. A., and Spirin, A. S. (1997) Cotranslational folding of globin. *J. Biol. Chem.* **272**, 10646-10651
8. Ellis, J. P., Bakke, C. K., Kirchdoerfer, R. N., Jungbauer, L. M., and Cavagnero, S. (2008) Chain dynamics of nascent polypeptides emerging from the ribosome. *ACS Chem. Biol.* **3**, 555-566
9. Lamprou, P., Kempe, D., Katranidis, A., Büldt, G., and Fitter, J. (2014) Nano-second dynamics of calmodulin and ribosome-bound nascent chains studied by time-resolved fluorescence anisotropy. *ChemBioChem* **15**, 977-985
10. Clark, P. L., and King, J. (2001) A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early *in vitro* refolding intermediates. *J. Biol. Chem.* **276**, 25411-25420
11. Cabrita, L. D., Hsu, S.-T. D., Launay, H., Dobson, C. M., and Christodoulou, J. (2009) Probing ribosome-nascent chain complexes produced *in vivo* by NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22239-22244
12. Eichmann, C., Preissler, S., Riek, R., and Deuerling, E. (2010) Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9111-9116
13. Kelkar, D. A., Khushoo, A., Yang, Z., and Skach, W. R. (2012) Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. *J. Biol. Chem.* **287**, 2568-2578
14. Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U. (1999) Co-translational domain folding as the structural basis for the rapid *de novo* folding of firefly luciferase. *Nat. Struct. Biol.* **6**, 697-705
15. Rutkowska, A., Beerbaum, M., Rajagopalan, N., Fiaux, J., Schmieder, P., Kramer, G., Oschkinat, H., and Bukau, B. (2009) Large-scale purification of ribosome-nascent chain complexes for biochemical and structural studies. *FEBS Lett.* **583**, 2407-2413

16. Khushoo, A., Yang, Z., Johnson, A. E., and Skach, W. R. (2011) Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1. *Mol. Cell* **41**, 682-692
17. Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012) Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* **48**, 63-74
18. Kudlicki, W., Kitaoka, Y., Odom, O. W., Kramer, G., and Hardesty, B. (1995) Elongation and folding of nascent ricin chains as peptidyl-tRNA on ribosomes: the effect of amino acid deletions on these processes. *J. Mol. Biol.* **252**, 203-212
19. Evans, M. S., Sander, I. M., and Clark, P. L. (2008) Cotranslational folding promotes β -helix formation and avoids aggregation *in vivo*. *J. Mol. Biol.* **383**, 683-692
20. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I., and Bustamante, C. (2011) The Ribosome modulates nascent protein folding. *Science* **334**, 1723-1727
21. Cao, Y., and Li, H. (2011) Dynamics of protein folding and cofactor binding monitored by single-molecule force spectroscopy. *Biophys. J.* **101**, 2009-2017
22. Hamm, H. H., and Decker, K. (1978) FAD Is covalently attached to peptidyl-tRNA during cell-free synthesis of 6-hydroxy-D-nicotine oxidase. *Eur. J. Biochem.* **92**, 449-454
23. Brandsch, R., Bichler, V., Mauch, L., and Decker, K. (1993) Cysteine to serine replacements in 6-hydroxy-D-nicotine oxidase - Consequences for enzyme-activity, cofactor incorporation, and formation of high-molecular-weight protein complexes with molecular chaperones (GroEL). *J. Biol. Chem.* **268**, 12724-12729
24. Macheroux, P., Kappes, B., and Ealick, S. E. (2011) Flavogenomics – a genomic and structural view of flavin-dependent proteins. *FEBS J.* **278**, 2625-2634
25. Bollen, Y. J. M., Westphal, A. H., Lindhoud, S., van Berkel, W. J. H., and van Mierlo, C. P. M. (2012) Distant residues mediate picomolar binding affinity of a protein cofactor. *Nature Comm.* **3**, 1010
26. Alagaratnam, S., van Pouderoyen, G., Pijning, T., Dijkstra, B. W., Cavazzini, D., Rossi, G. L., van Dongen, W. M. A. M., van Mierlo, C. P. M., van Berkel, W. J. H., and Canters, G. W. (2005) A crystallographic study of Cys69Ala flavodoxin II from *Azotobacter vinelandii*: Structural determinants of redox potential. *Protein Sci.* **14**, 2284-2295
27. Caetano-Anolles, G., Kim, H. S., and Mittenthal, J. E. (2007) The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9358-9363
28. Pierella Karlusich, J. J., Ceccoli, R. D., Grana, M., Romero, H., and Carrillo, N. (2015) Environmental selection pressures related to iron utilization are involved in the loss of the flavodoxin gene from the plant genome. *Genome Biol. Evol.* **7**, 750-767
29. Pierella Karlusich, J. J., Lodeyro, A. F., and Carrillo, N. (2014) The long goodbye: the rise and fall of flavodoxin during plant evolution. *J. Exp. Bot.* **65**, 5161-5178
30. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8411-8416
31. Garcin, E. D., Bruns, C. M., Lloyd, S. J., Hosfield, D. J., Tiso, M., Gachhui, R.,

Stuehr, D. J., Tainer, J. A., and Getzoff, E. D. (2004) Structural basis for isozyme-specific regulation of electron transfer in nitric-oxide synthase. *J. Biol. Chem.* **279**, 37918-37927

32. Bollen, Y. J. M., Sanchez, I. E., and van Mierlo, C. P. M. (2004) Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin. *Biochemistry* **43**, 10475-10489

33. Bollen, Y. J. M., Kamphuis, M. B., and van Mierlo, C. P. M. (2006) The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4095-4100

34. Lindhoud, S., Westphal, A. H., Borst, J. W., and van Mierlo, C. P. M. (2012) Illuminating the off-pathway nature of the molten globule folding intermediate of an α - β parallel protein. *PLoS ONE* 10.1371/journal.pone.0045746

35. Lindhoud, S., Westphal, A. H., Visser, A. J. W. G., Borst, J. W., and van Mierlo, C. P. M. (2012) Fluorescence of Alexa fluor dye tracks protein folding. *PLoS ONE* 10.1371/journal.pone.0046838

36. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2009) Noncooperative formation of the off-pathway molten globule during folding of the α - β parallel protein apoflavodoxin. *J. Am. Chem. Soc.* **131**, 2739-2746

37. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2008) Extensive formation of off-pathway species during folding of an α - β parallel protein is due to docking of (non)native structure elements in unfolded molecules. *J. Am. Chem. Soc.* **130**, 16914-16920

38. Nabuurs, S. M., Westphal, A. H., aan den Toorn, M., Lindhoud, S., and van Mierlo, C. P. M. (2009) Topological switching between an α - β parallel protein and a remarkably helical molten globule. *J. Am. Chem. Soc.* **131**, 8290-8295

39. Nabuurs, S. M., and van Mierlo, C. P. M. (2010) Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of α - β parallel protein flavodoxin. *J. Biol. Chem.* **285**, 4165-4172

40. Bollen, Y. J. M., and van Mierlo, C. P. M. (2005) Protein topology affects the appearance of intermediates during the folding of proteins with a flavodoxin-like fold. *Biophys. Chem.* **114**, 181-189

41. Nakatogawa, H., and Ito, K. (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**, 629-636

42. Schaffitzel, C., and Ban, N. (2007) Generation of ribosome nascent chain complexes for structural and functional studies. *J. Struct. Biol.* **158**, 463-471

43. Hansson, M. D., Rzeznicka, K., Rosenbäck, M., Hansson, M., and Sirijovski, N. (2008) PCR-mediated deletion of plasmid DNA. *Anal. Biochem.* **375**, 373-375

44. Steensma, E., Heering, H. A., Hagen, W. R., and Van Mierlo, C. P. (1996) Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur. J. Biochem.* **235**, 167-172

45. van Mierlo, C. P. M., van Dongen, W. M. A. M., Vergeldt, F., van Berkel, W. J. H., and Steensma, E. (1998) The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate. *Prot. Sci.* **7**, 2331-2344

46. Evans, M. S., Ugrinov, K. G., Frese, M. A., and Clark, P. L. (2005) Homogeneous stalled ribosome nascent chain complexes produced *in vivo* or *in vitro*. *Nature methods* **2**, 757-762
47. Waudby, C. A., Launay, H., Cabrita, L. D., and Christodoulou, J. (2013) Protein folding on the ribosome studied using NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **74**, 57-75
48. Moore, S. D., and Sauer, R. T. (2007) The tmRNA system for translational surveillance and ribosome rescue. *Ann. Rev. Biochem.* **76**, 101-124
49. Cymer, F., and von Heijne, G. (2013) Cotranslational folding of membrane proteins probed by arrest-peptide-mediated force measurements. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 14640-14645
50. Bollen, Y. J. M., Nabuurs, S. M., van Berkel, W. J. H., and van Mierlo, C. P. M. (2005) Last in, first out. The role of cofactor binding in flavodoxin folding. *J. Biol. Chem.* **280**, 7836-7844
51. Nabuurs, S. M., de Kort, B. J., Westphal, A. H., and van Mierlo, C. P. (2010) Non-native hydrophobic interactions detected in unfolded apoflavodoxin by paramagnetic relaxation enhancement. *Europ. Biophys. J.* **39**, 689-698
52. Laptinok, S. P., Visser, N. V., Engel, R., Westphal, A. H., van Hoek, A., van Mierlo, C. P., van Stokkum, I. H., van Amerongen, H., and Visser, A. J. (2011) A general approach for detecting folding intermediates from steady-state and time-resolved fluorescence of single-tryptophan-containing proteins. *Biochemistry* **50**, 3441-3450
53. Steensma, E., and van Mierlo, C. P. M. (1998) Structural characterisation of apoflavodoxin shows that the location of the stable nucleus differs among proteins with a flavodoxin-like topology. *J. Mol. Biol.* **282**, 653-666
54. Lindhoud, S., van den Berg, W. A. M., van den Heuvel, R. H. H., Heck, A. J. R., van Mierlo, C. P. M., and van Berkel, W. J. H. (2012) Cofactor binding protects flavodoxin against oxidative stress. *PLoS ONE* 10.1371/journal.pone.0041363.t002
55. Kjeldgaard, N. O., and Gausing, K. (1974) Regulation of biosynthesis of ribosomes. in *Ribosomes* (Nomura, M., Tissieres, A., and Lengyel, P. eds.), Cold Spring Harbor Laboratory Press, U.S. . pp 369-392
56. Visser, N. V., Westphal, A. H., van Hoek, A., van Mierlo, C. P. M., Visser, A. J. W. G., and van Amerongen, H. (2008) Tryptophan-tryptophan energy migration as a tool to follow apoflavodoxin folding. *Biophys. J.* **95**, 2462-2469
57. Borruel, A., Peleato, M. L., Orera, V. M., Gómez-Moreno, C., and Fillat, M. F. (1994) Electron paramagnetic resonance as a tool for monitoring overexpression in *Escherichia coli* of fully functional flavodoxin. *Anal. Biochem.* **218**, 255-258
58. Caldinelli, L., Iametti, S., Barbiroli, A., Fessas, D., Bonomi, F., Piubelli, L., Molla, G., and Pollegioni, L. (2008) Relevance of the flavin binding to the stability and folding of engineered cholesterol oxidase containing noncovalently bound FAD. *Protein Sci.* **17**, 409-419
59. Netzer, W. J., and Hartl, F. U. (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388**, 343-349
60. Rodrigues, J. V., Henriques, B. J., Lucas, T. G., and Gomes, C. M. (2012) Cofactors and metabolites as protein folding helpers in metabolic diseases. *Curr. Top. Med.*

Chem. **12**, 2546-2559

SUPPLEMENTARY INFORMATION

MATERIALS AND METHODS

RNase A degradation of RNCs. FL_{RNC} was produced and purified according to Material and Methods section 2.2. After use of the Superdex75 column to separate FL_{RNC} from FL_p, the FL_{RNC}-containing fractions were concentrated and incubated with RNase A (50 µg/ml) and 25 mM EDTA for 1 hour at 37 °C. The resulting sample was spun down for 10 minutes at 18000 RCF and the supernatant was subsequently run on a Superdex75 column while fractions were collected during the run. The fractions were characterized using a 15% SDS-PAGE gel and subsequently blotted with anti-flavodoxin.

Elution of ribosomes versus released nascent chains. To verify the elution patterns of ribosomes versus FL_p, first empty ribosomes (purchased from New England Biolabs (NEB)) were run on a Superdex75 column. Subsequently FL_p (purified in the experiment described above) was mixed with the empty ribosomes and this mixture was run over the Superdex75 column. Lastly purified FL_p was run over the same Superdex75 column.

Expression of terminally truncated variants of flavodoxin without fusion to the Smt3-domain and the stalling sequence. The pUC18 plasmid containing the flavodoxin gene was used to construct two protein variants, 1-174 and 1-169, in which the stop codon has been moved N-terminally. These protein variants mimic the -5_p (1-174) and -10_p (1-169) molecules, respectively, but lack the StrepII-tag, Smt3-domain and SecM sequence. Both protein variants (named -5 and -10, respectively), along with full-length flavodoxin (FL) and empty vector pUC18 (pUC) were expressed in *E. coli* (strain TG2) grown on TB or minimal medium. Half of the corresponding cultures were induced with IPTG for a period of 16 hours. Gel samples were taken before and after induction. The cells in these samples were lysed and separated into supernatant and pellet fractions by centrifugation. Supernatant and pellet samples were analysed by running them on a 15% SDS-PAGE gel and subsequent blotting with anti-flavodoxin.

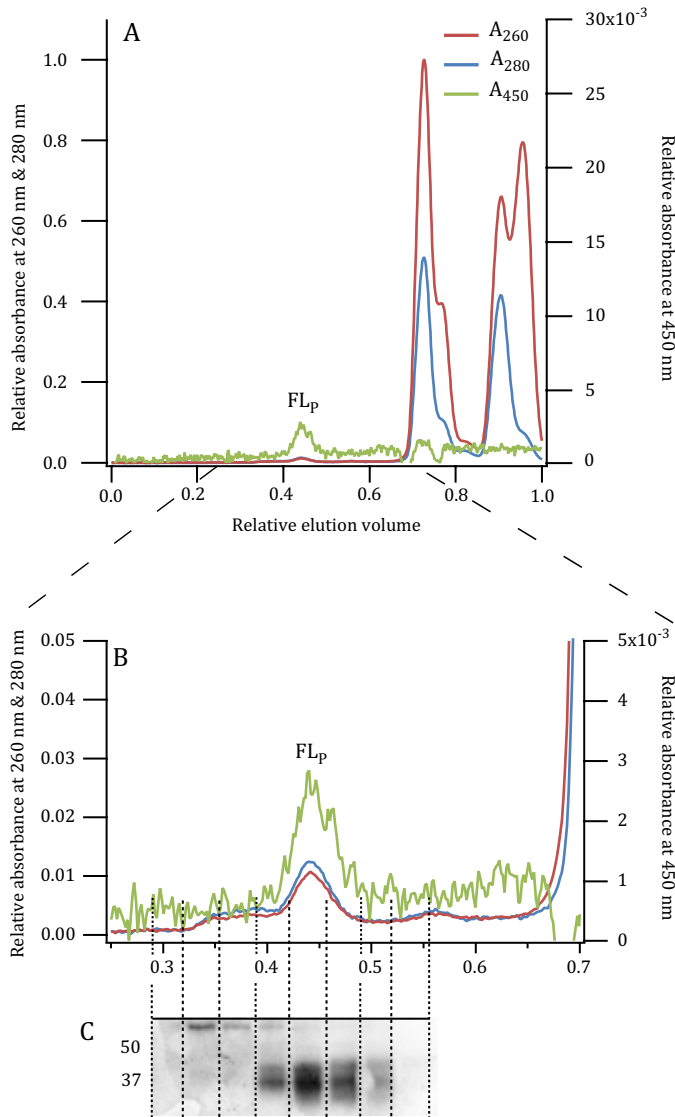


Figure S1 Disruption of FL_{RNC} by RNase A results in elution of nascent FL at relative volume of FL_p . Shown are Superdex75 elution profiles, with all absorbance traces normalized relative to the maximum absorbance at 260 nm. Absorbance at 450 nm tracks FMN incorporation. Elution volumes are normalized to the total column volume V_t . *A*, FL_{RNC} incubated for 1 hour with RNase A. Due to degradation of rRNA, the RNC peak has vanished, resulting in several smaller peaks corresponding to ribosomal proteins and rRNA. *B*, Expansion of (*A*) reveals the presence of a peak with absorbance at 450 nm, which elutes at the same relative volume as FL_p . Dashed vertical lines designate collected fractions during the run. *C*, Collected fractions are analysed by Western blot probed with anti-flavodoxin. Fractions corresponding to the peak with absorbance at 450 nm indeed contain FL_p .

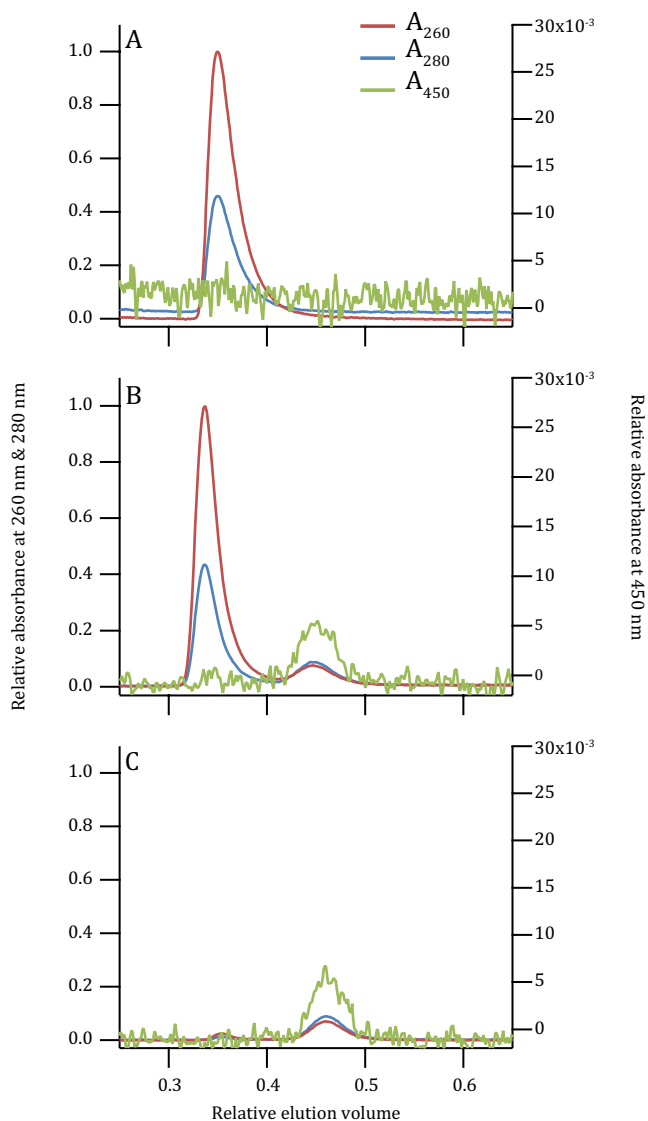


Figure S2 FL_p does not co-elute with intact, empty ribosomes. Shown are Superdex75 elution profiles, with all absorbance traces normalized relative to the maximum absorbance at 260 nm of empty ribosomes. Absorbance at 450 nm tracks FMN incorporation. Elution volumes are normalized to the total column volume V_t . *A*, Empty ribosomes (purchased from NEB). *B*, Mixture of empty ribosomes and purified FL_p. *C*, Purified FL_p.

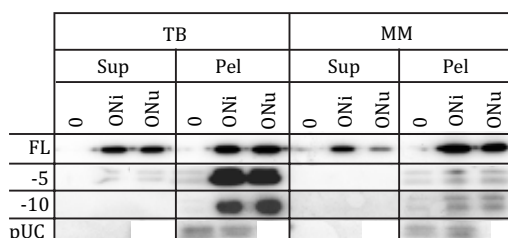


Figure S3 Expression of terminally truncated variants of flavodoxin. Shown are Western blots probed with anti-flavodoxin. Three variants of flavodoxin were produced. The first one is full-length protein (FL). In the other two variants the regular stop codon has been moved N-terminally to create protein variants 1-174 (-5) and 1-169 (-10). As a control, the empty vector containing no flavodoxin gene (pUC) was also tested. All four constructs were transformed into *E. coli* (strain TG2) and grown either in nutrient-rich medium (TB) or minimal medium (MM). Samples were taken before induction (0), after induction for 16 hours (ONi) or after growth for 16 hours without induction (ONu). Subsequently, cells of the samples were lysed and both supernatant (Sup) and pellet (Pel) fractions were analysed on Western blot. Due to a leaky promoter in the vector, protein is produced in un-induced samples. For full-length flavodoxin, protein is found in supernatant and pellet fractions of cells grown on both TB and MM medium. -5 and -10 are detected only in the pellet of cells grown on TB medium. Due to lower expression of the -5 and -10 protein in cells grown on MM medium compared to on TB medium, the protease machinery is able to degrade all produced protein in the first situation, but not in the second one. As expected, the empty vector does not show production of flavodoxin, though there is some cross-reactivity of the flavodoxin antibody with *E. coli* proteins in the fractions.

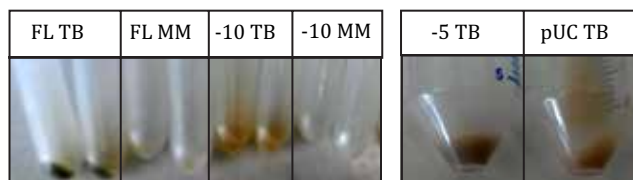
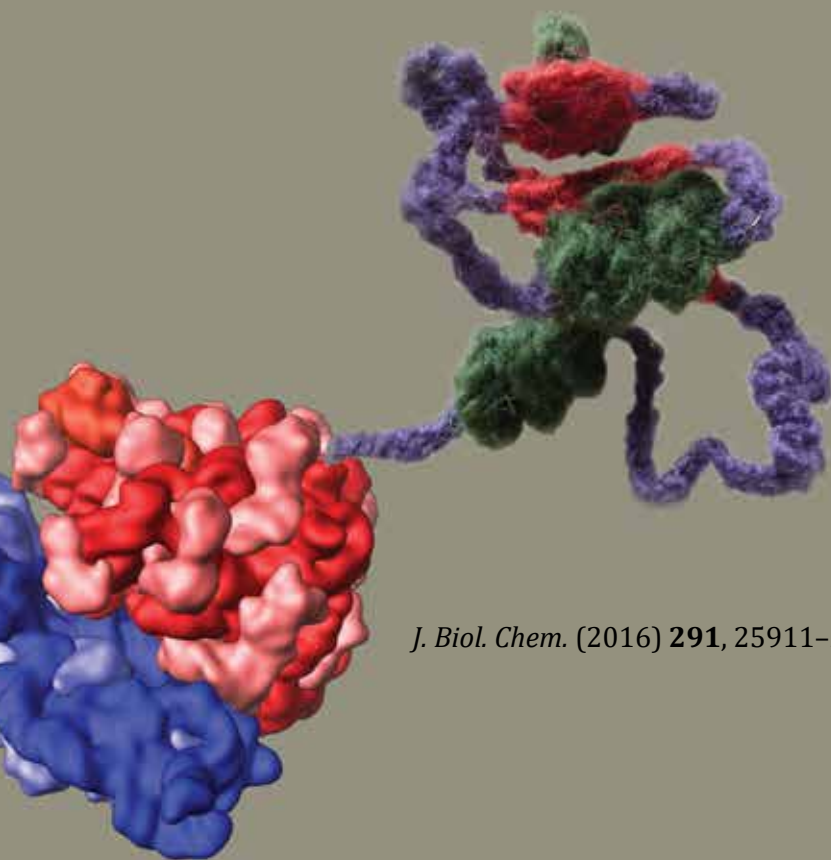


Figure S4 FMN content of overexpressed flavodoxin is reflected in the colouring of pelleted cells. Three variants of flavodoxin (FL, -5 and -10) and a vector containing no flavodoxin gene (pUC) were transformed in *E. coli* (strain TG2) and grown and induced for 16 hours. Subsequently, cells were pelleted and checked for their colour. Expression of FL in TB results in a blue-greenish colour due to the semiquinone form of FMN. TB and MM denote the medium in which the cells were induced, i.e., nutrient-rich (TB) or minimal medium (MM). FL and -5 induced in TB show a darker blue-greenish coloured cell pellet than -10 (which is incapable of binding FMN) or pUC (which does not express flavodoxin), indicating binding of FMN. FL has a darker coloured cell pellet than -5, signifying higher FMN content. Due to lower growth in MM medium, the cell pellet in these samples is not large enough for the colour change between FL and -10 to be visible.

The ribosome restrains molten globule formation in stalled nascent flavodoxin

Joseline A. Houwman, Estelle André, Adrie H. Westphal, Willem J.H. van
Berkel and Carlo P.M. van Mierlo

Laboratory of Biochemistry, Wageningen University, the Netherlands



J. Biol. Chem. (2016) **291**, 25911–25920

Folding of proteins usually involves intermediates, of which an important type is the molten globule (MG). MGs are ensembles of interconverting conformers that contain (non-)native secondary structure and lack the tightly packed tertiary structure of natively folded globular proteins. Whereas MGs of various purified proteins have been probed to date, no data is available on their presence and/or effect during protein synthesis. To study whether MGs arise during translation, we use ribosome-nascent chain (RNC) complexes of the electron-transfer protein flavodoxin. Full-length, isolated flavodoxin, which contains a non-covalently bound flavin mononucleotide (FMN) as cofactor, acquires its native α/β parallel topology via a folding mechanism that contains an off-pathway intermediate with molten globular characteristics. Extensive population of this MG state occurs at physiological ionic strength for apoflavodoxin variant F44Y, in which a phenylalanine at position 44 is changed to a tyrosine. Here, we show for the first time that ascertaining the binding rate of FMN as a function of ionic strength can be used as a tool to determine the presence of the off-pathway MG on the ribosome. Application of this methodology to F44Y apoflavodoxin RNCs shows that at physiological ionic strength the ribosome influences formation of the off-pathway MG and forces the nascent chain towards the native state.

INTRODUCTION

Proteins need to fold into their correct native conformations to perform their functions. Folding progresses from unfolded to natively folded protein and often involves intermediates (1,2). Some of these intermediates are dead-ends within the folding funnel (i.e., local minima in free energy) and must therefore (partially) unfold before the productive folding pathway can be followed (3). Molten globules are a specific type of folding intermediates, which contain secondary structure elements but lack the tight tertiary packing found in natively folded proteins. Molten globular intermediates can reside on or off the folding pathway to native protein. MGs are prone to aggregation and consequently are implicated in various diseases (4).

Since the first description of MGs (5) much research has focused on determining conditions at which MGs form. Several proteins form these intermediates under mildly alkaline or acidic conditions (6-9). Nowadays, many proteins are thought to fold via MG intermediates (10-13). For example, it has been postulated that the occurrence of MG species is necessary during insertion of proteins into membranes or through membrane pores (14-16). Such insertions often happen cotranslationally, i.e., while the ribosome synthesizes the protein concerned. It has been shown that proteins can fold cotranslationally and sample intermediate folding states (17-23), which might include MGs.

Determining proteins' ability to form cotranslational MGs would be a stepping-stone in elucidating protein folding in the cell, and potentially contributes to understanding the initiation of protein aggregation. Characterization of MGs is complicated by their often transient nature, generally low presence at equilibrium, conformational heterogeneity and tendency to aggregate (24). The presence of MGs can be revealed through use of extrinsic dyes, such as thioflavin T (ThT), as these dyes become highly fluorescent upon binding to exposed hydrophobic residues (25). Other methods for MG detection include circular dichroism and intrinsic tryptophan fluorescence (26-28). However, these methodologies are unsuitable for studying cotranslational MG formation, as ribosome-nascent chain complexes (RNCs) not only contain the emerging polypeptide, but also more than 50 ribosomal proteins. Thus, detection of MG formation on the ribosome is notoriously difficult.

To investigate whether nascent chains form MGs, we utilize a 179-residue flavodoxin from *Azotobacter vinelandii*. Flavodoxin contains a non-covalently bound flavin mononucleotide (FMN) as cofactor and is involved in electron transport in the nitrate reduction cycle. The native protein consists of five α -helices sandwiching a central parallel β -sheet, which is an α/β parallel topology (Fig. 1A) (29). According to SCOP classifications in the Protein Data Bank (PDB), approximately twenty five percent of proteins share this topology. The flavodoxin-like fold is an ancestral fold (30) and is archetypal for the class of $\alpha\beta\alpha$ sandwiches. It has been well established that proteins with a flavodoxin-like architecture fold *in vitro* through involvement of an off-pathway intermediate (11,31-37). Upon folding, the large majority of unfolded flavodoxin molecules temporarily misfold. This misfolding yields an off-pathway MG that is prone to aggregation (38). In contrast to native protein, this MG contains no β -sheet and is helical (Fig. 1B) (33,39,40). Under steady-state conditions the MG is in equilibrium with native apoflavodoxin. Before formation of native apoflavodoxin can occur the MG has to (partially) unfold. This unfolding is the rate-limiting step in folding of native apoflavodoxin. The cofactor FMN only binds to the flavin-binding site of natively folded apoflavodoxin (41). This binding can be followed in time by

fluorescence spectroscopy, as FMN's fluorescence becomes severely quenched upon binding to native apoflavodoxin (42,43).

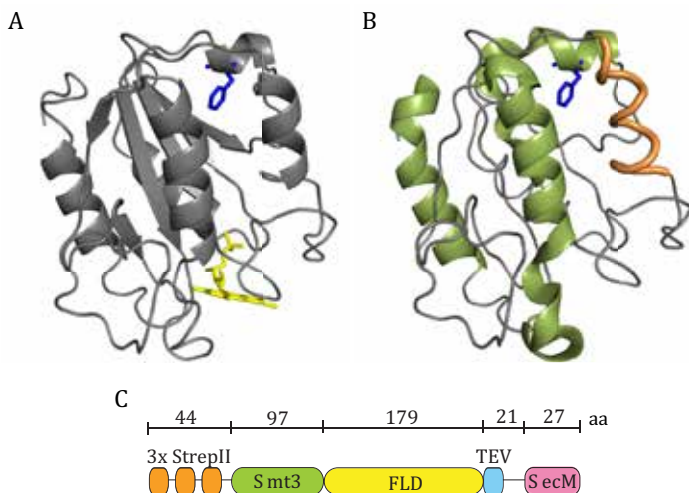


Figure 1 Differences between flavodoxin's native and MG structure and the construct used for RNC production. *A*, cartoon drawing of native flavodoxin (PDB entry 1YOB (29)) with FMN coloured yellow. Residue F44 (blue) is shown in stick representation. *B*, cartoon drawing of the four transiently structured regions of the off-pathway MG, which include α -helices (coloured blue) and structure that is neither α -helix nor β -strand (coloured orange) (33). These regions dock non-natively and form the core of apoflavodoxin's MG (33,35,39,40). The cartoon is a representation of structural elements in the MG, not of their packing. *C*, the construct for RNC production contains a triple N-terminal StrepII-tag (orange), an Smt3-domain (green) fused to flavodoxin (yellow), a recognition site for TEV protease (blue) and a linker that spans the ribosomal exit tunnel and concludes with the SecM stalling sequence (magenta).

Upon mixing FMN with apoflavodoxin under conditions where MG is present, cofactor binding and accompanied fluorescence quenching should be delayed, as the MG needs to unfold before native apoflavodoxin can form. This delayed FMN quenching is used in this study to reveal whether the ribosome modulates formation of molten globular apoflavodoxin. Previously, we demonstrated that once apoflavodoxin has been entirely synthesized and is exposed outside the ribosome, the protein is natively folded and capable of binding FMN (44). To demonstrate whether MGs form during translation, we take advantage of the following folding characteristic of flavodoxin variant F44Y (39): the apo-form of this variant switches from natively folded to the off-pathway MG upon decreasing ionic strength to physiological values (39,40). Due to the higher thermodynamic stability of the holo-form of this protein, once apoflavodoxin binds FMN this topological switching no longer occurs (45). We prepared stalled RNC complexes that expose the entire F44Y flavodoxin protein outside the ribosomal exit tunnel. The translational stalling of the ribosomes is caused by an *Escherichia coli* SecM derived peptide (46,47), which we attached at the C-terminus of flavodoxin. After purification of F44Y apoflavodoxin RNCs, we determined the corresponding FMN binding rates and thus the influence of the ribosome on MG formation.

MATERIAL AND METHODS

Protein expression and purification. To avoid covalent dimerization of purified *A. vinelandii* (apo)flavodoxin, the single cysteine at position 69 was substituted by an alanine (50). The C69A variant is similar to wild-type (apo)flavodoxin (49,50). An additional F44Y mutation was introduced (39), and this protein variant is referred to as F44Y (apo)flavodoxin. Both constructs were transformed in *Escherichia coli* (strain TG2) and purified as described previously (50). Apoprotein was prepared according to established protocols (44). Apoprotein concentrations were determined by titration with known amounts of FMN. Purified proteins in 100 mM potassium pyrophosphate (KPPi) buffer, pH 6.0, were flash frozen in liquid nitrogen and stored at -80 °C.

Due to cold denaturation, some F44Y apoflavodoxin deteriorates into soluble molecules that are incapable of binding FMN upon thawing. These molecules cannot be separated from FMN binding competent apoflavodoxin, thus the latter concentration was determined by titration with FMN in 100 mM KPPi, pH 6.0. Also, as much as possible, freshly prepared F44Y apoflavodoxin was used in experiments.

Influence of temperature on full-length, isolated protein. Thermal unfolding of F44Y apoflavodoxin was probed by ThT fluorescence using a Cary Eclipse spectrophotometer (Varian). Excitation was at 445 nm and emission was measured at 485 nm. Temperature was increased from 15 to 45 °C at a rate of 1 °C/min. The excitation and emission slits were set to bandwidths of 5 and 10 nm, respectively, and the PMT voltage was 900 V. 3.3 µM of F44Y apoflavodoxin and 56 µM of ThT (Sigma) were used in 10 or 100 mM KPPi, pH 6.0.

Engineering RNC variants. Plasmids for RNC production contain a sequence coding for a triple N-terminal StrepII-tag, an Smt3 domain, a TEV-site, a linker and a SecM stalling motif (Fig. 1C). The flavodoxin sequence (C69A or F44Y) was inserted in the multiple cloning site (MCS) present between the TEV-site and the linker. Ulp1 protease specifically recognises the Smt3 domain and cleaves down-stream this domain, thereby producing nascent flavodoxin chains with Ala-1 as authentic N-terminus, if required. The recognition site for TEV protease enables release of nascent protein from the ribosome. The plasmid was a kind gift of Professor Elke Deuerling (University of Konstanz, Germany) (56).

Expression and purification of RNCs *in vivo* and *in vitro*. Plasmids were transformed in *E. coli* strain BL21(DE3) Δ *tig::kan* for expression of RNCs *in vivo*. All RNC variants were prepared as described previously (44). After growth of *E. coli* in Terrific Broth (TB) medium, cells were transferred to minimal (M9) medium for expression of RNCs at either 37 or 15 °C for 1 hour. RNCs were purified using Strep-Tactin Sepharose (IBA GmbH) chromatography and subsequent size-exclusion (Superdex75) chromatography. RNCs and released protein chains were purified and stored in buffer I (50 mM HEPES-KOH, 100 mM potassium acetate, 15 mM magnesium acetate, 1 mM dithiothreitol (DTT), pH 7.4). After purification, the apo-form of released protein chains of F44Y flavodoxin was prepared in 100 mM KPPi, pH 6.0 as described in (44).

For *in vitro* expression of RNCs, the PURExpress® Δ (aa, tRNA) In Vitro Protein Synthesis kit from New England Biolabs was used. Instructions of the manufacturer were followed. After mixing the kit's components with the F44Y flavodoxin containing RNC plasmid, the reaction was allowed to proceed for 2 hours at 37 °C. To

purify RNCs produced *in vitro*, the reaction mixture was loaded onto a Strep-Tactin sepharose column and eluted as previously described in buffer I (44). Eluted RNCs were concentrated with a 10-kDa spin filter (EMD Millipore) at 5,500 g. Concentrated samples were frozen in liquid nitrogen and stored at -80 °C.

Samples of each purification step were analysed by Coomassie-stained SDS-PAGE or Western blotting using either flavodoxin antibody raised in rabbits (Eurogentec) followed by anti-rabbit HRP-IG (Eurogentec) or StrepMAB classic (IBA GmbH) against the N-terminal StrepII-tag.

FMN fluorescence. To establish FMN content of RNCs and released constructs, FMN titration or TCA precipitation was performed as described in (44). To determine the MG nature of RNCs of F44Y protein and of released protein, binding of FMN to these molecules was followed in time. All flavin fluorescence experiments were measured on a Cary Eclipse spectrophotometer (Varian) at 25 °C.

Samples for FMN titration contained 0.5 μ M of either RNC or released protein produced *in vivo* in buffer I. Samples were titrated with a FMN solution of 5 μ M to determine their cofactor-binding capacity using flavin fluorescence. Excitation was at 450 nm, with emission recorded between 500 and 600 nm. For each titration point the average of five scans was taken. Excitation and emission slits were set to a bandwidth of 5 nm and PMT voltage was set to 900 V.

Samples for TCA precipitation contained 50 nM of RNC or released protein produced *in vivo* in buffer I. To establish the total amount of FMN bound to protein, the samples were precipitated by adding 3 % (w/v) TCA. Precipitate was spun down for 10 min at 21,000 g at 4 °C and supernatant was carefully pipetted off to measure its flavin fluorescence. A 50 nM solution of FMN in 3 % (w/v) TCA in buffer I was used as reference. Excitation was at 450 nm, with emission recorded between 500 and 600 nm. Each sample was scanned five times and the average was used for calculations. Excitation and emission slits were set to a bandwidth of 10 or 20 nm, respectively. PMT voltage was set to 970 V.

To determine the rate of FMN binding to C69A and F44Y variants of apoflavodoxin, *E. coli* ribosomes (NEB), RNC or released protein, samples were made of each with protein concentrations ranging from 50 to 147 nM. Samples were measured in stirred fluorescence cuvettes for five minutes before and after adding FMN to a final concentration of 24 nM. Excitation was at 450 nm and emission was monitored at 540 nm. Excitation and emission slits were set to bandwidths of 10 and 20 nm, respectively, and PMT voltage was 970 V. The averaging time of the measurement was 0.1 s. The time between addition of FMN and the start of fluorescence detection was approximately 2 s. Buffers used in these experiments were 10 mM KPPI, pH 6.0; 100 mM KPPI pH 6.0; buffer I or buffer I with a final concentration of 290 mM NaCl. The latter buffer has an ionic strength equal to that of 100 mM KPPI, pH 6.0.

RESULTS

Dependence of apoflavodoxin's off-pathway MG on ionic strength and temperature. Decreasing salt concentration is a well-known manner to destabilize folded proteins (48). Using this approach, in F44Y apoflavodoxin samples we can shift the equilibrium between natively folded protein and MG to the latter species (39). Fluorescence anisotropy shows that lowering ionic strength from 100 to 10 mM

KPPi (i.e., 345 to 75 mM equivalent NaCl units) at 25 °C forces F44Y apoprotein from native apoflavodoxin to MG (Fig. 2A). The MG state is characterized by fluorescence anisotropy that is higher than of native apoflavodoxin. In contrast, at high salt concentrations, like for example 100 mM KPPi, F44Y apoflavodoxin is predominantly in its native state. For C69A apoflavodoxin, which is similar to wild-type apoflavodoxin (49,50), no salt-dependent switch between folding states is observed (Fig. 2A), as its native state is considerably more stable than native F44Y apoflavodoxin (39).

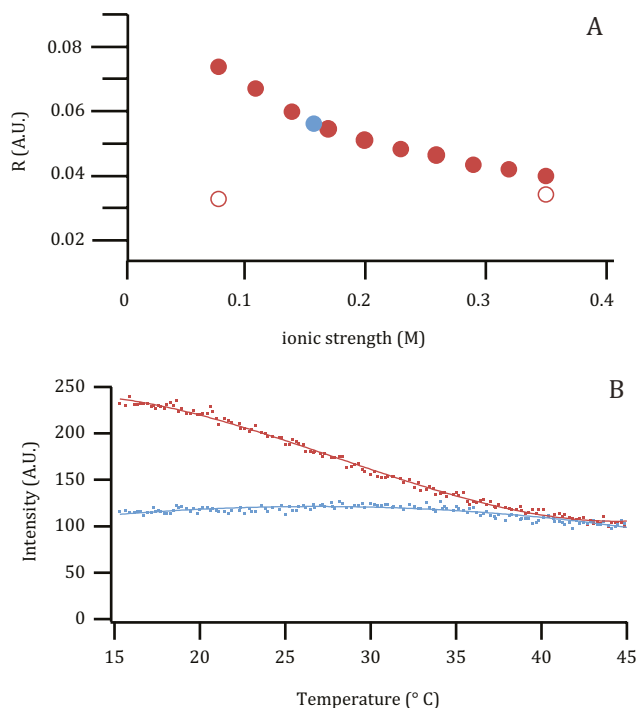


Figure 2 Effects of ionic strength or temperature on F44Y apoflavodoxin. *A*, ionic strength dependency of fluorescence anisotropy of F44Y (closed red circles) and C69A apoflavodoxin (open red circles) (graph adapted from (39)). Upon decreasing ionic strength F44Y apoflavodoxin forms a MG, which is characterized by fluorescence anisotropy that is higher than of native apoflavodoxin. Expected anisotropy of F44Y apoflavodoxin in Buffer I is shown as a blue circle. *B*, temperature-dependence of ThT binding to F44Y apoflavodoxin, as followed by fluorescence emission at 485 nm. Protein is in 10 (red) or 100 mM (blue) KPPi, pH 6.0. ThT binding to the MG of F44Y apoflavodoxin increases ThT fluorescence. Due to unfolding of the MG with increasing temperature, ThT releases and its fluorescence decreases. Protein concentration is 3.3 μ M and the heating rate is 1 °C/min.

The formation of the off-pathway MG not only depends on ionic strength but also on temperature. To follow the temperature-dependent existence of this MG we added the extrinsic dye ThT. This reporter molecule is commonly employed to follow amyloid fibril formation, as it has a high affinity for β -sheets. Upon binding to fibrillar structures, ThT shows a fluorescence increase at 480 nm upon excitation at 450 nm. Binding of ThT and subsequent fluorescence increase is not only restricted to fibrillar structures, but can also occur due to the presence of non β -sheet cavities,

such as may form in MGs (25). Another extrinsic dye commonly used to probe MGs is 1-anilino-8-naphthalene sulfonate (ANS), however, in the case of F44Y apoflavodoxin, its fluorescence does not follow the unfolding of the MG, whereas ThT fluorescence does (data not shown).

Fig. 2B shows how temperature affects F44Y apoflavodoxin at high and low ionic strengths (i.e., 345 and 75 mM equivalent NaCl units, respectively). Temperature was raised from 15 to 45 °C, upon which F44Y apoflavodoxin unfolds (39). Temperature was not raised further to avoid formation of aggregates, which would lead to a rise in ThT fluorescence. For protein at low ionic strength and low temperature, ThT fluorescence is high (Fig. 2B), because F44Y apoflavodoxin is predominantly in the MG state. Upon increasing temperature, ThT fluorescence decreases non-cooperatively, because the MG unfolds gradually and binding of ThT diminishes. This gradual, non-cooperative unfolding is a typical feature of apoflavodoxin's MG (35,37). Tryptophan fluorescence shows that the thermal unfolding transition of native F44Y apoflavodoxin in 100 mM KPPi ranges from about 20 to 40 °C, with a thermal midpoint of 32.9 ± 0.3 °C (39). Above approximately 40 °C, ThT fluorescence of protein at low and high salt concentrations is similar, because at both conditions the protein is now unfolded. At high salt concentration a slight hump in ThT fluorescence is observed in the thermal unfolding transition of native protein (Fig. 2B). This hump arises because the MG state of F44Y apoflavodoxin populates. At the thermal midpoint of native apoflavodoxin (i.e., 32.9 ± 0.3 °C) ThT fluorescence due to MG population is low, because the MG is gradually unfolded.

Fig. 2B implies that at low salt concentration the MG of F44Y apoflavodoxin is maximally populated at about 15 °C. This corroborates our previous findings that the off-pathway MG of F44Y apoflavodoxin is structured in the absence of denaturant (35,40).

FMN binding rate as a tool to reveal the presence of apoflavodoxin's MG.

FMN is essential for flavodoxin's role in electron transport. This cofactor binds to native apoflavodoxin through a very specific combination and geometry of aromatic and hydrogen bond interactions (41,45). In native apoflavodoxin the flavin-binding site is flexible, whereas upon cofactor incorporation it becomes rather rigid (45,51,52). Cofactor binding not only influences the binding pocket itself but also affects residues far away, resulting in picomolar binding affinity of FMN (45). As a consequence, flavodoxin is thermodynamically very stable and has to first release FMN before it can unfold (41,45). Accordingly, FMN binding to native apoflavodoxin is the last step during folding of flavodoxin. FMN neither acts as a nucleation site for folding nor does it interact with folding intermediates of apoflavodoxin, including the discussed off-pathway MG (41).

Binding of FMN to native apoflavodoxin can be followed by FMN fluorescence, as FMN incorporation into native apoflavodoxin severely quenches its fluorescence. We demonstrate this phenomenon in Fig. 3A, which shows the time-dependent decrease of FMN fluorescence upon cofactor binding to C69A apoflavodoxin in 100 and 10 mM KPPi at 25 °C. The rates of FMN binding at both salt concentrations are identical, because C69A apoflavodoxin is natively folded under these circumstances (39).

In contrast to C69A apoflavodoxin, the rates of FMN binding to F44Y apoflavodoxin in 100 and 10 mM KPPi differ drastically (Fig. 3B). This dissimilarity arises because F44Y apoflavodoxin in 100 mM KPPi is natively folded, whereas in 10 mM

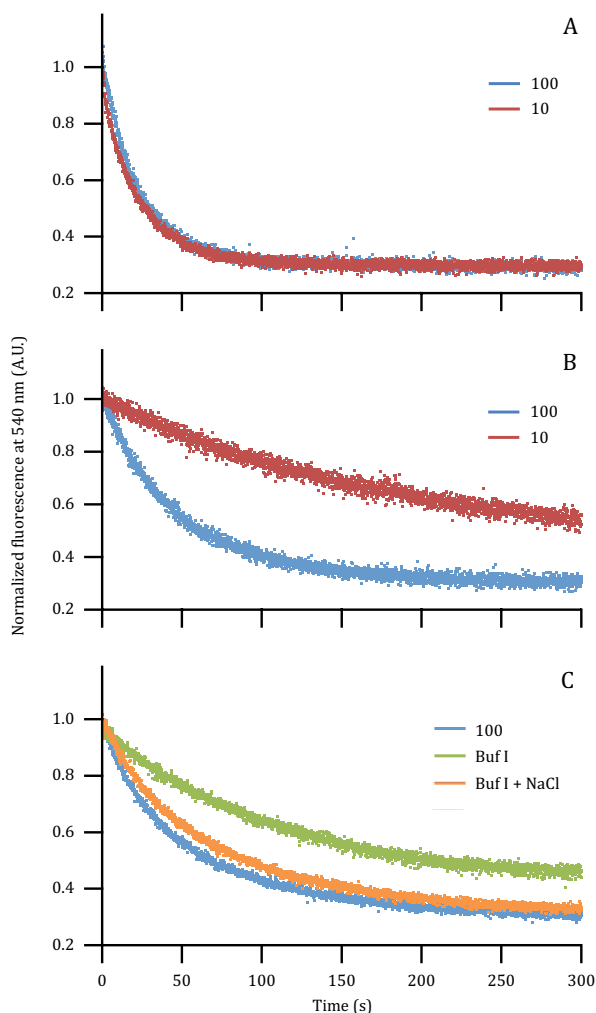


Figure 3 The rate with which FMN binds to apoflavodoxin depends on the presence of molten globular protein and thus on the salt concentration. To follow quenching of flavin fluorescence, FMN is added to protein at 25 °C. Shown are normalized FMN fluorescence traces. *A*, C69A apoflavodoxin in 100 (blue) and 10 (red) mM KPPi, pH 6.0. C69A protein is natively folded under both conditions and thus binds FMN with the same rate. Concentrations for protein and FMN are 140 and 25 nM, respectively. *B*, F44Y apoflavodoxin in 100 (blue) and 10 (red) mM KPPi, pH 6.0. In 100 mM KPPi, F44Y protein is natively folded, whereas in 10 mM KPPi the F44Y variant pre-dominantly populates the MG state, which delays FMN binding as the MG needs to unfold before native apoflavodoxin can form. Concentrations for F44Y apoflavodoxin and FMN are 71 and 25 nM, respectively. *C*, F44Y apoflavodoxin in 100 mM KPPi (blue) and in buffer I containing 290 mM NaCl (orange). Protein is natively folded under these conditions and thus binds FMN with similar rate. F44Y apoflavodoxin in buffer I (green) populates the MG state, thereby delaying FMN binding. Concentrations of F44Y apoflavodoxin and FMN are 62 and 25 nM respectively.

KPPi the protein extensively populates the off-pathway MG. As this MG needs to unfold before productive folding can take place, FMN binding is delayed significantly when compared to F44Y apoflavodoxin in 100 mM KPPi and C69A apoflavodoxin in 10 or 100 mM KPPi, respectively. Thus, ascertaining the rate of FMN binding to apoflavodoxin as a function of ionic strength is a suitable tool to detect the presence of apoflavodoxin's off-pathway MG.

To preserve ribosomal integrity during purification of RNCs we use buffer I (50 mM HEPES-KOH, 100 mM potassium acetate, 15 mM magnesium acetate, 1 mM dithiothreitol (DTT), pH 7.4), which contains magnesium to hold the two ribosomal subunits together. At the ionic strength of buffer I, which is similar to that of the cellular environment, the MG state of F44Y apoflavodoxin populates (Fig. 2A). We therefore assessed the rate of FMN binding to F44Y apoflavodoxin in buffer I (Fig. 3C). When compared to F44Y apoflavodoxin in 100 mM KPPi, the FMN binding rate is considerably reduced in buffer I, indicating the presence of significant amounts of MG. To confirm that ionic strength also affects the equilibrium between MG and native apoflavodoxin in buffer I, 290 mM NaCl was added to buffer I, which would approximate the ionic strength of 100 mM KPPi. Indeed, the FMN binding rate increases upon addition of salt to buffer I and becomes comparable to the rate found for native apoflavodoxin.

Production and purification of RNCs of F44Y flavodoxin *in vivo*. To assess whether MGs form during translation we produced RNCs of C69A and F44Y flavodoxin in *E. coli* (called C69A_{RNC} and F44Y_{RNC}, respectively), using the construct of Fig. 1C. Production at 37 °C and purification of C69A_{RNC} have been described elsewhere (44). In case of F44Y_{RNC} we use the same procedure, except that now RNCs are induced in *E. coli* at either 37 or 15 °C. Lowering temperature to 15 °C maximizes population of the MG state of F44Y apoflavodoxin, whereas at 37 °C the protein is unfolded (Fig. 2B).

While the SecM sequence allows for relatively tight stalling (46), some release of nascent chains happens during their production in *E. coli*, as we observed previously for RNCs of C69A flavodoxin (44). This release is most probably due to the physical force that is exerted as soon as the nascent chain folds (53,54). Binding of FMN to natively folded nascent chains does not dissociate nascent flavodoxin from the ribosome (44). To separate released protein (called C69A_p or F44Y_p) from RNCs (i.e., C69A_{RNC} or F44Y_{RNC}), we utilize size-exclusion chromatography. The absorbance ratio A260/A280 enables distinction between RNCs and released flavodoxin, as this ratio is ~2 for RNCs, while A260/A280 is <1 for released protein. The RNCs elute in the void volume of the Superdex75 column, whereas released protein elutes later. Fig. 4A illustrates this separation of C69A_{RNC} and C69A_p through size exclusion. A similar distribution is observed for the F44Y flavodoxin variant grown at 37 °C (Fig. 4B), showing that also in case of this protein variant release of nascent chains happens. However, an additional third elution peak is now present. This extra peak arises from the protein fragment consisting of the triple StrepII-tag and Smt3-domain (Fig. 4B,C), which we refer to as Smt3-fragment. The identity of the 18-kDa fragment is confirmed by Western blot probed with antibodies against StrepII-tag (Fig. 4D). Because the remaining C-terminal fragments do not contain a StrepII-tag anymore, they are not present after the purification step using a Strep-Tactin column. We previously observed the same degradation product for C-terminally shortened nascent chains of C69A flavodoxin. These shortened proteins are less stable relative to the

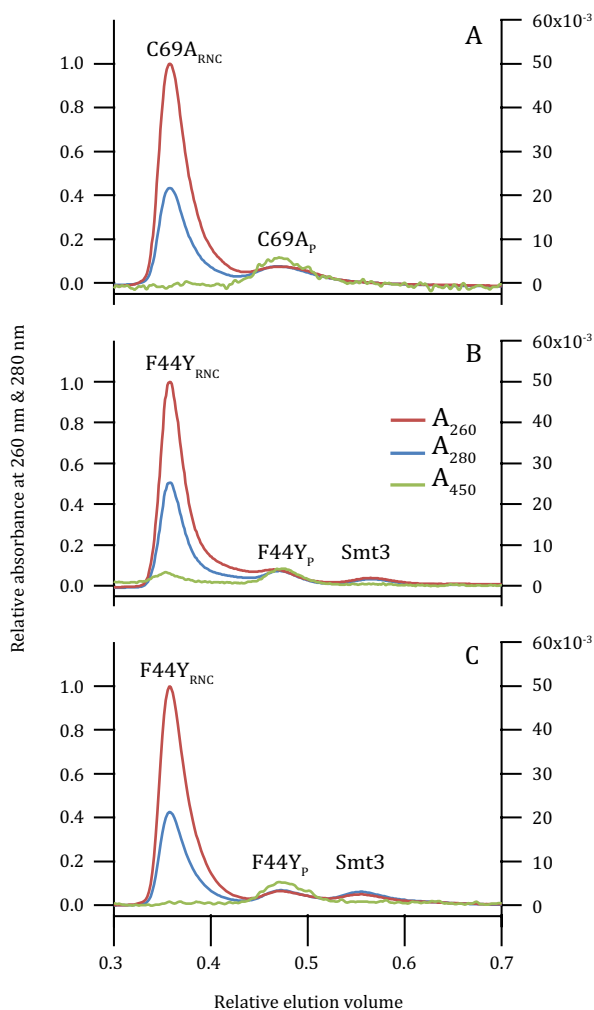


Figure 4 Comparison of release and degradation of C69A and F44Y flavodoxin nascent chains produced in *E. coli* under several growth conditions. Shown are Superdex75 elution profiles, with absorbance at 260 (red), 280 (blue) and 450 nm (green). All absorbance traces are normalized relative to the maximum absorbance at 260 nm of the respective RNCs. Absorbance at 450 nm tracks FMN incorporation. Elution volumes are normalized to the total column volume V_t . All constructs are induced for 2.5 h in *E. coli* Δ *tig::kan* in minimal medium at either 37 °C (A, B) or 15 °C (C). A, use of C69A flavodoxin construct produces C69A_{RNC} and C69A_p. B, use of F44Y flavodoxin construct produces F44Y_{RNC} and F44Y_p. Proteolytic degradation of F44Y protein construct results in the presence of the fragment triple StreptII-tag - Smt3 (labelled Smt3). C, use of F44Y flavodoxin construct at 15 °C produces F44Y_{RNC} and F44Y_p with the same degradation pattern (i.e., Smt3) as observed in (B). D, Western blot of the fractions labelled F44Y_p and Smt3 probed with antibodies against the N-terminal StreptII-tag.

full-length C69A protein (44), and thus Fig. 4B,C indicates decreased stability of also F44Y_{RNC} and/or F44Y_p show decreased stability when produced at 15 or 37 °C. Whether intracellular degradation of the F44Y construct occurs while it is bound to the ribosome and/or while it has been released is unknown. We can only distinguish between both species after their separation using size exclusion chromatography. As by then degradation has already happened, we are unable to determine which species degraded and formed the observed Smt3-fragment.

Production of the F44Y flavodoxin construct at 15 °C does not alter the elution peak arising from the Smt3-fragment (Fig. 4C). Thus, a temperature change of 37 to 15 °C has no effect on the amount of *in vivo* proteolytic degradation of F44Y_{RNC} and/or F44Y_p, suggesting that the respective proteolytic susceptibility of nascent chain and released protein is similar at both temperatures. The ratio between F44Y_p and F44Y_{RNC} is calculated taking into account the differences in extinction coefficients at 280 nm of RNCs and released protein. At both temperatures the ratio is around 70, which is comparable to the ratio we have found previously for a C-terminally shortened nascent chain of C69A flavodoxin (44).

All purified F44Y_p contains FMN. In a previous study we showed that when FMN binds to a released, C-terminally shortened C69A flavodoxin construct this construct is protected against intracellular proteolysis, due to the increased thermodynamic stability conferred by incorporated FMN (44). As a consequence, all of these released constructs were saturated with FMN, because the unstable apo forms are proteolytically degraded in *E. coli* (causing the appearance of the Smt3-fragment). In contrast, for C69A_p, which is full-length protein construct, the FMN content varies between 40 to 100 %, depending on the *E. coli* batch used for purification (44). Due to its increased thermodynamic stability compared to C-terminally shortened constructs, the apo-form of C69A_p is not proteolytically degraded. To verify in the current study whether the elution peak labelled F44Y_p arises from protein saturated with cofactor, we titrated this fraction with FMN (Fig. 5A). Due to quenching of FMN fluorescence upon cofactor binding to apoprotein, the slope of fluorescence data arising from FMN bound to protein is less steep than the slope resulting from free FMN. This characteristic is widely used to demonstrate FMN binding (42,43,45). Fig. 5A reveals similar slopes for FMN fluorescence titration data of buffer and F44Y_p. This observation shows that F44Y_p binds no additional FMN.

To determine the ratio of FMN to apoprotein, we added trichloroacetic acid (TCA) to F44Y_p. This dissociates the cofactor from the protein and after centrifugation we measured FMN fluorescence of the supernatant. This experiment demonstrates equimolar presence of FMN and apoprotein (Fig. 5B). Full saturation of F44Y_p with FMN was seen independent of production temperature (i.e., 15 or 37 °C). Finally, we followed FMN fluorescence in time upon addition of the cofactor to F44Y_p (Fig. 5C). If the sample would contain any F44Y_p in its apo-form, one should detect a decrease in FMN fluorescence upon FMN binding. However, no such decrease was observed and thus no cofactor binding happened.

In conclusion, binding of FMN protects destabilized apoflavodoxin variants against the action of intracellular proteases, and thus all F44Y_p we purify contains FMN.

F44Y_{RNC} isolated from *E. coli* is saturated with FMN. To determine the extent of cofactor incorporation into F44Y_{RNC} we performed a titration of the purified nascent

protein with FMN. Fig. 5A shows that no added FMN binds to these RNCs. TCA precipitation of RNCs shows that this observation arises because RNCs of F44Y flavodoxin contain FMN (Fig. 5B), regardless of whether they are produced by *E. coli* at 15 or 37 °C. Following FMN fluorescence in time upon addition of the cofactor to F44Y_{RNC} (Fig. 5C) shows only a marginal decrease in FMN fluorescence, which corresponds to maximally 5 % of F44Y_{RNC} being in the apo-form. In contrast to the observed nearly full saturation of F44Y_{RNC} with FMN, we previously demonstrated that C69A_{RNC} is purified from *E. coli* as apoprotein. Observation of an altering slope in FMN fluorescence titration data shows that C69A_{RNC} binds FMN (Fig. 5A). This difference in FMN occupation of F44Y_{RNC} and C69A_{RNC} and the observation that F44Y_p is fully saturated with FMN, suggests that when the apo form of F44Y protein is present in *E. coli*, it is not as stable in the cellular environment of *E. coli* as C69A apoprotein is. Therefore,

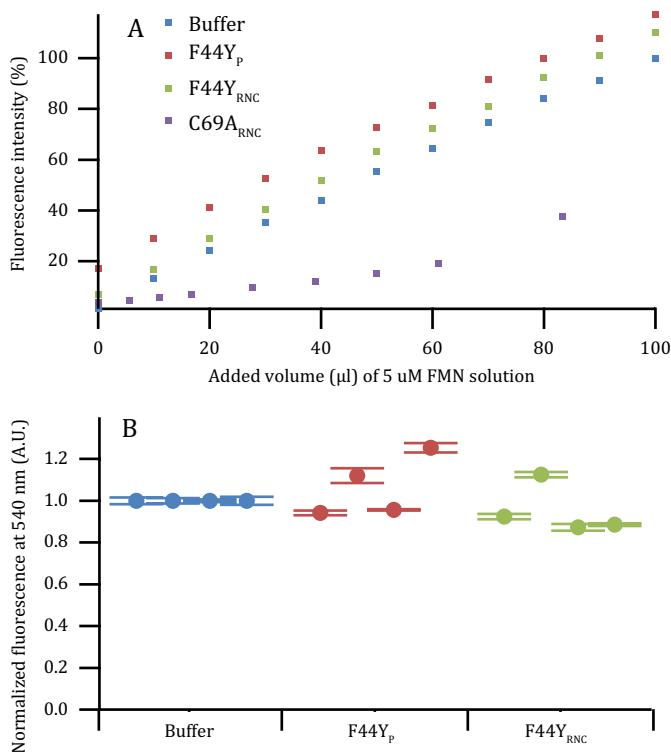
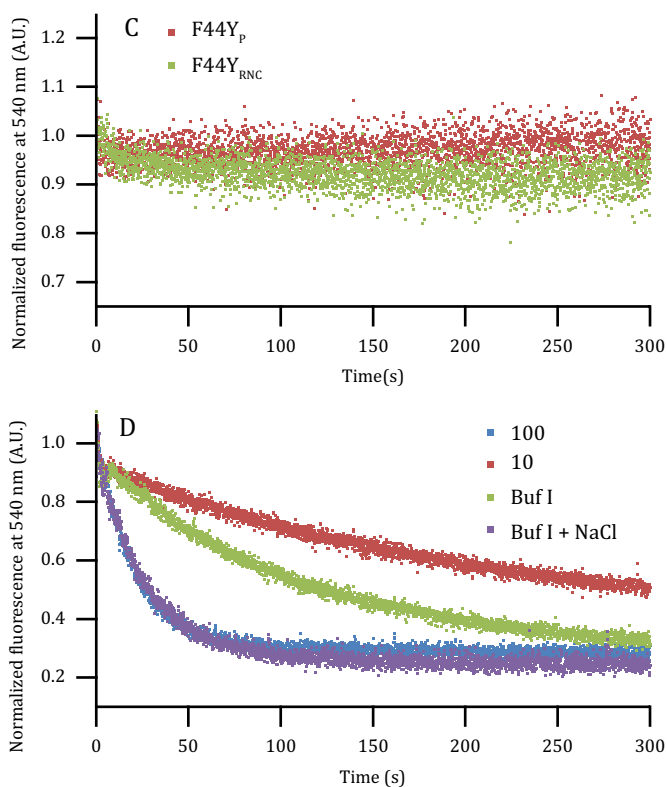


Figure 5 RNCs of F44Y flavodoxin construct and released protein construct isolated from *E. coli* are saturated with FMN. *A*, titration of buffer I, F44Y_p, F44Y_{RNC} and C69A_{RNC} with FMN. Observation of an altering slope in FMN fluorescence titration data reveals cofactor binding, as is the case for C69A_{RNC}. F44Y_p and F44Y_{RNC} show no change in slope of FMN fluorescence titration data. All fluorescence is normalized to the fluorescence of the end-point of the titrated buffer sample. Concentration of protein and RNC is 0.5 μM. Data of C69A_{RNC} are adapted from (44). *B*, fluorescence of FMN in supernatants of F44Y_p (green) and F44Y_{RNC} (red) after TCA precipitation. The average fluorescence of TCA samples of four purifications is shown, together with the corresponding standard deviation. Before addition of TCA, concentration of F44Y_p and F44Y_{RNC} was 50 nM. Fluorescence is normalized to fluorescence of buffer I containing 50 nM FMN (blue).

only the holo-forms of RNCs and released F44Y protein are purified.

The domains of F44Y_p do not affect MG formation of apoflavodoxin. To assess whether the triple Strep-tag, the Smt3-domain, the linker and the SecM sequence (Fig. 1C) influence MG formation of F44Y apoflavodoxin, we removed FMN from purified F44Y_p according to a previously published procedure (44). Subsequently, the FMN binding trace of this refolded apo-F44Y_p was measured in various buffers (Fig. 5D). Comparison of the refolded apo-F44Y_p traces with those of F44Y apoflavodoxin (Fig. 3C) shows that for both proteins a change in ionic strength has a similar effect on FMN binding. Increasing KPPi concentration from 10 to 100 mM, or addition of 290 mM NaCl to buffer I, leads to an increase in FMN binding rate. This observation is due to increased population of the native state of apoflavodoxin. Upon addition of salt (buffer I with 290 mM NaCl) the FMN binding rate of apo-F44Y_p is similar to the



C, FMN binding traces of F44Y_p and F44Y_{RNC} isolated from *E. coli*. FMN fluorescence is normalized. F44Y_p, F44Y_{RNC} and FMN concentrations are 50, 50 and 25 nM, respectively. D, the triple Strep-tag, Smt3-domain, the linker and the SecM sequence of the construct used for RNC production do not influence the formation of the off-pathway MG of apoflavodoxin. Shown are FMN binding traces of apo-F44Y_p in 100 mM KPPi (blue), 10 mM KPPi (red), Buffer I (green) and Buffer I + 290 mM NaCl (purple). F44Y_p and FMN concentrations are 147 and 25 nM, respectively.

rate observed for apo-F44Y_p in 100 mM KPPi, as both proteins are natively folded. Thus, presence of the triple Strep-tag, the Smt3-domain, the linker and the SecM sequence of the construct used for RNC production do not influence formation of the off-pathway MG of apoflavodoxin or impair FMN binding.

The ribosome forces entirely synthesized and fully exposed F44Y apoflavodoxin towards the native state. Because no apo-F44Y_{RNC} could be obtained from *E. coli* due to proteolytic degradation, we adopted an *in vitro* transcription/translation approach to generate nascent F44Y constructs in their apo form. We used a highly pure *in vitro* protein synthesis kit (PURExpress® Δ (aa, tRNA), New England Biolabs) to preclude presence of FMN in the translation reaction. We avoided usage of *E. coli* extracts for *in vitro* protein synthesis, as FMN is difficult to remove from these extracts. The highly purified *in vitro* protein synthesis kit does not contain chaperones that interact with nascent chains, such as Trigger Factor and DnaK. *In vitro* protein synthesis yields a very low percentage of stalled ribosomes, as can be inferred from Fig. 6A. Shown is a Coomassie stained gel of a sample from an *in vitro* translation reaction before and after purification with Strep-Tactin column chromatography. The gel of the sample before purification shows protein bands originating from ribosomal proteins. Also bands arising from non-ribosomal components of the *in vitro* translation kit, like aminoacyl-tRNA synthases, can be seen. These non-ribosomal components can be removed by Strep-Tactin column chromatography. In Fig. 6A the position of the band corresponding to F44Y protein construct is indicated. The identity of this band is verified by Western blots probed with antibodies against flavodoxin and StrepII-tag (Fig. 6B). Before purification of the *in vitro* translation mixture this band is barely visible, as most ribosomes are devoid of stalled nascent chains. After Strep-Tactin chromatography purification the intensity of this band is comparable to those of ribosomal proteins, because in stalled RNCs, all proteins are present in equimolar quantities.

In contrast, F44Y flavodoxin nascent chains produced in *E. coli* are released from the ribosome in considerable quantities, as described above (Fig. 4B, C). Interestingly, the amount of F44Y_p produced *in vitro* is very small, because, as mentioned, Fig. 6A shows that the Coomassie stained bands of F44Y flavodoxin RNCs are all of comparable intensities. In case of nascent chain release, the released construct would have been purified together with the remaining RNCs by Strep-Tactin chromatography. Essentially, F44Y flavodoxin construct would be overproduced and would therefore no longer be equimolar to ribosomal proteins, which is clearly not the case (Fig. 6A).

As discussed, determination of the FMN binding rate enables detection of the presence of apoflavodoxin's MG. This methodology is thus suitable to investigate formation of molten globular nascent apoflavodoxin, provided that ribosomes themselves do not bind FMN. To check the latter, FMN was added to purified *E. coli* ribosomes and its fluorescence followed in time. As can be inferred from Fig. 6C, FMN does not associate with ribosomes, since no change in flavin fluorescence is observed.

We assessed the rate of FMN binding to F44Y_{RNC} in buffer I and in this buffer with 290 mM NaCl. Fig. 6D shows that this increase in salt concentration does not affect the rate of FMN binding to F44Y_{RNC}. In contrast, such salt concentration change of buffer I considerably increases the rate of FMN binding to F44Y apoflavodoxin (Fig. 3C) and to the apo-form of F44Y_p (Fig. 5D), as F44Y apoflavodoxin switches from MG

to native protein. Because F44Y_{RNC} binds FMN rapidly at both salt concentrations, apoflavodoxin in F44Y_{RNC} must be natively folded under both conditions. We note that another implication of this observation is that the sample of F44Y_{RNC} cannot contain released F44Y flavodoxin construct, because if this would be the case the mentioned increase in salt concentration would lead to a change in FMN binding rate. Such a change is not seen in Fig. 6D thereby corroborating the above-mentioned equimolar presence of nascent chain and ribosomal proteins. At low ionic strength, F44Y apoflavodoxin and the apo-form of F44Y_p are molten globular, but, remarkably, apoflavodoxin in F44Y_{RNC} is natively folded. Apparently, the ribosome modulates flavodoxin folding and forces F44Y apoflavodoxin that is entirely synthesized and exposed outside the ribosome, to which it is stalled by an artificial linker containing the SecM sequence, towards the native state. To our knowledge, this is the first time the effect of the ribosome on formation of MGs during protein synthesis has been determined.

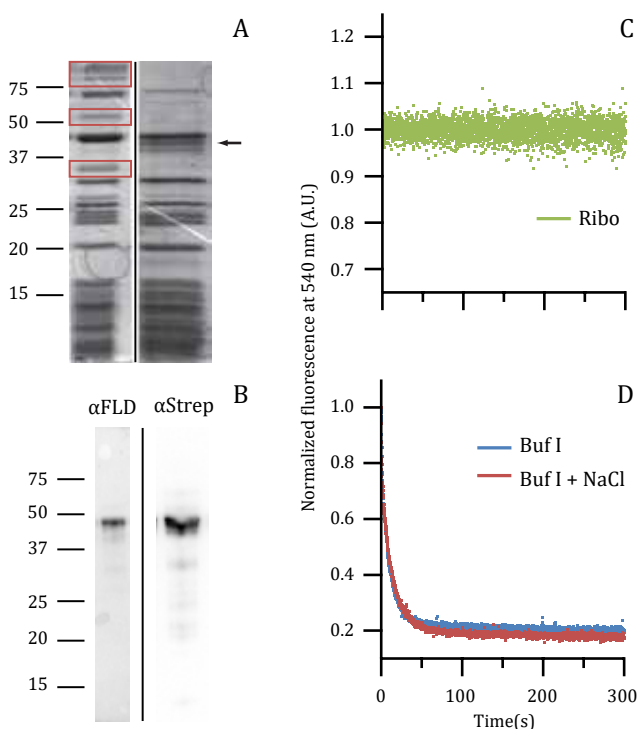


Figure 6 Apoflavodoxin in F44Y_{RNC} is natively folded at low and high salt concentrations. *A*, Coomassie-stained gels of samples taken from an *in vitro* translation reaction of F44Y flavodoxin construct before (1) and after (2) purification with Strep-Tactin affinity chromatography. An arrow labels the position of F44Y flavodoxin construct and red boxes enclose the bands of the non-ribosomal components of the *in vitro* translation reaction. *B*, Western blots of sample (2), probed with antibodies against flavodoxin (α FLD) or N-terminal StrepII-tag (α Strep), respectively. *C*, FMN fluorescence shows that ribosomes do not bind FMN. Concentrations of ribosomes and FMN are 50 and 25 nM, respectively. *D*, FMN binding traces of F44Y_{RNC} in buffer I (blue) and in this buffer with 290 mM NaCl (red). Concentrations of F44Y_{RNC} and FMN are 50 and 25 nM, respectively.

DISCUSSION

Determination of the rate of flavin binding is an innovative approach that makes it possible to investigate the presence of MGs during translation of flavin binding proteins. Using this methodology we demonstrated that the ribosome forces nascent, fully exposed, F44Y apoflavodoxin towards the native state. Confinement of MG formation during translation is an important observation that emphasizes differences between folding *in vivo* and *in vitro*.

Due to the presence of ribosomal RNA surrounding the end of the exit tunnel, this part of the ribosome area has a considerable negative surface charge (55). Though some proteins only seem to randomly interact with the ribosomal surface (56), others interact to such an extent that local motions of compact domains are constrained (57) or folding/unfolding transitions are slowed down (58). These interactions likely arise due to attraction of positively charged residues and simultaneous repulsion of negatively charged residues of the nascent chain by the negatively charged ribosomal outer surface. Because the population of natively folded apoflavodoxin within F44Y_{RNC} at low salt concentration (i.e., in buffer I) is higher than is the case for F44Y apoflavodoxin in buffer I, the ribosome restrains formation of the off-pathway MG. Upon release of the nascent chain this effect is negated, as F44Y_p, a construct which was stalled to the ribosome before its release due to physical exertion, can form the MG state upon lowering ionic strength (Fig. 5D). It is tempting to speculate that the ribosome potentially mimics the effects of high ionic strength, thereby forcing molten globular apoflavodoxin to the native state. Native apoflavodoxin has a net charge of -13 at neutral pH (59). Possibly, the conformational space of unfolded nascent apoflavodoxin is restricted due to electrostatic repulsion of the nascent chain by the ribosomal surface, leading to entropic stabilization of native protein at physiological ionic strength.

We note that the RNC construct we use creates a somewhat artificial situation compared to translation of solely the flavodoxin gene. Due to presence of the SecM sequence and attached linker (Fig. 1C), RNCs are produced with the entire flavodoxin domain exposed outside the ribosome. In contrast, when a ribosome reaches the stop codon and terminates the nascent chain during translation of solely the flavodoxin domain, approximately 30 to 40 residues of the protein are still buried in the exit tunnel. These residues need to traverse this tunnel before full-length protein emerges from the ribosome. We showed based on investigating C-terminally shortened flavodoxin RNCs that exposure of the five C-terminal residues of flavodoxin is essential for the nascent chain to be able to attain its native fold (44). Thus, while the 30 to 40 C-terminal residues of flavodoxin are inside the exit tunnel, the part of the protein that is already outside the exit tunnel is unable to fold natively.

It would be interesting to verify whether our finding that the ribosome forces flavodoxin towards the native state also applies to the folding of flavodoxin-like domains within multi-domain proteins like cytochrome P450 reductase and nitric oxide synthase (60,61). As each domain of such a protein is translated the nascent chain remains tethered to the ribosome, thus providing the ribosome ample opportunity to influence the folding of an already translated flavodoxin-like domain.

ACKNOWLEDGEMENTS

This work was supported by an ECHO (711.011.007) grant of NWO to CPMvM.

REFERENCES

1. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnel, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195
2. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19
3. Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: polypeptide conformations on competing pathways. *Arch. Biochem. Biophys.* **469**, 100-117
4. Dobson, C. M. (2003) Protein folding and misfolding. *Nature* **426**, 884-890
5. Ohgushi, M., and Wada, A. (1983) 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21-24
6. Kuroda, Y., Kidokoro, S., and Wada, A. (1992) Thermodynamic characterization of cytochrome c at low pH. Observation of the molten globule state and of the cold denaturation process. *J. Mol. Biol.* **223**, 1139-1153
7. Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995) Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5446-5450
8. Shastry, M. C., and Udgaonkar, J. B. (1995) The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates. *J. Mol. Biol.* **247**, 1013-1027
9. Arai, M., and Kuwajima, K. (2000) Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209-282
10. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., and Razgulyaev, O. I. (1990) Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **262**, 20-24
11. Bollen, Y. J. M., and van Mierlo, C. P. M. (2005) Protein topology affects the appearance of intermediates during the folding of proteins with a flavodoxin-like fold. *Biophys. Chem.* **114**, 181-189
12. Bhattacharyya, S., and Varadarajan, R. (2013) Packing in molten globules and native states. *Curr. Opin. Struct. Biol.* **23**, 11-21
13. Baldwin, R. L., and Rose, G. D. (2013) Molten globules, entropy-driven conformational change and protein folding. *Curr. Opin. Struct. Biol.* **23**, 4-10
14. van der Goot, F. G., González-Mañas, J. M., Lakey, J. H., and Pattus, F. (1991) A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature* **354**, 408-410
15. van der Goot, F. G., Lakey, J. H., and Pattus, F. (1992) The molten globule intermediate for protein insertion or translocation through membranes. *Trends Cell Biol.* **2**, 343-348
16. Lally, E. T., Hill, R. B., Kieba, I. R., and Korostoff, J. (1999) The interaction between RTX toxins and target cells. *Trends Microbiol.* **7**, 356-361
17. Ellis, J. P., Bakke, C. K., Kirchdoerfer, R. N., Jungbauer, L. M., and Cavagnero, S. (2008) Chain dynamics of nascent polypeptides emerging from the ribosome. *ACS Chem. Biol.* **3**, 555-566

18. Evans, M. S., Sander, I. M., and Clark, P. L. (2008) Cotranslational folding promotes β -helix formation and avoids aggregation *in vivo*. *J. Mol. Biol.* **383**, 683-692
19. Cabrita, L. D., Hsu, S.-T. D., Launay, H., Dobson, C. M., and Christodoulou, J. (2009) Probing ribosome-nascent chain complexes produced *in vivo* by NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22239-22244
20. Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012) Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* **48**, 63-74
21. Kelkar, D. A., Khushoo, A., Yang, Z., and Skach, W. R. (2012) Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. *J. Biol. Chem.* **287**, 2568-2578
22. Lamprou, P., Kempe, D., Katranidis, A., Büldt, G., and Fitter, J. (2014) Nano-second dynamics of calmodulin and ribosome-bound nascent chains studied by time-resolved fluorescence anisotropy. *ChemBioChem* **15**, 977-985
23. Holtkamp, W., Kokic, G., Jäger, M., Mittelstaet, J., Komar, A. A., and Rodnina, M. V. (2015) Cotranslational protein folding on the ribosome monitored in real time. *Science* **350**, 1104-1107
24. Jaenicke, R., and Seckler, R. (1997) Protein misassembly *in vitro*. *Adv. Protein Chem.* **50**, 1-59
25. Hawe, A., Sutter, M., and Jiskoot, W. (2008) Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* **25**, 1487-1499
26. Engelhard, M., and Evans, P. A. (1996) Experimental investigation of side-chain interactions in early folding intermediates. *Fold. Des.* **1**, R31-R37
27. Plaxco, K. W., and Dobson, C. M. (1996) Time-resolved biophysical methods in the study of protein folding. *Curr. Opin. Struct. Biol.* **6**, 630-636
28. Chakraborty, S., Ittah, V., Bai, P., Luo, L., Haas, E., and Peng, Z. (2001) Structure and dynamics of the α -lactalbumin molten globule: fluorescence studies using proteins containing a single tryptophan residue. *Biochemistry* **40**, 7228-7238
29. Alagaratnam, S., van Pouderoyen, G., Pijning, T., Dijkstra, B. W., Cavazzini, D., Rossi, G. L., van Dongen, W. M. A. M., van Mierlo, C. P. M., van Berkel, W. J. H., and Canters, G. W. (2005) A crystallographic study of Cys69Ala flavodoxin II from *Azotobacter vinelandii*: Structural determinants of redox potential. *Protein Sci.* **14**, 2284-2295
30. Caetano-Anollés, G., Kim, H. S., and Mittenthal, J. E. (2007) The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9358-9363
31. Bollen, Y. J. M., Sanchéz, I. E., and van Mierlo, C. P. M. (2004) Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin. *Biochemistry* **43**, 10475-10489
32. Bollen, Y. J. M., Kamphuis, M. B., and van Mierlo, C. P. M. (2006) The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4095-4100
33. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2008) Extensive formation of off-pathway species during folding of an α - β parallel protein is due to

docking of (non)native structure elements in unfolded molecules. *J. Am. Chem. Soc.* **130**, 16914-16920

34. Visser, N. V., Westphal, A. H., van Hoek, A., van Mierlo, C. P. M., Visser, A. J. W. G., and van Amerongen, H. (2008) Tryptophan-tryptophan energy migration as a tool to follow apoflavodoxin folding. *Biophys. J.* **95**, 2462-2469

35. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2009) Noncooperative formation of the off-pathway molten globule during folding of the α - β parallel protein apoflavodoxin. *J. Am. Chem. Soc.* **131**, 2739-2746

36. Lindhoud, S., Westphal, A. H., Visser, A. J. W. G., Borst, J. W., and van Mierlo, C. P. M. (2012) Fluorescence of Alexa fluor dye tracks protein folding. *PLoS ONE* **10**.1371/journal.pone.0046838

37. Lindhoud, S., Pirchi, M., Westphal, A. H., Haran, G., and van Mierlo, C. P. M. (2015) Gradual folding of an off-pathway molten globule detected at the single-molecule level. *J. Mol. Biol.* **427**, 3148-3157

38. Engel, R., Westphal, A. H., Huberts, D. H., Nabuurs, S. M., Lindhoud, S., Visser, A. J., and van Mierlo, C. P. (2008) Macromolecular crowding compacts unfolded apoflavodoxin and causes severe aggregation of the off-pathway intermediate during apoflavodoxin folding. *J. Biol. Chem.* **283**, 27383-27394

39. Nabuurs, S. M., Westphal, A. H., aan den Toorn, M., Lindhoud, S., and van Mierlo, C. P. M. (2009) Topological switching between an α - β parallel protein and a remarkably helical molten globule. *J. Am. Chem. Soc.* **131**, 8290-8295

40. Nabuurs, S. M., and van Mierlo, C. P. M. (2010) Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of α - β parallel protein flavodoxin. *J. Biol. Chem.* **285**, 4165-4172

41. Bollen, Y. J. M., Nabuurs, S. M., van Berkel, W. J. H., and van Mierlo, C. P. M. (2005) Last in, first out. The role of cofactor binding in flavodoxin folding. *J. Biol. Chem.* **280**, 7836-7844

42. Edmondson, D. E., and Tollin, G. (1971) Flavin-protein interactions and the redox properties of the Shethna flavoprotein. *Biochemistry* **10**, 133-145

43. Taylor, M. F., Boylan, M. H., and Edmondson, D. E. (1990) *Azotobacter vinelandii* flavodoxin: purification and properties of the recombinant, dephospho form expressed in *Escherichia coli*. *Biochemistry* **29**, 6911-6918

44. Houwman, J. A., Westphal, A. H., van Berkel, W. J. H., and van Mierlo, C. P. M. (2015) Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome. *Biochim. Biophys. Acta* **1854**, 1317-1324

45. Bollen, Y. J. M., Westphal, A. H., Lindhoud, S., van Berkel, W. J. H., and van Mierlo, C. P. M. (2012) Distant residues mediate picomolar binding affinity of a protein cofactor. *Nature Comm.* **10**.1038/ncomms2010

46. Nakatogawa, H., and Ito, K. (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**, 629-636

47. Schaffitzel, C., and Ban, N. (2007) Generation of ribosome nascent chain complexes for structural and functional studies. *J. Struct. Biol.* **158**, 463-471

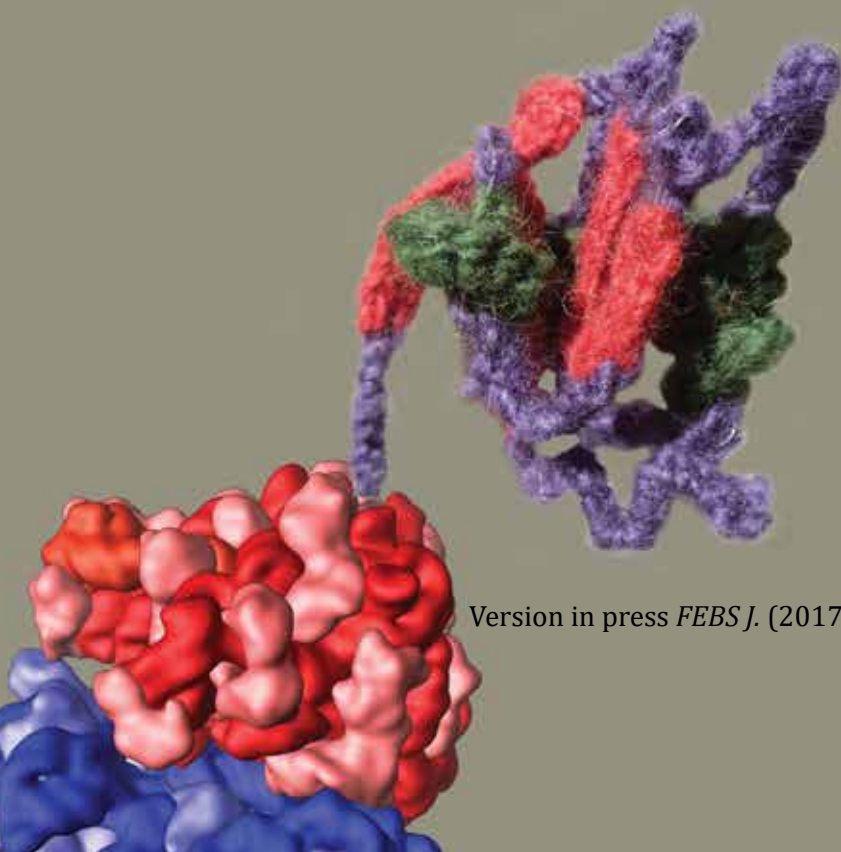
48. Timasheff, S. N., and Arakawa, T. (1989) *Stabilization of protein structure by solvents*, IRL Press, Oxford

49. Steensma, E., Heering, H. A., Hagen, W. R., and Van Mierlo, C. P. M. (1996) Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur. J. Biochem.* **235**, 167-172
50. van Mierlo, C. P. M., van Dongen, W. M. A. M., Vergeldt, F., van Berkel, W. J. H., and Steensma, E. (1998) The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate. *Protein Sci.* **7**, 2331-2344
51. Lindhoud, S., van den Berg, W. A. M., van den Heuvel, R. H. H., Heck, A. J. R., van Mierlo, C. P. M., and van Berkel, W. J. H. (2012) Cofactor binding protects flavodoxin against oxidative stress. *PLoS ONE* 10.1371/journal.pone.0041363.t002
52. Steensma, E., and van Mierlo, C. P. M. (1998) Structural characterisation of apoflavodoxin shows that the location of the stable nucleus differs among proteins with a flavodoxin-like topology. *J. Mol. Biol.* **282**, 653-666
53. Cymer, F., Hedman, R., Ismail, N., and von Heijne, G. (2015) Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. *J. Biol. Chem.* **290**, 10208-10215
54. Goldman, D. H., Kaiser, C. M., Milin, A., Righini, M., Tinoco, I., Jr., and Bustamante, C. (2015) Ribosome. Mechanical force releases nascent chain-mediated ribosome arrest *in vitro* and *in vivo*. *Science* **348**, 457-460
55. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905-920
56. Eichmann, C., Preissler, S., Riek, R., and Deuerling, E. (2010) Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9111-9116
57. Ellis, J. P., Culviner, P. H., and Cavagnero, S. (2009) Confined dynamics of a ribosome-bound nascent globin: Cone angle analysis of fluorescence depolarization decays in the presence of two local motions. *Protein Sci.* **18**, 2003-2015
58. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I., Jr., and Bustamante, C. (2011) The ribosome modulates nascent protein folding. *Science* **334**, 1723-1727
59. Klugkist, J., Voorberg, J., Haaker, H., and Veeger, C. (1986) Characterization of three different flavodoxins from *Azotobacter vinelandii*. *Eur. J. Biochem.* **155**, 33-40
60. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8411-8416
61. Garcin, E. D., Bruns, C. M., Lloyd, S. J., Hosfield, D. J., Tiso, M., Gachhui, R., Stuehr, D. J., Tainer, J. A., and Getzoff, E. D. (2004) Structural basis for isozyme-specific regulation of electron transfer in nitric-oxide synthase. *J. Biol. Chem.* **279**, 37918-37927

Folding of proteins with a flavodoxin-like architecture

Joseline A. Houwman and Carlo P.M. van Mierlo

Laboratory of Biochemistry, Wageningen University, the Netherlands



Version in press *FEBS J.* (2017)

The flavodoxin-like fold is a protein-architecture that can be traced back to the universal ancestor of the three kingdoms of life. Many proteins share this α - β parallel topology and hence it is highly relevant to illuminate how they fold. Here, we review experiments and simulations concerning the folding of flavodoxins and CheY-like proteins, which share the flavodoxin-like fold. These polypeptides tend to temporarily misfold during unassisted folding to their functionally active forms. This susceptibility to frustration is caused by the more rapid formation of an α -helix compared to a β -sheet, particularly when a parallel β -sheet is involved. As a result, flavodoxin-like proteins form intermediates that are off-pathway to native protein and several of these species are molten globules. Experiments suggest that the off-pathway species are of helical nature and that flavodoxin-like proteins have a non-conserved transition state that determines the rate of productive folding. Folding of flavodoxin from *Azotobacter vinelandii* has been investigated extensively, enabling a schematic construction of its folding energy landscape. It is the only flavodoxin-like protein of which cotranslational folding has been probed. New insights that emphasize differences between *in vivo* and *in vitro* folding energy landscapes are emerging: the ribosome modulates molten globule formation in nascent apoflavodoxin and forces this polypeptide towards the native state.

INTRODUCTION: FOLDING OF A-B PARALLEL PROTEINS, WITH A FOCUS ON THE FLAVODOXIN-LIKE FOLD

Protein structures are classified into topologies according to the relationship between sequential ordering of secondary structure elements and their spatial organization. Proteins that share the α - β parallel topology are characterised by a central parallel β -sheet surrounded by α -helices on both sides. This topology is a common structural design in nature. Various functionally and sequentially unrelated proteins with such a topology temporarily misfold during unassisted *in vitro* folding to their native state, yielding off-pathway folding intermediates. These species have been observed for α - β parallel proteins with for example TIM barrel, α/β hydrolase, P-loop containing nucleoside triphosphate hydrolase, or flavodoxin-like folds (1-10).

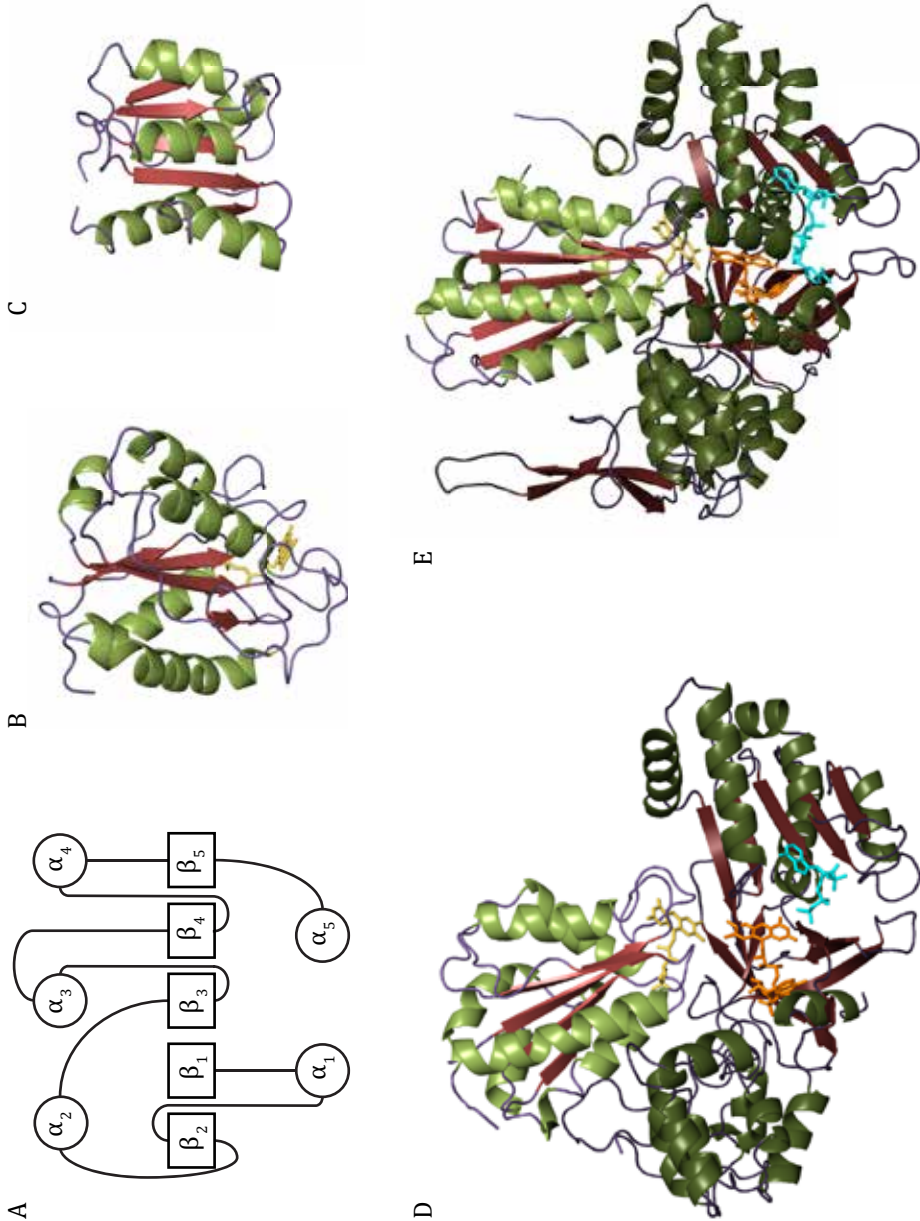
In this review we discuss the folding of proteins with a flavodoxin-like fold (Fig. 1A), which includes the families of flavodoxin- and CheY-related proteins, according to Structural Classification of Proteins (SCOP) in the Protein Data Bank (PDB). Besides other flavodoxins, we particularly focus on how flavodoxin from *A. vinelandii* (Fig. 1B) folds, as its folding behaviour is one of the best characterised to date. We also evaluate folding of the following CheY-related proteins: the bacterial response regulator CheY (Fig. 1C), the sporulation response regulator Spo0F, and the N-terminal receiver domain of nitrogen regulator protein NtrC (NT-NtrC).

FLAVODOXINS: FUNCTION, EVOLUTION AND CATEGORIZATION

Flavodoxins are monomeric, single domain flavoproteins that are involved in electron transfer between various physiological redox partners. To shuttle electrons, flavodoxins contain a non-covalently bound flavin mononucleotide (FMN). Phylogenomic analysis shows that flavodoxin is evolutionary one of the earliest proteins, just as ferredoxin is. Both proteins are rather promiscuous in their choice of redox partners and evolved in the anaerobic environment, preceding the advent of oxygenic photosynthesis (11). Before O_2 levels rose, ferredoxin was the obligatory redox protein. Once O_2 started to form, and bioavailability of iron became limiting, there was an intense selective pressure to replace oxidant-sensitive, iron-dependent proteins by oxidant-resistant, iron-free iso-functional counterparts. As flavodoxin satisfies both requirements, it temporarily took over ferredoxin's role (11,12).

The flavodoxin-like fold represents an architecture that can be traced back to the universal ancestor of all three kingdoms of life (11). Despite its ancient origin and low susceptibility to O_2 , flavodoxin itself has been lost in most eukaryotic lineages, including the entire plant and animal kingdom (12). Cyanobacteria and algae thriving in iron-deficient oceanic environments usually contain the flavodoxin gene. It is assumed that plants evolved from macro-algae that already lacked flavodoxin, because they thrived in an iron-rich, freshwater habitat with no need to back up ferredoxin functions and therefore no selective pressure to keep the flavodoxin gene (11). Flavodoxin is not found in mitochondria, though it is present in alphaproteobacteria, the suspected original endosymbiont. When both flavodoxin and ferredoxin are present in the same genome, like in many prokaryotes and algae, stress conditions such as iron starvation induce flavodoxin expression. Because electron transfer is essential for survival, flavodoxin largely replaces the roles of ferredoxin under these circumstances (11). Though flavodoxin is not found in the higher kingdoms of life, it has become incorporated in important multi-domain enzymes such as cytochrome

Figure 1 Schematic representations of flavodoxin-like fold containing proteins. FMN, FAD and NADPH are coloured yellow, orange and cyan respectively. Domains that are not flavodoxin-like are coloured darkly. *A*, Topology of the flavodoxin-like fold. α -Helices are represented as circles and β -strands as squares. In long-chain flavodoxins the β 5-strand is split into β 5a and β 5b. *B*, Flavodoxin from *A. vinelandii* (PDB entry 1YOB (152)). *C*, CheY from *E. coli* (PDB entry 3CHY) (198). *D*, NADPH cytochrome P450 reductase from *Rattus norvegicus* (PDB entry 1AMO (13)). *E*, Monomer of the dimeric reductase module of nitric oxide synthase from *Rattus norvegicus* (PDB entry 1TLL, strand A (14)).



P450 reductase (13) and nitric oxide synthase (14) (Figs. 1D,E).

Depending on methods used to determine evolution of protein architectures, the flavodoxin-like fold (c.23) is the sixth fold that arose (15,16). Interestingly, six out of the nine most ancient folds belong to the α/β class of proteins and more than half of the studied molecular fossils contain cofactors. Indeed, a flavodoxin made of only “early” amino acid residues is a functional protein with effective electron transfer capacity (17). Homology exists between the flavodoxin-like fold and the $(\beta/\alpha)_8$ -barrel fold (c.1 or TIM-barrel) (18), which is considered to be the second oldest protein fold (16). The flavodoxin-like fold is one of two preferred folds that bind FMN, and the other is the TIM-barrel (19).

Flavodoxins are categorized as short- or long-chain ones, depending on the presence of an insertion of about 20 amino acid residues in the last (i.e., fifth) β -strand (Figs. 1A, 2). The function of this loop has not yet been determined, though it is suspected to be involved in binding of partner proteins (20). The loop also seems to increase the binding affinity of FMN (20) and it is speculated that it is involved in formation of folding intermediates (21). Phylogenetic analyses suggests that short-chain flavodoxins derived from long-chain ones (22), although resolving this issue is complicated by several horizontal gene-transfer events between organisms (12). Only short-chain flavodoxins are found in firmicutes (Gram-positive bacteria), whereas cyanobacteria and algae exclusively synthesize long-chain flavodoxins (23). Gram-negative bacteria can contain flavodoxins from either category and may even contain both short- and long-chain ones. In *Escherichia coli*, for example, there are four genes predicted to code for flavodoxins. Two of these proteins belong to the long-chain category (i.e., FldA and FldB) and the other to the short-chain one (i.e., MioC and YqcA) (24).

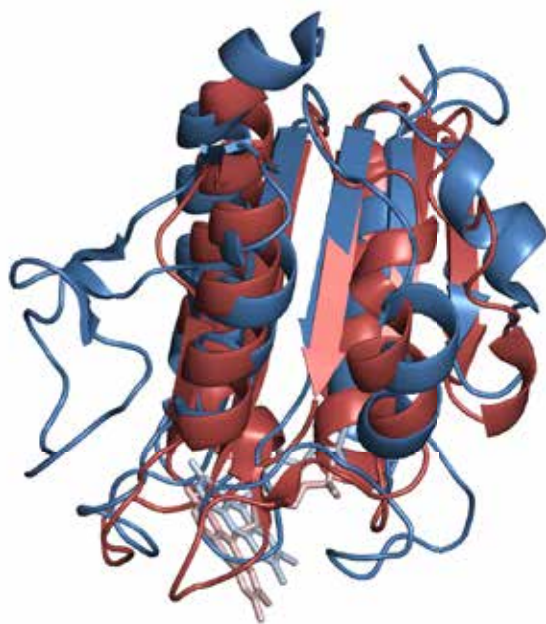


Figure 2 Three-dimensional structures of long- and short-chain flavodoxins. Shown are a long-chain protein from *A. vine-landii* (blue, PDB entry 1YOB) and a short-chain one from *D. vulgaris* (cyan, PDB entry 1J8Q). The loop shown on the left divides the fifth β -strand of the long-chain flavodoxin into $\beta 5a$ and $\beta 5b$. FMN is shown in stick representation.

Because the flavodoxin-like fold is widespread in the protein universe, illumination of its folding is highly relevant. Folding of flavodoxins from *Anabaena*, *A. vinelandii*, *Desulfovibrio desulfuricans*, *D. vulgaris*, and *Helicobacter pylori* has been investigated with differing levels of depth. Of these proteins, the ones from *Desulfovibrio* are short-chain flavodoxins.

ENERGY LANDSCAPES OF PROTEIN FOLDING AND ROLE OF INTERMEDIATES

How are proteins able to form a wide variety of three-dimensional structures from differing amino acid sequences? How can similar sequences fold to disparate structures? Two principles are essential for solving these challenges (25-29): Anfinsen's thermodynamic hypothesis and Levinthal's paradox. The thermodynamic hypothesis states that, at least for small globular proteins, the native structure is only determined by its amino acid sequence (25). In the protein's relevant physiological environment the native structure is a unique, stable conformation at the kinetically accessible minimum of free energy. Levinthal calculated the time it would take for a 100-residue protein to randomly search all possible conformations and found that this would last longer than the lifetime of the universe (30,31). Paradoxically, most small proteins fold to their native structure within micro- to milli-seconds, and even folding of large multi-domain proteins rarely exceeds folding times of minutes (32). This issue can be resolved if the conformational search is not random, but involves intermediates in which local interactions form that direct the subsequent folding of the polypeptide chain and thereby accelerate the folding reaction (30,31). Both Anfinsen's thermodynamic hypothesis and Levinthal's paradox are integrated in the funnel-like energy landscape for protein folding (26,33).

All possible conformations of a single polypeptide can be visualized forming a folding funnel (26,33). This funnel depicts the free energy of folding as a function of conformational entropy. The rim of the funnel is the starting area for folding (i.e., the unfolded state), which is not just one defined protein structure. Instead, the unfolded state comprises a conformational ensemble of rapidly interconverting unfolded structures, which is characterized by large entropy. Starting from the unfolded state, the peptide chain searches for the native state, which is the lowest point in the folding energy landscape, as the idealized, smooth funnel of Fig. 3A shows (25). A myriad of interactions is involved in folding an unfolded protein, such as establishment of hydrophobic contacts, van der Waals interactions and formation of intra-molecular electrostatic connections and hydrogen bonds. Lowering of the chain entropy also plays a role. Proteins appear to firstly develop secondary structures in the chain (such as turns and helices), which may interact transiently, followed by growth into more global structures (29).

Formation of intermediates gives rise to a rugged folding surface, reflecting the presence of local energetic minima (Fig. 3B). Most intermediates are transient and only few of them have sufficient stability to be detectable. These species can be either on- or off-pathway to the native structure. On-pathway ones pre-dominantly contain native-like interactions in their structured parts. In contrast, off-pathway species have significant non-native, misfolded structure and are trapped in the corresponding energetic minima. To return to the productive folding ensemble, these trapped species have to overcome significant energy barriers, which can be achieved by their (partial) unfolding. Intermediates that are off-pathway to the productive folding route can be visualized as residing in a "moated" landscape (Fig. 3C). Folding

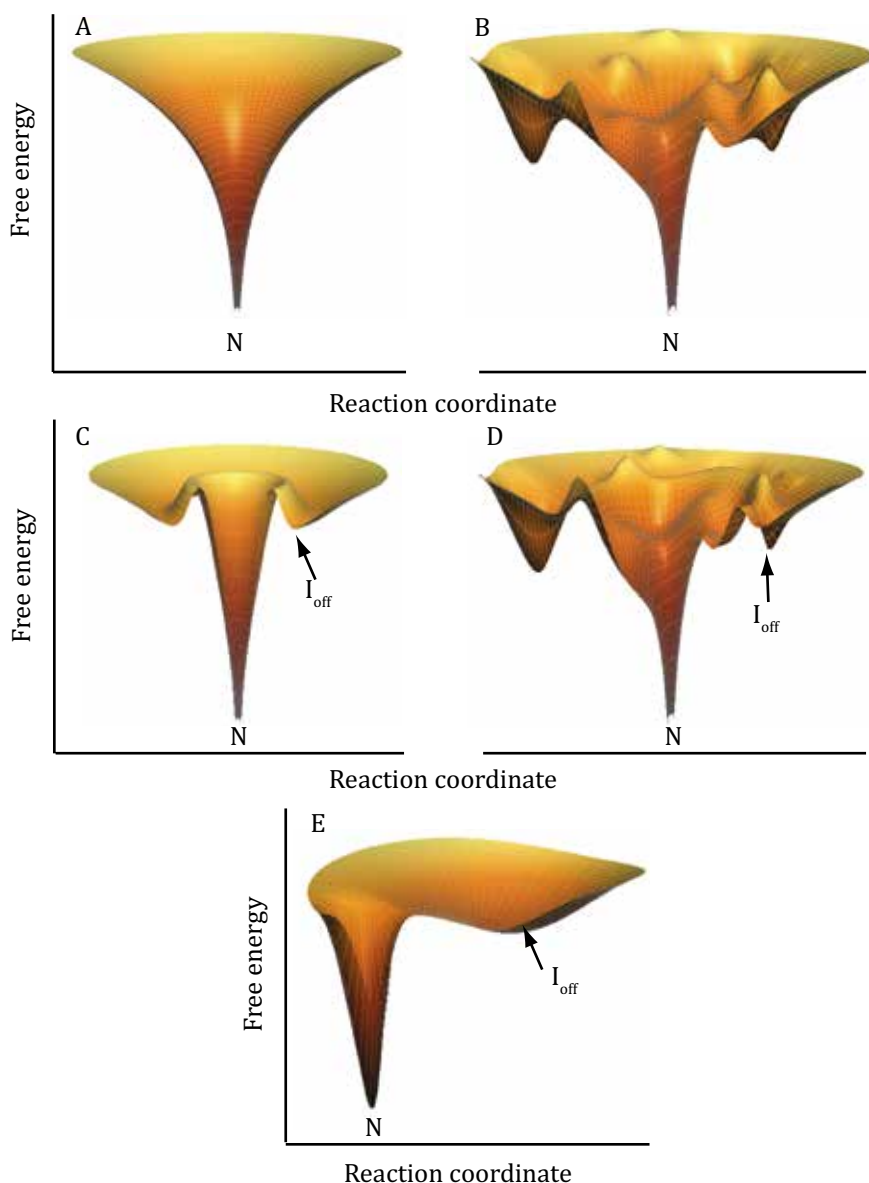


Figure 3 Energy surfaces of protein folding. Free energies are shown vertically and reaction coordinates horizontally. Unfolded protein structures reside at the rim of each folding funnel. N indicates native protein and I_{off} signifies off-pathway folding intermediate. A, Example of a smooth energy landscape, through which a polypeptide chain is effectively funnelled to its native structure. B, In a rugged folding funnel a protein can form intermediates on its way to the native state. C, In a moated folding funnel a protein can become kinetically trapped in a trench, resulting in formation of an off-pathway intermediate. D, Example of a folding funnel that is rugged and contains a moat, i.e., it contains on- and off-pathway intermediates. E, Simplified alternative energy landscape of a protein that can temporarily form an off-pathway intermediate.

of certain proteins involves both on- and off-pathway species, because the corresponding energy landscape is complex (Fig. 3D). It is rugged and contains a moat. Figure 3D suggests that both intermediates are on direct routes to the native state, as they reside within the same funnel. To better visualise the distinct character of both species, Fig. 3E positions the off-pathway intermediate on the bottom of a separate trough that is connected to an on-pathway funnel.

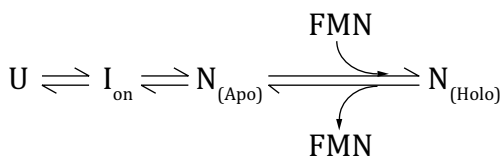
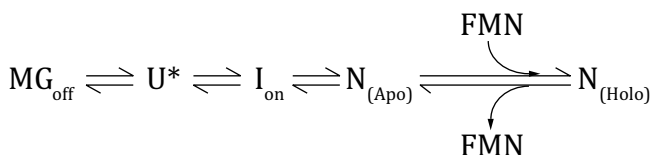
MOLTEN GLOBULAR FOLDING INTERMEDIATES

An important folding intermediate is the molten globule (MG), which was first described for the protein α -lactalbumin (34-36). MGs are characterized by a substantial amount of secondary structure, yet without the tertiary side-chain packing of natively folded protein. Furthermore, MGs are relatively compact (typical radius increase compared to native protein of about 10 to 30 %), possess a loosely packed hydrophobic core and expose hydrophobic patches to the solvent (36,37). MGs are prone to aggregation and consequently are implicated in various diseases (38). *In vitro*, several proteins form MGs under mildly alkaline or acidic conditions, including α -lactalbumin, apomyoglobin and cytochrome c (34,35,39-44). These MGs all contain native-like secondary structure and packing, causing them to be on-pathway to the native state. In contrast, off-pathway MGs contain non-native secondary structure and/or packing. Nowadays, many proteins are thought to fold via MGs (1,45-47). For example, MG species are necessary during insertion of proteins into membranes or during translocation (48-50). Such insertions often happen cotranslationally, i.e., while the ribosome synthesizes the polypeptide concerned. Proteins can fold cotranslationally and sample intermediate folding states (51-57), which might include MGs.

EXPLORATION OF THE FOLDING OF FLAVODOXIN FROM *A. VINELANDII*

We first consider the folding of flavodoxin from *A. vinelandii*. This bacterium contains three flavodoxins, of which the long-chain flavodoxin II is most abundant under nitrogen fixing conditions (58). Folding of both the apo- and the holo-form of this 179-residue protein has been probed by using guanidinium hydrochloride (GuHCl) as denaturant. A wide variety of techniques were used, including (time-resolved) fluorescence, far-UV circular dichroism (CD), nuclear magnetic resonance (NMR) spectroscopy, hydrogen/deuterium exchange, electron paramagnetic resonance spectroscopy, (single-molecule) Förster resonance energy transfer, stopped-flow, paramagnetic relaxation, and FMN binding kinetics (5,59-79). These experiments show that apoflavodoxin folds autonomously and spontaneously to its native state (5) and that the subsequent binding of FMN is the last step in folding of flavodoxin (65). The three-dimensional structures of both native proteins are nearly identical, except for considerable flexibility in the flavin-binding region of apoflavodoxin (61). This flexibility is probably required for apoflavodoxin to capture FMN out of solution. After binding FMN with nanomolar affinity, the protein relaxes to an energetically even more favourable state with picomolar binding affinity (74). As a result, many amino acid residues dispersed throughout the structure of flavodoxin become stabilized against unfolding (62,74). The stability of flavodoxin is so high that FMN needs to be released first before global unfolding of the protein can occur (65). Characterisation of how flavodoxin kinetically folds *in vitro* shows that two parallel fold-

ing routes are accessible to apoflavodoxin molecules with all prolyl peptide bonds in their native trans configuration (5,65). These routes, of which one is on- and the other off-pathway, are visualized in the energy landscape of Fig. 4. This landscape depicts a funnel to the native state and a trough in which a misfolded intermediate temporarily resides. Folding in the funnel and trough are described by schemes A and B, respectively:

**Scheme A****Scheme B**

Scheme A represents the approximately 10 % of unfolded molecules (U; rim of the funnel towards native protein) that directly follow the productive folding route to native apoflavodoxin (Apo; bottom of funnel). On this path, intermediate I_{on} is an obligatory, high-energy on-pathway species that rapidly converts to native apoflavodoxin (5). I_{on} resides in a moat in the funnel towards native protein. Due to its instability, I_{on} is not observed during denaturant-dependent equilibrium folding of the protein. This productive folding of 10 % of unfolded molecules happens on the microsecond timescale. Subsequently, flavodoxin (Holo; not depicted in Fig. 4) forms upon binding of FMN to native apoflavodoxin. Around 90 % of unfolded molecules (U^* ; the outer edge of the off-pathway trough) misfold within milliseconds and form the off-pathway intermediate MG_{off} (scheme B). This relatively stable species is also detected during denaturant-dependent equilibrium folding and is molten globular. MG_{off} is positioned at the bottom of the off-pathway trough. Thermal unfolding of apoflavodoxin at equilibrium also proceeds through formation of a MG-like intermediate (61). MG_{off} needs to unfold significantly (i.e., to U^*), which lasts about a second, before folding to the native state takes place. This productive folding occurs within microseconds and again involves I_{on} (5,65). Thus, all folding apoflavodoxin molecules, whether they reside on productive or non-productive pathways, ultimately pass through I_{on} before reaching the native state.

Native apoflavodoxin occasionally forms partially unfolded forms (PUFs). Four PUFs have been identified, which unfold subglobally in a cooperative manner (66). In PUF1 part of the long loop is unfolded and in PUF2 another part of this loop is unfolded. In PUF3 mainly the 24 N-terminal residues are unfolded. In PUF4 the core of the protein is unfolded, except for helix $\alpha 4$. Both PUF1 and PUF2 are unfolding excursions that start from native apoflavodoxin but do not continue to the unfolded state, and do not reside on the productive folding route. PUF4 is an unfolding excursion of MG_{off} and PUF3 likely is as well (66,72). PUF1 to PUF3 represent misfolded

conformations. All PUFs are off the productive folding route and demonstrate that the energy landscape of apoflavodoxin folding is even more rugged than depicted in Fig. 4.

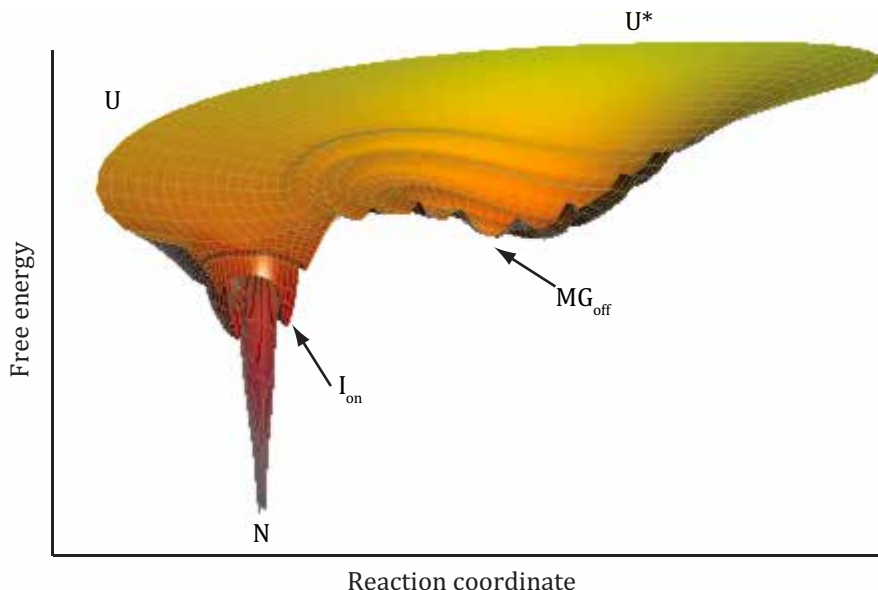


Figure 4 Schematic folding energy landscape of *A. vinelandii* apoflavodoxin. The left funnel shows the on-pathway route to native protein (N), on which an on-pathway intermediate (I_{on}) resides. The trough on the right displays formation of apoflavodoxin's off-pathway MG (MG_{off}). This molten globular species folds gradually, as exemplified by the presence of shallow moats. For simplicity's sake, the rugged character of MG_{off} folding, as evidenced by the detection of partially unfolded forms PUF3 and PUF4 (66), is not depicted in the MG_{off} trough. Unfolding excursions PUF1 and PUF2 that start from native protein are also omitted. Only native apoflavodoxin binds FMN, which leads to considerable stabilisation of the protein and deepening of the corresponding funnel (not shown). In reality, the ratio of the circumferences of the rims of the on- and off-pathway funnels is about 1 to 9. This ratio reflects that 10 % of unfolded molecules directly follow the productive folding route to native apoflavodoxin, whereas 90 % of unfolded molecules misfold and temporarily form MG_{off} (5). Apoflavodoxin is in 100 mM potassium pyrophosphate, pH 6.0, at 25 °C.

Unfolded apoflavodoxin is not a featureless statistical coil, but instead contains four transiently ordered regions (67). Three of these regions are α -helices, whereas the fourth adopts non-native structure that is neither α -helix nor β -strand (Fig. 5) (67). During folding, these structured elements interact and subsequently form the ordered core of MG_{off} (67,70). This propensity to interact is visualized in Fig. 4 and in scheme B by U^* . Non-native docking of the α -helices in MG_{off} prevents formation of the parallel β -sheet of native apoflavodoxin (67,70-72). Thus, the source for MG_{off} formation is situated in the unfolded state. The off-pathway intermediate has a drastically different architecture compared with native protein: it is largely α -helical and contains no β -sheet (69) and is slightly expanded compared to native apoflavodoxin (59). This α -helical MG acts as a trap and needs to unfold significantly in order to embark on a route to native α - β parallel protein.

While formation of native apoflavodoxin is highly cooperative (5,61,64,78), conversion of unfolded protein into MG_{off} is a non-cooperative, gradual process that simultaneously involves separate regions within apoflavodoxin (70,78). This suggests an energy landscape of MG folding with many barriers (visualised by shallow moats in the off-pathway trough of Fig. 4). *In vitro*, folding of the C-terminal part of MG_{off} precedes folding of the N-terminal part and this intermediate gradually compacts due to progressive extension of its ordered core (76,78). This non-cooperative folding happens because helices involve relatively short-range interactions and the folding of one helix does not need to affect the folding of other helices (70). Indeed, thermodynamically derived phase diagrams indicate that the transition from an unfolded species to a MG is a gradual, second-order-like process (80,81). Ultimately, after folding, the helical off-pathway MG of apoflavodoxin is almost entirely structured (70,78).



Figure 5 Model of the off-pathway MG of flavodoxin, in which the four, transiently structured regions detected in unfolded protein are highlighted. Helical parts also present in native protein are coloured green and regions of the unfolded protein that adopt α -helical structure not found in native protein are coloured blue; the orange element is structured, yet is neither α -helix nor β -strand in unfolded protein. These regions dock non-natively and form the core of apoflavodoxin's MG (67,69,70,72). The cartoon shows the structural elements of the off-pathway MG, but their relative positioning is unknown.

FOLDING OF OTHER FLAVODOXINS

Anabaena flavodoxin. In native apoflavodoxin from *Anabaena* PCC 7119 the cofactor-binding site is flexible and FMN binds preferentially to native apo-protein (82-84). Full release of cofactor happens upon unfolding of this 169-residue long-chain flavodoxin (85). Both urea- and GuHCl-dependent equilibrium folding of apoflavodoxin are two-state (83,85,86). During kinetic folding an intermediate transiently accumulates, whose unfolding is rate-limiting (3). Whether this off-pathway intermediate is a MG is unknown. Phi-value analysis of mutations (87) revealed that the transition state that separates unfolded and native apo-protein is diffuse, and various interactions seem to be similarly important for rate-limiting barrier crossing (88). The folding nucleus appears to be formed by the packing of helices α_3 , α_4 , and α_5 onto strands β_3 and β_4 .

In an apoflavodoxin variant whose stability has been reduced by excision of the

C-terminal helix, the resulting protein fragment adopts a molten globular conformation at pH 7 (89,90). Its far-UV CD spectrum resembles the one of full-length, molten globular apoflavodoxin at acidic pH (86). The structure of this MG is homogeneously weakened compared to the one of full-length native protein (90). The MG can be stabilized by increasing the helical propensity of helical regions of full-length native apoflavodoxin, suggesting preservation of helices in this MG (90).

Thermal unfolding of *Anabaena* apoflavodoxin at equilibrium follows a three-state mechanism in which an intermediate, with spectroscopic properties of a MG, populates (85,91,92). A low-resolution structure of this species has been inferred from Phi-value analysis (92). Large parts of the intermediate seem to have native-like topology, with somewhat weakened native interactions, but a 40-residue region is unfolded (85,91,92). The thermal intermediate neither resembles the MG that accumulates at low pH, nor the transition state for productive folding, nor the MG fragment. Whether this species is similar to the transient intermediate observed during kinetic folding of apoflavodoxin is unknown (93).

Apoflavodoxin variant F98N, in which the substitution of phenylalanine to asparagine destabilizes native protein, exhibits spectroscopic properties at 25 °C similar to those of the thermal intermediate of wild-type protein (93). For F98N apoflavodoxin the MG is the only species that populates significantly at 25 °C. The solution structure of this MG, which is less compact than native protein (94), comprises all 5 α -helices and β -strands 1 to 4 and 5a of native apoflavodoxin as well as their packing. Strand β_4 and helices α_4 and α_5 are shorter than in native protein (93). The loop that splits the fifth β -strand of apoflavodoxin into β_{5a} and β_{5b} is unstructured, just like the FMN binding loops encompassing residues 57 – 63 and 90 – 100 (92,93). These three segments are in contact in native protein and form a distinct region. Whether the thermal intermediate is on- or off-pathway to native apoflavodoxin has not been probed.

***H. pylori* flavodoxin.** This 164-residue protein is essential to *H. pylori*, the bacterium responsible for gastric ulcers. Hence, this long-chain flavodoxin is a target for drug discovery (95-97). Again, FMN only binds to native apoflavodoxin, which strongly stabilises the protein (98), thereby enabling it to be functional between pH 2 and 10 (99). Apoflavodoxin (100) and flavodoxin (97) are structurally similar, but in apo-protein the loops binding the isoalloxazine are flexible (100). This flavodoxin strongly resembles *A. vinelandii* and *Anabaena* flavodoxins, but its FMN binding site is more accessible (97,100). As a result, inhibitors can bind specifically to *H. pylori* flavodoxin and thereby disrupt its electron transfer function (96,101).

The folding characteristics of *H. pylori* (apo)flavodoxin have been studied by thermal unfolding and by lowering pH (98,99,102). Both methods result in formation of intermediates, yet with very different features. Thermal unfolding of (apo) flavodoxin involves two equilibrium intermediates (102). One species is similar to the thermal intermediate observed for *Anabaena* apoflavodoxin. Compared to native protein, this species has a slight increase in solvent-exposure of tryptophans and preserves most of the secondary structure of the native protein (around 80 % of native helix content). The other intermediate more closely resembles unfolded protein and has lost the tertiary contacts of native protein and a large part of its secondary structure, but at least one of its two tryptophans is partially buried (102).

Under strongly acidic conditions, *H. pylori* apoflavodoxin forms a MG that is slightly expanded compared to native protein. It has high helical content, displays exposed hydrophobic patches, and is devoid of, or maybe has very weakened, tertiary native interactions (98). This MG is structurally different from the two thermal intermediates identified at neutral pH (99,102). Whereas under acidic conditions the MG of *H. pylori* apoflavodoxin is soluble and monomeric (98), the MG of *Anabaena* apoflavodoxin aggregates massively (86). For *H. pylori* apoflavodoxin at neutral pH, the ensemble of molecules comprises native protein, the MG discovered by acid unfolding and the two intermediates revealed by thermal unfolding of the protein (99). The latter three species are present at relatively low mole fractions, with the MG state hardly populated at all (99). As no studies are reported of the folding kinetics of *H. pylori* flavodoxin, the on- or off-pathway nature of its intermediates is unidentified.

***D. desulfuricans* flavodoxin.** This 148-residue protein is a short-chain flavodoxin and is very sensitive to buffer and ionic strength conditions (103). A single folding study of *D. desulfuricans* (apo)flavodoxin from ATTC strain 27774 has been reported, which suggests that during its GuHCl-dependent equilibrium folding an intermediate populates. This species has a more solvent exposed tryptophan than native protein and judged by far-UV CD ellipticity at 220 nm it has native-like secondary structure (104). Folding of apoflavodoxin from *D. desulfuricans* strain ATTC 29577 (75 % sequence identity with ATTC 27774 protein) has been studied more extensively. This protein unfolds in a two-state fashion in denaturant-dependent equilibrium folding (105) and possibly populates intermediate states at low concentrations of denaturant (106). Thermal unfolding of apo- and holo-protein shows formation of an equilibrium intermediate. This species has native-like CD signal at 222 nm and reduced fluorescence compared to native protein and resembles the thermal intermediate of *Anabaena* apoflavodoxin (107), suggesting that it is a MG. Kinetic folding of *D. desulfuricans* apoflavodoxin studied by stopped-flow involves an intermediate that forms during the mixing dead-time. This 'burst-phase' species has a significant amount of secondary structure and appears to be off-pathway to the native state (105,108,109). However, kinetic analyses that conclusively verify the role of this intermediate still need to be done.

For Y98A and W60A variants of *D. desulfuricans* flavodoxin the apo-protein first has to fold to its native state before it can bind FMN (108). For wild-type flavodoxin (ATTC strains 27774 and 29577) it was suggested that FMN binding has no effect on its GuHCl-dependent equilibrium folding (104,105). However, FMN speeds up the folding of this protein (105). The researchers involved handled the resulting thermodynamic conflict by assuming that unfolded wild-type protein binds FMN with nanomolar affinity (104,105), which is a remarkable assumption. Indeed, with urea, FMN dissociates prior to protein unfolding and urea-unfolded apoflavodoxin does not interact with FMN (106).

Phi-value analysis revealed that the transition state for productive folding is diffuse and involves partially formed interactions throughout the polypeptide (109), as observed for *Anabaena* apoflavodoxin (88). In the transition state most β -strands on both sides of β -strand 3 must align properly with native-like interactions before final folding to native apoflavodoxin can occur (109).

Coarse-grained simulations can provide information about early folding events and intermediates. In such simulations of apoflavodoxin folding (109), interactions are represented by a Gō-like potential, in which only residue pairs that are in contact in the native state experience a pairwise attractive force. This procedure ignores non-native interactions, and thus is intrinsically limited, but can show frustration of another variety, called ‘topological frustration’. Topological frustration arises when native interactions form in the incorrect order or when competing folding pathways are present. The simulations suggest two competing folding nuclei at opposite sides of the central β -strand 3 of apoflavodoxin (centred on β -strands 1 and 4, respectively). The β -strand pair that interacts first among the five strands leads to dominance of the corresponding nucleation site and causes early misfolding (109). This topological frustration may contribute to temporary formation of an off-pathway species. It is unknown whether the *in vitro* kinetic intermediate contains the β -sheet structure identified for the *in silico* folding species. Neither is it known whether this kinetic intermediate is a MG.

Protein		Folding intermediate(s)	On- and/or off-pathway species	MG observed	Transition state determined	Simulations of folding
Flavodoxin	<i>A. vinelandii</i>	yes	on and off	yes	yes	n.d.
	<i>Anabaena</i>	yes	off	yes	yes	n.d.
	<i>H. pylori</i>	yes	n.d.	yes	n.d.	n.d.
	<i>D. desulfuricans</i>	yes	off	yes	yes	yes
	<i>D. vulgaris</i>	n.d.	n.d.	n.d.	n.d.	n.d.
CheY-like	CheY	yes	off	yes	yes	yes
	Spo0F	yes	off	n.d.	yes*	yes
	NT-NtrC	yes	off	n.d.	yes*	yes

Table 1 Folding of flavodoxin-like proteins

n.d.: not determined.

*: derived from simulation.

***D. vulgaris* (Hildenborough) flavodoxin.** This 148-residue, short-chain flavodoxin strongly resembles flavodoxin from *D. desulfuricans* (110,111). During its unfolding, dissociation into apo-protein and FMN is the rate-determining step and in unfolded protein FMN is fully dissociated (112). Binding of FMN is the last step during folding. The fluorescence changes observed during urea-dependent equilibrium folding of apo- and holo-protein fit two-state models, but involvement of intermediates cannot be excluded as only one spectroscopic technique was used (112). No kinetic folding data have been reported of this flavodoxin.

Table 1 summarises the folding of flavodoxins and shows that these proteins tend to form off-pathway intermediates during their unassisted folding *in vitro* to their functionally active forms. Several of these species are MGs.

FOLDING OF PROTEINS THAT SHARE THE FLAVODOXIN-LIKE FOLD: THE CHEY-LIKE SUPERFAMILY

CheY. Considerable effort has been put into elucidating the folding of CheY from *E. coli*, which is a 129-residue, single-domain, response regulator involved in chemotaxis. CheY becomes activated upon phosphorylation of D57 and has intrinsic auto-dephosphorylation capacity. Its temperature-dependent folding proceeds through a highly populated intermediate state, which is a MG dimer (113). Urea-dependent equilibrium folding of CheY monitored by fluorescence and far-UV CD fits a two-state model. However, ANS binding suggests that a MG-like folding state populates (113). Characterisation of the folding of the F14N/V83T variant of CheY by NMR spectroscopy also shows involvement of a MG, which can associate to a dimeric MG (114). It has been proposed that the monomeric MG can sequentially unfold into another intermediate I*, in which the first half of the polypeptide chain remains non-native and collapsed, and the second half is unfolded (114). In contrast, native protein unfolds highly cooperatively (114). Unfolded F14N/V83T protein in 5 M urea has residual structure (115). ^1H NMR chemical shift values indicate that two segments of unfolded CheY form helices. The 70 N-terminal residues of unfolded CheY have restricted mobility compared to the rest of the polypeptide. This collapsed region is stabilized by non-local interactions (115). It is tempting to speculate that formation of structured sequence segments in unfolded CheY leads to establishment of I*.

To characterize the transition state on the productive pathway to native protein, use was made of Phi-value analysis. This procedure identified two folding subdomains: an N-terminal one that is highly structured in the transition state and an unstructured C-terminal one (116). Formation of the 70-residue N-terminal subdomain is rate limiting and serves as the nucleus for subsequent condensation of the C-terminal subdomain (116,117). Intermediate I* resembles this transition state (114).

Upon folding, CheY collapses to a burst-phase intermediate that possesses significant stability and secondary structure (117,118). Analysis of the folding kinetics suggests that all five helices of CheY are formed in this sub-millisecond species (118). The far-UV CD spectrum of this intermediate resembles that of the native state. However, the packing of aromatic side chains might not be identical (9). Neither can it be excluded that this species lacks a β -sheet. CheY folds through two parallel pathways, which are defined by the state of isomerization of a proline in the active site, which is *cis* in native protein. Proline isomerization in the highly structured burst-phase species differs between both pathways. However, formation of this intermediate is not a consequence of proline isomerization (9). The most likely folding model places the burst-phase species off-pathway to the native state. Each intermediate has to at least partially unfold before productive folding occurs. The extent of secondary structure formation and side-chain packing in these off-pathway intermediates are consistent with a MG-like state (9).

Coarse-grained simulations were used to further explore CheY folding. In these

simulations the protein is represented as a string of C_α beads that interact via a Gō-like potential. One simulation shows formation of a misfolded intermediate, which has all five helices present in native CheY, but lacks its central β -parallel sheet. This misfolded species has to at least partially unfold before the native state is reached (119). More recent “flavoured” Gō-model simulations, in which the interaction energies of side-chain native contacts were scaled according to their abundance in the PDB, also show the existence of a frustrated state. This topologically frustrated species is a consequence of competition for native van der Waals contacts between the N- and C-terminal subdomains of CheY, leading to premature docking of these domains (120). Secondary structure elements α_2 , β_3 , α_3 and β_4 of native protein are involved in formation of this intermediate (9,120). It unfolds before productive folding in the N-subdomain occurs, leading to native protein. The experimentally observed, off-pathway intermediate seems consistent with early topological frustration during folding of CheY (9,120).

An alternative source for frustration during protein folding that leads to off-pathway species could be the formation of an unproductive local-in-sequence cluster of side chains of isoleucine, leucine, and valine (ILV) (9,121,122). CheY has two ILV clusters, one on either side of the central β -sheet, each serving to fuse the surface helices to each other and to the β -sheet. The smaller cluster contains 10 side chains and primarily links α_2 - β_3 - α_3 - β_4 on one face of the parallel β -sheet of native CheY; the larger cluster contains 15 side chains and links the β -sheet to α_1 and α_5 (122,123). The smaller cluster seems crucial to the off-pathway reaction (9).

Aspects of the models for premature docking of subdomains and for formation of an unproductive ILV cluster appear to describe the folding of CheY. The subdomain model fulfils the requirement for an intact N-terminal domain to access the native conformation. The ILV-cluster model captures best the initial formation of an off-pathway intermediate (122). The events that lead to this frustrated species comprise accumulation of substructures favoured by low-contact-order nonpolar interactions in the polypeptide. This leads to burial of surface areas, maximization of participation of aliphatic side chains in one of the two ILV clusters, and reduction in chain entropy penalty. Ultimately, CheY arrives in the lowest free energy minimum of its energy landscape of folding, which is the native state (122).

Spo0F and NT-NtrC. A combined experimental and simulation analysis was done on the folding of Spo0F from *Salmonella typhimurium* and NT-NtrC from *Bacillus subtilis*, which are 124-residue response-regulator proteins that have low sequence similarity (123). Both proteins are allosterically activated through phosphorylation of D54. “Flavoured” Gō-model-like simulations show that in the productive folding transition state for non-phosphorylated Spo0F and phosphorylated NT-NtrC the N-terminal subdomain is partially structured whereas the C-terminal one is not, just as observed for CheY (123). A similar result was obtained for non-phosphorylated NT-NtrC using a topology-based, coarse grained approach based on a variational model of protein folding, which shows that the N-terminal half of the protein appears to fold earlier than the C-terminal half (124). This folding scenario was also derived by using a coarse-grained, structure-based Hamiltonian model (125). Remarkably, for phosphorylated NT-NtrC the latter two simulations suggest that the C-terminal half of the protein folds first (124,125), contradicting the simulations of (123).

Kinetic folding experiments show that both Spo0F and NT-NtrC fold via parallel pathways through highly structured sub-millisecond intermediates before accessing their *cis* prolyl peptide bond-containing native conformations. Global analysis of the data favours an off-pathway folding mechanism for both proteins (123). Sequence-sensitive Gō-model simulations suggest that frustration in the folding of Spo0F leads to the appearance of an off-pathway species, reflecting competition for intra-subdomain van der Waals contacts between its N- and C-terminal subdomains. Secondary structure elements α_3 , β_4 , α_4 and β_5 of native protein (i.e., the C-subdomain of Spo0F) are involved in formation of this frustrated state. Local-in-sequence clusters of side chains of ILV stabilize the intermediate. Comparable observations were made for CheY folding. In case of NT-NtrC, the simulations fail to detect an off-pathway species. Experimental kinetic folding data show that the free-energy landscapes for folding of NT-NtrC and Spo0F are in many ways similar to the one of CheY (9,123).

Table 1 shows that off-pathway intermediate formation is characteristic for the folding of CheY-like proteins. Just as for flavodoxins, differences in the stability of these intermediates, and hence in their populations, may be explained by sequence-specific interactions within these species.

FLAVODOXIN-LIKE PROTEINS TEND TO TEMPORARILY MISFOLD DURING UNASSISTED FOLDING *IN VITRO* AND HAVE NON-CONSERVED TRANSITION STATES OF FOLDING

Flavodoxin-like proteins are susceptible to frustration in the early stages of folding, because an α -helix forms much more quickly than a parallel β -sheet. α -Helix formation is rapid due to the highly local character of interactions that produce this secondary structure. In contrast, many intervening residues separate the amino acids that constitute the strands of a parallel β -sheet and thus its formation is relatively slow. Since in Gō-like simulations only residue pairs that are in contact in the native state experience attractive forces, application of this model to flavodoxin-like proteins should produce comparably structured off-pathway intermediates. However, we note that slightly differing simulations, including those in which heterogeneity of the native contact energies is added to incorporate sequence effects, produce markedly dissimilar intermediates. For *D. desulfuricans* apoflavodoxin the misfolded species contains β -sheet structure (109). For CheY the off-pathway intermediate is reported to be helical, lacking the central parallel β -sheet of native protein (119). However, other simulations show that secondary structure elements α_2 , β_3 , α_3 and β_4 of native CheY seem to be involved in formation of this species (9,120). For Spo0F, simulations demonstrate that formation of this intermediate engages secondary structure elements α_3 , β_4 , α_4 and β_5 of native protein. Apparently, variations in sequence seem to modulate formation of these folding intermediates (122,123).

It is important to experimentally characterize the structural features of the discussed off-pathway intermediates in order to verify whether theory correctly predicts their conformations. In addition, characterization of the unfolded states of flavodoxin-like proteins is essential to comprehend how these misfolded species form, as residual structure in unfolded protein facilitates formation of folding intermediates. For *A. vinelandii* flavodoxin (67,70-72) and CheY (115) the unfolded protein has

been experimentally probed at the residue level. Transiently ordered regions exist in both unfolded proteins and largely comprise α -helical structures. Besides native also non-native α -helical structure may develop (Fig. 5). Gō-like simulations cannot predict such non-native structure. In flavodoxin-like proteins, hydrophobic interactions of side chains pack α -helices onto the parallel β -sheet. This typical feature probably contributes to the observed misfolding, as upon folding the α -helices fold first and can subsequently dock onto one another through hydrophobic interactions. This docking prevents formation of the parallel β -sheet and the resulting misfolded intermediate needs to unfold considerably before folding to native protein can take place, explaining why this species is off-pathway, as experimentally shown for *A. vinelandii* flavodoxin (67,70-72,76,78). Indeed, this off-pathway species is largely α -helical (69). The off-pathway intermediate of CheY also appears to be helical, as all five native α -helices are formed and as stabilization of any of its α -helices slows refolding at low denaturant concentrations (118). Experiments thus strongly suggest that the off-pathway intermediate of flavodoxin-like proteins is of helical nature. Future experimental efforts, using for example Fourier transform infrared spectroscopy, should reveal whether the misfolded species of *D. desulfuricans* apoflavodoxin and Spo0F really contain β -sheet structure, as simulations seem to suggest.

Current knowledge about the kinetics and energetics of protein folding largely stems from *in vitro* studies and from simulations. These studies revealed that the topology of a native protein influences the rate with which the native state is formed *in vitro*. This implies that the transition state that determines the rate of folding resembles the native state of a protein. Interactions involving a few key residues force a folding protein to adopt a rudimentary native-like architecture. Once the correct topology has been achieved, the native structure will then almost invariably be generated during the final stages of folding (126-130). Phi-value analysis shows that the productive folding transition state of CheY is polarized with a structured N-terminal subdomain and an unstructured C-terminal one (116). Simulations reveal transition states for non-phosphorylated Spo0F and phosphorylated NT-NtrC similar to the one of CheY (123). This indicates that the N-subdomain serves as the folding nucleus for these proteins. In contrast, the transition states of *Anabaena* and *D. desulfuricans* apoflavodoxins are diffuse with partial formation of many inter-residue interactions (88,109). Thus, the productive folding transition state of a flavodoxin-like protein seems to differ from one protein to another. Apparently, both chain topology and amino acid sequence contribute to defining the structure of this transition state and the folding energy landscapes of flavodoxin-like proteins.

In summary, flavodoxin-like proteins tend to temporarily misfold during their unassisted folding *in vitro* and have a non-conserved transition state that determines their rate of productive folding

PROTEIN FOLDING *IN VIVO*

While most protein folding experiments have been done *in vitro* (Fig. 6A), the circumstances under which folding occurs in the cellular environment differ dramatically (Fig. 6B).

Firstly, the conditions under which most folding landscapes are probed *in vitro* are physiologically speaking irrelevant. Illustrative in this regard is the use of high concentrations of GuHCl or urea (although data can be extrapolated to zero molar

denaturant). More gentle techniques to study protein folding are alteration of pH or temperature. Yet, even results obtained by these methods need to be carefully considered before they can be judged to be physiologically relevant. For example, how pertinent are folding intermediates found at pH 4, such as MGs, when the cytoplasmic environment is buffered at approximately pH 7 (131)? Concerning proteins that originate from endothermic organisms: how significant are folding intermediates found at 20 or 50 °C? Ultimately, existence of folding intermediates *in vitro* does not necessarily imply occurrence of the same or similar species and their consequences *in vivo*.

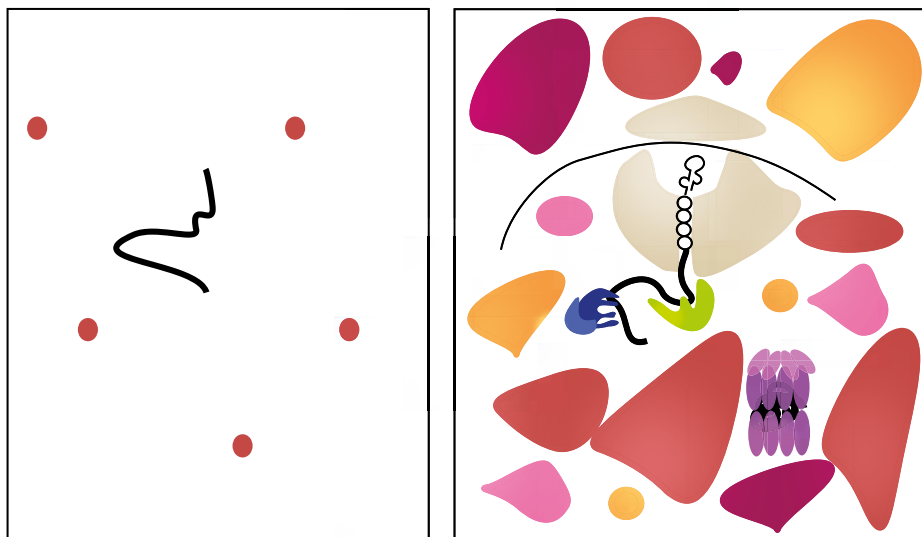


Figure 6 Comparing protein folding *in vitro* and *in vivo*. *A*, Folding *in vitro* involves the entire polypeptide sequence (i.e., full-length protein (black)) in a dilute environment of interactors (red), the majority of which is similar in physical properties like size, charge and hydrophobicity. *B*, Folding *in vivo* starts during translation of the nascent chain (black) by the ribosome (beige). Chaperones like Trigger factor (green) and DnaK/DnaJ (blue) can interact with the nascent chain, whereas chaperones like GroEL/GroES (purple) act on the released polypeptide. The cellular environment is crowded with many macromolecules (various colours), which span the entire range of physical properties.

Secondly, compared to most *in vitro* folding experiments, the cellular environment is immensely crowded (Fig. 6B), with estimations of the concentration of proteins and other macromolecules ranging up to 300 - 400 g/L (132). This crowding can lead to a significant excluded volume effect, which in turn can influence the type and amount of interactions a polypeptide can form (133,134). However, folding *in vitro* is typically done using dilute buffers with micro-molar protein concentrations (Fig. 6A). For *A. vinelandii* apoflavodoxin, conditions that mimic macromolecular crowding inside cells lead to compaction of unfolded protein and slight stabilization of native protein (59). Crowding causes severe aggregation of its off-pathway MG (59). In case of *H. pylori* apoflavodoxin, crowding slightly stabilizes both native and molten globular protein, but does not significantly stabilize the native state relative to its MG (99). Remarkably, crowding increases the midpoint of thermal unfolding

of apoflavodoxin from *D. desulfuricans* by up to 20 °C (103). In contrast to other apoflavodoxins, the stability of this protein is very sensitive to buffer composition. Its thermal midpoint of unfolding can decrease as much as 25 °C upon changing from phosphate to HEPES buffer (103). Crowding makes the ensemble of unfolded conformations of *D. desulfuricans* apoflavodoxin less expanded, resulting in a folding funnel that apparently is smoother and narrower (135). Upon changing the crowding agent the folding mechanism of this protein seems to be modulated differently, and folding routes experiencing topological frustrations might be either enhanced or relieved (136). The influence of crowding on the folding CheY-like proteins has not been reported. For various other proteins the effects of artificial crowders, such as Ficoll, glycerol, dextrans and others, and occasionally of similar kinds of crowders put together, have also been probed (for an extensive review see (137)). However, these crowders differ from those in a cell, in which they span broad spectra of size, hydrophobicity and charge, and are present simultaneously. Artificial crowders thus possess shortcomings and their usefulness in predicting effects of crowding *in vivo* is contested (138).

Thirdly, a major difference between folding *in vitro* and *in vivo* is the circumstance under which a protein is introduced in either environment. During *in vitro* folding a full-length polypeptide, or sometimes fragments of it, is studied (Fig. 6A). The protein is chemically or thermally unfolded and subsequently it can in principle sample all of its folding states while progressing to the native state. *In vivo*, however, a polypeptide can start to fold as soon as it becomes produced by the ribosome (Fig. 6B).

Ribosomes play a crucial role as a central hub in coordinating protein quality control (139,140). They sense the nature of the emerging polypeptide, recruit protein folding and translocation factors, and integrate mRNA and nascent chain quality control. Ribosomes can control the cellular abundance of proteins at the translational level and thus are a significant constituent of protein homeostasis (139,141-144). In growing cells, most ribosomes are active in translation and contain an emerging polypeptide. Whereas ribosomal structures have been elucidated in atomic detail (see e.g. (145-149)), relatively little is known about the conformational events polypeptide chains undergo while they are produced.

Upon addition of amino acid residues, the growing nascent chain gradually emerges from the ribosomal exit tunnel. This tunnel is between 80 and 100 Å long and 10 to 20 Å wide (150). Depending on nascent chain conformation, the exit tunnel can hold between 30 (extended) to 70 (α-helical) amino acid residues. Folding or misfolding of the nascent chain may already start in the exit tunnel (151-154) and in most cases commences while the nascent chain gradually emerges from it (51-53,55,155-163). As synthesis of the nascent chain is orders of magnitude slower than it takes a protein to fold (i.e., seconds versus micro- to milliseconds), nascent chains can already sample parts of their conformational space before the next amino acid residue is added. This phenomenon, in combination with gradual emergence of the polypeptide from the exit tunnel, might impose conformational constraints on the folding energy landscape of the protein involved. Folding during synthesis on the ribosome leads to constant remodelling of the folding energy landscape as translation proceeds. Matters are complicated even further by the various manners in which translational speed can be influenced. A well-known one is the effect of codon usage (164). The resulting ribosome pausing affects cotranslational folding

(165-168). Folding of the mRNA sequence that is being translated also influences translation velocity (169).

The ribosome itself can also modulate protein folding, for example through interaction of the nascent chain with the exit tunnel or, upon its emergence from the tunnel, with the ribosomal surface (57,170,171). Nascent chains can adopt distinct conformations within this tunnel. Formation of α -helices, α - and β -hairpins and even native structure for some small proteins (< 10 kDa), but also non-native conformation, can occur near the tunnel exit, where the tunnel widens to form a vestibule (57,172-174). The tunnel prevents proper tertiary folding of larger protein domains and precludes the C-terminal residues of the nascent chain from participating in long-range interactions (175). As a consequence, productive folding can occur only after a complete protein or a domain has come out of the ribosome (176). Once a nascent chain emerges it's folding can be influenced by transient electrostatic interactions with the ribosomal surface. As the exterior of the ribosome consists for a large part of rRNA, it is negatively charged in many areas. Negative charge also surrounds the exit tunnel, thereby providing an environment that is very different from the cytoplasm (177). This charge may influence protein folding, as it attracts positively charged residues of the nascent chain and repels negatively charged ones (170). It also restricts the dynamics of nascent chains according to their charge (177). Due to these electrostatic interactions, nascent chains may be protected from misfolding or aggregation.

Methodologies like epitope recognition, enzymatic activity, cofactor binding, NMR and fluorescence spectroscopy indicate structural ordering and acquisition of activity of polypeptides once they arrive outside the exit tunnel (51-53,55,155-163). Cotranslational folding has been observed for all α (51), all β (53,55,159,162) and α/β proteins (54,178,179). Various nascent chains form native- or non-native-like intermediates (51-53,55,159,180-182). In the next section we discuss the only flavodoxin-like protein where cotranslational folding has been probed: *A. vinelandii* flavodoxin.

Lastly, upon emergence from the exit tunnel, cellular proteins can start to interact with the nascent chain and affect it's folding (Fig. 6B). Some of the first proteins that meet the nascent chain are processing proteins such as peptide deformylase, which removes the formyl moiety from the N-terminal formyl-methionine. Another processing protein is methionine aminopeptidase, which hydrolyses the N-terminal methionine for more than 50 % of nascent chains (183). Other processes that can happen during translation include nascent chain modification through glycosylation, phosphorylation, etcetera's, and/or translocation of the nascent chain into membranes, like the endoplasmic reticulum or the plasma membrane. Chaperone proteins that hover at the exit tunnel facilitate nascent chain folding just outside this tunnel (184,185). They prevent protein aggregation and/or stimulate folding towards the native state. Chaperones often interact with co-chaperones (186). After release of a protein from the ribosome, its folding may also be affected by chaperones. Because chaperones can be purified, their individual effects on protein folding can be studied with relative ease *in vitro*. As a cell contains many different chaperones, with estimations ranging to hundreds for eukaryotic cells (187), the total effect of the chaperone network on protein folding cannot be determined in a test tube. To date, no study of the influences of chaperones on the folding of flavodoxin-like proteins has been reported.

These differences between *in vitro* and *in vivo* protein folding show that a demand exists to obtain a molecular description of how flavodoxin-like proteins fold in a cell. In the following, as a first step towards this goal, we discuss the cotranslational folding of flavodoxin.

THE RIBOSOME AFFECTS FLAVODOXIN FOLDING

Various strategies exist to study folding of a protein while it emerges from the ribosome. To characterize nascent chain folding of *A. vinelandii* flavodoxin, the approach we followed is to produce, purify and characterize stably arrested ribosome-nascent chain complexes (RNCs). RNCs consist of the entire 70S ribosome with a polypeptide that is stalled through tight interaction with the exit tunnel. Production of these RNCs happens in *E. coli* (179) or by using an *in vitro* protein synthesis kit (188). To achieve translational stalling, a sequence derived from *E. coli* SecM (Secretion Monitor protein) is attached to the C-terminus of the translating nascent protein (189,190). In addition, the construct used contains a linker that connects SecM and flavodoxin and spans the length of the exit tunnel, thereby entirely exposing the flavodoxin polypeptide outside the ribosome (179,188). Use is made of constructs that lack zero, five or ten amino acid residues at the C-terminus of stalled flavodoxin. This procedure allows mimicking of late stages during protein translation and enables one to obtain “snapshots” of cotranslational protein folding. Due to physical forces on the nascent chain (191,192), the flavodoxin construct is sometimes released from the RNC (179). Comparison of released protein and RNCs allows assessment of the influence the ribosome has on apoflavodoxin folding.

FMN only binds to the flavin-binding site of full-length, isolated *A. vinelandii* apo-flavodoxin when the protein is natively folded (65). We addressed cofactor binding to nascent flavodoxin and showed that FMN does not bind to apoflavodoxin during translation (179). The cofactor can only bind to the protein once it is entirely synthesized and exposed outside the ribosome, because then apoflavodoxin becomes natively folded (Fig. 7A). Even when incomplete apoflavodoxin that lacks only its 5 C-terminal residues emerges from the ribosomal exit tunnel it cannot incorporate FMN, because it forms a non-native intermediate, the conformation of which is presently unknown. In contrast, the corresponding released, shortened protein product does bind the cofactor (Fig. 7B). These differences in FMN binding capacities show that the ribosome affects nascent flavodoxin folding and cofactor binding. As a result, binding of cofactor to released full-length protein is the last step in the production of this flavoprotein in the cell (Figs. 7A,B) (179). The longer the chain of nascent apoflavodoxin is, the more stabilized the protein becomes and the less prone apo-protein is to intracellular proteolytic degradation. Upon incorporation of FMN the corresponding holo-proteins become protected against degradation (Figs. 7A,B) (179).

To further understanding of flavodoxin folding *in vivo*, we investigated whether the ribosome modulates formation of apoflavodoxin's off-pathway MG (188). To this end, the F44Y mutation was introduced into an RNC construct that exposes the entire nascent chain outside the exit tunnel. Full-length, isolated F44Y apoflavodoxin switches from natively folded to off-pathway MG upon decreasing ionic strength to physiological values (69), thereby avoiding the use of denaturant, which could adversely affect ribosomal integrity. The F44Y mutation introduces an extra oxygen atom into a hydrophobic pocket of native apoflavodoxin, causing considerable de-

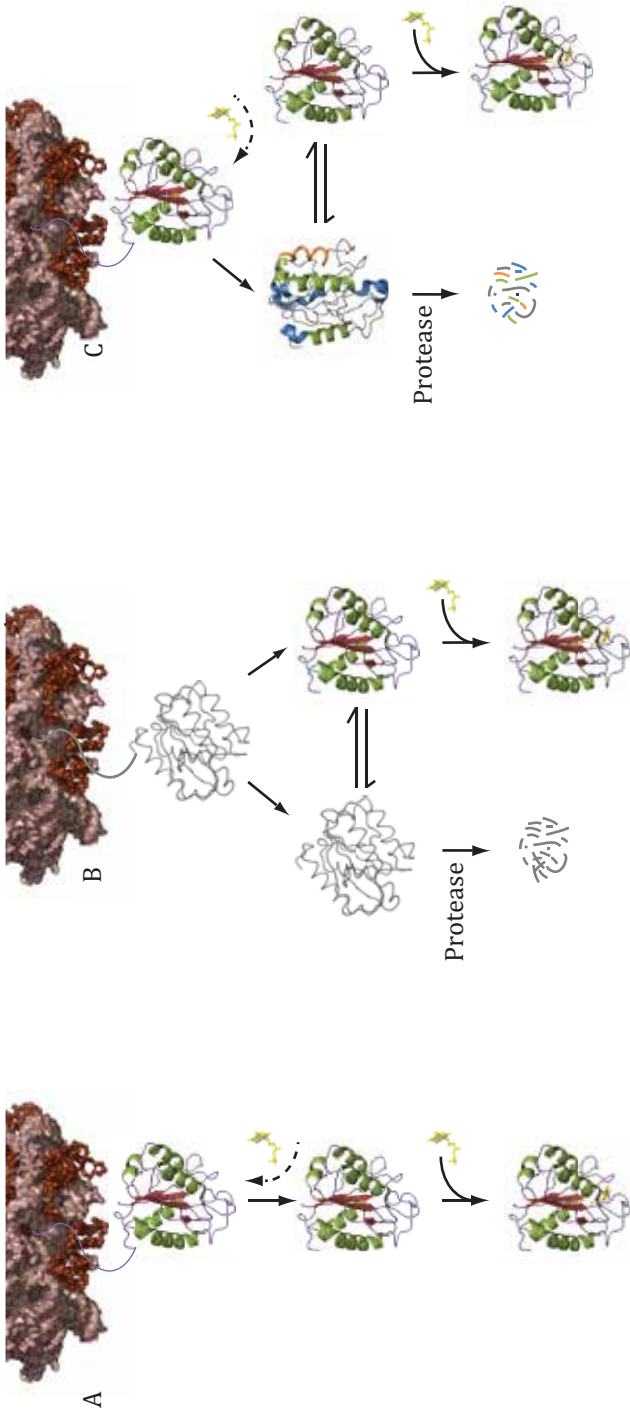


Figure 7 Model showing cotranslational folding of apoflavodoxin and subsequent binding of cofactor. The ribosome and FMN are coloured brown and yellow, respectively. Depending on its folding state, apoflavodoxin is either coloured (native, colour scheme of Fig. 1) or grey (non-native). The colour scheme of Figure 5 is used to depict apoflavodoxin in its off-pathway MG state. *A*, Full-length RNCs of apoflavodoxin can fold to the native state while stalled to the ribosome and are capable of binding FMN. Most nascent chains, however, are released as apoprotein, which subsequently binds FMN. *B*, RNCs that are C-terminally shortened by five amino acid residues do not acquire apoflavodoxin's native fold while stalled to the ribosome. Upon release, the non-natively folded protein is either degraded by proteases or attains its native fold and subsequent binding of FMN protects it against proteolytic degradation. *C*, When the entire F44Y apoflavodoxin protein is exposed outside the ribosome it becomes natively folded and may thus bind FMN. Upon its release as apoflavodoxin, this protein variant pre-dominantly forms the off-pathway MG at the ionic strength that prevails in the cytosol. Molten globular F44Y apoflavodoxin is proteolytically degraded in the cell. Binding of FMN to native apoflavodoxin protects F44Y protein against forming MG and proteolytic degradation (179,188).

stabilization (69,72). It was verified that the SecM sequence, the linker and other components of the RNC construct do not influence formation of the off-pathway MG of F44Y apoflavodoxin or impair FMN binding. In addition, it has been shown that FMN does not associate with ribosomes (188). As the off-pathway MG needs to unfold before native apoflavodoxin can form, the rate of cofactor binding is delayed compared to the situation where only natively folded apo-protein is present. Thus, ascertaining the rate of FMN binding as a function of ionic strength is a suitable tool to detect the presence of apoflavodoxin's off-pathway MG on the ribosome (188). Because F44Y apoflavodoxin RNCs bind FMN rapidly at both physiological and at high salt concentration, the fully exposed polypeptide must be natively folded under both conditions. Therefore, the ribosome modulates MG formation and forces nascent apoflavodoxin towards the native state (Fig. 7C) (188). This confinement of MG formation is an important observation that emphasizes differences between folding energy landscapes *in vivo* and *in vitro*. Possibly, electrostatic repulsion of the nascent chain (apoflavodoxin has a net charge of -13 at neutral pH) by the negatively charged ribosomal surface restricts the conformational space of unfolded protein, leading to entropic stabilization of native protein at physiological ionic strength (188). When release of the nascent chain occurs at physiological ionic strength, released F44Y apoflavodoxin construct predominantly forms the off-pathway MG, which is unstable and thus susceptible to intracellular proteolytic degradation (Fig. 7C) (188). Because molten globular F44Y apoflavodoxin is in equilibrium with native protein, binding of FMN withdraws this latter state from the equilibrium and protects the protein against the action of proteases (Fig. 7C).

FUTURE PERSPECTIVES

Our understanding of the folding of isolated flavodoxin-like proteins in a test tube, including the influence of macromolecular crowding, has increased considerably during the past decades. Also, the first facets of cotranslational folding of flavodoxin have been illuminated. Various aspects of the folding of flavodoxin-like proteins still require clarification, some of which are highlighted below.

Experimental determination of folding landscapes is a challenge and therefore relatively little experimental knowledge exists about them (29). Developments in biophysical techniques, including single-molecule ones, need to be applied to determine, refine and compare folding energy landscapes of various flavodoxin-like proteins and obtain more insight into off-pathway intermediate formation. For example, it is currently unknown when off-pathway intermediates of flavodoxin-like proteins unfold and embark on productive folding routes, which specific intramolecular interactions need to be disrupted. Mechanical unfolding and refolding by using optical tweezers may provide insight into this issue and can also be applied to further our understanding of the cotranslational folding of these proteins (170).

Separate domains of multi-domain proteins are expected to fold independently of each other during translation (181). It would be interesting to verify whether the ribosome also forces the folding of a flavodoxin-like domain within multi-domain proteins, like cytochrome P450 reductase and nitric-oxide synthase (Figs. 1D,E), towards the native state. When a flavodoxin-like domain is non-C-terminally located, it is tethered to the ribosome during further translation of the multi-domain protein involved. Thus, once this domain emerges from the exit tunnel, the ribosome

has ample opportunity to influence its folding. If the ribosome pushes the flavodoxin-like domain towards the native state it should be able to bind FMN, as demonstrated for *A. vinelandii* apoflavodoxin (179). Cotranslational incorporation of flavin would impede proteolytic degradation of nascent multi-domain chains that contain a flavodoxin-like domain, due to increased stability conferred by cofactor binding. This could increase protein production efficiency (179). For several flavoproteins involved in diseases, a beneficial effect is observed upon supplementing the respective flavin (193).

While at equilibrium the ribosome restrains off-pathway MG formation in stalled RNC complexes that expose the entire F44Y flavodoxin protein outside the exit tunnel, comparable confinement may not happen during translation. For example, nascent flavodoxin that lacks five C-terminal amino acid residues adopts a non-native structure (179). It would be fascinating to elucidate how non-native interactions evolve during translation. Recently, by exploiting a reconstituted *in vitro* translation system combined with labelling of nascent chains at defined positions with bright fluorophores, nascent protein folding could be monitored in real-time by Förster resonance energy transfer (57). Application of this methodology can shed light onto folding of flavodoxin-like protein during translation and reveal whether temporary misfolding and formation of MGs occurs.

Isolated, full-length flavodoxin-like proteins tend to form off-pathway intermediates during folding *in vitro*. While this observation does not signify that they also form misfolded species cotranslationally, once released from the restraining influence of the ribosome these proteins may temporarily adopt misfolded structures in the cellular environment. Nevertheless, many flavodoxin-like proteins can be overexpressed in their native form, suggesting involvement of chaperones. Hence, a strong demand exists to clarify the influences chaperones, like Trigger factor, DnaK/DnaJ and GroEL/GroES, have on the folding of flavodoxin-like proteins.

The ultimate challenge is to resolve the folding of flavodoxin-like proteins *in vivo*. Though progress has been made in characterising protein conformations at the atomic level in the crowded milieu of live cells (194-197), structural resolution of how proteins attain their native architecture in such an environment is still a formidable challenge. Surprises are likely to emerge about the folding of flavodoxin-like proteins in a living cell.

ACKNOWLEDGEMENTS

We thank Adrie Westphal, dr. Simon Lindhoud and prof. dr. Willem van Berkel for helpful comments on the manuscript. This work was supported by an ECHO (711.011.007) grant of the Netherlands Organisation of Scientific Research (NWO) to C.v.M.

REFERENCES

1. Bollen, Y. J. M., and van Mierlo, C. P. M. (2005) Protein topology affects the appearance of intermediates during the folding of proteins with a flavodoxin-like fold. *Biophys. Chem.* **114**, 181-189
2. Bilsel, O., Zitzewitz, J. A., Bowers, K. E., and Matthews, C. R. (1999) Folding mechanism of the α -subunit of tryptophan synthase, an α/β barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry* **38**, 1018-1029
3. Fernández-Recio, J., Genzor, C. G., and Sancho, J. (2001) Apoflavodoxin folding mechanism: an α/β protein with an essentially off-pathway intermediate. *Biochemistry* **40**, 15234-15245
4. Forsyth, W. R., and Matthews, C. R. (2002) Folding mechanism of indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*: a test of the conservation of folding mechanisms hypothesis in $(\beta\alpha)_8$ barrels. *J. Mol. Biol.* **320**, 1119-1133
5. Bollen, Y. J. M., Sánchez, I. E., and van Mierlo, C. P. M. (2004) Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin. *Biochemistry* **43**, 10475-10489
6. Forsyth, W. R., Bilsel, O., Gu, Z., and Matthews, C. R. (2007) Topology and sequence in the folding of a TIM barrel protein: global analysis highlights partitioning between transient off-pathway and stable on-pathway folding intermediates in the complex folding mechanism of a $(\beta\alpha)_8$ barrel of unknown function from *B. subtilis*. *J. Mol. Biol.* **372**, 236-253
7. Gu, Z., Zitzewitz, J. A., and Matthews, C. R. (2007) Mapping the structure of folding cores in TIM barrel proteins by hydrogen exchange mass spectrometry: the roles of motif and sequence for the indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*. *J. Mol. Biol.* **368**, 582-594
8. Otzen, D. E., Giehm, L., Baptista, R. P., Kristensen, S. R., Melo, E. P., and Pedersen, S. (2007) Aggregation as the basis for complex behaviour of cutinase in different denaturants. *Biochim. Biophys. Acta - Proteins Proteom.* **1774**, 323-333
9. Kathuria, S. V., Day, I. J., Wallace, L. A., and Matthews, C. R. (2008) Kinetic traps in the folding of beta alpha-repeat proteins: CheY initially misfolds before accessing the native conformation. *J. Mol. Biol.* **382**, 467-484
10. Lorenz, T., and Reinstein, J. (2008) The influence of proline isomerization and off-pathway intermediates on the folding mechanism of eukaryotic UMP/CMP kinase. *J. Mol. Biol.* **381**, 443-455
11. Pierella Karlusich, J. J., Lodeyro, A. F., and Carrillo, N. (2014) The long goodbye: the rise and fall of flavodoxin during plant evolution. *J. Exp. Bot.* **65**, 5161-5178
12. Pierella Karlusich, J. J., Ceccoli, R. D., Grana, M., Romero, H., and Carrillo, N. (2015) Environmental selection pressures related to iron utilization are involved in the loss of the flavodoxin gene from the plant genome. *Gen. Biol. Evol.* **7**, 750-767
13. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8411-

8416

14. Garcin, E. D., Bruns, C. M., Lloyd, S. J., Hosfield, D. J., Tiso, M., Gachhui, R., Stuehr, D. J., Tainer, J. A., and Getzoff, E. D. (2004) Structural basis for isozyme-specific regulation of electron transfer in nitric-oxide synthase. *J. Biol. Chem.* **279**, 37918-37927

15. Caetano-Anolles, G., Kim, H. S., and Mittenthal, J. E. (2007) The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9358-9363

16. Ma, B.-G., Chen, L., Ji, H.-F., Chen, Z.-H., Yang, F.-R., Wang, L., Qu, G., Jiang, Y.-Y., Ji, C., and Zhang, H.-Y. (2008) Characters of very ancient proteins. *Biochem. Biophys. Res. Comm.* **366**, 607-611

17. Lu, M.-F., Ji, H.-F., Li, T.-X., Kang, S.-K., Zhang, Y.-J., Zheng, J.-F., Tian, T., Jia, X.-S., Lin, X.-M., and Zhang, H.-Y. (2013) Reconstructing a flavodoxin oxidoreductase with early amino acids. *Int. J. Mol. Sci.* **14**, 12843-12852

18. Farias-Rico, J. A., Schmidt, S., and Hoecker, B. (2014) Evolutionary relationship of two ancient protein superfolds. *Nat. Chem. Biol.* **10**, 710-715

19. Macheroux, P., Kappes, B., and Ealick, S. E. (2011) Flavogenomics - a genomic and structural view of flavin-dependent proteins. *FEBS J.* **278**, 2625-2634

20. López-Llano, J., Maldonado, S., Bueno, M., Lostao, A., Angeles-Jimenez, M., Lillo, M. P., and Sancho, J. (2004) The long and short flavodoxins - I. The role of the differentiating loop in apoflavodoxin structure and FMN binding. *J. Biol. Chem.* **279**, 47177-47183

21. López-Llano, J., Maldonado, S., Jain, S., Lostao, A., Godoy-Ruiz, R., Sanchez-Ruiz, J. M., Cortijo, M., Fernandez-Recio, J., and Sancho, J. (2004) The long and short flavodoxins - II. The role of the differentiating loop in apoflavodoxin stability and folding mechanism. *J. Biol. Chem.* **279**, 47184-47191

22. Sancho, J. (2006) Flavodoxins: sequence, folding, binding, function and beyond. *Cell. Mol. Life Sci.* **63**, 855-864

23. Perez-Dorado, I., Bortolotti, A., Cortez, N., and Hermoso, J. A. (2013) Structural and phylogenetic analysis of *Rhodobacter capsulatus* NifF: uncovering general features of nitrogen-fixation (nif)-flavodoxins. *Int. J. Mol. Sci.* **14**, 1152-1163

24. Birch, O. M., Hewitson, K. S., Fuhrmann, M., Burgdorf, K., Baldwin, J. E., Roach, P. L., and Shaw, N. M. (2000) MioC is an FMN-binding protein that is essential for *Escherichia coli* biotin synthase activity *in vitro*. *J. Biol. Chem.* **275**, 32277-32280

25. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230

26. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19

27. Jackson, S. E. (1998) How do small single-domain proteins fold? *Fold. Des.* **3**, R81-91

28. Dill, K. A., Ozkan, S. B., Weikl, T. R., Chodera, J. D., and Voelz, V. A. (2007) The protein folding problem: when will it be solved? *Curr. Opin. Struct. Biol.* **17**, 342-346

29. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years

Science **338**, 1042-1046

30. Levinthal, C. (1968) Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44-45
31. Levinthal, C. (1969) How to fold graciously. in *Mössbauer Spectroscopy in Biological Systems* (Munck, J. T. P. D. a. E. ed., University of Illinois Press, Allerton House, Monticello, Illinois
32. Zou, T. S., Williams, N., Ozkan, S. B., and Ghosh, K. (2014) Proteome folding kinetics is limited by protein halflife. *PLoS ONE* **9**, e112701
33. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnel, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195
34. Kuwajima, K., Nitta, K., Yoneyama, M., and Sugai, S. (1976) Three-state denaturation of α -lactalbumin by guanidine hydrochloride. *J. Mol. Biol.* **106**, 359-373
35. Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S., and Ptitsyn, O. B. (1981) α -Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett.* **136**, 311-315
36. Ohgushi, M., and Wada, A. (1983) 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21-24
37. Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* **6**, 87-103
38. Dobson, C. M. (2003) Protein folding and misfolding. *Nature* **426**, 884-890
39. Kuroda, Y., Kidokoro, S., and Wada, A. (1992) Thermodynamic characterization of cytochrome c at low pH. Observation of the molten globule state and of the cold denaturation process. *J. Mol. Biol.* **223**, 1139-1153
40. Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995) Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5446-5450
41. Colón, W., and Roder, H. (1996) Kinetic intermediates in the formation of the cytochrome c molten globule. *Nat. Struct. Biol.* **3**, 1019-1025
42. Jamin, M., and Baldwin, R. L. (1998) Two forms of the pH 4 folding intermediate of apomyoglobin. *J. Mol. Biol.* **276**, 491-504
43. Arai, M., and Kuwajima, K. (2000) Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209-282
44. Mok, K. H., Nagashima, T., Day, I. J., Hore, P. J., and Dobson, C. M. (2005) Multiple subsets of side-chain packing in partially folded states of α -lactalbumins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8899-8904
45. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., and Razgulyaev, O. I. (1990) Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **262**, 20-24
46. Baldwin, R. L., and Rose, G. D. (2013) Molten globules, entropy-driven conformational change and protein folding. *Curr. Opin. Struct. Biol.* **23**, 4-10
47. Bhattacharyya, S., and Varadarajan, R. (2013) Packing in molten globules

and native states. *Curr. Opin. Struct. Biol.* **23**, 11-21

48. van der Goot, F. G., Gonzalez-Manas, J. M., Lakey, J. H., and Pattus, F. (1991) A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature* **354**, 408-410

49. Ren, J. H., Kachel, K., Kim, H., Malenbaum, S. E., Collier, R. J., and London, E. (1999) Interaction of diphtheria toxin T domain with molten globule-like proteins and its implications for translocation. *Science* **284**, 955-957

50. Benke, S., Roderer, D., Wunderlich, B., Nettels, D., Glockshuber, R., and Schuler, B. (2015) The assembly dynamics of the cytolytic pore toxin ClyA. *Nat. Comm.* **6**, 6198

51. Ellis, J. P., Bakke, C. K., Kirchdoerfer, R. N., Jungbauer, L. M., and Cavagnero, S. (2008) Chain dynamics of nascent polypeptides emerging from the ribosome. *ACS Chem. Biol.* **3**, 555-566

52. Evans, M. S., Sander, I. M., and Clark, P. L. (2008) Cotranslational folding promotes β -helix formation and avoids aggregation *in vivo*. *J. Mol. Biol.* **383**, 683-692

53. Cabrita, L. D., Hsu, S.-T. D., Launay, H., Dobson, C. M., and Christodoulou, J. (2009) Probing ribosome-nascent chain complexes produced *in vivo* by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22239-22244

54. Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012) Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* **48**, 63-74

55. Kelkar, D. A., Khushoo, A., Yang, Z., and Skach, W. R. (2012) Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. *J. Biol. Chem.* **287**, 2568-2578

56. Lamprou, P., Kempe, D., Katranidis, A., Büldt, G., and Fitter, J. (2014) Nano-second dynamics of calmodulin and ribosome-bound nascent chains studied by time-resolved fluorescence anisotropy. *ChemBioChem* **15**, 977-985

57. Holtkamp, W., Kokic, G., Jager, M., Mittelstaet, J., Komar, A. A., and Rodnina, M. V. (2015) Cotranslational protein folding on the ribosome monitored in real time. *Science* **350**, 1104-1107

58. Klugkist, J., Voorberg, J., Haaker, H., and Veeger, C. (1986) Characterization of three different flavodoxins from *Azotobacter vinelandii*. *Eur. J. Biochem.* **155**, 33-40

59. Engel, R., Westphal, A. H., Huberts, D. H. E. W., Nabuurs, S. M., Lindhoud, S., Visser, A. J. W. G., and van Mierlo, C. P. M. (2008) Macromolecular crowding compacts unfolded apoflavodoxin and causes severe aggregation of the off-pathway intermediate during apoflavodoxin folding. *J. Biol. Chem.* **283**, 27383-27394

60. Steensma, E., Heering, H. A., Hagen, W. R., and Van Mierlo, C. P. M. (1996) Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur. J. Biochem.* **235**, 167-172

61. van Mierlo, C. P. M., van Dongen, W. M. A. M., Vergeldt, F., van Berkel, W. J. H., and Steensma, E. (1998) The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate. *Protein Sci.* **7**, 2331-2344

62. Steensma, E., Nijman, M. J., Bollen, Y. J., de Jager, P. A., van den Berg, W. A., van Dongen, W. M., and van Mierlo, C. P. (1998) Apparent local stability of the secondary structure of *Azotobacter vinelandii* holoflavodoxin II as probed by hydrogen exchange: implications for redox potential regulation and flavodoxin folding. *Protein Sci.* **7**, 306-317
63. van Mierlo, C. P. M., and Steensma, E. (2000) Protein folding and stability investigated by fluorescence, circular dichroism (CD), and nuclear magnetic resonance (NMR) spectroscopy: the flavodoxin story. *J. Biotechnol.* **79**, 281-298
64. van Mierlo, C. P. M., van den Oever, J. M. P., and Steensma, E. (2000) Apoflavodoxin (un)folding followed at the residue level by NMR. *Protein Sci.* **9**, 145-157
65. Bollen, Y. J. M., Nabuurs, S. M., van Berkel, W. J. H., and van Mierlo, C. P. M. (2005) Last in, first out. The role of cofactor binding in flavodoxin folding. *J. Biol. Chem.* **280**, 7836-7844
66. Bollen, Y. J. M., Kamphuis, M. B., and van Mierlo, C. P. M. (2006) The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4095-4100
67. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2008) Extensive formation of off-pathway species during folding of an alpha-beta parallel protein is due to docking of (non)native structure elements in unfolded molecules. *J. Am. Chem. Soc.* **130**, 16914-16920
68. Visser, N. V., Westphal, A. H., van Hoek, A., van Mierlo, C. P. M., Visser, A. J. W. G., and van Amerongen, H. (2008) Tryptophan-tryptophan energy migration as a tool to follow apoflavodoxin folding. *Biophys. J.* **95**, 2462-2469
69. Nabuurs, S. M., Westphal, A. H., aan den Toorn, M., Lindhoud, S., and van Mierlo, C. P. M. (2009) Topological switching between an alpha-beta parallel protein and a remarkably helical molten globule. *J. Am. Chem. Soc.* **131**, 8290-8295
70. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2009) Noncooperative formation of the off-pathway molten globule during folding of the alpha-beta parallel protein apoflavodoxin. *J. Am. Chem. Soc.* **131**, 2739-2746
71. Nabuurs, S. M., de Kort, B. J., Westphal, A. H., and van Mierlo, C. P. M. (2010) Non-native hydrophobic interactions detected in unfolded apoflavodoxin by paramagnetic relaxation enhancement. *Eur. Biophys. J.* **39**, 689-698
72. Nabuurs, S. M., and van Mierlo, C. P. M. (2010) Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of alpha-beta parallel protein flavodoxin. *J. Biol. Chem.* **285**, 4165-4172
73. Laptinok, S. P., Visser, N. V., Engel, R., Westphal, A. H., van Hoek, A., van Mierlo, C. P. M., van Stokkum, I. H. M., van Amerongen, H., and Visser, A. J. W. G. (2011) A general approach for detecting folding intermediates from steady-state and time-resolved fluorescence of single-tryptophan-containing proteins. *Biochemistry* **50**, 3441-3450
74. Bollen, Y. J. M., Westphal, A. H., Lindhoud, S., van Berkel, W. J. H., and van Mierlo, C. P. M. (2012) Distant residues mediate picomolar binding affinity of a protein cofactor. *Nat. Comm.* **3**, 1010
75. Lindhoud, S., Westphal, A. H., Visser, A. J. W. G., Borst, J. W., and van Mierlo,

C. P. M. (2012) Fluorescence of Alexa Fluor dye tracks protein folding. *PLoS ONE* **7**, e46838

76. Lindhoud, S., Westphal, A. H., Borst, J. W., and van Mierlo, C. P. M. (2012) Illuminating the off-pathway nature of the molten globule folding intermediate of an alpha-beta parallel protein. *PLoS ONE* **7**, e45746

77. Lindhoud, S., Westphal, A. H., van Mierlo, C. P. M., Visser, A. J. W. G., and Borst, J. W. (2014) Rise-time of FRET-acceptor fluorescence tracks protein folding. *Int. J. Mol. Sci.* **15**, 23836-23850

78. Lindhoud, S., Pirchi, M., Westphal, A. H., Haran, G., and van Mierlo, C. P. M. (2015) Gradual folding of an off-pathway molten globule detected at the single-molecule level. *J. Mol. Biol.* **427**, 3148-3157

79. van Son, M., Lindhoud, S., van der Wild, M., van Mierlo, C. P. M., and Huber, M. (2015) Double electron-electron spin resonance tracks flavodoxin folding. *J. Phys. Chem. B* **119**, 13507-13514

80. Shakhnovich, E. I., and Finkelstein, A. V. (1989) Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers* **28**, 1667-1680

81. Finkelstein, A. V., and Shakhnovich, E. I. (1989) Theory of cooperative transitions in protein molecules. II. Phase diagram for a protein molecule in solution. *Biopolymers* **28**, 1681-1694

82. Lostao, A., El Harrous, M., Daoudi, F., Romero, A., Parody-Morreale, A., and Sancho, J. (2000) Dissecting the energetics of the apoflavodoxin-FMN complex. *J. Biol. Chem.* **275**, 9518-9526

83. Langdon, G. M., Jimenez, M. A., Genzor, C. G., Maldonado, S., Sancho, J., and Rico, M. (2001) *Anabaena* apoflavodoxin hydrogen exchange: On the stable exchange core of the alpha/beta(21345) flavodoxin-like family. *Proteins* **43**, 476-488

84. Lostao, A., Daoudi, F., Irún, M. P., Ramón, Á., Fernández-Cabrera, C., Romero, A., and Sancho, J. (2003) How FMN binds to *Anabaena* apoflavodoxin - A hydrophobic encounter at an open binding site. *J. Biol. Chem.* **278**, 24053-24061

85. Campos, L. A., and Sancho, J. (2006) Native-specific stabilization of flavodoxin by the FMN cofactor: Structural and thermodynamical explanation. *Proteins* **63**, 581-594

86. Genzor, C. G., Beldarrain, A., Gomez-Moreno, C., Lopez-Lacomba, J. L., Cortijo, M., and Sancho, J. (1996) Conformational stability of apoflavodoxin. *Protein Sci.* **5**, 1376-1388

87. Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122-126

88. Bueno, M., Ayuso-Tejedor, S., and Sancho, J. (2006) Do proteins with similar folds have similar transition state structures? A diffuse transition state of the 169 residue apoflavodoxin. *J. Mol. Biol.* **359**, 813-824

89. Maldonado, S., Jimenez, M. A., Langdon, G. M., and Sancho, J. (1998) Cooperative stabilization of a molten globule apoflavodoxin fragment. *Biochemistry* **37**, 10589-10596

90. López-Llano, J., Campos, L. A., Bueno, M., and Sancho, J. (2006) Equilibrium phi-analysis of a molten globule: The 1-149 apoflavodoxin fragment. *J. Mol. Biol.* **356**, 354-366
91. Irun, M. P., Garcia-Mira, M. M., Sanchez-Ruiz, J. M., and Sancho, J. (2001) Native hydrogen bonds in a molten globule: The apoflavodoxin thermal intermediate. *J. Mol. Biol.* **306**, 877-888
92. Campos, L. A., Bueno, M., Lopez-Llano, J., Jimenez, M. A., and Sancho, J. (2004) Structure of stable protein folding intermediates by equilibrium phi-analysis: The apoflavodoxin thermal intermediate. *J. Mol. Biol.* **344**, 239-255
93. Ayuso-Tejedor, S., Angarica, V. E., Bueno, M., Campos, L. A., Abián, O., Bernado, P., Sancho, J., and Jiménez, M. A. (2010) Design and structure of an equilibrium protein folding intermediate: A hint into dynamical regions of proteins. *J. Mol. Biol.* **400**, 922-934
94. Ayuso-Tejedor, S., Garcia-Fandino, R., Orozco, M., Sancho, J., and Bernadó, P. (2011) Structural analysis of an equilibrium folding intermediate in the apoflavodoxin native ensemble by small-angle X-ray scattering. *J. Mol. Biol.* **406**, 604-619
95. Cremades, N., Bueno, M., Toja, M., and Sancho, J. (2005) Towards a new therapeutic target: *Helicobacter pylori* flavodoxin. *Biophys. Chem.* **115**, 267-276
96. Cremades, N., Velazquez-Campoy, A., Martinez-Julvez, M., Neira, J. L., Perez-Dorado, I., Hermoso, J., Jimenez, P., Lanás, A., Hoffman, P. S., and Sancho, J. (2009) Discovery of specific flavodoxin inhibitors as potential therapeutic agents against *Helicobacter pylori* infection. *ACS Chem. Biol.* **4**, 928-938
97. Freigang, J., Diederichs, K., Schafer, K. P., Welte, W., and Paul, R. (2002) Crystal structure of oxidized flavodoxin, an essential protein in *Helicobacter pylori*. *Protein Sci.* **11**, 253-261
98. Cremades, N., Bueno, M., Neira, J. L., Velazquez-Campoy, A., and Sancho, J. (2008) Conformational stability of *Helicobacter pylori* flavodoxin - Fit to function at pH 5. *J. Biol. Chem.* **283**, 2883-2895
99. Cremades, N., and Sancho, J. (2008) Molten globule and native state ensemble of *Helicobacter pylori* flavodoxin: Can crowding, osmolytes or cofactors stabilize the native conformation relative to the molten globule? *Biophys. J.* **95**, 1913-1927
100. Martinez-Julvez, M., Cremades, N., Bueno, M., Perez-Dorado, I., Maya, C., Cuesta-Lopez, S., Prada, D., Falo, F., Hermoso, J. A., and Sancho, J. (2007) Common conformational changes in flavodoxins induced by FMN and anion binding: The structure of *Helicobacter pylori* apoflavodoxin. *Prot. Struct. Funct. Bioinf.* **69**, 581-594
101. Galano, J. J., Alias, M., Perez, R., Velazquez-Campoy, A., Hoffman, P. S., and Sancho, J. (2013) Improved flavodoxin inhibitors with potential therapeutic effects against *Helicobacter pylori* infection. *J. Med. Chem.* **56**, 6248-6258
102. Cremades, N., Velazquez-Campoy, A., Freire, E., and Sancho, J. (2008) The flavodoxin from *Helicobacter pylori*: Structural determinants of thermostability and FMN cofactor binding. *Biochemistry* **47**, 627-639
103. Stagg, L., Zhang, S. Q., Cheung, M. S., and Wittung-Stafshede, P. (2007) Molecular crowding enhances native structure and stability of α/β protein flavodoxin. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18976-18981

104. Apiyo, D., Guidry, J., and Wittung-Stafshede, P. (2000) No cofactor effect on equilibrium unfolding of *Desulfovibrio desulfuricans* flavodoxin. *BBA-Prot. Struct. Mol. Enzymol.* **1479**, 214-224
105. Apiyo, D., and Wittung-Stafshede, P. (2002) Presence of the cofactor speeds up folding of *Desulfovibrio desulfuricans* flavodoxin. *Protein Sci.* **11**, 1129-1135
106. Muralidhara, B. K., and Wittung-Stafshede, P. (2005) FMN binding and unfolding of *Desulfovibrio desulfuricans* flavodoxin: "Hidden" intermediates at low denaturant concentrations. *Biochim. Biophys. Acta* **1747**, 239-250
107. Muralidhara, B. K., and Wittung-Stafshede, P. (2004) Thermal unfolding of apo and holo *Desulfovibrio desulfuricans* flavodoxin: Cofactor stabilizes folded and intermediate states. *Biochemistry* **43**, 12855-12864
108. Muralidhara, B. K., Rathinakumar, R., and Wittung-Stafshede, P. (2006) Folding of *Desulfovibrio desulfuricans* flavodoxin is accelerated by cofactor fly-casting. *Arch. Biochem. Biophys.* **451**, 51-58
109. Stagg, L., Samiotakis, A., Homouz, D., Cheung, M. S., and Wittung-Stafshede, P. (2010) Residue-specific analysis of frustration in the folding landscape of repeat beta/alpha protein apoflavodoxin. *J. Mol. Biol.* **396**, 75-89
110. Watt, W., Tulinsky, A., Swenson, R. P., and Watenpaugh, K. D. (1991) Comparison of the crystal structures of a flavodoxin in its three oxidation states at cryogenic temperatures. *J. Mol. Biol.* **218**, 195-208
111. Romero, A., Caldeira, J., Legall, J., Moura, I., Moura, J. J. G., and Romao, M. J. (1996) Crystal structure of flavodoxin from *Desulfovibrio desulfuricans* ATCC 27774 in two oxidation states. *Eur. J. Biochem.* **239**, 190-196
112. Nuallain, B. O., and Mayhew, S. G. (2002) A comparison of the urea-induced unfolding of apoflavodoxin and flavodoxin from *Desulfovibrio vulgaris*. *Eur. J. Biochem.* **269**, 212-223
113. Filimonov, V. V., Prieto, J., Martinez, J. C., Bruix, M., Mateo, P. L., and Serrano, L. (1993) Thermodynamic analysis of the chemotactic protein from *Escherichia coli*, CheY. *Biochemistry* **32**, 12906-12921
114. Garcia, P., Serrano, L., Rico, M., and Bruix, M. (2002) An NMR view of the folding process of a CheY mutant at the residue level. *Structure* **10**, 1173-1185
115. Garcia, P., Serrano, L., Durand, D., Rico, M., and Bruix, M. (2001) NMR and SAXS characterization of the denatured state of the chemotactic protein CheY: Implications for protein folding initiation. *Protein Sci.* **10**, 1100-1112
116. Lopez-Hernandez, E., and Serrano, L. (1996) Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2. *Fold. Des.* **1**, 43-55
117. Munoz, V., Lopez, E. M., Jager, M., and Serrano, L. (1994) Kinetic characterization of the chemotactic protein from *Escherichia coli*, CheY. Kinetic analysis of the inverse hydrophobic effect. *Biochemistry* **33**, 5858-5866
118. Lopez-Hernandez, E., Cronet, P., Serrano, L., and Munoz, V. (1997) Folding kinetics of CheY mutants with enhanced native α -helix propensities. *J. Mol. Biol.* **266**, 610-620
119. Clementi, C., Nymeyer, H., and Onuchic, J. N. (2000) Topological and energet-

ic factors: what determines the structural details of the transition state ensemble and “en-route” intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **298**, 937-953

120. Hills, R. D., Jr., and Brooks, C. L., 3rd. (2008) Subdomain competition, cooperativity, and topological frustration in the folding of CheY. *J. Mol. Biol.* **382**, 485-495

121. Wu, Y., Vadrevu, R., Kathuria, S., Yang, X., and Matthews, C. R. (2007) A tightly packed hydrophobic cluster directs the formation of an off-pathway sub-millisecond folding intermediate in the α subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol.* **366**, 1624-1638

122. Nobrega, R. P., Arora, K., Kathuria, S. V., Graceffa, R., Barrea, R. A., Guo, L., Chakravarthy, S., Bilsel, O., Irving, T. C., Brooks, C. L., III, and Matthews, C. R. (2014) Modulation of frustration in folding by sequence permutation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10562-10567

123. Hills, R. D., Jr., Kathuria, S. V., Wallace, L. A., Day, I. J., Brooks, C. L., 3rd, and Matthews, C. R. (2010) Topological frustration in beta alpha-repeat proteins: Sequence diversity modulates the conserved folding mechanisms of alpha/beta/alpha sandwich proteins. *J. Mol. Biol.* **398**, 332-350

124. Tripathi, S., and Portman, J. J. (2013) Allostery and folding of the N-terminal receiver domain of protein NtrC. *J. Phys. Chem. B* **117**, 13182-13193

125. Itoh, K., and Sasai, M. (2009) Multidimensional theory of protein folding. *J. Chem. Phys.* **130**, 145104

126. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.* **25**, 331-339

127. Baker, D. (2000) A surprising simplicity to protein folding. *Nature* **405**, 39-42

128. Fersht, A. R. (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: Contact order, chain topology, stability, and the extended nucleus mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1525-1529

129. Vendruscolo, M., Paci, E., Dobson, C. M., and Karplus, M. (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature* **409**, 641-645

130. Fersht, A. R., and Daggett, V. (2002) Protein folding and unfolding at atomic resolution. *Cell* **108**, 573-582

131. Madshus, I. H. (1988) Regulation of intracellular pH in eukaryotic cells. *Biochem. J.* **250**, 1-8

132. Zimmerman, S. B., and Trach, S. O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol.* **222**, 599-620

133. Ellis, R. J. (2001) Macromolecular crowding: Obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597-604

134. Zhou, H. X., Rivas, G. N., and Minton, A. P. (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Ann. Rev. Biophys.* **37**, 375-397

135. Stagg, L., Christiansen, A., and Wittung-Stafshede, P. (2011) Macromolecular crowding tunes folding landscape of parallel α/β protein, apoflavodoxin. *J. Am. Chem. Soc.* **133**, 646-648

136. Homouz, D., Stagg, L., Wittung-Stafshede, P., and Cheung, M. S. (2009) Macromolecular crowding modulates folding mechanism of alpha/beta protein apoflavodoxin. *Biophys. J.* **96**, 671-680

137. Kuznetsova, I. M., Turoverov, K. K., and Uversky, V. N. (2014) What macromolecular crowding can do to a protein. *Int. J. Mol. Sci.* **15**, 23090-23140

138. Rivas, G., and Minton, A. P. (2016) Macromolecular crowding *in vitro*, *in vivo*, and in between. *Trends Biochem. Sci.* **41**, 970-981

139. Pechmann, S., Willmund, F., and Frydman, J. (2013) The ribosome as a hub for protein quality control. *Mol. Cell* **49**, 411-421

140. Brandman, O., and Hegde, R. S. (2016) Ribosome-associated protein quality control. *Nat. Struct. Mol. Biol.* **23**, 7-15

141. Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Ann. Rev. Biochem.* **78**, 959-991

142. Schwanhauser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* **473**, 337-342

143. Barna, M. (2013) Ribosomes take control. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9-10

144. Lee, A. S., Burdeinick-Kerr, R., and Whelan, S. P. (2013) A ribosome-specialized translation initiation pathway is required for cap-dependent translation of vesicular stomatitis virus mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 324-329

145. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905-920

146. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonnrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature* **407**, 327-339

147. Harms, J., Schlutzenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* **107**, 679-688

148. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**, 827-834

149. Selmer, M., Dunham, C. M., Murphy, F. V. T., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942

150. Wilson, D. N., and Beckmann, R. (2011) The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Curr. Opin. Struct. Biol.* **21**, 274-282

151. Woolhead, C. A., McCormick, P. J., and Johnson, A. E. (2004) Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **116**, 725-736
152. Alagaratnam, S., van Pouderoyen, G., Pijning, T., Dijkstra, B. W., Cavazzini, D., Rossi, G. L., van Dongen, W. M. A. M., van Mierlo, C. P. M., van Berkel, W. J. H., and Canters, G. W. (2005) A crystallographic study of Cys69Ala flavodoxin II from *Azotobacter vinelandii*: Structural determinants of redox potential. *Protein Sci.* **14**, 2284-2295
153. Bhushan, S., Gartmann, M., Halic, M., Armache, J. P., Jarasch, A., Mielke, T., Berninghausen, O., Wilson, D. N., and Beckmann, R. (2010) α -Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* **17**, 313-317
154. Marino, J., von Heijne, G., and Beckmann, R. (2016) Small protein domains fold inside the ribosome exit tunnel. *FEBS Lett.* **590**, 655-660
155. Fedorov, A. N., Friguet, B., Djavadi-Ohanian, L., Alakhov, Y. B., and Goldberg, M. E. (1992) Folding on the ribosome of *Escherichia coli* tryptophan synthase β -subunit nascent chains probed with a conformation-dependent monoclonal antibody. *J. Mol. Biol.* **228**, 351-358
156. Komar, A. A., Kommer, A., Krashennnikov, I. A., and Spirin, A. S. (1997) Cotranslational folding of globin. *J. Biol. Chem.* **272**, 10646-10651
157. Netzer, W. J., and Hartl, F. U. (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388**, 343-349
158. Nicola, A. V., Chen, W., and Helenius, A. (1999) Co-translational folding of an alphavirus capsid protein in the cytosol of living cells. *Nat. Cell. Biol.* **1**, 341-345
159. Clark, P. L., and King, J. (2001) A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early *in vitro* refolding intermediates. *J. Biol. Chem.* **276**, 25411-25420
160. Hsu, S. T., Fucini, P., Cabrita, L. D., Launay, H., Dobson, C. M., and Christodoulou, J. (2007) Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16516-16521
161. Hsu, S. T., Cabrita, L. D., Fucini, P., Christodoulou, J., and Dobson, C. M. (2009) Probing side-chain dynamics of a ribosome-bound nascent chain using methyl NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 8366-8367
162. Eichmann, C., Preissler, S., Riek, R., and Deuerling, E. (2010) Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9111-9116
163. Khushoo, A., Yang, Z., Johnson, A. E., and Skach, W. R. (2011) Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1. *Mol. Cell* **41**, 682-692
164. Quax, T. E. F., Claassens, N. J., Söll, D., and van der Oost, J. (2015) Codon bias as a means to fine-tune gene expression. *Mol. Cell* **59**, 149-161
165. Chadani, Y., Niwa, T., Chiba, S., Taguchi, H., and Ito, K. (2016) Integrated *in vivo* and *in vitro* nascent chain profiling reveals widespread translational pausing. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E829-838
166. Komar, A. A. (2009) A pause for thought along the co-translational folding

pathway. *Trends Biochem. Sci.* **34**, 16-24

167. Zhang, G., and Ignatova, Z. (2011) Folding at the birth of the nascent chain: Coordinating translation with co-translational folding. *Curr. Opin. Struct. Biol.* **21**, 25-31

168. Sander, I. M., Chaney, J. L., and Clark, P. L. (2014) Expanding Anfinsen's principle: Contributions of synonymous codon selection to rational protein design. *J. Am. Chem. Soc.* **136**, 858-861

169. Gloge, F., Becker, A. H., Kramer, G., and Bukau, B. (2014) Co-translational mechanisms of protein maturation. *Curr. Opin. Struct. Biol.* **24**, 24-33

170. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I., Jr., and Bustamante, C. (2011) The ribosome modulates nascent protein folding. *Science* **334**, 1723-1727

171. Cabrita, L. D., Cassaignau, A. M., Launay, H. M., Waudby, C. A., Wlodarski, T., Camilloni, C., Karyadi, M. E., Robertson, A. L., Wang, X., Wentink, A. S., Goodsell, L. S., Woolhead, C. A., Vendruscolo, M., Dobson, C. M., and Christodoulou, J. (2016) A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. *Nat. Struct. Mol. Biol.* **23**, 278-285

172. Kosolapov, A., and Deutsch, C. (2009) Tertiary interactions within the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* **16**, 405-411

173. Tu, L. W., and Deutsch, C. (2010) A folding zone in the ribosomal exit tunnel for Kv1.3 helix formation. *J. Mol. Biol.* **396**, 1346-1360

174. Nilsson, O. B., Hedman, R., Marino, J., Wickles, S., Bischoff, L., Johansson, M., Muller-Lucks, A., Trovato, F., Puglisi, J. D., O'Brien, E. P., Beckmann, R., and von Heijne, G. (2015) Cotranslational protein folding inside the ribosome exit tunnel. *Cell Rep.* **12**, 1533-1540

175. Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009) The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat. Struct. Mol. Biol.* **16**, 589-597

176. Hartl, F. U., and Hayer-Hartl, M. (2009) Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 574-581

177. Knight, A. M., Culviner, P. H., Kurt-Yilmaz, N., Zou, T., Ozkan, S. B., and Cavagnero, S. (2013) Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem. Biol.* **8**, 1195-1204

178. Rutkowska, A., Beerbaum, M., Rajagopalan, N., Fiaux, J., Schmieder, P., Kramer, G., Oschkinat, H., and Bukau, B. (2009) Large-scale purification of ribosome-nascent chain complexes for biochemical and structural studies. *FEBS Lett.* **583**, 2407-2413

179. Houwman, J. A., Westphal, A. H., van Berkel, W. J. H., and van Mierlo, C. P. M. (2015) Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome. *BBA-Proteins Proteom* **1854**, 1317-1324

180. Kudlicki, W., Kitaoka, Y., Odom, O. W., Kramer, G., and Hardesty, B. (1995) Elongation and folding of nascent ricin chains as peptidyl-tRNA on ribosomes: The effect of amino acid deletions on these processes. *J. Mol. Biol.* **252**, 203-212

181. Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U. (1999) Co-translational domain folding as the structural basis for the rapid *de novo* folding

of firefly luciferase. *Nat. Struct. Biol.* **6**, 697-705

182. Chow, C. C., Chow, C., Raghunathan, V., Huppert, T. J., Kimball, E. B., and Cavagnero, S. (2003) Chain length dependence of apomyoglobin folding: Structural evolution from misfolded sheets to native helices. *Biochemistry* **42**, 7090-7099

183. Jha, S., and Komar, A. A. (2011) Birth, life and death of nascent polypeptide chains. *Biotechnol. J.* **6**, 623-640

184. Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324-332

185. Preissler, S., and Deuerling, E. (2012) Ribosome-associated chaperones as key players in proteostasis. *Trends Biochem. Sci.* **37**, 274-283

186. Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z. Y., Larsen, B., Choi, H., Berger, B., Gingras, A. C., and Lindquist, S. (2014) A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* **158**, 434-448

187. Brehme, M., Voisine, C., Rolland, T., Wachi, S., Soper, J. H., Zhu, Y., Orton, K., Vilella, A., Garza, D., Vidal, M., Ge, H., and Morimoto, R. I. (2014) A chaperome sub-network safeguards proteostasis in aging and neurodegenerative disease. *Cell Rep.* **9**, 1135-1150

188. Houwman, J. A., André, E., Westphal, A. H., van Berkel, W. J. H., and van Mierlo, C. P. M. (2016) The ribosome restrains molten globule formation in stalled nascent apoflavodoxin. *J. Biol. Chem.* **291**, 25911-25920

189. Schaffitzel, C., and Ban, N. (2007) Generation of ribosome nascent chain complexes for structural and functional studies. *J. Struct. Biol.* **158**, 463-471

190. Nakatogawa, H., and Ito, K. (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**, 629-636

191. Cymer, F., Hedman, R., Ismail, N., and von Heijne, G. (2015) Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. *J. Biol. Chem.* **290**, 10208-10215

192. Goldman, D. H., Kaiser, C. M., Milin, A., Righini, M., Tinoco, I., Jr., and Bustamante, C. (2015) Ribosome. Mechanical force releases nascent chain-mediated ribosome arrest *in vitro* and *in vivo*. *Science* **348**, 457-460

193. Rodrigues, J. V., Henriques, B. J., Lucas, T. G., and Gomes, C. M. (2012) Cofactors and metabolites as protein folding helpers in metabolic diseases. *Curr. Top. Med. Chem.* **12**, 2546-2559

194. Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., Takeuchi, T., Futaki, S., Ito, Y., Hiroaki, H., and Shirakawa, M. (2009) High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106-109

195. Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., Mikawa, T., Walchli, M., Smith, B. O., Shirakawa, M., Gunttert, P., and Ito, Y. (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102-105

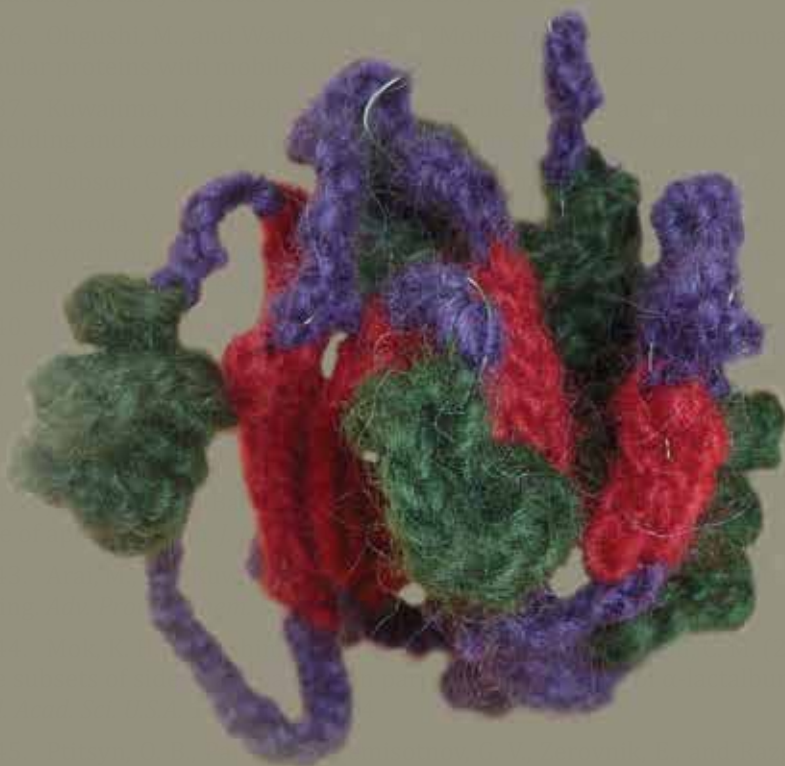
196. Renault, M., Tommassen-van Boxtel, R., Bos, M. P., Post, J. A., Tommassen, J., and Baldus, M. (2012) Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4863-4868

197. Hamatsu, J., O'Donovan, D., Tanaka, T., Shirai, T., Hourai, Y., Mikawa, T., Ikeya, T., Mishima, M., Boucher, W., Smith, B. O., Laue, E. D., Shirakawa, M., and Ito, Y. (2013) High-resolution heteronuclear multidimensional NMR of proteins in living insect cells using a baculovirus protein expression system. *J. Am. Chem. Soc.* **135**, 1688-1691

198. Volz, K., and Matsumura, P. (1991) Crystal structure of *Escherichia coli* CheY refined at 1.7-Å resolution. *J. Biol. Chem.* **266**, 15511-15519

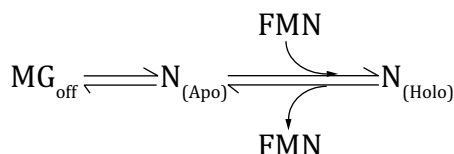
6

General Discussion



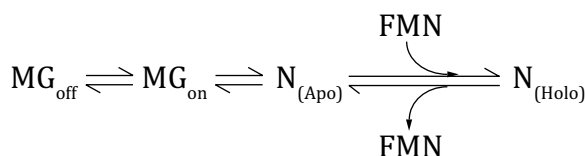
This thesis explores the influence the ribosome has on cotranslational folding of flavodoxin and on formation of apoflavodoxin's off-pathway molten globule (MG_{off}), a folding state that has been extensively characterized *in vitro* (1-9). Using the experiments described in the previous chapters, we gained important insights into folding of apoflavodoxin and its MG.

In **Chapter 2** we show that the model (scheme 1) used to describe the equilibrium folding of apoflavodoxin *in vitro* (9-11) requires revision.

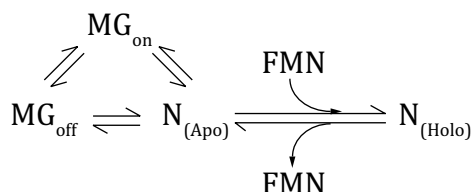


Scheme 1

Besides MG_{off} another MG exists, namely one that is on-pathway to native protein and likely has a high degree of native structure. Two possible folding models can describe the folding through MG_{on} : a linear one (scheme 2) or a triangular one (scheme 3).



Scheme 2



Scheme 3

Under physiological conditions *in vitro* both MGs coexist for the F44Y variant of apoflavodoxin, indicating that they should also be concurrently present in the cellular environment. Due to cellular factors such as chaperones or macromolecular crowding these MGs might display different behaviour *in vivo* compared to *in vitro*. Preliminary experiments, not reported in this thesis, show that prokaryotic chaperone Trigger factor as well as eukaryotic chaperone Hsp90 seem to interact with F44Y apoflavodoxin at physiological ionic strength (i.e., when both MGs are present). We could not yet determine whether these chaperones have a preference for a specific MG or interact with both of them. F44Y apoflavodoxin is thus an interesting candidate to study the effect of chaperones on MG formation *in vitro* and *in vivo*.

Another reason why F44Y apoflavodoxin is an interesting candidate for future chaperone research *in vitro* is its propensity to aggregate at biologically relevant protein concentrations (12). Upon overproduction of this flavodoxin variant in *E. coli*, significant amounts of this protein end up in the insoluble cellular fraction. Also, F44Y apoflavodoxin has decreased global stability compared to C69A protein and thus is more susceptible to degradation by proteases (4). Both properties make F44Y a suitable protein for studying the effects of chaperones, especially in combination with C69A apoflavodoxin as a stable and non-aggregating control.

In **Chapters 3** and **4** we investigate cotranslational folding of C-terminally shortened apoflavodoxin constructs and of F44Y apoflavodoxin, respectively. It has previously been demonstrated that FMN binding is the last step during flavodoxin folding *in vitro* (13). In **Chapter 3** we show that the ribosome affects apoflavodoxin folding and thus prevents FMN binding to apoflavodoxin during translation. Consequently, binding of cofactor to released protein is the last step in production of this flavoprotein *in vivo* (14). In the case of stalled ribosomes, FMN binding can also occur when the completely synthesized nascent chain is fully exposed outside the exit tunnel. We show that in this situation binding of FMN stabilizes the fully exposed nascent chain of apoflavodoxin against proteolytic degradation. We demonstrate that at physiological ionic strength the ribosome restrains formation of the off-pathway MG in F44Y RNCs by forcing the nascent chain towards the native state. This phenomenon is most likely caused by restriction of apoflavodoxin's conformational space due to electrostatic repulsion by the highly negatively charged ribosomal surface (**Chapter 4**).

Though the ribosome restrains formation of the off-pathway MG of F44Y apoflavodoxin, it is still interesting to investigate the effect of chaperones, like for example Trigger factor, on the folding of this nascent protein. Trigger factor has been shown to possess a "holdase" functionality (15), which essentially keeps the protein in an unfolded state. Would this also be the case for F44Y apoflavodoxin, even though the ribosome forces the nascent chain towards the native state?

In **Chapter 3** we illustrate that in *E. coli* production of FMN lags well behind production of nascent chains and/or released apoflavodoxin. As a result, in case of destabilized protein such as F44Y apoflavodoxin, large amounts of this protein end up in cellular aggregates or become degraded by intracellular proteases, before F44Y protein can incorporate the stability-enhancing FMN. Thus, compared to C69A flavodoxin, less soluble F44Y flavodoxin can be purified from *E. coli*. Yet, still a significant amount of F44Y holo-protein is obtained from *E. coli*, as intracellular binding of FMN to native F44Y apoflavodoxin stabilizes the protein against both processes. FMN acts like a chemical chaperone, which protects apoflavodoxin against degradation and aggregation. It does not, however, act as a chaperone in the sense that it affects folding, as FMN binds only to natively folded apo-protein.

This latter finding may have implications for large-scale production of flavoproteins. It is well known that flavin incorporation stabilizes flavoproteins. As flavin production lags behind protein production, higher protein yields may thus occur in *E. coli* strains that have a high flavin production capability. Most research on increasing flavin production has been done using *Bacillus subtilis*, as it is one of the

organisms for biological production of riboflavin on an industrial scale (16,17). By expressing part of the *B. subtilis* *rib* operon the production of flavin in *E. coli* could be increased 3.6 fold (18). Increasing production of the enzymes riboflavin kinase and FAD synthase could boost conversion of riboflavin to respectively FMN and FAD. Would expression of unstable flavoproteins in such a strain increase their yield? Or would induction of the flavin operon before induction of protein expression also be a way to increase protein yield?

As discussed, **Chapter 4** reports that in case of F44Y apoflavodoxin the ribosome seems to force full-length nascent chain to the native state. However, this observation is based on ascertaining the rate by which FMN binds to the nascent chain. As **Chapter 2** shows, determination of this rate cannot distinguish between the presence of the on-pathway MG or natively folded apoflavodoxin. Whether both folding species exist on the ribosome is still unknown. While distinguishing between both folding states is difficult, NMR spectroscopy might be able to do so in future experiments. MGs have a reduced NMR signal compared to natively folded protein due to exchange between different conformations on the micro- to millisecond timescale.

While the experiments of **Chapters 3** and **4** increase our knowledge about the effect the ribosome has on protein folding, they still do not fully probe protein folding in a cell. The construct used for production of RNCs creates a somewhat artificial situation compared to translation of solely the flavodoxin gene. During the latter situation, translation termination and concurrent release of nascent chains occurs while approximately 30 to 40 C-terminal amino acids are still buried in the exit tunnel. These residues need to traverse this tunnel before full-length protein emerges from the ribosome. Situations like in **Chapters 3** and **4** where protein is stalled to the ribosome with the entire flavodoxin domain exposed outside the exit tunnel do not occur during translation of solely the flavodoxin gene. Despite these restrictions, demonstration of for example the importance of the five C-terminal amino acids for cofactor binding to apoflavodoxin remains valid. Binding of FMN to translating apoflavodoxin can thus only occur after the entire nascent chain has traversed the exit tunnel and the protein has obtained its native fold, as shown in **Chapter 3**.

With regard to the non-observed formation of the off-pathway MG during translation of the flavodoxin gene (**Chapter 4**) we have additional circumstantial evidence from previous work. Using fluorescence resonance energy transfer (FRET), it was discovered that folding of the C-terminal part of the off-pathway MG precedes folding of its N-terminal part (7). As cotranslational folding proceeds from the N- to the C-terminus, misfolding of apoflavodoxin's nascent chain may possibly be prevented before the entire protein is exposed outside the ribosome. Nevertheless, it is unknown where in the C-terminal part the folding of the off-pathway MG starts, thus it cannot be ruled out that this species may form during translation of the apoflavodoxin gene. Still, the influence of the ribosomal surface on apoflavodoxin's conformational space would be in effect, thereby potentially restraining formation of the off-pathway MG.

In **Chapter 5**, the folding of proteins with a flavodoxin-like fold, including flavodoxin- and CheY-related protein families, is reviewed. These proteins form mis-

folded intermediates, and several of these species are molten globular. Experiments suggest that these off-pathway intermediates are of helical nature. It would be interesting to investigate whether the ribosome also restrains the formation of folding intermediates of flavodoxins other than the one from *A. vinelandii* and of CheY, Spo0F and NtrC.

Phylogenetic analysis shows that all flavodoxins studied concerning their folding belong to different classes of bacteria, with *A. vinelandii* belonging to the gammaproteobacteria, *D. desulfuricans* to the deltaproteobacteria, *H. pylori* to the epsilonproteobacteria and *Anabaena* to the cyanobacteria (19). These classes occupy distinct environmental niches, as exemplified by the case of *H. pylori*, which can survive in the acidic environment of the stomach. The question arises whether the minor differences in folding behaviour between the various proteins as discussed in **Chapter 5** is due to adaption to these environmental niches. It might thus be appealing to investigate whether flavodoxins from the same bacterial class fold in similar ways. Until now no flavodoxin from an algal species has been studied for its folding behaviour. As one of the few eukaryotic classes that retain the flavodoxin protein, studying the folding of algal flavodoxin may give us insights into the evolution of this protein in eukaryotes and bacteria. Additionally one can query whether all flavodoxins fold in the same manner cotranslationally, in other words does the ribosome negate the differences in folding behaviour found *in vitro*? Does the ribosome restrain formation of the MGs of other flavodoxins, as it does for *A. vinelandii* flavodoxin? Or does adaptation to different environmental niches also lead to different cotranslational folding behaviour? And even if the ribosome does restrain MG formation for other flavodoxins, upon release from the ribosome, F44Y apoflavodoxin reverts back into MG (**Chapter 4**). Thus, while the ribosome may temporarily restrict MG formation of apoflavodoxin, without sufficient intracellular cofactor, the produced protein may relapse into the MG state.

Though our studies revealed the first glimpses of translation and folding of flavodoxin, and much still needs to be disclosed, they offer insight into nascent chain folding of multi-domain proteins that contain a domain with a flavodoxin-like topology. This architecture is part of several important multi-domain proteins, such as cytochrome P450 reductase and nitric oxide synthase (20,21). Translation of the C-terminal domains of such proteins continues when their N-terminal domains have already fully emerged from the ribosomal exit tunnel. Our findings imply that an N-terminal domain with a flavodoxin-like topology can already fold in a multi-domain protein while translation of the other domains still takes place. As proteins with a flavodoxin-like topology seem to fold via involvement of off-pathway MGs (22), the same may hold true for flavodoxin-like domains in multi-domain proteins. Hence, if this domain is N-terminally located, the ribosome may force this domain towards the native state, just as is the case for *A. vinelandii* flavodoxin (**Chapter 4**), and thereby protect the nascent chain against degradation or aggregation. Subsequent binding of FMN to an N-terminally located flavodoxin-like domain even further stabilizes the nascent chain (**Chapter 3**). It would be interesting to determine whether there is a correlation between positioning of cofactor-binding domains in multi-domain proteins and their stability during synthesis. If they predominantly locate in the N-terminal part of multi-domain proteins this protection against degradation and aggregation may be an important factor during cotranslational folding of these

proteins. One such example is the group of P450 fusion proteins, in which the N-terminal domain is in all but one case either a heme-binding P450 or a flavin-binding domain (23).

The field of cotranslational folding develops quickly, with new techniques facilitating breakthroughs. In a recent pioneering study, cotranslational folding has been studied in real time by incorporating fluorescent dyes in the nascent chain and following the resulting FRET signal during *in vitro* translation (24). It would be an exciting perspective to compare the results presented in this thesis with cotranslational folding of flavodoxin in real time, particularly in perspective of the meticulous probing by FRET of the *in vitro* folding of dye-labelled flavodoxin (7-9).

Another, less complicated, approach to study cotranslational folding is following proteolytic degradation of RNCs as a function of urea concentration (25). This could be a relatively quick method to further our understanding of the folding of apo-flavodoxin with differing nascent chain lengths. Besides the -5 and -10 constructs reported in **Chapter 3**, we have a library of further shortened nascent flavodoxin constructs, which are interesting candidates to investigate. While the proteolytic degradation technique cannot provide the time-resolution of for example real-time FRET, it would allow for a first selection of potentially interesting constructs concerning their folding, before further in depth investigations.

In vitro investigations of full-length, isolated proteins are important to improve our understanding of their folding, which is of benefit for example in industrial settings where purified proteins are used. A strong desire exists to investigate the conformational events a protein undergoes while it folds in a cell, events that have hardly been probed to date. This thesis shows that differences exist between protein folding *in vitro* and *in vivo*, for example because the ribosome can affect this process. Exploration of polypeptide folding on the ribosome, and how this process is assisted by chaperones, contributes to a molecular description of protein folding in the cell. Cotranslational folding is suspected to be critical for many newly synthesized polypeptides in order to prevent inter- or intra-molecular misfolding and aggregation (26,27). Increasing our knowledge of nascent chain folding is relevant to comprehend aspects of the molecular basis for folding defects. These defects underlie protein malfunction and protein homeostasis (28) failures that can ultimately manifest themselves as a wide variety of diseases (29-33).

One of the findings in this thesis that has potential applications is the role of FMN as a chemical chaperone. As humans lack the ability to synthesize riboflavin, which is the precursor to FMN and FAD, this cofactor must either be obtained through the diet or through uptake from synthesis by the gut microbiome (34,35). In human cells, riboflavin is then converted into FMN and FAD by respectively riboflavin kinase and FAD synthase (36). In the endocrine disorders hypothyroidism and adrenal insufficiency, this conversion is impaired (36). Many flavoproteins are essential proteins in cells and are amongst others involved in the biosynthesis of coenzymes, such as folate, coenzyme A and pyridoxal 5'-phosphate (PLP), and of various steroid-derived hormones (37). About two-thirds of all human flavoproteins may have mutations that are associated with diseases (37) and for several of these mutations it is known that they decrease cofactor affinity (38), such as for methylenetetrahydrofolate re-

ductase (39). In some cases of reduced cofactor affinity, supplementation of flavin upon onset of disease has already been shown to have a beneficial effect (38,40).

Binding of their respective flavin may not only stabilize a flavoprotein, but may also lead to dimerization to functionally active protein (41). Proteins that bind flavins covalently may require posttranslational modifications to complete the flavinylation, such as is the case for succinate dehydrogenase (42,43). Flavins can therefore act in many different ways to increase stability and activity of their binding partners, though the exact timing of cofactor binding is unknown for most flavoproteins (i.e. whether this happens co- or post-translationally). This thesis shows that the pharmacological chaperone function of flavins may already operate at the cotranslational level, as binding of FMN stabilizes a nascent chain and thereby protects it from degradation. Increasing the amount of available flavin is therefore a valid strategy to ameliorate diseases caused by inefficient cofactor binding of flavoproteins. Besides dietary supplementation, cofactor levels may be augmented by increasing the levels of riboflavin kinase and/or FAD synthase or by finding ways to boost synthesis by the gut microbiome. Cotranslational flavin binding is thus the first step for production of a stable and active protein and it may well be a general mechanism for flavoproteins to increase their stability during synthesis.

REFERENCES

1. Bollen, Y. J. M., Kamphuis, M. B., and van Mierlo, C. P. M. (2006) The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4095-4100
2. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2008) Extensive Formation of Off-Pathway Species during Folding of an alpha-beta Parallel Protein Is Due to Docking of (Non)native Structure Elements in Unfolded Molecules. *J. Am. Chem. Soc.* **130**, 16914-16920
3. Nabuurs, S. M., Westphal, A. H., and van Mierlo, C. P. M. (2009) Noncooperative Formation of the Off-Pathway Molten Globule during Folding of the alpha-beta Parallel Protein Apoflavodoxin. *J. Am. Chem. Soc.* **131**, 2739-2746
4. Nabuurs, S. M., Westphal, A. H., aan den Toorn, M., Lindhoud, S., and van Mierlo, C. P. M. (2009) Topological Switching between an alpha-beta Parallel Protein and a Remarkably Helical Molten Globule. *J. Am. Chem. Soc.* **131**, 8290-8295
5. Nabuurs, S. M., de Kort, B. J., Westphal, A. H., and van Mierlo, C. P. M. (2010) Non-native hydrophobic interactions detected in unfolded apoflavodoxin by paramagnetic relaxation enhancement. *Eur. Biophys. J.* **39**, 689-698
6. Nabuurs, S. M., and van Mierlo, C. P. M. (2010) Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of alpha-beta parallel protein flavodoxin. *J. Biol. Chem.* **285**, 4165-4172
7. Lindhoud, S., Westphal, A. H., Borst, J. W., and van Mierlo, C. P. M. (2012) Illuminating the Off-Pathway Nature of the Molten Globule Folding Intermediate of an alpha-beta Parallel Protein. *PLoS ONE* **7** 10.1371/journal.pone.0045746
8. Lindhoud, S., Westphal, A. H., van Mierlo, C. P., Visser, A. J., and Borst, J. W. (2014) Rise-time of FRET-acceptor fluorescence tracks protein folding. *Int. J. Mol. Sci.* **15**, 23836-23850
9. Lindhoud, S., Pirchi, M., Westphal, A. H., Haran, G., and van Mierlo, C. P. M. (2015) Gradual folding of an off-pathway molten globule detected at the single-molecule level. *J. Mol. Biol.* **427**, 3148-3157
10. van Mierlo, C. P. M., van Dongen, W. M. A. M., Vergeldt, F., van Berkel, W. J. H., and Steensma, E. (1998) The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate. *Protein Sci.* **7**, 2331-2344
11. Bollen, Y. J. M., Sanchez, I. E., and van Mierlo, C. P. M. (2004) Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin. *Biochemistry* **43**, 10475-10489
12. Engel, R., Westphal, A. H., Huberts, D. H., Nabuurs, S. M., Lindhoud, S., Visser, A. J., and van Mierlo, C. P. (2008) Macromolecular crowding compacts unfolded apoflavodoxin and causes severe aggregation of the off-pathway intermediate during apoflavodoxin folding. *J. Biol. Chem.* **283**, 27383-27394
13. Bollen, Y. J. M., Nabuurs, S. M., van Berkel, W. J. H., and van Mierlo, C. P. M. (2005) Last in, first out. *J. Biol. Chem.* **280**, 7836-7844
14. Houwman, J. A., Westphal, A. H., van Berkel, W. J., and van Mierlo, C. P. (2015) Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome.

BBA-Proteins Proteom. **1854**, 1317-1324

15. Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012) Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* **48**, 63-74

16. McAnulty, M. J., and Wood, T. K. (2014) YeeO from *Escherichia coli* exports flavins. *Bioengineered* **5**, 386-392

17. Abbas, C. A., and Sibirny, A. A. (2011) Genetic control of biosynthesis and transport of riboflavin and flavin nucleotides and construction of robust biotechnological producers. *Microbiol. Mol. Biol. Rev.* **75**, 321-360

18. Rabinovich, P., Beburow, M., Linevich, Z., Bandrin, S., and Stepanov, A. (1978) Expression of *Bacillus subtilis* riboflavin operon in *Escherichia coli* cells in the form of hybrid plasmids pPR1 and pPR2. *Doklady Akademii nauk SSSR* **238**, 1459

19. Perez-Dorado, I., Bortolotti, A., Cortez, N. s., and Hermoso, J. A. (2013) Structural and phylogenetic analysis of *Rhodobacter capsulatus* Niff: uncovering general features of nitrogen-fixation (nif)-flavodoxins. *Int. J. Mol. Sci.* **14**, 1152-1163

20. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8411-8416

21. Garcin, E. D., Bruns, C. M., Lloyd, S. J., Hosfield, D. J., Tiso, M., Gachhui, R., Stuehr, D. J., Tainer, J. A., and Getzoff, E. D. (2004) Structural basis for isozyme-specific regulation of electron transfer in nitric-oxide synthase. *J. Biol. Chem.* **279**, 37918-37927

22. Bollen, Y. J. M., and van Mierlo, C. P. M. (2005) Protein topology affects the appearance of intermediates during the folding of proteins with a flavodoxin-like fold. *Biophys. Chem.* **114**, 181-189

23. Guengerich, F. P., and Munro, A. W. (2013) Unusual cytochrome P450 enzymes and reactions. *J. Biol. Chem.* **288**, 17065-17073

24. Holtkamp, W., Kokic, G., Jager, M., Mittelstaet, J., Komar, A. A., and Rodnina, M. V. (2015) Cotranslational protein folding on the ribosome monitored in real time. *Science* **350**, 1104-1107

25. Samelson, A. J., Jensen, M. K., Soto, R. A., Cate, J. H. D., and Marqusee, S. (2016) Quantitative determination of ribosome nascent chain stability. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13402-13407

26. Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324-332

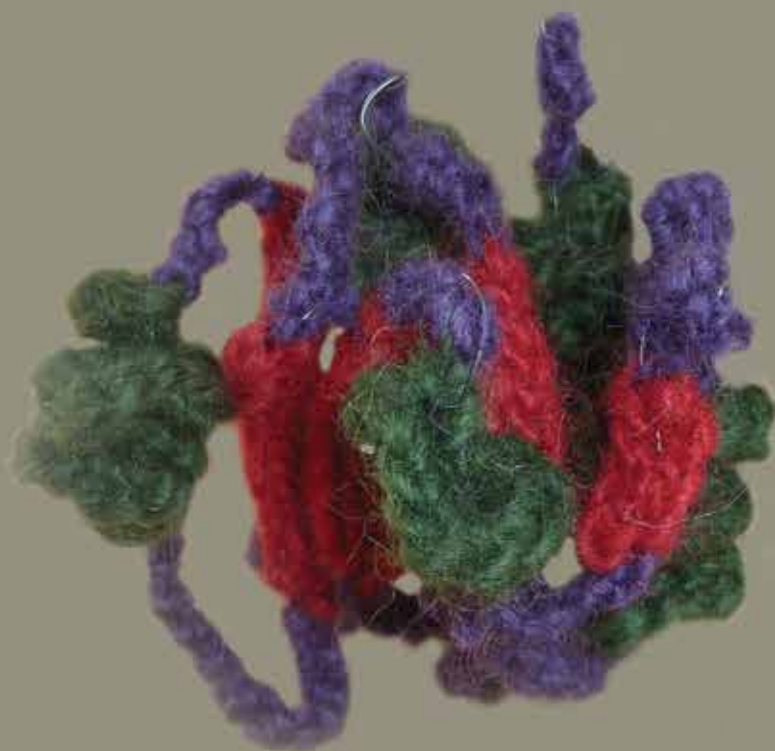
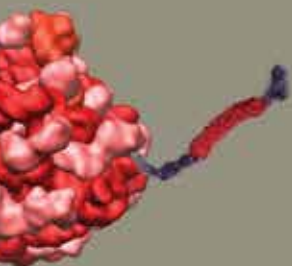
27. Eichmann, C. d., Preissler, S., Riek, R., and Deuerling, E. (2010) Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9111-9116

28. Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Ann. Rev. Biochem.* **78**, 959-991

29. Pechmann, S., Willmund, F., and Frydman, J. (2013) The ribosome as a hub for protein quality control. *Mol. Cell* **49**, 411-421

30. Dobson, C. M. (2003) Protein folding and misfolding. *Nature* **426**, 884-890
31. Luheshi, L. M., Crowther, D. C., and Dobson, C. M. (2008) Protein misfolding and disease: from the test tube to the organism. *Curr. Opin. Chem. Biol.* **12**, 25-31
32. Jucker, M., and Walker, L. C. (2013) Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45-51
33. Woerner, A. C., Frottin, F. D. R., Hornburg, D., Feng, L. R., Meissner, F., Patra, M., Tatzelt, J. R., Mann, M., Winklhofer, K. F., and Hartl, F. U. (2016) Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. *Science* **351**, 173-176
34. Magnusdottir, S. A., Ravcheev, D., de Crocy-Lagard, V. R., and Thiele, I. (2015) Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front. Gen.* **6**, 148
35. LeBlanc, J. G., Milani, C., de Giori, G. S., Sesma, F., Van Sinderen, D., and Ventura, M. (2013) Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr. Opin. Biotech.* **24**, 160-168
36. Said, H., and Ross, C. (2011) Riboflavin. *Modern nutrition in health and disease* **11**, 325-330
37. Lienhart, W.-D., Gudipati, V., and Macheroux, P. (2013) The human flavoproteome. *Arch. Biochem. Biophys.* **535**, 150-162
38. Ames, B. N., Elson-Schwab, I., and Silver, E. A. (2002) High-dose vitamin therapy stimulates variant enzymes with decreased coenzyme binding affinity (increased Km): relevance to genetic disease and polymorphisms. *Am. J. Clin. Nutr.* **75**, 616-658
39. Yamada, K., Chen, Z., Rozen, R., and Matthews, R. G. (2001) Effects of common polymorphisms on the properties of recombinant human methylenetetrahydrofolate reductase. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14853-14858
40. V Rodrigues, J., J Henriques, B., G Lucas, T., and M Gomes, C. (2012) Cofactors and metabolites as protein folding helpers in metabolic diseases. *Curr. Top. Med. Chem.* **12**, 2546-2559
41. Berkel, W. J., Benen, J. A., and Snoek M. C. (1991) On the FAD induced dimerization of apolipoamide dehydrogenase from *Azotobacter vinelandii* and *Pseudomonas fluorescens*. *Eur. J. Biochem.* **197**, 769-779
42. Robinson, K. M., and Lemire, B. D. (1996) Covalent attachment of FAD to the yeast succinate dehydrogenase flavoprotein requires import into mitochondria, presequence removal, and folding. *J. Biol. Chem.* **271**, 4055-4060
43. Hefti, M. H., Vervoort, J., and van Berkel, W. J. H. (2003) Deflavination and reconstitution of flavoproteins. *Eur. J. Biochem.* **270**, 4227-4242

Summary



ENGLISH SUMMARY

During and after their translation by the ribosome, folding of polypeptides to biologically active proteins is of vital importance for all living organisms. Gaining knowledge about nascent chain folding is required to enhance our understanding of protein folding in the cell. This in turn allows us to obtain insights into factors responsible for protein misfolding, aggregation, and, potentially, for numerous devastating pathologies.

In **Chapter 1** the model protein flavodoxin is introduced. Also theories about protein folding are presented, which led to the concept of the “folding energy landscape”. Flavodoxin folds via a misfolded off-pathway intermediate, which is molten globular and forms extensively during its refolding *in vitro*.

In **Chapter 2** we show that flavodoxin also populates an on-pathway molten globule during its folding. In the F44Y variant of apoflavodoxin, lowering the ionic strength induces the off-pathway molten globule state. By adding the cofactor FMN, we could follow aspects of the folding of this protein, as off-pathway molten globular flavodoxin first has to unfold and subsequently refold before FMN can bind. Thus, presence of the off-pathway molten globule retards FMN binding. We determined the presence of the off-pathway molten globule at decreasing ionic strengths with cofactor binding kinetics and polarized time-resolved tryptophan fluorescence spectroscopy. Comparison of both data sets revealed the presence of another, concurrently present molten globule. This species is most likely on-pathway to native protein. To our knowledge this is the first time that two concurrent molten globules have been discovered that reside on folding routes of decidedly different nature (i.e., on- and off-pathway ones).

While much work has been done on the folding of flavodoxin *in vitro*, the next step is to elucidate how this protein folds *in vivo*. In **Chapter 3** the first insights into cotranslational flavodoxin folding are presented. By using ribosomal nascent chains (RNCs) we could determine that when flavodoxin is fully exposed outside the ribosome it can bind its cofactor. However, while its five C-terminal amino acids are still sequestered in the ribosomal exit tunnel, the protein is in a non-native state and cannot bind FMN. Thus the last step in production of this flavoprotein *in vivo* is the binding of cofactor.

Chapter 4 reveals the influence of the ribosome on formation of the off-pathway molten globule of flavodoxin. By using RNCs of the F44Y variant of apoflavodoxin, we proved that the ribosome restrains formation of this molten globule. This discovery was possible by exploiting the findings of **Chapter 2** and **Chapter 3**, namely that cofactor binding kinetics slow down when off-pathway molten globule is present and that a fully exposed, natively folded flavodoxin nascent chain binds FMN. For F44Y RNCs no retardation in FMN binding occurs, whereas cofactor binding slows down

in case of isolated, full-length F44Y in the molten globule state. Thus the ribosome restrains formation of molten globules in stalled nascent flavodoxin. This is possibly due to electrostatic repulsion of the nascent chain by the ribosomal surface, as both are negatively charged, leading to entropic stabilization of native protein at physiological ionic strength.

In **Chapter 5** we review experiments and simulations concerning the folding of flavodoxins and CheY-like proteins, which share the flavodoxin-like fold. These proteins form intermediates that are off-pathway to native protein and several of these species are molten globules. This susceptibility to frustration is caused by the more rapid formation of an α -helix compared to a β -sheet, particularly when a parallel β -sheet is involved. The experimentally characterized off-pathway intermediates seem to be of α -helical nature. We discuss the probing of the cotranslational folding of flavodoxin as a first step towards a molecular description of how flavodoxin-like proteins fold *in vivo*.

Finally, **Chapter 6** touches upon the implications of our findings and possible applications in biotechnology, health and disease. A finding that has potential application is the role FMN has as a chemical chaperone. This chemical chaperone can already work at the cotranslational level, as binding of FMN stabilizes a nascent chain and thereby protects the nascent chain against degradation.

NEDERLANDSE SAMENVATTING

Vouwing van polypeptides tot biologisch actieve eiwitten is van groot belang voor alle levende organismen. Dit vouwingsproces gebeurt tijdens en na het maken van de polypeptide keten door ribosomen (dit proces heet translatie). Inzicht verkrijgen in de vouwing van eiwitten tijdens translatie van belang voor het begrijpen van eiwitvouwing in de cel. Deze kennis helpt ons om meer te weten te komen over de verschillende factoren die verantwoordelijk zijn voor het verkeerd vouwen van eiwitten en de vaak daarbij voorkomende aggregatie verschijnselen. Deze factoren zijn potentieel betrokken bij verscheidene ziektebeelden, zoals Alzheimer en Parkinson.

In **Hoofdstuk 1** wordt het model-eiwit flavodoxine geïntroduceerd, alsmede de eiwitvouwing theorie van de “vouwings-trechter” (folding funnel). Flavodoxine vouwt via een tussen vorm die verkeerd gevouwen is, een zogenaamde “off-pathway” intermediair. Deze intermediair vouwt als een “molten globule”, dat wil zeggen met secundaire structuur, maar zonder de pakking van een tertiaire structuur zoals in goed gevouwen globulaire eiwitten. Deze molten globule vormt in grote getale tijdens het hervouwen van flavodoxine *in vitro*.

In **Hoofdstuk 2** laten we zien dat flavodoxine ook via een “on-pathway” intermediair vouwt en dat deze intermediair ook een molten globule is. In tegenstelling tot de off-pathway intermediair, is de on-pathway wel groten deels correct gevouwen. In de F44Y variant van apoflavodoxine kan de vorming van de off-pathway molten globule geïnduceerd worden door de zout sterkte (en daarmee de ionische sterkte) te verlagen. Door het toevoegen van de cofactor van flavodoxine, FMN, konden we aspecten van de vouwing volgen. De off-pathway molten globule moet namelijk eerst ontvouwen en correct hervouwen voordat FMN kan binden. Als de off-pathway molten globule dus aanwezig is, wordt de binding van FMN vertraagd. We hebben de aanwezigheid van de off-pathway molten globule bij verschillende ion sterktes bepaald door middel van cofactor bindings-kinetiek en gepolariseerde tryptofaan fluorescentie spectroscopie. Door beide data sets te vergelijken kon de aanwezigheid van een tweede molten globule worden aangetoond. Deze tweede intermediair is gelijktijdig met de off-pathway molten globule aanwezig en is hoogstwaarschijnlijk on-pathway. Voor zover wij weten is dit de eerste keer dat twee gelijktijdige molten globules zijn aangetoond, die op verschillende vouwings-routes liggen (namelijk off-pathway en on-pathway).

Hoewel er al veel werk gedaan is op het gebied van vouwing van flavodoxine *in vitro*, is het nog onbekend hoe dit eiwit vouwt *in vivo*. In **Hoofdstuk 3** worden de eerste resultaten op het gebied van co-translationele vouwing van flavodoxine gepresenteerd. Door middel van “ribosomal nascent chains” (RNCs, eiwitketens die nog vast zitten aan het ribosoom) konden we bepalen dat flavodoxine zijn cofactor FMN kan binden als het helemaal te voor schijn is gekomen uit het ribosoom. Echter, wanneer zijn laatste vijf C-terminale aminozuur residuen nog in de exit tunnel van het ribosoom verblijven, kan flavodoxine niet vouwen naar zijn natieve toestand en kan FMN niet binden. De binding van FMN aan natief gevouwen flavodoxine is dus de laatste stap in de productie van dit flavo-eiwit.

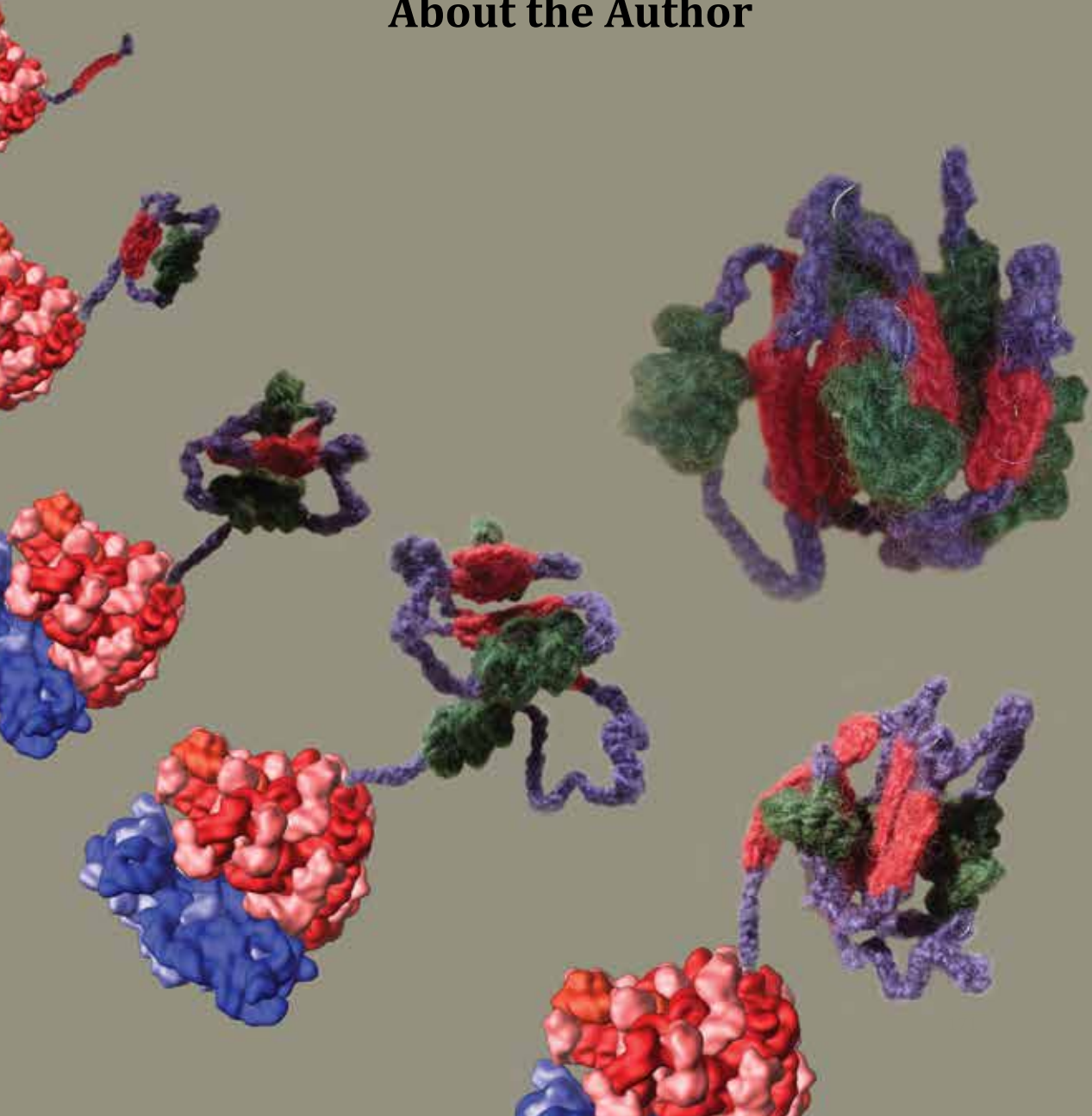
Hoofdstuk 4 beschrijft de invloed van het ribosoom zien op de formatie van de

off-pathway molten globule van flavodoxine. Door gebruik te maken van RNCs van de F44Y variant van apoflavodoxine konden we aantonen dat het ribosoom deze formatie verhindert. Deze ontdekking werd mogelijk gemaakt door onze bevindingen uit **Hoofdstuk 2** en **Hoofdstuk 3**, namelijk dat de kinetiek van cofactor binding vertraagd wordt wanneer de off-pathway molten globule aanwezig is en dat volledig uit de exit tunnel gekomen flavodoxine natief vouwt en FMN kan binden. Voor F44Y RNCs zien we geen vertraging in de binding van FMN, terwijl dit wel het geval is voor losgelaten F44Y onder dezelfde omstandigheden. Zodoende beïnvloedt het ribosoom de formatie van molten globules in flavodoxine dat nog vast zit aan het ribosoom. Waarschijnlijk komt deze beïnvloeding door elektrostatische repulsie van het flavodoxine polypeptide door het oppervlak van het ribosoom, aangezien beide negatief geladen zijn. Dit zorgt voor stabilisatie van het natief gevouwen eiwit bij fysiologisch relevante ionische sterktes.

In **Hoofdstuk 5** bediscussiëren we experimenten en simulaties betreffende de vouwing van verschillende flavodoxines en CheY-eiwitten, die eenzelfde soort vouwing hebben als flavodoxine. Deze eiwitten vormen allen intermediären die off-pathway zijn tot de natieve toestand en waarvan sommige gekarakteriseerd zijn als molten globule. Deze off-pathway intermediären vormen vanwege de snellere vouwing van α -helixen ten opzichte van β -sheets, met name wanneer een parallelle β -sheet moet worden gevormd. De experimenteel gekarakteriseerde off-pathway intermediären lijken allemaal alleen α -helices te hebben en geen β -structuren. Verder bespreken we de experimenten inzake de co-translationale vouwing van flavodoxine, welke een eerste stap zijn op weg naar een omschrijving van flavodoxine vouwing *in vivo*.

Ten slotte worden in **Hoofdstuk 6** de implicaties van onze resultaten en mogelijke industriële of farmaceutische toepassingen besproken. Een mogelijke applicatie zou het gebruik van FMN als chemische “chaperone” zijn. Dit chemische chaperone zou al op het co-translationale niveau kunnen werken, aangezien binding van FMN het flavodoxine polypeptide stabiliseert en zodoende beschermt tegen degradatie.

About the Author



ACKNOWLEDGEMENTS

En dan nu de laatste pagina's die door iedereen als eerste worden gelezen:

Allereerst mijn co-promotor Carlo: bedankt voor alle vrijheid die je me de afgelopen jaren hebt gegeven. Ook al was het soms misschien iets te veel van het goede (deadlines!) ik heb me er wel door tot een zelfstandige wetenschapper kunnen ontwikkelen. Je enthousiasme over mijn onderwerp was altijd groot en je positiviteit (ook als er we weer eens minder aardige referees tegen kwamen) heb ik zeer gewaardeerd.

Mijn promotor Willem, dankzij jou ben ik stage gaan lopen in Manchester bij een van jouw oud-PhD's wat ook de eerste stap bleek te zijn in de richting van mijn nieuwe baan. In jouw functie als "stok-achter-de-deur" heb je zeker in het laatste jaar de druk er op proberen te houden. Dank je wel voor alle hulp voor, tijdens en na de PhD! Tot ziens op CHAINS!

Tijdens het experimenteren voor mijn proefschrift ben ik geholpen door zes fantastische studenten, die allemaal erg veel werk hebben verzet! Pim, Sebastian, Sanne, Estelle, Wendy en Femke, heel erg bedankt voor al jullie hulp!

Antsje en Mieke, heel erg bedankt dat jullie mijn paranimfen wilden zijn! Antsje, dank dat je mijn kantoormaatje bent geweest en altijd in was om het kantoor wat op te vrolijken met hamsters in de boom en andere gekke dingen (wat zou er met de boom gebeurd zijn?). Mieke, twee maanden geleden was ik jouw paranimf, nu ben jij de mijne en zo is het cirkeltje weer rond! Meiden, ik vind het geweldig dat we de afgelopen 11 (!) jaar zoveel hebben kunnen delen.

Natuurlijk ook de rest van Enzymes@work bedankt: Adrie, Tom, Gudrun! Het was altijd erg gezellig op het lab en in het kantoor. Adrie, ook al zet je mijn auto op de handrem zonder dat ik het door heb, het zij je vergeven voor alle hulp in het lab. Tom, geen idee hoe je je hele PhD zo onverstoorbaar kunt blijven, had je niet wat mijn kant op kunnen sturen? Gudrun, even if you forget to come and drink tea with me in Amsterdam, you're still always welcome in Zeist for games and drinks (and probably also food, but Pascal is the cook).

Of course I also want to thank the rest of Biochemistry, you have made the past years very "gezellig" with many activities (the PhD-trip, watching bad movies, eating way too much unhealthy food at Gudrun's, WE-day...) and parties (weddings, PhD-parties, defences...). Good luck with all the remaining PhD's and everything else.

Martina and Stefan, thank you for the nice collaboration on the effect of chaperone Hsp90 on folding of flavodoxin. Martina, good luck with finishing your own PhD!

PhD-klaaggroep, we zijn straks bijna allemaal PhD af! Dat hadden we ons vijf jaar geleden denk ik nog niet zo kunnen voorstellen. Bas, Henriette, Mieke, Leonie en Antsje, ook al hebben we dan straks niks meer te klagen over onze PhD's, we kunnen het groepje vast hernoemen naar Werk-klaaggroep ;).

Pauline, dankjewel voor alle kopjes thee die we hebben gedronken als we het allebei even nodig hadden. En natuurlijk voor de uren in de sportschool tijdens yoga en pilates. Sorry dat je nu op zoek moet naar een nieuwe achterkant om naar te kijken tijdens de les!

Lieve spelletjes-fanaten dank jullie wel voor alle dagen/middagen/avonden afleiding die jullie mij bezorgd hebben!

Caro, ook al leek het er zeven jaar geleden nog op dat we allebei een heel andere loopbaan zouden hebben, opeens waren we allebei in het PhD-traject beland. En jij doet het dan ook nog eens met een kleintje er bij! Mocht je de komende jaren willen klagen, je weet waar ik woon ;).

Wilhelmien en Peter, dank voor alle hulp van de zomer met het huis! Dat jullie zo veel mee konden helpen heeft me een hoop rust gegeven om te kunnen schrijven. Harold en Daphne, ook al viel ik soms op de bank in slaap, de vrijdagavonden zijn wel altijd erg gezellig.

Pap, mam, goed voorbeeld doet volgen zeggen ze! En blijkbaar werkt dat, want nu heeft de meerderheid van ons gezin een PhD ;). Pap, dank voor de mathematische hulp bij het maken van de energielandschappen die in drie hoofdstukken voorkomen en het kritisch commentaar op mijn eerste artikel. Mam, ook al heb ik er het afgelopen jaar eigenlijk niet over willen praten, het werd wel altijd heel erg gewaardeerd dat je er naar vroeg. Meinou en Arfor, dank voor alle hulp die jullie me gaven om te ontspannen afgelopen zomer, zonder jullie zou de tuin nog steeds betegeld zijn!

Lieverd, het laatste dankwoord behoort jou toe. Dank je wel voor alle K's die je de afgelopen jaren aan me hebt gegeven (kopjes thee, koken en knuffels)! Ik ben zeker het afgelopen jaar niet altijd de gezelligste persoon in huis geweest, maar je hebt me altijd gesteund als ik het even niet meer wist. Ik heb er heel veel zin in om over twee weken met jou te vieren dat we al meer dan negen jaar samen op pad zijn en dat nog vele jaren vol willen houden!

LIST OF PUBLICATIONS

Leferink, N. G. H., Antonyuk, S. V., Houwman, J.A., Scrutton, N.S., Eady, R.R., Hasnain, S.S., (2014), Impact of residues remote from the catalytic centre on enzyme catalysis of copper nitrite reductase, *Nature communications* **5**, DOI:10.1038/ncomms5395

Houwman, J.A., Westphal, A.H., van Berkel, W.J.H., van Mierlo, C. P. M., (2015), Stalled flavodoxin binds its cofactor while fully exposed outside the ribosome, *BBA-Proteins Proteom.* **1854**, 1317-1324, DOI: 10.1016/j.bbapap.2015.06.004

Houwman, J. A., André, E., Westphal, A.H., van Berkel, W. J. H., van Mierlo, C. P. M., (2016), The ribosome restrains molten globule formation in stalled nascent flavodoxin, *J. Biol. Chem.* **291**, 25911-25920, DOI: 10.1074/jbc.M116.756205

Houwman, J. A., and van Mierlo, C. P. M., (2017), Folding of proteins with a flavodoxin-like architecture, *FEBS J.*, DOI: 10.1111/febs.14077 (*in press*)

Houwman, J. A., Westphal, A. H., Visser, A. J. W. G., Borst, J. W., van Mierlo, C. P. M., (2017), Concurrent on- and off-pathway molten globules under physiological conditions (*submitted*)

CURRICULUM VITAE

Joseline Houwman was born on October the 22nd 1988 in Ruedersdorf, Germany (former DDR). Shortly afterwards she moved to the Netherlands, where she spent most of her formative years, except for a 2-year stay in Oxford, UK. After completing her secondary education at the Stedelijk Gymnasium Nijmegen, she started her Bachelor in Molecular Life Sciences at Wageningen University in 2006. In 2009, during her BSc, she completed a minor in Geology at the University of Reykjavik, Iceland. There she followed courses in Volcanology and Glaciology. Back in the Netherlands, Joseline focused more on proteins and biochemistry. She received her BSc in 2010 and started her Master in Molecular Life Sciences with a specialization in Biological Chemistry in the same year. During her Masters she did a thesis on evolution of antibiotic resistance at the Laboratory of Genetics at Wageningen University. Afterwards she went to the Scrutton group at the Manchester Institute of Biotechnology at the University of Manchester, UK. There she investigated the order and timing of the reaction mechanism of a nitrogen reductase enzyme. This research led to a publication in Nature Communications. Joseline graduated from her Masters in 2011 and in December of the same year started a PhD in the Laboratory of Biochemistry at Wageningen University.



Joseline is currently working as a Research Technician in the Biocatalysis group of Dr. Francesco Mutti at the University of Amsterdam.

OVERVIEW OF COMPLETED TRAINING ACTIVITIES



Discipline specific activities

Courses

Magnetic resonance II and IIIa		
- Radboud University	Nijmegen (NL)	2012
3DM – Bio-Product	Nijmegen (NL)	2012
Stability, Folding, and Misfolding of Recombinant Proteins		
- ESSIB	Milano (IT)	2014
Design of biocatalysis: concepts, methods and applications		
- ESSIB	Stuttgart (DE)	2015

Conferences

The translating ribosome: towards mature proteins		
- Jacques Monod Conference – CNRS	Roscoff (FR)	2012*
Protein Folding in the Cell – FASEB	Saxton's River (US)	2014
Molecular chaperones: From molecules to cells and misfolding diseases – EMBO	Crete (GR)	2015*
NWO Study group meetings	Veldhoven (NL)	2012-2013
Molecular Chaperone Club		
- Dutch Chaperone Group	Utrecht (NL)	2013-2014*
CHAINS – chemistry as innovating science		
- NWO Chemical Sciences	Veldhoven (NL)	2015*-2016#

General courses

VLAG PhD Week (VLAG)	Baarlo (NL)	2012
Mini symposium “How to write a world-class paper”	Wageningen (NL)	2013
Scientific Writing (WGS)	Wageningen (NL)	2014
Career Orientation (WGS)	Wageningen (NL)	2015
Mobilising your - scientific - network (WGS)	Wageningen (NL)	2015

Optional activities

Preparation of research proposal	Wageningen (NL)	2011
Biochemistry PhD study tour (organizing committee)	UK	2013
Biochemistry weekly group meetings and colloquia	Wageningen (NL)	2011-2016
Wageningen PhD Council	Wageningen (NL)	2013-2015

* poster presentation, # oral presentation

The research presented in this thesis was performed at the Laboratory of Biochemistry, Wageningen University & Research, and was financially supported by the Netherlands Organization for Scientific Research (NWO; ECHO-711.011.007)

Financial support from the Laboratory of Biochemistry, Wageningen University for printing this thesis is gratefully acknowledged.

Cover design and layout by Joseline Houwman

Ribosome Shape by Vossman / Wikimedia Commons / CC-BY-SA-3.0

Printed by GVO drukkers & vormgevers, Ede, the Netherlands

