

# Surfing the free energy landscape of flavodoxin folding

**Promotor:**

Prof. dr. S.C. de Vries  
hoogleraar Biochemie  
Wageningen Universiteit

**Co-promotor:**

Dr. C.P.M. van Mierlo  
universitair docent  
Laboratorium voor Biochemie  
Wageningen Universiteit

**Promotiecommissie:**

Dr. J. Balbach, Universität Bayreuth, Deutschland  
Prof. Dr. S.M. van der Vies, Vrije Universiteit Amsterdam  
Prof. Dr. A.E. Mark, Universiteit Groningen  
Prof. Dr. J. de Vlieg, Universiteit Nijmegen  
Prof. Dr. H. van Amerongen, Wageningen Universiteit  
Dr. N. van Nuland, Universiteit Utrecht

Dit onderzoek is uitgevoerd binnen de onderzoekschool VLAG.

# Surfing the free energy landscape of flavodoxin folding

Yves J.M. Bollen

Proefschrift  
ter verkrijging van de graad van doctor  
op gezag van de rector magnificus  
van Wageningen Universiteit,  
Prof. dr. ir. L. Speelman,  
in het openbaar te verdedigen  
op dinsdag 15 juni 2004  
des namiddags te vier uur in de Aula.

Bollen, Yves Jean Marie

Surfing the free energy landscape of flavodoxin folding  
Thesis, Wageningen. – With ref. – With summary in Dutch  
ISBN 90-8504-064-7

Copyright © 2004 by Yves Bollen  
All rights reserved

## Contents

Abbreviations	6
1. Introducton	7
2. Formation of on- and off-pathway intermediates in the folding kinetics of <i>Azotobacter vinelandii</i> apoflavodoxin	17
3. Is the presence of intermediates observed in the folding kinetics of <i>Azotobacter vinelandii</i> apoflavodoxin governed by protein topology?	51
4. Partially unfolded forms of <i>Azotobacter vinelandii</i> apoflavodoxin identified by native state hydrogen exchange and NMR spectroscopy are off the productive folding route	65
5. Last in, first out: the role of cofactor binding in <i>Azotobacter vinelandii</i> flavodoxin folding	95
6. The influence of cofactor binding on local unfolding of <i>Azotobacter vinelandii</i> flavodoxin as determined by H/D exchange and NMR spectroscopy	123
Summary	151
Samenvatting	155
Nawoord	159
Curriculum vitae	161
List of publications	165

## Abbreviations

CD	Circular Dichroism
DSS	3-(trimethylsilyl)-1-propane sulfonic acid
FMN	Flavin Mononucleotide
GuDCI	deuterated Guanidinium Chloride
GuHCl	Guanidine Hydrochloride
HSQC	Heteronuclear Single Quantum Coherence spectroscopy
NMR	Nuclear Magnetic Resonance
pD	pH meter reading of a deuterated solution
RET	Resonance Energy Transfer

# 1 Introduction

Proteins consist of linear chains of amino acids with a specific sequence that differs per protein. In order to become functional, proteins usually fold to a compact three-dimensional structure that is dictated by the amino acid sequence. This 3D-structure is thermodynamically the most stable of all conformations that a protein chain can adopt under native circumstances. The precise details of the 3D-structure of a protein, and thus ultimately its sequence, determines its function. Proteins are involved in virtually all processes in living organisms. For example, proteins can catalyse a chemical reaction, sense the invasion of a hostile organism, maintain the cell structure or control the expression of certain genes.

The function of most proteins involves interactions with other molecules like cofactors, substrates, proteins, DNA or RNA. To be able to interact with these molecules, most proteins have to be dynamic in their native state. Consequently, proteins are only marginally stable and a dynamic equilibrium between folded, unfolded and partially folded protein states constantly exists. These states are populated according to Boltzmann's distribution law. In addition, the marginal stability of proteins is also required as many proteins in a cell are only required for a restricted period of time. A cell needs to dispose such proteins when their functions are no longer required and degrades them by using proteolytic enzymes. The latter process is facilitated when the target proteins have a not too high stability. This may be another reason why nature has selected proteins that are only marginally stable under native conditions.

### **Protein folding *in vitro***

*In vitro*, under native conditions most unfolded proteins spontaneously fold within milliseconds to seconds to their native state without any assistance. Despite more than forty years of research, it is still not well understood how proteins are able to fold so rapidly. Various small proteins (< 100 amino acids) are shown to fold *in vitro* in a single step to the native state (Jackson 1998). However, many small proteins populate transient intermediates (Tanford et al. 1973; Kuwajima et al. 1987; Matouschek et al. 1990; Jennings and Wright 1993). Some of these intermediates are obligatory structures on the direct folding route to the native state, whereas other intermediates are off the direct folding route. The latter intermediates have to unfold to produce native protein. Unfortunately, in most cases the kinetic role of the observed intermediate can not be determined (Baldwin 1996; Roder and Colon 1997).

### **Protein folding *in vivo***

The interior of a cell is a crowded environment. The effective protein concentration inside a typical cell is estimated to be approximately 300 mg/ml (Zimmerman and Trach 1991). Because of the high concentration of proteins and other macromolecules, the interior of a cell does not behave as an ideal solution, which influences the folding behaviour of proteins (Minton 2000). Compactly folded states are stabilised relative to (partially) unfolded states due to the limited space available in the cell (van den Berg et al. 2000). Another consequence of the high protein concentration in a cell is the increased probability of aggregation of (partially) unfolded proteins (van den Berg et al. 1999).



The synthesis of a typical protein in a cell lasts several minutes (Thulasiraman et al. 1999), whereas its folding takes place on a time-scale of milliseconds to seconds. As a consequence, the N-terminal part of the protein, which is synthesised first, may start to fold while the C-terminal part is still being synthesised. This may lead to misfolding (Creighton 1990; Jaenicke 1991), a problem that is not faced during *in vitro* folding studies. Cells have developed a system to avoid the formation of protein aggregates (Frydman 2001), allowing misfolded proteins to fold correctly. This system consists of protein chaperones which are able to bind either nascent polypeptides on the ribosome or misfolded proteins in the cell.

### **Folding-related diseases**

Despite the presence of chaperones in cells, aggregation of (partially) unfolded proteins still occurs. Such protein aggregates play an important role in a number of diseases, like Alzheimer's disease, Parkinson's disease, type II diabetes and prion-related diseases like Scrapie, Creutzfeldt-Jakob disease, and bovine spongiform encephalopathy (BSE) or mad cow disease. These diseases are characterised by the deposition of fibrillar structures of proteins known as amyloids (Dobson 2002).

Generally, amyloids consist of a single type of protein and are stabilised by an extended  $\beta$ -sheet structure formed by the protein backbone. Consequently, proteins with a wide variety of amino acid sequences can form amyloid structures, and the structure of an amyloid is independent of the amino acid sequence of the protein it is made of. In fact, the fibrillar form can be regarded as the intrinsic polymer form of a polypeptide chain. Evolution has selected proteins that are able to avoid aggregation by rapidly forming a stable monomeric structure, i.e. the native state. However, even partial unfolding will expose significant parts of the protein's interior, allowing the protein to aggregate and convert into amyloid fibrils under appropriate conditions. The strong hydrogen bonds between protein molecules in the fibril make this process effectively irreversible (Dobson 2002).

Amyloid diseases are generally related to a mutation that either makes the protein less stable or enhances the propensity of a partially folded form of the protein to aggregate. As a consequence of evolutionary selection and the presence of chaperones, aggregation of proteins *in vivo* is extremely slow. One of the characteristics of amyloid-related diseases is the late onset, generally after the age of reproduction. Until less than a century ago, most people did not live long enough to develop amyloid diseases. Hence, evolution was unable to completely cancel out amyloid diseases.

### **Protein unfolding can be biologically important**

In general, the unfolding of proteins *in vivo* for an extended amount of time is undesirable due to the risk of aggregation. However, in some specific cellular processes unfolding of a protein is required. One such process is protein translocation across membranes. Many proteins present in organelles or in the exterior of cells are produced in the cytoplasm. These proteins need to be translocated across one or more membranes by specific protein transport systems. In most cases, the substrate protein needs to be unfolded to be transported (Matouschek 2003). Another cellular process that involves protein unfolding is

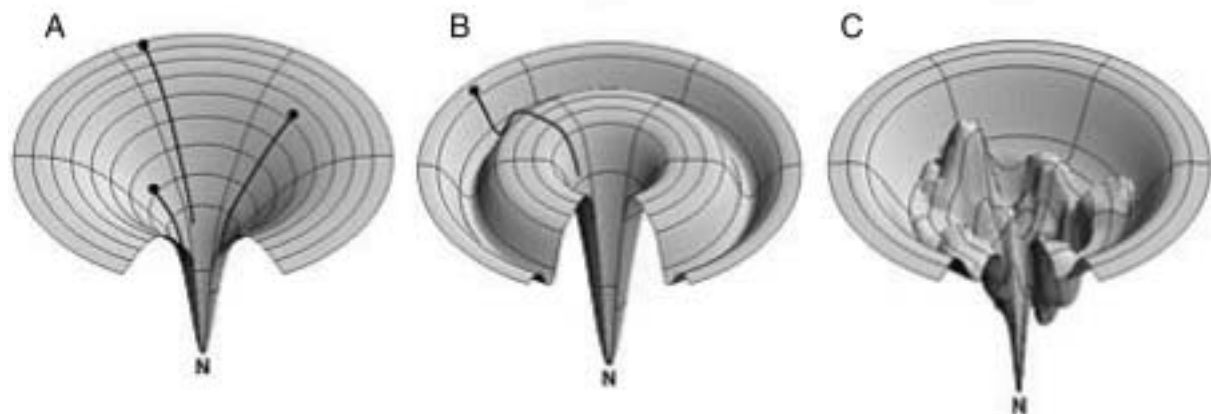
degradation of proteins by proteases, like the proteasome. In both translocation and degradation, the unfolding of the substrate protein appears to be induced by a pulling force acting at a specific location in the protein molecule. Consequently, the resistance of a protein against unfolding is in this case not determined by its global stability, but is instead determined by the local stability at the location where pulling takes place (Matouschek 2003).

For some proteins, an unfolding reaction is an essential part of their function. Photoactive yellow protein for example is a light-sensing protein present in photosynthetic bacteria. It has been demonstrated that the signalling state of the protein is partially unfolded. When photoactive yellow protein folds, this signalling state is an intermediate on its folding route (Lee et al. 2001). In this specific case, a folding intermediate, and not the native state, is the functional state of the protein.

An even more extreme example of non-native states that are biologically important is formed by a group of natively unfolded or intrinsically unstructured proteins that do not form a compact, stable and unique structure under native conditions. The evolutionary persistence of such proteins is strong evidence in favour of their biological importance, although not much is known about their function (Uversky 2002). Many of these proteins are believed to fold only upon binding to a specific ligand.

### The energy landscape of protein folding

Ever since Anfinsen and co-workers demonstrated that proteins can fold rapidly and spontaneously to their native state (Anfinsen 1973) it was realised that protein folding does not take place via a random sampling of the conformational space. Even for small proteins this process would take longer than the lifetime of the universe. To explain the fast folding behaviour of proteins, the concept of protein folding pathways arose, i.e. a protein folds along a defined pathway from the fully unfolded to the fully folded state. More recently, an important breakthrough was the realisation that there is not a single protein folding pathway, but that instead, a multidimensional energy landscape (Figure 1) better describes the folding process (Bryngelson et al. 1995; Dill and Chan 1997). In principle, there are many routes to



**Figure 1.** Schematic representations of folding free energy landscapes. (A) An idealised funnel landscape. (B) An energy landscape for the folding of a protein via a kinetic trap (i.e. an intermediate). (C) A rugged landscape with kinetic traps, energy barriers, and some narrow throughway paths to the native state. N represents the native state of a protein. The figures are taken from (Dill and Chan 1997).

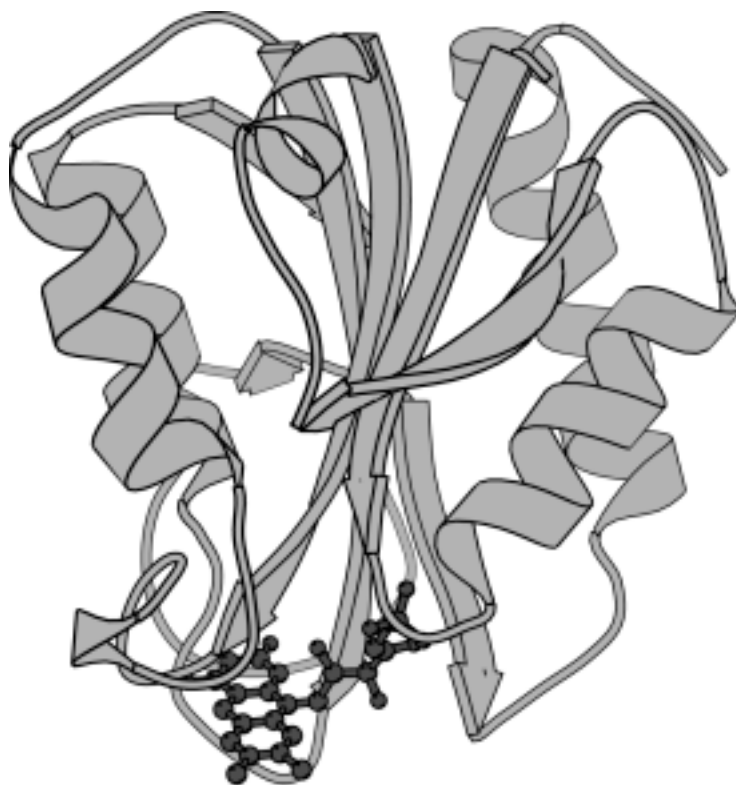
the native state and which pathway will be populated depends on the details of the system under investigation.

Although much has been learned about protein folding in the past decade, many mechanistic and dynamical aspects of this process remain uncharted. To further the theoretical insight into the principles that govern protein folding more experimental data are required. The study presented in this thesis aims to obtain insight into the free energy landscape for folding of a specific protein by using various spectroscopic techniques. The protein chosen is flavodoxin from *Azotobacter vinelandii*.

## Flavodoxin

In 1964, a Flavin Mononucleotide (FMN) containing protein was isolated from the obligate aerobe bacterium *Azotobacter vinelandii* (Shethna et al. 1964). Due to the initial difficulty to demonstrate the function of the protein, it has been known by several names, such as Shethna protein (Edmondson and Tollin 1971), *Azotobacter* free-radical protein (Hinkson and Bulen 1967) and Azotoflavin (Yoch et al. 1969). Based on structural similarities, the protein was believed to be a flavodoxin. However, it took several years before it was demonstrated that this protein could replace ferredoxin in the photosynthetic reduction of NADP<sup>+</sup> by spinach chloroplasts, the activity assay for flavodoxins used at the time (Van Lin and Bothe 1972). In contrast to the other flavodoxins known at the time, *A. vinelandii* flavodoxin's main function is not to replace ferredoxin under low iron conditions, but it is involved in the transport of electrons to nitrogenase, a nitrogen fixating enzyme (Yoch et al. 1969). Later, it was demonstrated that *A. vinelandii* produces three flavodoxins, with flavodoxin II being the most abundant under nitrogen fixating conditions (Klugkist et al. 1986).

Flavodoxins are small monomeric proteins (15 to 20 kDa). They adopt the  $\alpha$ - $\beta$  parallel or doubly-wound topology that consists of a central five-stranded parallel  $\beta$ -sheet surrounded on either side by  $\alpha$ -helices, as shown in Figure 2. This topology is the third most populated topology in the protein structure data bank (Brenner et al. 1997). In contrast to most other protein folds, it is shared by many (i.e. nine) protein superfamilies. These nine superfamilies exhibit little or no sequence homology and comprise a broad range of evolutionary unrelated proteins with different functions, like catalases, cutinases, esterases, chemotactic proteins and flavodoxins. The  $\alpha$ - $\beta$  parallel proteins offer a unique opportunity in the field of protein folding. Comparison of folding data of different members of this structural family enables the dissection of the influence of protein structure, which is conserved among members, and of protein sequence, which is highly divergent, on the folding behaviour of these proteins. Furthermore, as flavodoxins contain a tightly but non-covalently bound cofactor, the study of the influence of cofactor binding on the folding and stability of these proteins is possible.



**Figure 2.** Molscript cartoon drawing (Kraulis 1991) of the X-ray structure of *Azotobacter chroococcum* flavodoxin (Thorneley et al. 1994), the sequence of which is 95 % identical to *A. vinelandii* flavodoxin. The FMN cofactor is shown in dark grey ball-and-stick representation.

## Outline of the thesis

This thesis describes a study of the folding of *Azotobacter vinelandii* (strain ATCC 478) flavodoxin II. Flavodoxin is chosen as the representative of the group of  $\alpha$ - $\beta$  parallel proteins because the backbone  $^1\text{H}$ -,  $^{13}\text{C}$ - and  $^{15}\text{N}$ -NMR resonances of both holo- and apoflavodoxin have been assigned (Steensma et al. 1998; Steensma and van Mierlo 1998). This allows the use of NMR spectroscopy to study its folding and stability. Furthermore, apoflavodoxin, i.e. flavodoxin lacking the FMN cofactor, is known to populate a relatively stable folding intermediate during its equilibrium unfolding (van Mierlo et al. 1998; van Mierlo et al. 2000). This facilitates the study of the role of intermediates in protein folding.

Chapter 2 of this thesis describes a detailed study of the denaturant-induced equilibrium and kinetic (un)folding of apoflavodoxin, using fluorescence emission spectroscopy, fluorescence anisotropy and circular dichroism spectroscopy. Apoflavodoxin is shown to kinetically fold via two intermediates, according to the scheme:  $I_1 \rightleftharpoons \text{unfolded} \rightleftharpoons I_2 \rightleftharpoons \text{native}$ . One of these intermediates,  $I_1$ , is off-pathway and populates heavily during refolding and also populates during denaturant-induced equilibrium unfolding of apoflavodoxin. It is shown that  $I_1$  has to unfold before native apoflavodoxin can be formed. A minority of the folding apoflavodoxin molecules manages to circumvent this kinetic trap and folds more rapidly on a direct route to the native state. Intermediate  $I_2$  is shown to be on this direct folding route from unfolded to native apoflavodoxin. Its free energy is so high that it

never significantly populates, neither at equilibrium nor during kinetic folding. All folding apoflavodoxin molecules pass through this intermediate before reaching the native state.

The topology of a native protein is known to influence the rate with which it is formed. Does topology affect the appearance of folding intermediates and their specific role in kinetic folding as well? This question is addressed in Chapter 3 of this thesis by comparing the folding data obtained on apoflavodoxin from *A. vinelandii* with those available on two other  $\alpha$ - $\beta$  parallel proteins, i.e. a flavodoxin extracted from *Anabaena* and CheY from *Escherichia coli*, which is a chemotactic protein that is sequentially unrelated to flavodoxin. In addition, the results presented in Chapter 2 are compared to those obtained during a folding simulation of a protein that also adopts the  $\alpha$ - $\beta$  parallel topology. Based on the comparisons mentioned, it is shown that the appearance of folding intermediates in the class of  $\alpha$ - $\beta$  parallel proteins is apparently governed by protein topology.

In Chapter 4, amide hydrogen deuterium (H/D) exchange results detected by two-dimensional heteronuclear NMR spectroscopy are reported with *A. vinelandii* apoflavodoxin as the protein under investigation. It is shown that under native conditions the stabilities against local unfolding of all residues within apoflavodoxin are lower than the one determined for global unfolding of the protein. Only at denaturant concentrations exceeding 0.5 M, a few of the backbone amide protons of apoflavodoxin exchange via global protein unfolding. The use of native state H/D exchange leads to the identification of four partially unfolded forms of apoflavodoxin that are occasionally adopted by apoflavodoxin. In contrast to what has been generally proposed for partially unfolded forms, it is shown in Chapter 4 that these structures do not reside on productive folding and unfolding routes. Native state H/D exchange and classical denaturant-induced equilibrium and kinetic folding experiments together provide detailed information about the free energy landscape for protein folding, as is shown in Chapter 4 for apoflavodoxin folding.

In Chapter 5, the influence of the presence of the non-covalently bound flavin mononucleotide (FMN) cofactor on the global stability and on the kinetic folding of *A. vinelandii* holoflavodoxin are reported. It is shown that the stability of holoflavodoxin equals the sum of the global stability of native apoflavodoxin and the free energy associated with FMN binding to native apoflavodoxin. FMN binding to native apoflavodoxin occurs with two kinetically observable rate constants at all denaturant and protein concentrations studied. These two rate constants arise from two conformationally differing apoflavodoxin species, which most likely exist due to the binding of inorganic phosphate to the FMN phosphate binding site of a fraction of the *A. vinelandii* apoflavodoxin molecules. It is shown that excess FMN does not accelerate flavodoxin folding and that FMN does not act as a nucleation site for flavodoxin folding. During kinetic folding of holoflavodoxin formation of native apoflavodoxin precedes ligand binding. Even under strongly denaturing conditions, it is shown that global unfolding of holoflavodoxin occurs only after release of its FMN.

Finally, in Chapter 6 native state H/D exchange detected by NMR spectroscopy is used to probe the influence of FMN binding on the stability of flavodoxin against local, subglobal and global unfolding. It is shown that almost the entire flavodoxin backbone is substantially more rigid in holoflavodoxin than in apoflavodoxin. Only a few local unfolding

processes that occur at the periphery of flavodoxin opposite the FMN binding site are not affected by FMN binding. At least three out of the four partially unfolded forms that apoflavodoxin occasionally adopts under native conditions are inaccessible to holo flavodoxin. Holo flavodoxin can form these partially unfolded conformations only when FMN is released. Thus, it is shown that ligand binding influences the conformational dynamics of virtually the entire protein molecule.

Taking all observations together, a schematic representation of the free energy landscape of folding of *A. vinelandii* flavodoxin is presented that is consistent with the experimental data obtained.

## References

- Anfinsen, C. B. (1973). "Principles that govern the folding of protein chains." Science 181: 223-230.
- Baldwin, R. L. (1996). "On-pathway versus off-pathway folding intermediates." Fold Des 1(1): R1-8.
- Brenner, S. E., C. Chothia and T. J. Hubbard (1997). "Population statistics of protein structures: lessons from structural classifications." Curr Opin Struct Biol 7(3): 369-76.
- Bryngelson, J. D., J. N. Onuchic, N. D. Socci and P. G. Wolynes (1995). "Funnels, pathways, and the energy landscape of protein folding: a synthesis." Proteins 21(3): 167-95.
- Creighton, T. E. (1990). "Protein folding." Biochem J 270(1): 1-16.
- Dill, K. A. and H. S. Chan (1997). "From Levinthal to pathways to funnels." Nat Struct Biol 4(1): 10-9.
- Dobson, C. M. (2002). "Getting out of shape." Nature 418(6899): 729-30.
- Edmondson, D. E. and G. Tollin (1971). "Chemical and physical characterization of the Shethna flavoprotein and apoprotein and kinetics and thermodynamics of flavin analog binding to the apoprotein." Biochemistry 10(1): 124-32.
- Frydman, J. (2001). "Folding of newly translated proteins *in vivo*: the role of molecular chaperones." Annu Rev Biochem 70: 603-47.
- Hinkson, J. W. and W. A. Bulen (1967). "A free radical flavoprotein from *Azotobacter*. Isolation, crystallization, and properties." J Biol Chem 242(14): 3345-51.
- Jackson, S. E. (1998). "How do small single-domain proteins fold?" Fold des 3: R81-R91.
- Jaenicke, R. (1991). "Protein folding: local structures, domains, subunits, and assemblies." Biochemistry 30(13): 3147-61.
- Jennings, P. A. and P. E. Wright (1993). "Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin." Science 262(5135): 892-6.
- Klugkist, J., J. Voorberg, H. Haaker and C. Veeger (1986). "Characterization of three different flavodoxins from *Azotobacter vinelandii*." Eur J Biochem 155(1): 33-40.
- Kraulis, P. J. (1991). "MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures." J Appl Cryst 24: 946-950.
- Kuwajima, K., H. Yamaya, S. Miwa, S. Sugai and T. Nagamura (1987). "Rapid formation of secondary structure framework in protein folding studied by stopped-flow circular dichroism." FEBS Lett 221(1): 115-8.

- Lee, B. C., A. Pandit, P. A. Croonquist and W. D. Hoff (2001). "Folding and signaling share the same pathway in a photoreceptor." Proc Natl Acad Sci U S A 98(16): 9062-7.
- Matouschek, A. (2003). "Protein unfolding - an important process in vivo?" Curr Opin Struct Biol 13(1): 98-109.
- Matouschek, A., J. T. Kellis, Jr., L. Serrano, M. Bycroft and A. R. Fersht (1990). "Transient folding intermediates characterized by protein engineering." Nature 346(6283): 440-5.
- Minton, A. P. (2000). "Implications of macromolecular crowding for protein assembly." Curr Opin Struct Biol 10(1): 34-9.
- Roder, H. and W. Colon (1997). "Kinetic role of early intermediates in protein folding." Curr Opin Struct Biol 7(1): 15-28.
- Shethna, Y. I., P. W. Wilson, R. E. Hansen and H. Beinert (1964). "Identification by Isotopic Substitution of the Epr Signal at G = 1.94 in a Non-Heme Iron Protein from *Azotobacter*." Proc Natl Acad Sci U S A 52: 1263-71.
- Steensma, E., M. J. Nijman, Y. J. Bollen, P. A. de Jager, W. A. van den Berg, W. M. van Dongen and C. P. M. van Mierlo (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(2): 306-17.
- Steensma, E. and C. P. M. van Mierlo (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(3): 653-66.
- Tanford, C., K. C. Aune and A. Ikai (1973). "Kinetics of unfolding and refolding of proteins. 3. Results for lysozyme." J Mol Biol 73(2): 185-97.
- Thorneley, R. N. F., G. A. Ashby, M. H. Drummond, R. R. Eady, D. L. Hughes, G. Ford, P. M. Harrison, A. Shaw, R. L. Robson, J. Kazlauskaitė and H. A. O. Hill (1994). Flavodoxin and nitrogen fixation: Structure, electrochemistry and posttranslational modification by coenzyme A. Flavins and flavoproteins 1993. K. Yagi. Berlin, Walter de Gruyter & Co.: 343-354.
- Thulasiraman, V., C. F. Yang and J. Frydman (1999). "In vivo newly translated polypeptides are sequestered in a protected folding environment." Embo J 18(1): 85-95.
- Uversky, V. N. (2002). "What does it mean to be natively unfolded?" Eur J Biochem 269(1): 2-12.
- van den Berg, B., R. J. Ellis and C. M. Dobson (1999). "Effects of macromolecular crowding on protein folding and aggregation." Embo J 18(24): 6927-33.
- van den Berg, B., R. Wain, C. M. Dobson and R. J. Ellis (2000). "Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell." Embo J 19(15): 3870-5.
- Van Lin, B. and H. Bothe (1972). "Flavodoxin from *Azotobacter vinelandii*." Arch Mikrobiol 82(2): 155-72.
- van Mierlo, C. P. M., J. M. van den Oever and E. Steensma (2000). "Apoflavodoxin (un)folding followed at the residue level by NMR." Protein Sci 9(1): 145-57.

- van Mierlo, C. P. M., W. M. van Dongen, F. Vergeldt, W. J. van Berkel and E. Steensma (1998). "The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate." Protein Sci 7(11): 2331-44.
- Yoch, D. C., J. R. Benemann, R. C. Valentine and D. I. Arnon (1969). "The electron transport system in nitrogen fixation by *Azotobacter*. II. Isolation and function of a new type of ferredoxin." Proc Natl Acad Sci U S A 64(4): 1404-10.
- Zimmerman, S. B. and S. O. Trach (1991). "Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*." J Mol Biol 222(3): 599-620.



## 2 Formation of on- and off-pathway intermediates in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin<sup>\*</sup>

Yves J.M. Bollen, Ignacio E. Sánchez<sup>‡</sup>, and Carlo P.M. van Mierlo

The folding kinetics of the 179-residue *Azotobacter vinelandii* apoflavodoxin, which has an  $\alpha$ - $\beta$  parallel topology, have been followed by stopped-flow experiments monitored by fluorescence intensity and anisotropy. Single jump and interrupted refolding experiments show that the refolding kinetics involve four processes yielding native molecules. Interrupted unfolding experiments show that the two slowest folding processes are due to Xaa-Pro peptide bond isomerisation in unfolded apoflavodoxin. The denaturant dependence of the folding kinetics is complex. Under strongly unfolding conditions ( $> 2.5$  M GuHCl), single exponential kinetics are observed. The slope of the chevron plot changes between 3 and 5 M denaturant and no additional unfolding process is observed. This reveals the presence of two consecutive transition states on a linear pathway that surround a high-energy on-pathway intermediate. Under refolding conditions, two processes are observed for the folding of apoflavodoxin molecules with native Xaa-Pro peptide bond conformations, which implies the population of an intermediate. The slowest of these two processes becomes faster with increasing denaturant concentration, meaning that an unfolding step is rate-limiting for folding of the majority of apoflavodoxin molecules. It is shown that the intermediate that populates during refolding is off-pathway. The experimental data obtained on apoflavodoxin folding are consistent with the linear folding mechanism  $I_{\text{off}} \rightleftharpoons U \rightleftharpoons I_{\text{on}} \rightleftharpoons N$ , the off-pathway intermediate being the molten globule one that also populates during denaturant-induced equilibrium unfolding of apoflavodoxin.

---

<sup>\*</sup> A modified version of this Chapter is accepted for publication in Biochemistry.

<sup>‡</sup> Department of Biophysical Chemistry, Biozentrum der Universität Basel, Basel, Switzerland

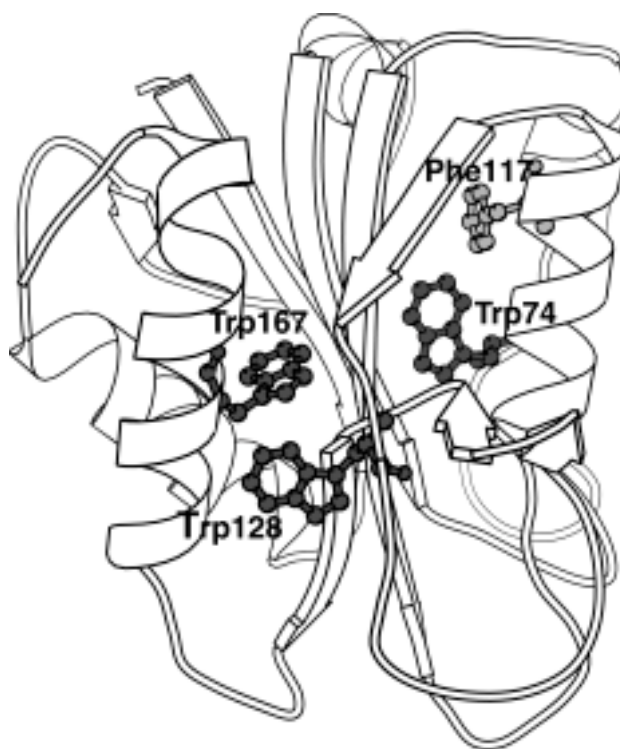
## Introduction

It is still not well-understood how an unfolded protein finds its unique native tertiary structure amongst an astronomically large number of possible conformations. The topology of a protein is proposed to control folding. A strong correlation has been found between topological characteristics and folding rates (Plaxco et al. 1998). In addition, the structure of the folding transition state can to some extent also be predicted from topology (Grantcharova et al. 2001). However, little is known about whether topology controls the formation of intermediates and the role they play in the kinetic folding mechanism (Clarke et al. 1999; Gunasekaran et al. 2001).

Various single-domain proteins form an intermediate structure before they fold to their native state (Tanford et al. 1973; Kuwajima et al. 1987; Matouschek et al. 1990; Jennings and Wright 1993). Whether folding intermediates are productive forms on the folding route or misfolded off-pathway species has been a subject of discussion for a long time (Baldwin 1996; Roder and Colon 1997). An evaluation of experimental data on 23 proteins that apparently fold in a two-state manner and display non-linear activation free-energy relationships has gained evidence for the involvement of obligatory high-energy on-pathway intermediates in protein folding (Sanchez and Kiefhaber 2003a). Recent insight suggests that whether an intermediate appears to be on or off the direct folding route depends to some extent on the experimental conditions used. There may be several parallel routes towards the native state, and relatively small changes in the environment or in the sequence of the protein may drastically alter the population of different parallel pathways (Kiefhaber 1995; Wildegger and Kiefhaber 1997; Bieri et al. 1999).

The study of the role of intermediates during folding is facilitated when these intermediates populate both kinetically and at equilibrium. Equilibrium population of an intermediate facilitates the determination of its structure, which can clarify why it is on or off the direct folding route under certain circumstances. This population is more likely to happen for large proteins (>100 residues), which therefore offer a special opportunity to study the role of intermediates in protein folding mechanisms.

The relatively large, 179-residue apoflavodoxin from *Azotobacter vinelandii* populates an intermediate with molten-globule like characteristics during denaturant- and thermally-induced equilibrium unfolding of the protein (van Mierlo et al. 1998; van Mierlo et al. 2000). Flavodoxins are flavoproteins that function as low-potential one-electron carriers and contain a noncovalently bound Flavin Mononucleotide (FMN) cofactor (Mayhew and Tollin 1992). Apoflavodoxin from *A. vinelandii* (i.e. flavodoxin in the absence of the FMN cofactor) is structurally identical to flavodoxin except for some dynamic disorder in the flavin-binding region (Steensma et al. 1998; Steensma and van Mierlo 1998).



**Figure 1.** MOLSCRIPT cartoon drawing (Kraulis 1991) of the X-ray structure of *Azotobacter chroococcum* flavodoxin (Thorneley et al. 1994), the sequence of which is 95 % identical to *A. vinelandii* flavodoxin. The side chains of the three tryptophan residues of *A. vinelandii* flavodoxin (dark grey) and of phenylalanine 117 (light grey) are shown in ball-and-stick representation.

Flavodoxins adopt the so-called doubly-wound or  $\alpha$ - $\beta$  parallel topology. They are all characterised by a five-stranded parallel  $\beta$ -sheet surrounded by  $\alpha$ -helices at either side of the sheet (Figure 1). The doubly wound topology is a rather popular fold: it belongs to the five most common observed folds, together with the TIM-barrel, Rossman, thiamin-binding and P-loop hydrolase folds (Gerstein 1997). In contrast to most protein folds, this topology is shared by many (i.e. nine) protein superfamilies (Brenner et al. 1997). These nine superfamilies exhibit little or no sequence similarity and comprise a broad range of unrelated proteins with different functions like catalases, chemotactic proteins, lipases, esterases, and flavodoxins.

Here, we present the experimental results of a kinetic folding study of *Azotobacter vinelandii* apoflavodoxin and discuss the influence of protein topology on the occurrence of kinetic folding intermediates. It is shown that kinetic folding of apoflavodoxin involves two intermediates. The change in denaturant accessibility upon unfolding of the first intermediate is 29 % of that of native apoflavodoxin, and the intermediate has to unfold before formation of the native state can occur. Instead, the second intermediate lies on a direct route from the unfolded to the native state of the protein and the change in denaturant accessibility upon unfolding of the second intermediate is approximately 80-90 % of that of native apoflavodoxin. It is shown that the off-pathway intermediate is the one that populates during equilibrium unfolding of apoflavodoxin. The kinetic results presented on apoflavodoxin

folding lead to the first experimental identification of both intermediates predicted theoretically to exist during the folding of an  $\alpha$ - $\beta$  parallel protein (Clementi et al. 2000).

## Materials and methods

### Materials

Guanidine hydrochloride (GuHCl, ultrapure) and potassium pyrophosphate were from Sigma (Bornem, Belgium).

### Protein expression and purification

The single cysteine residue 69 in wild-type *A. vinelandii* (strain ATCC 478) flavodoxin II was replaced by an alanine to avoid covalent dimerisation of apoflavodoxin. The mutant protein is largely similar to wild-type flavodoxin regarding both redox potential of the holoprotein and stability of the apoprotein (Steensma et al. 1996; van Mierlo et al. 1998). Recombinant *A. vinelandii* C69A holoflavodoxin was obtained and purified as described previously (Steensma et al. 1998; van Mierlo et al. 1998). Apoflavodoxin was subsequently prepared by trichloroacetic acid preparation (Edmondson and Tollin 1971; van Mierlo et al. 1998). Finally, apoflavodoxin molecules in an oligomeric state were removed via gel filtration on a Superdex 200 prep grade column (Pharmacia, Uppsala, Sweden).

### Fluorescence spectroscopy

Steady-state fluorescence measurements were done on a Cary eclipse fluorimeter equipped with a peltier accessory (Varian, Palo Alto, CA, USA). Excitation was at 280 nm with a slit of 5 nm; emission spectra were recorded from 320 to 360 nm in steps of 0.5 nm, using a slit of 2.5 nm. The temperature was 25 °C in all experiments; the protein concentration was 5.6  $\mu$ M in 100 mM potassium pyrophosphate, pH 6.0. All samples were equilibrated overnight at 25 °C.

### Circular dichroism (CD) spectroscopy

Steady-state CD measurements were performed on a Jasco J715 spectropolarimeter (Tokyo, Japan) equipped with a PTC-348WI peltier temperature control system. GuHCl unfolding samples were measured in a 1 mm quartz cuvette (Starna, Hainault, England) at 222, 225 and 255 nm, and averaged over 3 minutes per wavelength at a temperature of 25 °C. The ellipticity at 255 nm was subtracted from the other ellipticities as a baseline value. During all experiments the cell chamber was purged with dry nitrogen gas at a flow rate of 5 l/min. All samples were equilibrated overnight at 25 °C; the protein concentration was 5.6  $\mu$ M in 100 mM potassium pyrophosphate, pH 6.0.

### Fluorescence anisotropy

Fluorescence anisotropy was measured on a home-built fluorimeter equipped with two photomultipliers arranged in a T-format (Thorn EMI 9863QA/350, operating in photon-counting detection mode). Excitation light was generated by a 150 W short arc Xenon lamp

and the excitation wavelength of 300 nm was selected in a monochromator (Bausch & Lomb, Rochester, NY, USA) with a bandpass of 4.8 nm. Polariser were used in both the excitation light path (rotatable Glan Taylor polariser) and the emission light path (Polaroid, sheet). The emission light was selected with a 335 nm cut off filter (Schott). A blank measurement, containing all components except apoflavodoxin, was subtracted from each sample, and five measurements were averaged for each sample. The measurement chamber was thermostated at 25 °C by a circulating waterbath. All samples were equilibrated overnight at 25 °C; the protein concentration was 5.6 µM in 100 mM potassium pyrophosphate, pH 6.0.

### **Stopped-flow fluorescence spectroscopy**

Stopped-flow folding and unfolding experiments were performed on an Applied Photophysics (Leatherhead, UK) SX18-MV instrument with a 5 or a 20 µl cell. The pathlength of the observation chamber was 2 mm. Excitation was at 280 nm with a slit of 1.85 nm and a cut-off filter of 320 nm was used to select the emitted light. In unfolding experiments apoflavodoxin in 100 mM potassium pyrophosphate pH 6.0 was mixed 1:10 into the same buffer containing varying concentrations of GuHCl. In refolding experiments apoflavodoxin was first unfolded in 5.0 M GuHCl in 100 mM potassium pyrophosphate pH 6.0 for at least 30 minutes, and subsequently diluted 1:10 or 1:25 into the same buffer with a varying concentration of GuHCl. The final concentration of apoflavodoxin in the observation cell was 1 µM in all cases. At least four curves were averaged for each final denaturant concentration. The syringes and observation chamber were thermostated at 25 °C by a circulation waterbath.

#### *a. Refolding kinetics of freshly unfolded apoflavodoxin*

The influence of the presence of *cis* proline isomers on the faster folding rates was examined by refolding freshly unfolded protein. The same experimental set-up as described for the stopped-flow measurements was used, but now in the sequential mixing mode.

Apoflavodoxin was first unfolded at 3 M GuHCl in 100 mM potassium pyrophosphate, pH 6.0 by a 1:6 mixing step. After 622 ms, refolding was initiated by 1:6 mixing into the same buffer containing various concentrations of GuHCl. The final protein concentration in the observation chamber was always 1 µM. At least four curves were averaged for each final denaturant concentration.

#### *b. Interrupted unfolding of apoflavodoxin*

Interrupted unfolding was used to examine the origin of the different refolding processes. Apoflavodoxin was first unfolded at 3 M GuHCl for a varying time  $t_i$ , and then refolded by 1:6 mixing into native buffer, to yield a final GuHCl concentration of 0.5 M and a final protein concentration of 1 µM in 100 mM potassium pyrophosphate, pH 6.0. At least four curves were averaged for each  $t_i$ . The  $t_i$ -dependent amplitude of each individual refolding process gives the rate with which the species causing the respective refolding process is formed during the unfolding step at 3 M GuHCl.

Analogously, interrupted unfolding by manual mixing was used to examine the origin of the slowest folding process. A volume of 0.4 ml of a 30 µM apoflavodoxin stock solution

was mixed into 0.6 ml of a 5 M GuHCl solution. After a varying time  $t_i$ , 0.4 ml of this solution was mixed into a 3 ml fluorescence cuvette containing 2 ml of stirred buffer to result in a final protein concentration of 2  $\mu$ M and a final GuHCl concentration of 0.5 M in 100 mM potassium pyrophosphate, pH 6.0. Refolding traces of 100 seconds were recorded. All solutions were pre-equilibrated at 25 °C. Fluorescence traces were recorded on a Varian Cary Eclipse fluorimeter using an excitation wavelength of 280 nm, an excitation slit of 5 nm and an emission wavelength of 330 nm with a 20 nm slit.

### *c. Interrupted refolding of apoflavodoxin*

Interrupted refolding was used to examine whether all rates observed in refolding lead to native protein. Unfolded apoflavodoxin at 5 M GuHCl was first allowed to refold at 0.83 M GuHCl for a varying time  $t_i$  and subsequently unfolded at 5 M GuHCl. The buffer used was 100 mM potassium pyrophosphate pH 6.0. The amplitude corresponding to the observed unfolding process was used as a measure for the formation of native apoflavodoxin during the refolding time  $t_i$ . At least four curves were averaged for each  $t_i$ .

### **Stopped-flow fluorescence anisotropy**

Stopped-flow fluorescence anisotropy measurements were done on a BioLogic (Claix, France) SFM4 equipped with a 2x2 mm cuvette (FC-20), with the additional use of a photo-elastic modulator (Canet et al. 2001; Engel et al. 2002). Excitation was at 297 nm with a slit of 1 mm, resulting in a bandpass of 8 nm. A cut off filter of 335 nm (335FG01-25, Andover Corporation) was used to select the emission light. The solutions in the syringes as well as in the observation chamber were thermostated at 25 °C by a circulating waterbath. In the refolding experiment 10  $\mu$ M apoflavodoxin at 5 M GuHCl in 100 mM potassium pyrophosphate pH 6.0 was diluted ten fold into the same buffer without GuHCl, resulting in a final protein concentration of 1  $\mu$ M at 0.5 M GuHCl. The dead time was 14.8 ms and 14 curves were averaged. In the unfolding experiment 10  $\mu$ M apoflavodoxin in 100 mM potassium pyrophosphate pH 6.0 was diluted ten fold into the same buffer containing 5.0 M GuHCl, which results in a final protein concentration of 1  $\mu$ M and a GuHCl concentration of 4.5 M. 25 Curves were averaged.

### **Data analysis**

#### *Equilibrium (un)folding data*

Apoflavodoxin equilibrium (un)folding data as monitored by fluorescence emission intensity (at 340, 350 and 360 nm), circular dichroism (at 222 nm and 225 nm), and fluorescence anisotropy were globally fitted to a three-state model for equilibrium unfolding (equation 1), according to equations 2-6, using ProFit (Quantum Soft, Zürich):



$$K_{UI} = [I]/[U], K_{IN} = [N]/[I] \quad (2)$$

$$K_{ij}(D) = K_{ij}^0 \exp(m_{ij}^{eq}[D]) \quad (3)$$

$$\begin{aligned} f_U &= 1/(1 + K_{UI} + K_{UI}K_{IN}) \\ f_I &= K_{UI}/(1 + K_{UI} + K_{UI}K_{IN}) \\ f_N &= K_{UI}K_{IN}/(1 + K_{UI} + K_{UI}K_{IN}) \end{aligned} \quad (4)$$

$$Y^{obs} = (a_N + b_N[D]) \cdot f_N + (a_I + b_I[D]) \cdot f_I + (a_U + b_U[D]) \cdot f_U \quad (5)$$

$U$ ,  $I$ , and  $N$  represent the unfolded, intermediate and native state of the protein, respectively,  $K_{ij}$  the equilibrium constant of the  $i$ - $j$  equilibrium,  $K_{ij}^0$  the  $i$ - $j$  equilibrium constant at zero concentration denaturant,  $[D]$  the denaturant concentration,  $m_{ij}^{eq}$  the constant that describes the denaturant concentration-dependence of the equilibrium constant  $K_{ij}$ ,  $f_i$  the fractional population of state  $i$  at a certain denaturant concentration,  $Y^{obs}$  the observed spectroscopic signal,  $a_i$  the spectroscopic property of state  $i$  at zero concentration denaturant, and  $b_i$  the constant describing the denaturant concentration-dependence of the spectroscopic signal of state  $i$ . Each individual data point was weighted by the square of the corresponding standard error during the global fit procedure.

To incorporate the fluorescence anisotropy data in the global analysis procedure, a modification of equation 5 is required. Unlike fluorescence emission intensity and circular dichroism signals, fluorescence anisotropy signals do not linearly track the mole fractions of the three states of apoflavodoxin. This is due to the anisotropy signals being weighted by both the fraction of states involved and by the fluorescence quantum yield of each individual state (Eftink 1994). To describe the anisotropy signal adequately, the fraction of each species was multiplied with its intrinsic anisotropy (with corresponding denaturant concentration dependence) and weighted by its fluorescence intensity at 360 nm, according to equation 6.

$$Y_{anisotropy}^{obs} = \frac{\left( (a_{N,Fl} + b_{N,Fl}[D]) \cdot (a_{N,anis} + b_{N,anis}[D]) \cdot f_N \right.}{\left. + (a_{I,Fl} + b_{I,Fl}[D]) \cdot (a_{I,anis} + b_{I,anis}[D]) \cdot f_I \right.}{\left. + (a_{U,Fl} + b_{U,Fl}[D]) \cdot (a_{U,anis} + b_{U,anis}[D]) \cdot f_U \right)} \bigg/ \frac{\left( (a_{N,Fl} + b_{N,Fl}[D]) \right.}{\left. + (a_{I,Fl} + b_{I,Fl}[D]) \right.}{\left. + (a_{U,Fl} + b_{U,Fl}[D]) \right)} \quad (6)$$

with  $a_{i,Fl}$  being the intrinsic fluorescence intensity of species  $i$  at 360 nm,  $b_{i,Fl}$  the denaturant concentration dependence of  $a_{i,Fl}$ , and  $a_{i,anis}$  and  $b_{i,anis}$  the intrinsic anisotropy of state  $i$  and its corresponding denaturant concentration dependence, respectively.

The fluorescence emission spectrum of the equilibrium intermediate was calculated by fitting the equilibrium unfolding curves at all fluorescence emission wavelengths measured (320 – 360 nm) to equations 2 to 5 with the equilibrium constants  $K_{UI}$  and  $K_{IN}$  fixed to the values determined in the global fit mentioned before, using the program MatLab (Mathworks).

*Kinetic (un)folding data*

All kinetic fluorescence data and the  $t_i$ -dependent refolding and unfolding amplitudes in double-jump experiments were fitted to sums of exponential functions using ProFit (Quantum Soft, Zürich).

Of the four observed rate constants in stopped-flow (un)folding experiments, rate constant  $\lambda_2$  was fitted to the four-state linear folding mechanism described in Scheme 1.



In this linear scheme,  $U$  corresponds to unfolded protein,  $N$  to native protein,  $I_1$  and  $I_2$  to the off-pathway and on-pathway folding intermediate, respectively, and  $k_{ij}$  to the microscopic rate constant from state  $i$  to state  $j$ . Use is made of the analytical solution of a four-state folding mechanism as described by Ikai & Tanford (Ikai and Tanford 1973), and an exponential dependence of the microscopic rate constants on the denaturant concentration is assumed, as described by equation 7:

$$k_{ij} = k_{ij}^0 \exp(m_{ij}[D]) \quad (7)$$

with  $k_{ij}$  being the intrinsic rate constant for the transition from state  $i$  to  $j$  at a certain denaturant concentration  $[D]$ ,  $k_{ij}^0$  the rate constant  $k_{ij}$  at zero concentration denaturant, and  $m_{ij}$  the constant that describes the denaturant concentration dependence of  $k_{ij}$ . The rate constant  $k_{I_2 N}$  and the corresponding  $m_{I_2 N}$  were fixed to  $10^5$  and 0, respectively (see results for explanation).

**Results and Discussion****Equilibrium (un)folding of *A. vinelandii* apoflavodoxin involves a populated intermediate**

The equilibrium unfolding of *A. vinelandii* apoflavodoxin with GuHCl is followed by both circular dichroism (CD) and fluorescence emission spectroscopy. In Figure 2A the fluorescence emission data at 340 nm and the ellipticity values at 222 nm are shown, each analysed with a two-state model (Schmid 1992) for unfolding. The fit of the fluorescence data results in a global apoflavodoxin stability of  $6.24 \pm 0.25$  kcal/mol, with a corresponding  $m$ -value of  $-4.07 \pm 0.16$  kcal/mol M<sup>-1</sup>, resulting in a GuHCl concentration at the midpoint of unfolding ( $C_m$ ) of  $1.533 \pm 0.007$  M. The fit of the CD data results in a global apoflavodoxin stability of  $2.91 \pm 0.30$  kcal/mol, with a corresponding  $m$ -value of  $-1.69 \pm 0.14$  kcal/mol M<sup>-1</sup>, resulting in a  $C_m$  of  $1.72 \pm 0.05$  M GuHCl, values that are all significantly different from those determined on basis of the fluorescence data. The unfolding curves obtained by CD and fluorescence spectroscopy do not coincide, which is characteristic for the population of an intermediate in equilibrium apoflavodoxin unfolding, as reported previously (van Mierlo et

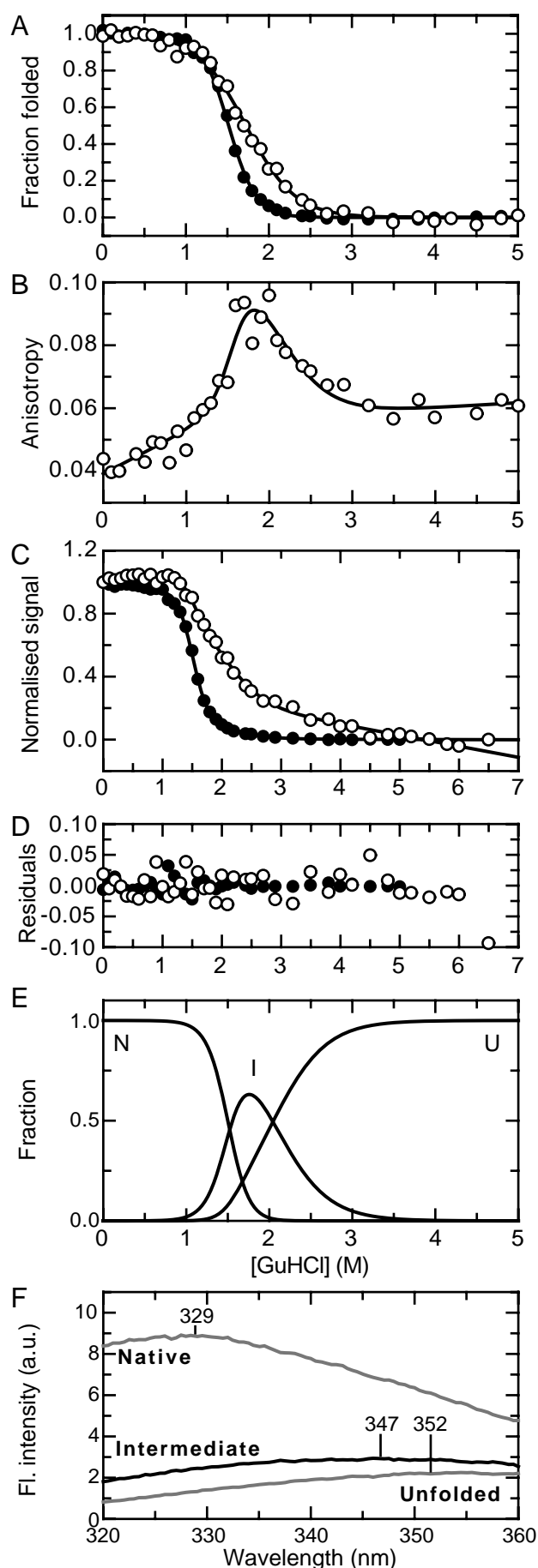


al. 1998). Analogously, thermal unfolding of apoflavodoxin results in apparent midpoint temperatures of  $48.6 \pm 0.1$  °C (fluorescence) and  $64.9 \pm 0.4$  °C (CD), a difference of 16.3 °C (van Mierlo et al. 1998).

The anisotropy of tryptophan fluorescence can also be used to monitor the equilibrium unfolding transition of a protein, and often gives quite different results compared to fluorescence intensity measurements (Otto et al. 1994; Beechem et al. 1995; Canet et al. 2001). The apoflavodoxin unfolding curve measured by anisotropy shown in Figure 2B is clearly biphasic, which illustrates the need of a folding intermediate to describe the equilibrium unfolding data of apoflavodoxin.

A three-state model (equations 1 to 6) is fitted globally to all apoflavodoxin equilibrium unfolding curves recorded, i.e. at two CD wavelengths (222 and 225 nm) and at three fluorescence intensity wavelengths (340, 350 and 360 nm) and obtained by fluorescence anisotropy. The denaturant-dependence of the spectroscopic parameters of the intermediate ( $b_i$  in equation 5) can not be accurately determined because the intermediate populates only in a small range of GuHCl concentrations and never populates for 100 %. Therefore,  $b_i$  is set to zero in the global fit procedure. The global fit of the three-state model to all apoflavodoxin (un)folding data is excellent (see e.g. Figures 2B and 2C) and the results are summarised in Table 1. In Figure 2C the results of the global fit to the denaturant-induced equilibrium unfolding data of apoflavodoxin as observed by CD at 222 nm and fluorescence intensity at 340 nm are shown. As can be judged from the residuals (Figure 2D) the equilibrium data shown in Figure 2C are well described by a three-state fit. In conclusion, at least one intermediate state populates in the transition zone of denaturant induced apoflavodoxin unfolding.

The stabilities against global unfolding of both native apoflavodoxin and of its equilibrium folding intermediate as presented in Table 1 are higher than those reported previously (van Mierlo et al. 1998). This is mainly caused by differences in the assumptions made about the spectroscopic properties of the apoflavodoxin equilibrium folding intermediate. In the previous analysis (van Mierlo et al. 1998), the fluorescence intensity of the intermediate is assumed to be equal to the one of the unfolded state and the CD ellipticity of the intermediate is set equal to the one of native apoflavodoxin. Due to the limited number of data points, the original data of van Mierlo et al. (van Mierlo et al. 1998) do not allow the sophisticated analysis presented here in which fluorescence intensities and CD ellipticities of the equilibrium folding intermediate are treated as free parameters.



**Figure 2.** Guanidine hydrochloride-induced equilibrium (un)folding of *A. vinelandii* apoflavodoxin. The protein concentration is 5.6  $\mu\text{M}$  in 100 mM potassium pyrophosphate pH 6.0, and the data are recorded at 25  $^{\circ}\text{C}$ . (A) The fluorescence emission intensity data recorded at 340 nm ( $\bullet$ ) and the circular dichroism data recorded at 222 nm (O) are each separately analysed according to a two-state model for (un)folding, and the resulting fractions of native apoflavodoxin molecules are shown. The curves do not coincide, as a folding intermediate populates during denaturant-induced apoflavodoxin (un)folding. (B) Guanidine hydrochloride unfolding profile of apoflavodoxin as measured by fluorescence anisotropy. The curve is clearly biphasic, due to the population of an intermediate during denaturant-induced apoflavodoxin (un)folding. (C) The fluorescence and CD data of apoflavodoxin (un)folding at 340 and 222 nm, respectively, are normalised such that the first data point has a value of 1 and the last data point has a value of 0. The solid lines in panels B and C represent the result of a global fit of a three state model for equilibrium (un)folding to the fluorescence data at 340, 350 and 360 nm, to the CD data at 222 and 225 nm, and to the anisotropy data in panel B, according to equations 1-6 (see materials and methods). (D) Residuals of the global fit of a three-state model for apoflavodoxin equilibrium (un)folding to the fluorescence data at 340 nm ( $\bullet$ ) and the CD data at 222 nm (O) shown in panel C. (E) Normalised population of native (N), intermediate (I) and unfolded (U) apoflavodoxin molecules as a function of the concentration GuHCl. (F) Computed fluorescence emission spectrum of the equilibrium folding intermediate of apoflavodoxin at 1.6 M GuHCl based on the global fit of a three-state model (equation 1) to the denaturant-induced equilibrium unfolding data of *A. vinelandii* apoflavodoxin. The fluorescence emission spectra of native apoflavodoxin and of unfolded apoflavodoxin, both at 1.6 M GuHCl, are shown in grey. The fluorescence emission maxima are shown for all three species.

The thermodynamic parameters resulting from the three-state fit to the apoflavodoxin (un)folding data are summarised in Table 1.

To investigate whether the ionic nature of GuHCl causes the observed population of an apoflavodoxin intermediate, the denaturant-induced equilibrium unfolding of apoflavodoxin has also been measured using urea as the denaturant and fluorescence intensity and CD as spectroscopic probes. The obtained urea-induced (un)folding curves are similar to the ones presented here (data not shown). As the urea-induced equilibrium unfolding curves of apoflavodoxin measured by fluorescence and CD also do not coincide, it is ruled out that the apoflavodoxin folding intermediate is observed due to a stabilisation caused by the salt-effect of GuHCl.

### Spectroscopic properties of the intermediate observed during equilibrium unfolding of *A. vinelandii* apoflavodoxin

In this section a low-resolution structural description of the equilibrium folding intermediate of apoflavodoxin is constructed based on the available spectroscopic data.

The fitted denaturant concentration dependencies of the stability of both native apoflavodoxin and of its equilibrium folding intermediate, together with the recorded fluorescence emission spectra between 320 and 360 nm at all GuHCl concentrations used, enables the calculation of the fluorescence emission spectrum of the equilibrium intermediate (Figure 2F). Despite the emission spectrum of the intermediate being quite similar to the emission spectrum of unfolded apoflavodoxin, which is shown in Figure 2F as a reference, they are not identical. The limited blue shift of the fluorescence emission maximum (5 nm) and increased fluorescence intensity of the intermediate relative to unfolded apoflavodoxin indicates that the tryptophan residues experience on average some shielding from the solvent. The change in denaturant accessibility upon unfolding of the equilibrium intermediate can be calculated from the reported equilibrium unfolding  $m$ -values (Myers et al. 1995) in Table 1 and is 29 % of that of the native protein.

The difference in ellipticity between the intermediate and the unfolded state of

**Table 1. Thermodynamic parameters obtained from a three-state fit ( $N \rightleftharpoons I \rightleftharpoons U$ ) to the GuHCl-induced equilibrium unfolding data of *A. vinelandii* apoflavodoxin<sup>a</sup>**

$\Delta G_{UI}$ (kcal/mol)	$3.74 \pm 0.49$	$m_{UI}^{eq}$ (kcal/mol M <sup>-1</sup> )	$-1.83 \pm 0.19$
$\Delta G_{IN}$ (kcal/mol)	$6.70 \pm 0.17$	$m_{IN}^{eq}$ (kcal/mol M <sup>-1</sup> )	$-4.40 \pm 0.11$
$\Delta G_{UN}$ (kcal/mol)	$10.45 \pm 0.52$	$m_{UN}^{eq}$ (kcal/mol M <sup>-1</sup> )	$-6.23 \pm 0.23$

<sup>a</sup> Equilibrium unfolding curves of *A. vinelandii* apoflavodoxin as obtained by CD at 222 and 225 nm, by fluorescence intensity at 340, 350, and 360 nm and by fluorescence anisotropy. The errors shown are standard deviations. U, I and N represent the unfolded, intermediate and native state of apoflavodoxin.  $\Delta G_{XY}$  is the difference in free energy between species X and Y at zero molar denaturant, and  $m_{XY}^{eq}$  is the dependence of  $\Delta G_{XY}$  on the denaturant concentration.

apoflavodoxin at 0 M GuHCl is 65 % of the corresponding ellipticity difference between native and unfolded apoflavodoxin at both 222 and 225 nm. Consequently, approximately 65 % of the native  $\alpha$ -helical content of apoflavodoxin is probably present in the equilibrium intermediate, as at these wavelengths the CD spectrum is especially sensitive to  $\alpha$ -helical structure.

Notably, the anisotropy of the apoflavodoxin folding intermediate (Figure 2B), which populates significantly between 1 and 3 M GuHCl is much higher than the anisotropy of the native state. Fluorescence anisotropy informs about the rotational freedom of the fluorophore involved. Tryptophan sidechains that can rotate freely, like in an unfolded protein, have an anisotropy value of approximately 0.06, depending on the wavelength of excitation and emission. The anisotropy of an immobilised tryptophan sidechain in a folded protein depends on the size of the protein. In case of apoflavodoxin, an anisotropy of 0.10 – 0.12 is expected when all of its three tryptophan sidechains are immobilised (Otto et al. 1994; Beechem et al. 1995; Canet et al. 2001). The observed fluorescence anisotropy of native apoflavodoxin is surprisingly low (i.e. 0.04, Figure 2B). In case of apoflavodoxin, no indications for a high tryptophan sidechain mobility in the native state are observed, since the hydrogen exchange rates of the tryptophan sidechain amides of apoflavodoxin correspond to local stabilities of 5 – 7 kcal/mol (Steensma and van Mierlo 1998). Apparently, a specific tertiary interaction involving one or more tryptophans causes the low anisotropy of native apoflavodoxin.

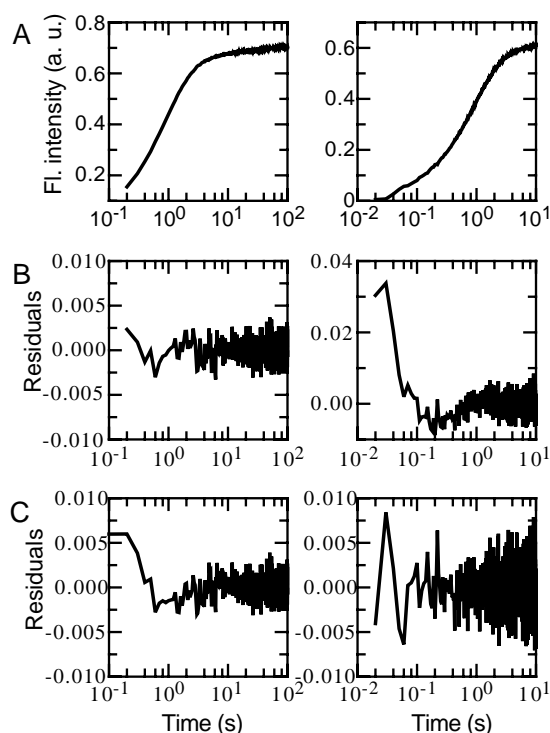
The fluorescence of *A. vinelandii* apoflavodoxin arises from its three tryptophan sidechains, which are shown in Figure 1. Trp 74 is located in  $\alpha$ -helix 3, and is 3.3 Å away from the sidechain of Phe 117 in  $\alpha$ -helix 4 of apoflavodoxin. The orientation of these aromatic rings relative to one another is close to perpendicular. The fluorescence of Trp 74 is probably quenched by the close-to-perpendicular phenyl ring, as observed in a similar situation for the protein FKBP59-I (Rouviere et al. 1997). Consequently, the emission spectrum is probably dominated by the two remaining tryptophans. The distance between Trp 128, which is close to  $\beta$ -strand 5a, and Trp 167, which is in helix 5, is 4.5 Å, and their relative orientation in apoflavodoxin is close to perpendicular as well. The two indole rings are well within the Förster distance for resonance energy transfer (RET) of one another (10 – 15 Å for tryptophans in a hydrophobic protein environment). RET between two tryptophans reduces the fluorescence anisotropy (Lakowicz 1999). Since the two indole rings are in close proximity and almost perpendicular to each other, this reduction in anisotropy would be efficient, explaining the low anisotropy observed for native apoflavodoxin. As a consequence, the equilibrium folding intermediate, as well as other compact structures formed during the folding of apoflavodoxin, are expected to have a higher anisotropy than both native and unfolded apoflavodoxin. In such intermediates the tryptophan side chains can be immobilised but the highly specific native tertiary interactions that cause the low native anisotropy are not likely. This feature explains the remarkable unfolding profile of apoflavodoxin as observed by fluorescence anisotropy (Figure 2B).

In summary, the apoflavodoxin equilibrium folding intermediate has a fluorescence emission spectrum which shows that its tryptophan sidechains are little shielded from the

solvent. As can be inferred from fluorescence anisotropy data, these tryptophan sidechains are immobilised but their immediate environment is not fully native. Based on CD data, approximately 65 % of the  $\alpha$ -helical content of native apoflavodoxin is present in the intermediate. The NMR cross peaks of the equilibrium folding intermediate of apoflavodoxin are extremely broadened, which suggests that this intermediate is highly dynamic on a millisecond to second time-scale (van Mierlo et al. 2000). All spectroscopic data put together show that the apoflavodoxin equilibrium folding intermediate has molten globule-like structural characteristics (Ptitsyn 1992).

### Kinetic characterisation of *A. vinelandii* apoflavodoxin folding by stopped-flow fluorescence shows the presence of an on-pathway and an off-pathway intermediate

#### a. Refolding and unfolding kinetics of apoflavodoxin monitored by intrinsic Trp fluorescence

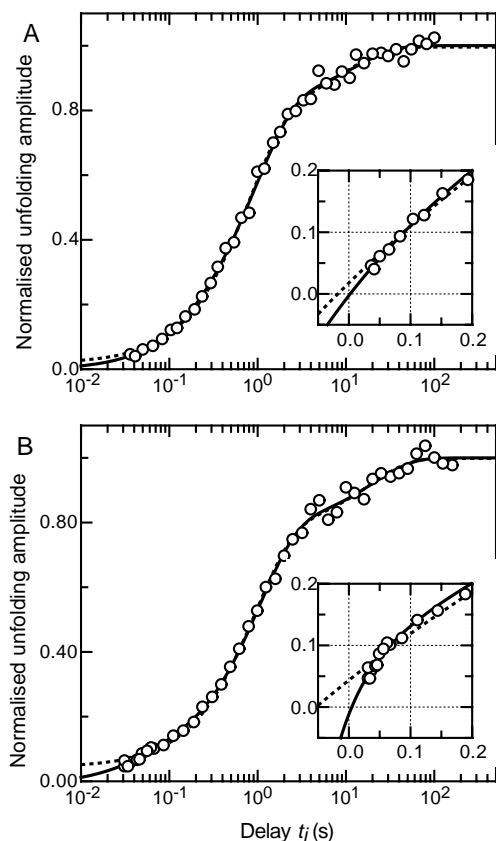


**Figure 3.** Stopped-flow refolding kinetics of *A. vinelandii* apoflavodoxin at 0.56 M GuHCl as observed by tryptophan fluorescence emission intensity. (A) Fluorescence data recorded for a period of 100 seconds (left) and 10 seconds (right). Residuals of a fit of a sum of three exponential equations to both data traces are shown in B. In C, the residuals of a global fit of a sum of four exponential functions to both the 100 s and 10 s data are shown, with  $\lambda_1 = 23.6 \pm 1.6 \text{ s}^{-1}$ ,  $\lambda_2 = 1.16 \pm 0.01 \text{ s}^{-1}$ ,  $\lambda_3 = 0.33 \pm 0.01 \text{ s}^{-1}$ , and  $\lambda_4 = 0.025 \pm 0.001 \text{ s}^{-1}$ . The data are plotted on a logarithmic time scale. The final protein concentration is 1  $\mu\text{M}$  in 100 mM potassium pyrophosphate pH 6.0, at 25  $^{\circ}\text{C}$ .

To investigate the role of the molten-globule like equilibrium intermediate in the kinetic folding of apoflavodoxin, a stopped-flow folding experiment is performed in 100 mM potassium pyrophosphate, pH 6.0, at 0.56 M GuHCl. Figure 3A shows the refolding of apoflavodoxin recorded by fluorescence spectroscopy for a period of 100 seconds and 10 seconds, respectively. The fluorescence emission increases with time, because unfolded apoflavodoxin has a lower emission than native protein (Figure 2F). Figure 3B shows the residuals of a fit of a sum of three exponentials to the data. Three exponentials suffice to describe the observed signal in the 100 s trace. However, the 10 second trace is not properly fitted with these three exponentials, which indicates that an additional process is present. As a consequence, a sum of four exponential curves is globally fitted to both 10 s and 100 s folding traces. The corresponding residuals of both curves at 0.56 M GuHCl (Figure 3C) show that the data are now correctly described by four folding processes (with folding rate constants  $\lambda_1$  to  $\lambda_4$ ).

### b. Formation of native molecules during apoflavodoxin folding

To determine whether all four observed processes produce native apoflavodoxin or whether some of the observed processes represent the formation of partially folded intermediates, an interrupted refolding experiment (Kiefhaber 1995) is done. In such an



**Figure 4.** Formation of native molecules during refolding of *A. vinelandii* apoflavodoxin. The normalised amplitude of the unfolding process of apoflavodoxin in an interrupted refolding experiment at 0.83 M GuHCl (A) and at 0.42 M GuHCl (B) is shown as a function of the refolding delay time  $t_i$ . All data are shown on a logarithmic time scale except in the insets, in which the data obtained during the first 0.2 seconds of interrupted refolding are shown on a linear time scale. The dashed line is the result of a fit of a sum of two exponentials to the data, and the solid line is the result of a fit of a sum of three exponentials to the data. The inset shows that the data are best described by three exponentials, and that in this fit the amplitude of the unfolding process is 0 at  $t_i = 0$ , as required. The rate constants and corresponding amplitudes are listed in Table 2. The final protein concentration is 1  $\mu$ M in 100 mM potassium pyrophosphate pH 6.0, at 25  $^{\circ}$ C. For further details see the Materials and Methods section.

experiment, unfolded apoflavodoxin is allowed to fold for a certain time  $t_i$ , and is subsequently unfolded in 5 M GuHCl and followed by fluorescence spectroscopy. The observed amplitude of the unfolding process, with a rate constant of 48  $\text{s}^{-1}$  in case of *A. vinelandii* apoflavodoxin, is then used as a measure for the amount of native apoflavodoxin that is formed during the refolding time  $t_i$ . Generally, folding intermediates unfold much faster than native proteins (Schmid 1992), and thus the apoflavodoxin folding intermediates formed during  $t_i$  are assumed not to contribute to the observed kinetic trace.

Figure 4 shows the amplitudes of the unfolding processes as a function of the delay time  $t_i$  at two different GuHCl concentrations. A fit of a sum of two exponential equations to these data extrapolates to  $1.8 \pm 0.6$  and  $4.4 \pm 0.4$  % of all molecules being native at  $t_i = 0$ , respectively, which would mean that an additional route to the native state is not accounted for in the fit. The data are best described by three exponentials, including a fast process. This fit extrapolates, within experimental error, to 0 % of all molecules being native at  $t_i = 0$ . The two fastest rate constants of the latter fit correspond to the folding rates  $\lambda_1$  and  $\lambda_2$  observed during conventional apoflavodoxin refolding at the respective GuHCl concentrations. This shows that native apoflavodoxin is formed in the two fastest processes observed. The third rate constant observed in the interrupted refolding experiment is in between  $\lambda_3$  and  $\lambda_4$  observed in the conventional refolding experiment. The low sensitivity of the interrupted refolding

experiment compared to the direct refolding traces prevents discrimination of the latter two slow processes. The interrupted refolding data at both GuHCl concentrations can be described well using the four respective rate constants observed in direct refolding of apoflavodoxin. We interpret this to imply that indeed native apoflavodoxin is formed in all four processes observed in the conventional refolding experiment.

All four apoflavodoxin folding processes observed represent a route to the native state without intermediates populating to a significant extent on the timescale investigated. This can be concluded as no lag is observed in the formation of native molecules during apoflavodoxin folding (Figure 4) (Kiefhaber 1995). The observed rate constants  $\lambda_1$  to  $\lambda_4$  and corresponding amplitudes that characterise apoflavodoxin folding, as observed by either Trp fluorescence changes during direct refolding experiments or by the time course of formation of native apoflavodoxin in interrupted refolding experiments, are summarised in Table 2.

**Table 2. Rate constants ( $\lambda_1$  to  $\lambda_4$ , in units of  $s^{-1}$ ) and corresponding amplitudes (A1 to A4, in percentages) obtained for the direct refolding and the interrupted refolding of *A. vinelandii* apoflavodoxin at 0.42 M and 0.83 M GuHCl (for further details see Materials and Methods).<sup>a</sup>**

	0.42 M GuHCl		0.83 M GuHCl	
	Direct ref.	Interrupted ref.	Direct ref.	Interrupted ref.
$\lambda_1$	$12 \pm 1$	$22 \pm 3$	$9.5 \pm 1.6$	$5.6 \pm 1.1$
A1	$5.9 \pm 0.3$	$8.4 \pm 0.3$	$3.2 \pm 0.3$	$9.1 \pm 1.5$
$\lambda_2$	$1.06 \pm 0.01$	$0.97 \pm 0.01$	$1.22 \pm 0.01$	$1.04 \pm 0.03$
A2	$76 \pm 1$	$71.0 \pm 0.6$	$74.2 \pm 0.5$	$73.8 \pm 1.4$
$\lambda_3$	$0.32 \pm 0.01$		$0.29 \pm 0.01$	
A3	$13.6 \pm 0.6$		$15.0 \pm 0.3$	
$\lambda_{3-4}$		$0.048 \pm 0.001$		$0.076 \pm 0.004$
A3-4		$20.6 \pm 0.3$		$17.4 \pm 0.5$
$\lambda_4$	$0.021 \pm 0.001$		$0.034 \pm 0.001$	
A4	$4.24 \pm 0.07$		$7.4 \pm 0.1$	

<sup>a</sup> In case of the interrupted refolding experiments, 3 exponentials describe the time-course of the data, with the third rate constant ( $\lambda_{3-4}$  with amplitude A3-4) being in between rate constants  $\lambda_3$  and  $\lambda_4$  observed in the conventional apoflavodoxin kinetic folding experiments. The errors shown are standard deviations. Refolding conditions are 1  $\mu$ M apoflavodoxin in 100 mM potassium pyrophosphate, pH 6.0, at 25 °C.

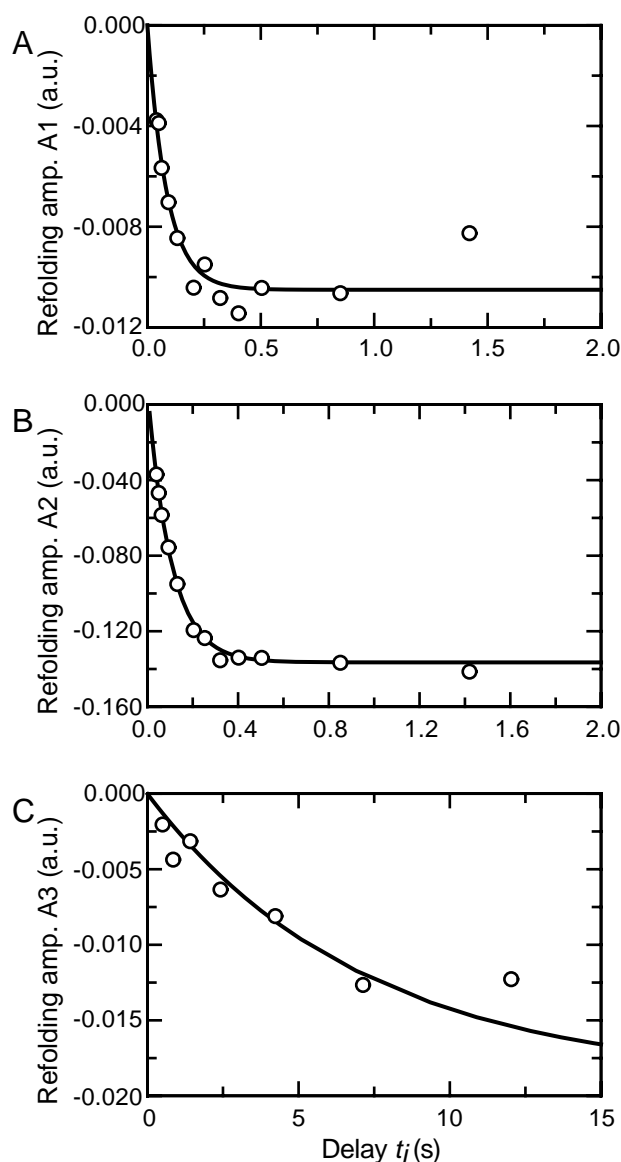
### c. Origin of the parallel apoflavodoxin folding routes

The most common origin of parallel folding routes in proteins is heterogeneity in the unfolded state caused by isomerisation around Xaa-Pro peptide bonds (Hagerman and Baldwin 1976; Schmid 1992) or Xaa-nonPro peptide bonds (Pappenberger et al. 2001). Proline peptide bond isomerisation is likely to play a role in *A. vinelandii* flavodoxin folding, because it contains five proline residues. All of the Xaa-Pro peptide bonds are in the *trans* conformation in the native structure. The percentage of *cis* peptide bonds of a certain proline residue in equilibrium unfolded protein depends mainly on the preceding amino acid (Reimer et al. 1998). In case of apoflavodoxin, based on its amino acid sequence, 33 % of the molecules are predicted to have at least one *cis* proline peptide bond in the equilibrium unfolded state.

To investigate the origin of the parallel folding routes during apoflavodoxin folding an interrupted unfolding experiment has been performed. Native apoflavodoxin is first unfolded in 3 M GuHCl for a varying delay time  $t_i$ , and subsequently refolded at 0.5 M GuHCl and followed by fluorescence spectroscopy. The dependence of the amplitudes of the four observed refolding processes (with rate constants  $\lambda_1$  to  $\lambda_4$ ) on  $t_i$  clarifies whether a specific refolding process originates from the folding mechanism of apoflavodoxin or from secondary processes in the unfolded protein, like Xaa-Pro peptide bond isomerisation. In the former case, the dependence of the amplitude of a specific refolding process on  $t_i$  equals the unfolding rate ( $9\text{ s}^{-1}$ ), whereas in the latter case the  $t_i$ -dependence of the amplitude is much slower (Reimer et al. 1998). Due to the large difference between the  $\lambda_1$  and  $\lambda_4$  rate constants it is not possible to resolve all refolding processes in one kinetic trace in the interrupted unfolding experiment. Traces of 5 seconds are recorded to sample  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$ , and the fourth process is examined separately in a manual mixing experiment, as the rate constant  $\lambda_4$  is sufficiently slow to do so.

In Figure 5 the refolding amplitudes corresponding to the refolding rate constants  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  obtained in the interrupted unfolding experiment are shown as a function of the unfolding time  $t_i$ . Both refolding amplitudes corresponding to  $\lambda_1$  and  $\lambda_2$  appear with rate constants that are in good agreement with the apoflavodoxin unfolding rate constant of  $9\text{ s}^{-1}$  under these conditions. This shows that the two faster folding processes represent two parallel folding routes accessible to apoflavodoxin molecules with all prolyl peptide bonds in the native *trans* conformation. In contrast, the refolding amplitude corresponding to rate constant  $\lambda_3$  appears with a rate of only  $0.14\text{ s}^{-1}$ , which is a very likely rate for proline peptide bond isomerisation (Reimer et al. 1998). The manual interrupted unfolding experiment shows that the  $\lambda_4$  refolding amplitude appears with a rate of  $0.015\text{ s}^{-1}$  (data not shown), which is compatible with proline peptide bond isomerisation as well.

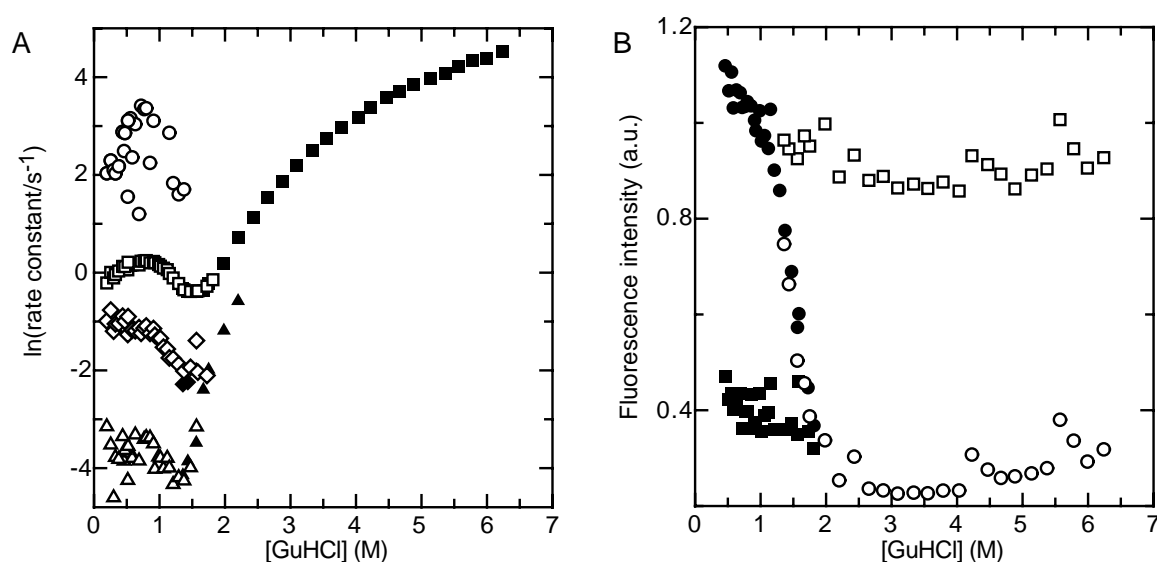




**Figure 5.** Dependence of the amplitudes associated with *A. vinelandii* apoflavodoxin folding rate constants  $\lambda_1$  to  $\lambda_3$  on the unfolding time  $t_i$  in an interrupted unfolding experiment. In this experiment, native apoflavodoxin is first unfolded in 3 M GuHCl for a varying time  $t_i$ , and the subsequent refolding at 0.5 M GuHCl is followed by fluorescence spectroscopy. Dependence on  $t_i$  of the refolding amplitude corresponding to folding rate constant  $\lambda_1$  ( $7.44 \text{ s}^{-1}$ ) (A), folding rate constant  $\lambda_2$  ( $1.21 \text{ s}^{-1}$ ) (B), and folding rate constant  $\lambda_3$  ( $0.30 \text{ s}^{-1}$ ) (C). A single exponential equation is fitted to all data and the corresponding fit is represented by a solid line. The amplitude corresponding to  $\lambda_1$  builds up with a rate of  $10.6 \pm 0.8 \text{ s}^{-1}$ ; in case of  $\lambda_2$  the rate of appearance of the corresponding amplitude is  $9.60 \pm 0.06 \text{ s}^{-1}$ ; in case of  $\lambda_3$  the rate of appearance of the corresponding amplitude is  $0.137 \pm 0.005 \text{ s}^{-1}$ . The final protein concentration is  $1 \text{ }\mu\text{M}$  in 100 mM potassium pyrophosphate, pH 6.0,  $25 \text{ }^\circ\text{C}$ .

#### d. GuHCl dependence of the apoflavodoxin folding and unfolding kinetics

To obtain more insight into the origin of the observed complexity in apoflavodoxin folding we performed kinetic refolding and unfolding experiments followed by fluorescence spectroscopy at different GuHCl concentrations. Three or four exponentials are required to describe the observed apoflavodoxin refolding kinetics up to 2 M GuHCl. In contrast, apoflavodoxin unfolding traces are in most cases described by a single exponential. Only in the transition zone of (un)folding, where both folding and unfolding processes significantly contribute to the observable rates, the slower folding processes are observed in kinetic unfolding traces. The fastest step observed in kinetic folding is not observed in the kinetic unfolding traces. The four observed rate constants for apoflavodoxin folding and unfolding are shown in Figure 6A as a chevron plot. Figure 6B shows that the complete difference in fluorescence intensity between native and unfolded apoflavodoxin is observed at all concentrations denaturant used in the kinetic experiments. This means that all processes with a significant corresponding change in fluorescence are observed in the stopped-flow experiments.



**Figure 6.** (A) GuHCl-dependence of the natural logarithm of the observed rate constants ( $\circ = \lambda_1$ ,  $\square = \lambda_2$ ,  $\diamond = \lambda_3$ ,  $\triangle = \lambda_4$ ) for *A. vinelandii* apoflavodoxin unfolding (filled symbols) and refolding (open symbols). (B) Initial ( $\square$ ) and final ( $\circ$ ) fluorescence intensity of the kinetic traces obtained for apoflavodoxin stopped-flow refolding (filled symbols) and unfolding (open symbols). The final protein concentration is 1  $\mu\text{M}$  in 100 mM potassium pyrophosphate pH 6.0, at 25  $^{\circ}\text{C}$ .

Of the four rate constants observed during apoflavodoxin folding (Figure 6) only  $\lambda_1$  and  $\lambda_2$  inform about the mechanism that describes apoflavodoxin folding, as the corresponding molecules are shown to have all peptide bonds in the native *trans* conformation. As discussed, folding rate constants  $\lambda_3$  and  $\lambda_4$  originate from Xaa-Pro peptide bond isomerisation and do not inform about the apoflavodoxin folding mechanism. Figure 6 is consistent with the latter as both  $\lambda_3$  and  $\lambda_4$  have a low GuHCl concentration dependence.

As a result, only the denaturant concentration-dependence of the  $\lambda_1$  and  $\lambda_2$  rate constants can be used to decipher the apoflavodoxin folding mechanism.

*e. Involvement of an on-pathway intermediate during apoflavodoxin folding*

Under strong denaturing conditions ( $> 2.5$  M GuHCl), a single unfolding rate constant is observed, which displays a smooth curvature as a function of the GuHCl concentration (Figure 6A). This curvature is explained by the presence of two consecutive transition states on one linear apoflavodoxin (un)folding route. At any denaturant concentration the transition state with the highest free energy determines the observed rate of folding and unfolding. At low concentrations GuHCl the most unfolded-like transition state TS1 is rate limiting for folding and unfolding. However, the free energies of both transition states have differing dependencies on the concentration denaturant. The transition state TS2 that is most native-like is destabilised most by the denaturant. As a consequence, upon increasing the denaturant concentration an interchange of the transition state that determines the rate-limiting step in apoflavodoxin unfolding occurs. As a result, the slope of the unfolding limb of the chevron plot alters (Figure 6A).

The presence of two transition states, the free energies of which have differing denaturant concentration dependencies, that are each positioned on a separate apoflavodoxin folding route cannot explain the observed curvature in the unfolding limb of  $\lambda_2$ . In this situation, an upward curvature of the unfolding limb would be observed as the unfolding molecules would follow the route with the lowest free energy barrier at all concentrations denaturant (Bachmann and Kiefhaber 2001; Sanchez and Kiefhaber 2003a). This is clearly not the case (Figure 6A).

Another alternative explanation for the observed curvature in the  $\lambda_2$  unfolding limb of the apoflavodoxin chevron plot could be a transition state, which separates N from U, that shifts gradually as a function of denaturant concentration. However, this phenomenon has been studied in detail recently and is shown to be highly unlikely (Sanchez and Kiefhaber 2003b).

Two consecutive transition states on a linear apoflavodoxin (un)folding route that links unfolded and native apoflavodoxin molecules explains the curvature of the  $\lambda_2$  unfolding limb. A minimum in the free energy between these two transition states must exist. In this minimum an on-pathway folding intermediate resides. The rate-limiting step during apoflavodoxin unfolding changes between 3 and 5 M GuHCl from the unfolding of this intermediate (i.e. I to U) at low denaturant concentrations to the formation of this intermediate (i.e. N to I) at high denaturant concentrations. Consequently, at intermediate denaturant concentrations (i.e. 2 to 4 M GuHCl) this intermediate is expected to transiently populate during kinetic apoflavodoxin unfolding. However, only a single unfolding process is observed at all denaturant concentrations above 2.5 M GuHCl (Figure 6A), and the fluorescence intensity does not change during the dead-time of the experiment (Figure 6B). No population of the on-pathway apoflavodoxin folding intermediate is thus observed. In case the fluorescence intensity of the on-pathway intermediate is very similar to the fluorescence intensity of unfolded apoflavodoxin, transient population of this intermediate

could be overlooked. However, the latter is unlikely as fluorescence anisotropy also shows no additional kinetic phases or dead-time processes during kinetic apoflavodoxin unfolding (data not shown).

In conclusion, the intermediate that causes the curvature observed in the unfolding limb of the  $\lambda_2$  chevron plot of apoflavodoxin must be a high-energy on-pathway folding intermediate, as it is not populated to a significant extent. Such high-energy on-pathway intermediates appear to play a role in the folding kinetics of many proteins (Bachmann and Kiefhaber 2001; Sanchez and Kiefhaber 2003a). The unstable, not detectably populated apoflavodoxin folding intermediate is surrounded by two transition states. One transition state separates the on-pathway intermediate from the native state and the other one separates the intermediate from the unfolded state of apoflavodoxin.

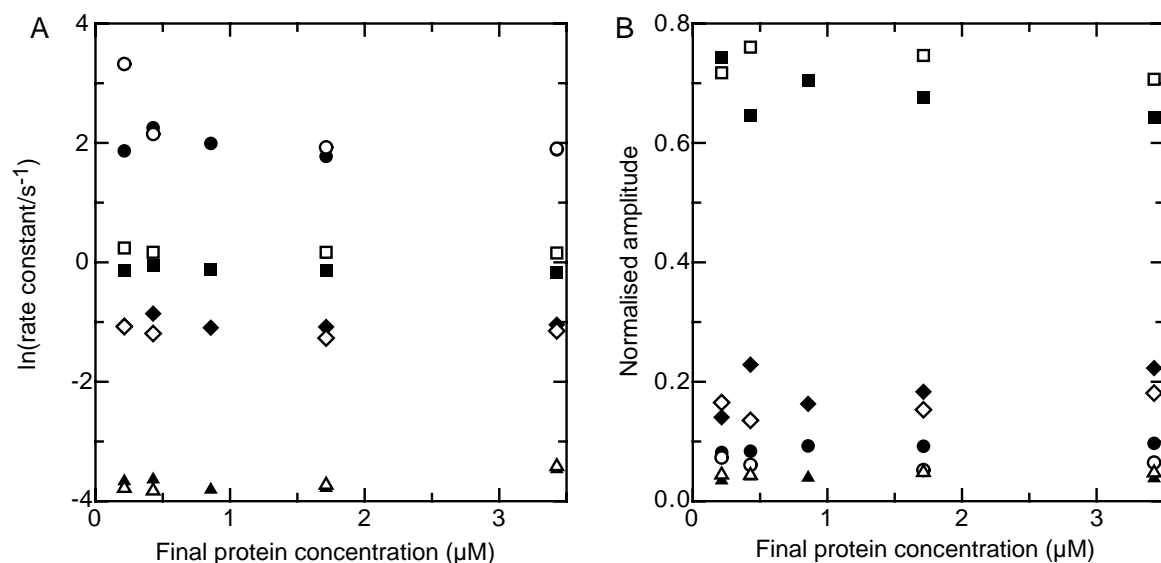
*f. Involvement of an off-pathway intermediate during apoflavodoxin folding*

Under conditions where apoflavodoxin is fully native ( $< 1$  M GuHCl), two folding rate constants (i.e.  $\lambda_1$  and  $\lambda_2$ ) are observed that do not originate from Xaa-Pro peptide bond isomerisations. The observation of these two folding rates implies that an intermediate populates significantly during refolding of apoflavodoxin. Is the intermediate that populates during refolding the same species as the one that causes the curvature in the unfolding limb of the chevron plot? If that would be the case, a three state protein folding model would be able to describe the observed apoflavodoxin  $\lambda_2$  chevron plot. However, the denaturant-dependence of  $\lambda_2$  can not be described by any three-state folding mechanism, be it linear with an on-pathway intermediate involved ( $U \rightleftharpoons I \rightleftharpoons N$ ), or linear with an off-pathway intermediate involved ( $I \rightleftharpoons U \rightleftharpoons N$ ), or triangular, with all three species (U, I, N) being connected (results not shown). Thus at least two folding intermediates must play a role during kinetic (un)folding of apoflavodoxin.

Additional evidence that the on-pathway intermediate does not populate significantly during apoflavodoxin refolding comes from the interrupted refolding experiment. In case formation of an on-pathway apoflavodoxin intermediate during folding is more rapid than its conversion to the native state, the intermediate accumulates and formation of native molecules is delayed. The kinetic formation of native molecules is then characterised by the presence of a lag phase, which is clearly not observed during the interrupted apoflavodoxin refolding experiment (Figure 4). The lack of an observable lag in the formation of native molecules implies that the on-pathway folding intermediate does not significantly populate during apoflavodoxin folding (Kiefhaber 1995). The on-pathway intermediate is besides being unstable under denaturing conditions ( $> 2.5$  M GuHCl) thus also unstable under native conditions. The intermediate that populates during apoflavodoxin refolding must be a different one than the on-pathway intermediate already discussed.

The folding limb of the  $\lambda_2$  chevron plot curves to such an extent that at low denaturant concentration the observed folding rate actually increases with increasing GuHCl concentration (Figure 6A). The latter is characteristic for an unfolding process. Dissociation of dimers or higher oligomeric states that may be transiently formed during apoflavodoxin refolding can be excluded as a source for the observed curvature, since variation of the final

protein concentration from 0.2 to 3.4  $\mu\text{M}$  did not significantly alter the observed rate constants nor their relative amplitudes during refolding (Figure 7).



**Figure 7.** Dependence of the natural logarithm of the observed rate constants ( $\circ = \lambda_1$ ,  $\square = \lambda_2$ ,  $\diamond = \lambda_3$ ,  $\Delta = \lambda_4$ ) (A) and of the corresponding amplitudes (B) on protein concentration at 0.27 M GuHCl (filled symbols) and at 0.59 M GuHCl (open symbols) for *A. vinelandii* apoflavodoxin folding. Folding is started from equilibrium unfolded apoflavodoxin in 3.0 M GuHCl. The buffer used is 100 mM potassium pyrophosphate pH 6.0, at 25 °C.

The curvature of the folding limb of the  $\lambda_2$  chevron plot is explained by the transient population of a folding intermediate which is off the productive folding route. The unfolding of this off-pathway intermediate is the rate-limiting step in the formation of native apoflavodoxin via the process corresponding to  $\lambda_2$ . In principle, the partial unfolding of an on-pathway intermediate could also be the rate-limiting process during protein folding. However, if such an intermediate transfers to native apoflavodoxin with rate constant  $\lambda_2$ , a delay in the formation of native apoflavodoxin would be observed, which is not the case.

The interrupted refolding experiment confirms that an off-pathway intermediate populates during apoflavodoxin folding. In case an off-pathway intermediate that needs to unfold en route to the native state populates, native protein forms kinetically with a sum of two exponentials. The latter is exactly as is observed in the interrupted refolding experiment, besides the observation of rate constants associated with peptide bond isomerisation processes. At low denaturant concentrations ( $< 0.75$  M GuHCl)  $\lambda_2$  represents the kinetic formation of native apoflavodoxin via an off-pathway intermediate, whereas  $\lambda_1$  at these concentrations reflects the direct kinetic formation of native apoflavodoxin by those molecules that circumvent the off-pathway intermediate. From the interrupted refolding experiment (Table 2) it can be inferred that  $10.6 \pm 0.4$  % and  $11 \pm 2$  % of the molecules with native proline peptide bonds fold via the direct path to the native state at 0.42 and 0.83 M GuHCl, respectively (as determined via  $A_1/(A_1+A_2)$  (Table 2)). The majority of the apoflavodoxin molecules thus folds via the off-pathway apoflavodoxin folding intermediate.

*g. Kinetic role of the equilibrium (un)folding intermediate of apoflavodoxin*

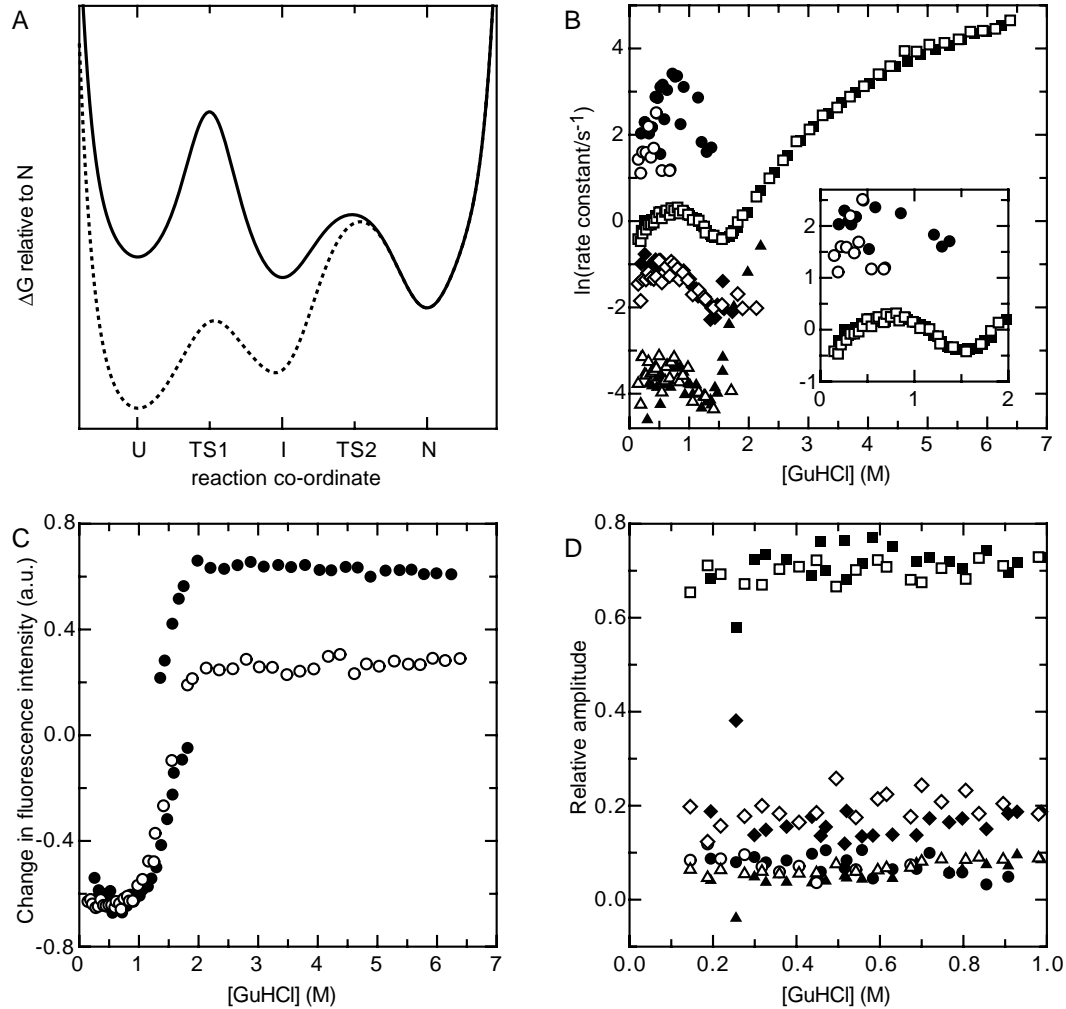
Are the intermediate that populates during equilibrium unfolding of apoflavodoxin and the kinetic on-pathway intermediate the same species? To answer this question, kinetic folding and unfolding experiments that start with apoflavodoxin in 1.6 M GuHCl are done. Under these conditions the *A. vinelandii* apoflavodoxin equilibrium intermediate is near to maximally populated (for 55 %, Figure 2E).

As discussed, the apoflavodoxin on-pathway intermediate is surrounded by two transition states: TS1 separates the on-pathway intermediate from the unfolded state and TS2 separates the intermediate from the native state of apoflavodoxin (Figure 8A). Below approximately 4 M GuHCl TS1 has the highest free energy of both transition states involved. TS1 is less structured than TS2. Consequently, an increase of the denaturant concentration stabilises TS1 more than TS2 and above approximately 4 M GuHCl TS2 becomes the rate-limiting transition state for apoflavodoxin unfolding. This causes the curvature in the unfolding limb of the chevron plot (Figure 6). Of the two transition states involved, TS2 can not be the one having the highest free energy below 4 M GuHCl, because if that would be the case increasing the GuHCl concentration above 4 M would not result in the observed change of the rate-limiting unfolding process.

In case the apoflavodoxin on-pathway intermediate is the equilibrium intermediate, what folding and unfolding kinetics are expected starting with apoflavodoxin in 1.6 M GuHCl? Below a final GuHCl concentration of approximately 4 M, the rate constant for folding of the on-pathway intermediate to native apoflavodoxin is determined by TS2 and the rate constant for folding from globally unfolded apoflavodoxin to native apoflavodoxin is determined by TS1 (Figure 8A). As under these circumstances TS1 has a higher free energy than TS2, the on-pathway intermediate folds faster to the native state than unfolded apoflavodoxin does. The latter kinetic behaviour is clearly not observed when starting with the equilibrium intermediate of apoflavodoxin populated: the observed folding rate constants are identical to the folding rate constants observed in the classical chevron experiment already discussed (Figure 8B). In addition, the total amplitude of the folding reaction is the same in both experiments (Figure 8C), as well as the relative amplitude of each of the four folding processes (Figure 8D).

Above a final GuHCl concentration of approximately 4 M, the rate constant for unfolding of the on-pathway intermediate is determined by TS1, whereas the rate constant for unfolding of native apoflavodoxin is determined by TS2, which at these denaturant concentrations has a higher free energy than TS1 (Figure 8A). Consequently, under these circumstances the on-pathway intermediate needs to unfold faster than native apoflavodoxin does. No such fast unfolding process is observed when the equilibrium folding intermediate is populated (Figure 8B). However, the equilibrium intermediate still could be the kinetic on-pathway intermediate but the unfolding of this equilibrium intermediate could be too fast to be detected. In that case, the amplitude associated with this unfolding process should decrease between 3 and 5 M GuHCl (as in this GuHCl concentration range the free energy of TS1 gradually becomes lower than the one of TS2), and then level off horizontally. However, no such change in amplitude is observed as well (Figure 8C). Rather, Figure 8C shows that

the amplitude of the unfolding process starting with the equilibrium intermediate populated is constant and drastically reduced compared to the situation in which unfolding starts with native apoflavodoxin. Both below and above 4 M GuHCl, unfolding of the equilibrium intermediate occurs within the dead-time of the stopped-flow unfolding experiment, and thus is faster than unfolding of native apoflavodoxin.



**Figure 8.** (A) Schematic free energy landscape for the folding of a hypothetical protein via an on-pathway intermediate. The on-pathway intermediate (I) is surrounded by two transition states: TS1 separates I from the unfolded state (U) and TS2 separates I from the native state of the protein (N). At low GuHCl concentrations (solid line) TS1 has the highest free energy of both transition states involved. Above a certain GuHCl concentration (dashed line) TS2 becomes the rate-limiting transition state for unfolding of the protein. (B) GuHCl-dependence of the natural logarithm of the observed rate constants ( $\circ = \lambda_1$ ,  $\square = \lambda_2$ ,  $\diamond = \lambda_3$ ,  $\Delta = \lambda_4$ ) for *A. vinelandii* apoflavodoxin unfolding and refolding. Open symbols: folding and unfolding are both started with apoflavodoxin in 1.6 M GuHCl, the concentration GuHCl at which 55% of the apoflavodoxin molecules is present as the equilibrium folding intermediate (see Figure 2E). Filled symbols are the data presented in Figure 6, folding is here started with unfolded apoflavodoxin (at 5 M GuHCl) and unfolding is here started with native apoflavodoxin (at 0 M GuHCl). The inset shows a zoom of  $\lambda_1$  and  $\lambda_2$  at GuHCl concentrations below 2 M. (C) GuHCl-dependence of the total fluorescence intensity change (i.e. the initial minus the final fluorescence intensity) during kinetic apoflavodoxin folding and unfolding starting with apoflavodoxin in 1.6 M GuHCl ( $\circ$ ). The GuHCl-dependence of the total change in fluorescence intensity during kinetic folding starting with unfolded apoflavodoxin (at 5 M GuHCl) and during kinetic unfolding starting with native apoflavodoxin (at 0 M GuHCl), is also shown ( $\bullet$ ). (D) GuHCl-dependence of the relative amplitudes corresponding to each of the four observed apoflavodoxin folding processes ( $\circ = A_1$ ,  $\square = A_2$ ,  $\diamond = A_3$ ,  $\Delta = A_4$ ). Open symbols represent amplitudes of folding processes starting with apoflavodoxin in 1.6 M GuHCl, whereas filled symbols represent amplitudes of folding processes starting with unfolded apoflavodoxin in 5 M GuHCl. The final protein concentration is 1  $\mu\text{M}$  in 100 mM potassium pyrophosphate pH 6.0, at 25  $^\circ\text{C}$ .

In conclusion, the experiments described show that the intermediate that is observed during denaturant-induced equilibrium unfolding of *A. vinelandii* apoflavodoxin can not be the on-pathway intermediate observed during kinetic apoflavodoxin folding.

Is the intermediate observed during denaturant-induced equilibrium unfolding of apoflavodoxin the kinetic off-pathway intermediate that populates during apoflavodoxin folding? If this is the case, its unfolding should be observed at all GuHCl concentrations above at which it maximally populates, i.e. above approximately 1.8 M GuHCl (Figure 2E). Indeed, a loss of the amplitude of the unfolding process is observed when unfolding starts with the equilibrium intermediate populated compared to when unfolding starts with native apoflavodoxin (Figure 8C).

Further confirmation of the equilibrium intermediate being the kinetic off-pathway intermediate comes from the following. The spectroscopic properties of the apoflavodoxin equilibrium intermediate are remarkably similar to those of the kinetic off-pathway folding intermediate. The intermediate that populates during denaturant-induced equilibrium unfolding of apoflavodoxin has a low fluorescence emission (Figure 2F). The kinetic apoflavodoxin off-pathway intermediate must also have a low fluorescence intensity, as is discussed below. The interrupted refolding experiment shows that approximately 90 % (i.e.  $A_2/(A_1+A_2)\%$ , Table 2) of the folding apoflavodoxin molecules with native Xaa-Pro peptide bonds populate the off-pathway intermediate, and form native apoflavodoxin with rate constant  $\lambda_2$ . The direct refolding experiment shows that the amplitude corresponding to  $\lambda_2$  is also approximately 90 % of the total amplitude caused by apoflavodoxin molecules with native Xaa-Pro peptide bonds (as can be inferred from  $A_1$  and  $A_2$  shown in Figure 8D). In this direct refolding experiment, the amplitude of  $\lambda_2$  corresponds to the change in fluorescence intensity. Consequently, the fluorescence intensity of the off-pathway intermediate must be similar to the fluorescence intensity of unfolded apoflavodoxin as is also observed for the equilibrium intermediate (Figure 2F). In addition, when kinetic folding of apoflavodoxin is followed by fluorescence anisotropy, large anisotropy is observed immediately after the dead-time of the experiment. This anisotropy subsequently decreases with rate constant  $\lambda_2$  (not shown). As the folding process with rate constant  $\lambda_2$  represents the unfolding of the off-pathway intermediate, the latter intermediate must have a high anisotropy. As discussed, a high anisotropy signal is also observed for the equilibrium (un)folding intermediate.

In conclusion, the results presented show that the *A. vinelandii* apoflavodoxin equilibrium folding intermediate and the intermediate that appears kinetically off the productive apoflavodoxin folding route most likely are the same species.

#### *h. Quantitative kinetic model for apoflavodoxin folding*

Based on the apoflavodoxin folding data presented here, a kinetic model for apoflavodoxin folding is constructed:

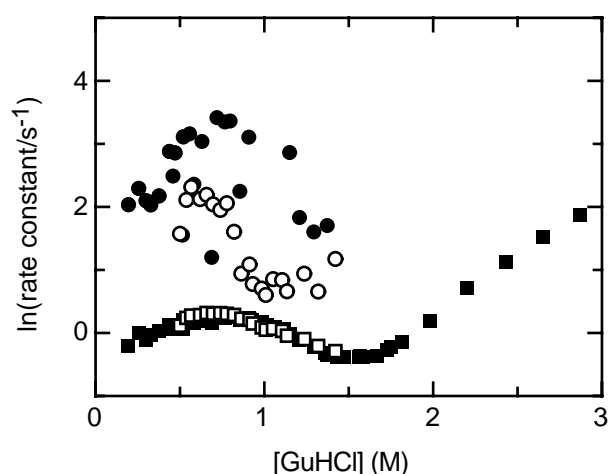




with  $I_1$  the off-pathway kinetic folding intermediate that is stable and also populates during equilibrium unfolding of apoflavodoxin,  $I_2$  the high-energy on-pathway kinetic folding intermediate,  $U$  the unfolded state, and  $N$  native apoflavodoxin. Scheme 1 is the most simple model containing two intermediates that is able to explain the observations made on apoflavodoxin folding.

In case of any four-state kinetic folding model, three observable rate constants are expected (Szabo 1969; Ikai and Tanford 1973; Hagerman and Baldwin 1976). However, since one of the four apoflavodoxin species is very unstable (i.e. the on-pathway intermediate), only two rate constants that inform about apoflavodoxin folding are observed (i.e.  $\lambda_1$  and  $\lambda_2$ ). The latter is highly homologous to the case of apparent two-state folding, where a single rate constant is observed with a curvature of the unfolding limb of the chevron plot, which is due to the presence of a high-energy on-pathway intermediate, as discussed in detail recently by Sánchez and Kiefhaber (Sanchez and Kiefhaber 2003a).

Normally, a kinetic protein folding model is fitted to the GuHCl-dependence of all observed kinetic folding rates and corresponding amplitudes that are not due to Xaa-Pro isomerisations, i.e.  $\lambda_1$  and  $\lambda_2$  in case of *A. vinelandii* apoflavodoxin folding. However, the fastest process observed during apoflavodoxin folding, with observed rate constant  $\lambda_1$ , displays a lot of scatter (Figure 6), which is mainly due to its small amplitude. The scatter may in part also be caused by contributions of additional processes with similar rates, like e.g. the rapid formation of intermediates with non-native peptide bond isomers. To investigate the influence of Xaa-Pro peptide bond isomerisation on  $\lambda_1$ , an experiment is performed in which freshly unfolded apoflavodoxin (made by unfolding the protein for 622 ms in 3 M GuHCl) is refolded at different GuHCl concentrations. Refolding of freshly



**Figure 9.** GuHCl-dependence of the natural logarithm of the observed rate constants for *A. vinelandii* apoflavodoxin folding (O =  $\lambda_1$ ,  $\square$  =  $\lambda_2$ ) obtained by starting with equilibrium unfolded apoflavodoxin (filled symbols) or by starting with freshly unfolded apoflavodoxin, made by unfolding apoflavodoxin for 622 ms in 3 M GuHCl (open symbols). The final protein concentration is 1  $\mu$ M in 100 mM potassium pyrophosphate pH 6.0, at 25 °C.

unfolded apoflavodoxin leads to an observed rate constant  $\lambda_1$  which differs from the one obtained by starting with equilibrium unfolded apoflavodoxin, as is shown in Figure 9. The observed rate constant  $\lambda_2$  of the major folding process obtained by refolding freshly unfolded apoflavodoxin is identical within error to  $\lambda_2$  obtained in the conventional refolding experiment in which equilibrium unfolded apoflavodoxin is used. The observed change in rate constant  $\lambda_1$  indeed shows that in the conventional refolding experiment  $\lambda_1$  is influenced by the formation of intermediates with incorrect proline isomers.

Unfortunately, even the refolding process starting from freshly unfolded apoflavodoxin results in a complex GuHCl-concentration dependence of the observed rate constant  $\lambda_1$ , which curves downwards upon increasing the denaturant concentration. This curvature can not be reproduced by a linear four-state model, and indicates further complexities in the apoflavodoxin folding mechanism that can not be resolved accurately. As in case of *A. vinelandii* apoflavodoxin  $\lambda_1$  can not be resolved, only one rate constant (i.e.  $\lambda_2$ ) is used in our analysis.

Scheme 1 is fitted to the denaturant-dependence of folding rate constant  $\lambda_2$ . Folding rate constant  $\lambda_2$  is chosen as it is the most informative of the four rate constants observed, as it has the largest amplitude, is observed at all denaturant concentrations, involves native proline peptide bond isomers and thus informs about the apoflavodoxin folding mechanism. It has a minimum around 1.55 M GuHCl and displays a complex denaturant-concentration dependence (Figure 6A). The amplitude  $A_2$  corresponding to  $\lambda_2$  is not used as it makes no sense to fit a folding model to a single amplitude. The validation of Scheme 1 is obtained in the previous sections by the qualitative analysis of the apoflavodoxin direct and interrupted refolding and unfolding data. Similarly, in case of hen egg-white lysozyme interrupted refolding experiments have also allowed the determination of a complex kinetic folding mechanism without the use of the corresponding kinetic amplitudes (Kiefhaber 1995; Wildegger and Kiefhaber 1997; Bieri et al. 1999).

Because only the GuHCl dependence of  $\lambda_2$  can be analysed, additional information is required (Ikai and Tanford 1973) to determine unambiguously all microscopic rates shown in Scheme 1. As discussed, the apoflavodoxin equilibrium (un)folding intermediate and the intermediate that appears kinetically off the productive folding route of apoflavodoxin most likely are the same species. We therefore restrained the fit of Scheme 1 to the kinetic apoflavodoxin (un)folding data by fixing  $K_{UI}$  and  $m_{UI}$  to the values for  $K_{UI}$  and  $m_{UI}$  obtained from the equilibrium (un)folding measurements (Table 1), according to equations 8 and 9:

$$\Delta G_{UI} = -RT \ln K_{UI} = -RT \ln \left( \frac{k_{UI}}{k_{IU}} \right) \quad (8)$$

$$m_{UI}^{eq} = m_{UI} - m_{IU} \quad (9)$$

Since  $I_2$  is shown to never populate significantly during apoflavodoxin folding,  $k_{I_2N}$  in Scheme 1 must be much larger than  $(k_{UI_2} + k_{I_2U})$ , and thus the latter will always limit the observed folding rate. Therefore  $k_{I_2N}$  can not be determined. The kinetic data can only inform about the relative free energies of the two transition states surrounding  $I_2$ , i.e. the ratio between  $k_{I_2N}$  and  $k_{I_2U}$ . For technical reasons, the ratio is not fitted directly, but instead  $k_{I_2N}$  is fixed to  $10^5$  to be sufficiently large not to influence the observable kinetics. With the assumptions described above, the number of degrees of freedom of the model (Scheme 1) is sufficiently reduced to yield non-redundant results when fitted to the  $\lambda_2$  chevron obtained for kinetic apoflavodoxin folding.

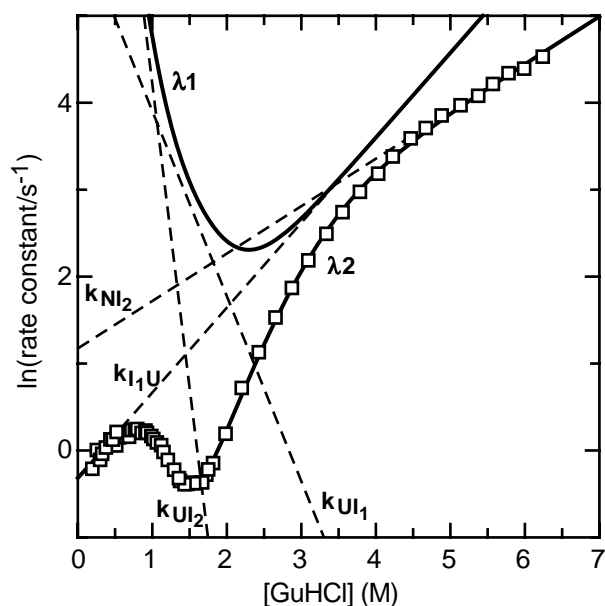
The results of the fit of Scheme 1 to  $\lambda_2$ , using the above assumptions, are summarised in Table 3 and are shown in Figure 10. The free energy landscape for apoflavodoxin folding derived from the results of the equilibrium and the kinetic folding studies, is shown schematically in Figure 11. Note that the on-pathway intermediate  $I_2$  resides in a valley in a region of high free energy at all denaturant concentrations, and that the time course of formation of  $I_2$  is always slower than the rate of its further conversion, either to the native or to the unfolded state. As the experimental results show,  $I_2$  never significantly populates, neither kinetically nor at equilibrium, but its presence influences the observed kinetics. Because it never significantly populates,  $I_2$  is classified as a “high-energy” species.

Because in the fit of Scheme 1 to the apoflavodoxin folding data the rate constant  $k_{I_2N}$  and the corresponding  $m$ -value  $m_{I_2N}$  are fixed to  $100000 \text{ s}^{-1}$  and  $0 \text{ kcal/mol}\cdot\text{M}^{-1}$ , respectively, the exact denaturant accessibility of  $I_2$  can not be calculated. However, the denaturant accessibilities of both transition states surrounding  $I_2$  can be calculated from the kinetic  $m$ -values and are expressed as  $\alpha$ -values, i.e. the ratio of the  $m$ -value of a transition state (TS) and the  $m$ -value of the native state, both relative to the unfolded state. The  $\alpha$  of TS1 is 0.75, and the  $\alpha$  of TS2 is 0.94. Because  $I_2$  must have a solvent accessibility that is in between those of the two transition states, it follows that its  $\alpha$  is approximately 0.8 - 0.9.

**Table 3. Parameters obtained via the fit of the analytical solution to the linear four-state folding scheme  $I_1 \rightleftharpoons \text{unfolded} \rightleftharpoons I_2 \rightleftharpoons \text{native}$  to the denaturant-dependence of rate constant  $\lambda_2$  of *A. vinelandii* apoflavodoxin.<sup>a</sup>**

$k_{I_1U}$	$0.733 \pm 0.002$	$m_{I_1U}$	$0.576 \pm 0.002$
$K_{UI}$	569	$m_{UI}^{eq}$	-1.83
$k_{UI_1}$	$417 \pm 1$	$m_{UI_1}$	$-1.256 \pm 0.002$
$k_{UI_2}$	$(7.31 \pm 0.07)10^4$	$m_{UI_2}$	$-4.12 \pm 0.01$
$k_{I_2U}/k_{I_2N}$	$(4.02 \pm 0.01)10^{-3}$	$m_{I_2U} - m_{I_2N}$	$1.064 \pm 0.001$
$k_{NI_2}$	$3.234 \pm 0.007$	$m_{NI_2}$	$0.3218 \pm 0.0002$
$\Delta G_{UN}$	$9.17 \pm 0.01$	$m_{UN}^{eq}$	$-5.508 \pm 0.001$

<sup>a</sup>  $K_{UI}$  is the equilibrium constant of the  $I$ - $U$  equilibrium at zero concentration denaturant,  $m_{UI}^{eq}$  is the constant which describes the denaturant concentration-dependence of  $K_{UI}$ ,  $k_{XY}$  is the intrinsic rate constant for the transition from state  $X$  to  $Y$  at zero concentration denaturant,  $m_{XY}$  is the factor which describes the denaturant concentration-dependence of the rate constant  $k_{XY}$ ,  $\Delta G_{UN}$  is the difference in free energy between  $N$  and  $U$ , and  $m_{UN}^{eq}$  describes the denaturant concentration-dependence of  $\Delta G_{UN}$ . Rate constants are in  $\text{s}^{-1}$ ,  $m$ -values are in  $\text{kcal/mol}\cdot\text{M}^{-1}$ . The errors are standard deviations. The equilibrium constant  $K_{UI}$  and the corresponding  $m$ -value  $m_{UI}^{eq}$  are fixed during the fit to the values derived from Table 1. Rate constant  $k_{UI_1}$  and  $m_{UI_1}$  are calculated using the equilibrium constant  $K_{UI}$  and the fitted rate  $k_{I_1U}$ , and their corresponding  $m$ -values, respectively. Rate constant  $k_{I_2N}$  is fixed to 100000, and its  $m$ -value  $m_{I_2N}$  to zero (see main text). The exact value of the rate constant  $k_{I_2U}$  can not be determined, the data can only inform about the ratio between  $k_{I_2N}$  and  $k_{I_2U}$  (see main text).



**Figure 10.** Result of the fit of the analytical solution to the linear four-state folding mechanism ( $I_1 \rightleftharpoons \text{unfolded} \rightleftharpoons I_2 \rightleftharpoons \text{native}$ ) to the GuHCl-dependence of the observed (un)folding rate constant  $\lambda_2$  ( $\square$ ) of *A. vinelandii* apoflavodoxin as presented in Figure 6. The solid lines represent the fitted denaturant-dependence of the observable rate constants  $\lambda_1$  and  $\lambda_2$ , whereas the dashed lines represent the fitted denaturant dependencies of some of the microscopic rate constants. The numerical results of this fit are summarised in Table 3.

Besides from denaturant-induced equilibrium unfolding experiments, the stability of the native state of apoflavodoxin relative to its unfolded state ( $\Delta G_{UN}$ ) can also be calculated from the kinetically determined folding and unfolding rate constants (Table 3), using an equation analogous to equation 8. The  $\Delta G_{UN}$  calculated from the kinetic results is 9.17 kcal/mol, which is 1.70 kcal/mol less than the  $\Delta G_{UN}$  derived from equilibrium unfolding measurements ( $10.87 \pm 0.52$  kcal/mol), taking into account the 0.42 kcal/mol due to the Xaa-Pro peptide bond isomerisations. Both  $\Delta G_{UN}$  values are reasonably similar and kinetic complexities that cannot be resolved by the current stopped-flow experiments most likely explain the difference in free energy values observed.

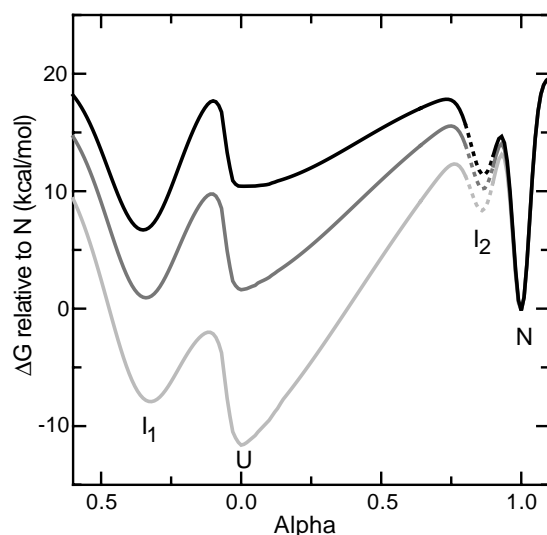
In conclusion, a linear four-state folding scheme  $I_1 \rightleftharpoons \text{unfolded} \rightleftharpoons I_2 \rightleftharpoons \text{native}$  and its corresponding free energy landscape is consistent with the experimental data obtained on *A. vinelandii* apoflavodoxin folding.

### On- and off-pathway intermediates in protein folding

The role of intermediates in kinetic protein folding differs between different theoretical models. In the framework model (Kim and Baldwin 1990), intermediates are supposed to effectively guide a folding protein molecule to the native state, whereas other models regard intermediates as structures that are kinetically trapped in a local minimum of the energy landscape (Bryngelson et al. 1995). Experimentally, many proteins have been shown to form intermediates, but whether these intermediates are productive or not is in most cases not resolved. Obligatory on-pathway intermediates have been shown to occur in 23 proteins that display apparent two-state folding kinetics (Sanchez and Kiefhaber 2003a). One protein that folds via intermediates the kinetic roles of which have been characterised thoroughly is lysozyme from hen egg-white. It can fold via three parallel routes to the native state, either via one of two on-pathway intermediates or via a direct pathway without detectable population of intermediates (Kiefhaber 1995; Wildegger and Kiefhaber 1997; Bieri et al. 1999). Apparently, the folding energy landscape for lysozyme is rough, including

distinct local minima. Another intermediate the kinetic role of which has been determined is the one formed by sunflower albumin 8 (SFA-8) (Pandya et al. 1999). This highly compact intermediate is off the direct folding pathway. The chevron plot of SFA-8 is at low denaturant concentration characterised by a refolding rate that increases with increasing denaturant concentration. It is suggested that the large number of exposed hydrophobic residues in native SFA-8 cause this initial misfolding (Pandya et al. 1999). Like in *A. vinelandii* apoflavodoxin folding, the corresponding positive slope indicates that an unfolding process is limiting the time course of formation of native protein under the conditions used.

Intermediates of both kinds, i.e. on- and off-pathway intermediates, have been identified in *A. vinelandii* apoflavodoxin folding in the present study. The presence of a high-energy on-pathway intermediate suggests that the transition region of the apoflavodoxin folding landscape is rough, containing at least two maxima and one distinct minimum. The off-pathway apoflavodoxin intermediate may be a misfolded structure as is also suggested for the SFA-8 intermediate. However, the structured part of the apoflavodoxin intermediate could also be native-like. In the latter case, the energy barrier that separates this intermediate from the native state is too high to be crossed under the experimental conditions used.



**Figure 11.** Schematic representation of the free energy landscape for *A. vinelandii* apoflavodoxin folding at 0 M (black), 1.6 M (dark grey) and 4 M GuHCl (light grey), derived from the results obtained from equilibrium and kinetic (un)folding studies of the protein. The horizontal axis represents the process co-ordinate, the vertical axis the free energy difference relative to native apoflavodoxin. The process co-ordinate scales to denaturant accessibility, because it is expressed as the  $\alpha$ -value, i.e. the ratio of the  $m$ -value of a folding species and the  $m$ -value of native apoflavodoxin. U and N represent unfolded and native apoflavodoxin, respectively;  $I_1$  and  $I_2$  are the two folding intermediates presented in Scheme 1. The off-pathway intermediate  $I_1$  is represented on the left-hand side of the unfolded state, whereas both the on-pathway intermediate  $I_2$  and native apoflavodoxin reside on the right-hand side of the unfolded state. The heights of the barriers between two species ( $i$  and  $j$ ) were calculated from the (un)folding rates according to  $\Delta G_{ij}^\ddagger = -RT \ln(k_{ij}/k_0)$  using a value for  $k_0$ , the factor describing the chain diffusion rate of a polypeptide, of  $10^8$  (Krieger et al. 2003). The exact position and the depth of the minimum in the free energy landscape in which  $I_2$  resides is unknown, and is therefore represented by a dashed line.

Discrimination between these two possibilities may occur once the characteristics of the structure of the intermediate are known. Because the off-pathway intermediate of *A. vinelandii* apoflavodoxin is relatively stable and populates for up to 63 % at equilibrium under the experimental conditions currently used, its structural characterisation stands a good chance. Small changes in environmental conditions such as temperature, ionic strength and pH can probably lead to a full population of the intermediate, facilitating the characterisation of its properties by for instance NMR spectroscopy. It will be especially interesting to correlate the folding kinetics of specific apoflavodoxin mutants with the structure of this off-pathway folding intermediate.

Additional theoretical and experimental investigation of the folding of other proteins that share the  $\alpha$ - $\beta$  parallel topology with *A. vinelandii* apoflavodoxin and of apoflavodoxin mutants is required. The results to be obtained for these topologically related proteins should contribute to a detailed understanding of the influence of the primary sequence on the structural and energetic characteristics of protein folding intermediates. Characterisation of the structural features of the equilibrium folding intermediate of *A. vinelandii* apoflavodoxin will shed light onto why it is kinetically off the direct folding pathway to native apoflavodoxin.

## Conclusion

The experimental data obtained on *A. vinelandii* apoflavodoxin folding show that two folding intermediates play a role. The large majority of the folding apoflavodoxin molecules form an intermediate that needs to unfold before becoming productive. All folding apoflavodoxin molecules pass through a second high energy folding intermediate before reaching the native state. The appearance of both kinetic folding intermediates seems to be governed by protein topology.

## References

- Bachmann, A. and T. Kiefhaber (2001). "Apparent two-state tendamistat folding is a sequential process along a defined route." *J. Mol. Biol.* **306**(2): 375-86.
- Baldwin, R. L. (1996). "On-pathway versus off-pathway folding intermediates." *Fold. Des.* **1**(1): R1-8.
- Beechem, J. M., M. A. Sherman and M. T. Mas (1995). "Sequential domain unfolding in phosphoglycerate kinase: fluorescence intensity and anisotropy stopped-flow kinetics of several tryptophan mutants." *Biochemistry* **34**(42): 13943-8.
- Bieri, O., G. Wildegger, A. Bachmann, C. Wagner and T. Kiefhaber (1999). "A salt-induced kinetic intermediate is on a new parallel pathway of lysozyme folding." *Biochemistry* **38**(38): 12460-70.
- Brenner, S. E., C. Chothia and T. J. Hubbard (1997). "Population statistics of protein structures: lessons from structural classifications." *Curr. Opin. Struct. Biol.* **7**(3): 369-76.
- Bryngelson, J. D., J. N. Onuchic, N. D. Socci and P. G. Wolynes (1995). "Funnels, pathways, and the energy landscape of protein folding: a synthesis." *Proteins* **21**(3): 167-95.

- Canet, D., K. Doering, C. M. Dobson and Y. Dupont (2001). "High-sensitivity fluorescence anisotropy detection of protein-folding events: application to alpha-lactalbumin." Biophys. J. **80**(4): 1996-2003.
- Clarke, J., E. Cota, S. B. Fowler and S. J. Hamill (1999). "Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway." Structure Fold. Des. **7**(9): 1145-53.
- Clementi, C., H. Nymeyer and J. N. Onuchic (2000). "Topological and energetic 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**(5): 937-53.
- Edmondson, D. E. and G. Tollin (1971). "Chemical and physical characterization of the Shethna flavoprotein and apoprotein and kinetics and thermodynamics of flavin analog binding to the apoprotein." Biochemistry **10**(1): 124-32.
- Eftink, M. R. (1994). "The use of fluorescence methods to monitor unfolding transitions in proteins." Biophys. J. **66**(2): 482-501.
- Engel, M. F., C. P. van Mierlo and A. J. Visser (2002). "Kinetic and structural characterisation of adsorption induced unfolding of bovine {alpha}-lactalbumin." J. Biol. Chem. **8**: 8.
- Gerstein, M. (1997). "A structural census of genomes: comparing bacterial, eukaryotic, and archaeal genomes in terms of protein structure." J. Mol. Biol. **274**(4): 562-76.
- Grantcharova, V., E. J. Alm, D. Baker and A. L. Horwich (2001). "Mechanisms of protein folding." Curr. Opin. Struct. Biol. **11**(1): 70-82.
- Gunasekaran, K., S. J. Eyles, A. T. Hagler and L. M. Gierasch (2001). "Keeping it in the family: folding studies of related proteins." Curr. Opin. Struct. Biol. **11**(1): 83-93.
- Hagerman, P. J. and R. L. Baldwin (1976). "A quantitative treatment of the kinetics of the folding transition of ribonuclease A." Biochemistry **15**(7): 1462-73.
- Ikai, A. and C. Tanford (1973). "Kinetics of unfolding and refolding of proteins. I. Mathematical analysis." J. Mol. Biol. **73**(2): 145-63.
- Jennings, P. A. and P. E. Wright (1993). "Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin." Science **262**(5135): 892-6.
- Kiefhaber, T. (1995). "Kinetic traps in lysozyme folding." Proc. Natl. Acad. Sci. USA **92**(20): 9029-33.
- Kim, P. S. and R. L. Baldwin (1990). "Intermediates in the folding reactions of small proteins." Annu. Rev. Biochem. **59**: 631-60.
- Kraulis, P. J. (1991). "MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures." J. Appl. Cryst. **24**: 946-950.
- Krieger, F., B. Fierz, O. Bieri, M. Drewello and T. Kiefhaber (2003). "Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding." J. Mol. Biol. **332**(1): 265-274.
- Kuwajima, K., H. Yamaya, S. Miwa, S. Sugai and T. Nagamura (1987). "Rapid formation of secondary structure framework in protein folding studied by stopped-flow circular dichroism." FEBS Lett. **221**(1): 115-8.
- Lakowicz, J. R. (1999). Principles of fluorescence spectroscopy. New York, Kluwer Academic/ Plenum Publishers (p. 456 and 459).
- Matouschek, A., J. T. Kellis, Jr., L. Serrano, M. Bycroft and A. R. Fersht (1990). "Transient folding intermediates characterized by protein engineering." Nature **346**(6283): 440-5.
- Mayhew, S. G. and G. Tollin (1992). General properties of flavodoxins. Chemistry and biochemistry of flavoenzymes. F. Müller. Boca Raton, CRC Press. **3**: 389-426.

- Myers, J. K., C. N. Pace and J. M. Scholtz (1995). "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding." Protein Sci. **4**(10): 2138-48.
- Otto, M. R., M. P. Lillo and J. M. Beechem (1994). "Resolution of multiphasic reactions by the combination of fluorescence total-intensity and anisotropy stopped-flow kinetic experiments." Biophys. J. **67**(6): 2511-21.
- Pandya, M. J., P. B. Williams, C. E. Dempsey, P. R. Shewry and A. R. Clarke (1999). "Direct kinetic evidence for folding via a highly compact, misfolded state." J. Biol. Chem. **274**(38): 26828-37.
- Pappenberger, G., H. Aygun, J. W. Engels, U. Reimer, G. Fischer and T. Kiefhaber (2001). "Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics." Nat. Struct. Biol. **8**(5): 452-8.
- Plaxco, K. W., K. T. Simons and D. Baker (1998). "Contact order, transition state placement and the refolding rates of single domain proteins." J. Mol. Biol. **277**(4): 985-94.
- Ptitsyn, O. B. (1992). The molten globule state. Protein folding. T. E. Creighton. New York, W. H. Freeman and Company: 243-300.
- Reimer, U., G. Scherer, M. Drewello, S. Kruber, M. Schutkowski and G. Fischer (1998). "Side-chain effects on peptidyl-prolyl cis/trans isomerisation." J. Mol. Biol. **279**(2): 449-60.
- Roder, H. and W. Colon (1997). "Kinetic role of early intermediates in protein folding." Curr. Opin. Struct. Biol. **7**(1): 15-28.
- Rouviere, N., M. Vincent, C. T. Craescu and J. Gallay (1997). "Immunosuppressor binding to the immunophilin FKBP59 affects the local structural dynamics of a surface beta-strand: time-resolved fluorescence study." Biochemistry **36**(24): 7339-52.
- Sanchez, I. E. and T. Kiefhaber (2003a). "Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding." J. Mol. Biol. **325**(2): 367-76.
- Sanchez, I. E. and T. Kiefhaber (2003b). "Hammond behavior versus ground state effects in protein folding: evidence for narrow free energy barriers and residual structure in unfolded states." J. Mol. Biol. **327**(4): 867-84.
- Schmid, F. X. (1992). Kinetics of unfolding and refolding of single-domain proteins. Protein folding. T. E. Creighton: 197-241.
- Steensma, E., H. A. Heering, W. R. Hagen and C. P. Van Mierlo (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**(1-2): 167-72.
- Steensma, E., M. J. Nijman, Y. J. Bollen, P. A. de Jager, W. A. van den Berg, W. M. van Dongen and C. P. van Mierlo (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**(2): 306-17.
- Steensma, E. and C. P. van Mierlo (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**(3): 653-66.
- Szabo, Z. G. (1969). Kinetic Characterization of complex reaction systems. Comprehensive chemical kinetics. C. H. Bamford and C. F. H. Tipper. Amsterdam, Elsevier publishing company. **2**: 1-80.
- Tanford, C., K. C. Aune and A. Ikai (1973). "Kinetics of unfolding and refolding of proteins. 3. Results for lysozyme." J. Mol. Biol. **73**(2): 185-97.
- Thorneley, R. N. F., G. A. Ashby, M. H. Drummond, R. R. Eady, D. L. Hughes, G. Ford, P. M. Harrison, A. Shaw, R. L. Robson, J. Kazlauskaitė and H. A. O. Hill (1994).



- Flavodoxin and nitrogen fixation: Structure, electrochemistry and posttranslational modification by coenzyme A. Flavins and flavoproteins 1993. K. Yagi. Berlin, Walter de Gruyter & Co.: 343-354.
- van Mierlo, C. P., J. M. van den Oever and E. Steensma (2000). "Apoflavodoxin (un)folding followed at the residue level by NMR." Protein Sci. **9**(1): 145-57.
- van Mierlo, C. P., W. M. van Dongen, F. Vergeldt, W. J. van Berkel and E. Steensma (1998). "The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate." Protein Sci. **7**(11): 2331-44.
- Wildegger, G. and T. Kiefhaber (1997). "Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate." J. Mol. Biol. **270**(2): 294-304.



### 3 Is the presence of intermediates observed in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin governed by protein topology?

Yves J.M. Bollen and Carlo P.M. van Mierlo

The topology of a native protein influences the rate with which it is formed. Does topology affect the appearance of folding intermediates and their specific role in kinetic folding as well? This question is addressed by comparing the folding data obtained on apoflavodoxin from *Azotobacter vinelandii* with those available on two other  $\alpha$ - $\beta$  parallel proteins, i.e. *Anabaena* apoflavodoxin and CheY. Two kinetic folding intermediates, one on-pathway and the other off-pathway, seem to be present during the folding of proteins with an  $\alpha$ - $\beta$  parallel topology. Simulation of the folding of CheY by Clementi et al. (2000) shows the involvement of two intermediates with characteristics that resemble those of the two intermediates experimentally observed during *A. vinelandii* apoflavodoxin folding (Chapter 2 of this thesis). The appearance of folding intermediates in the class of  $\alpha$ - $\beta$  parallel proteins is apparently governed by protein topology.

## Introduction

There are strong indications (Goldberg 1999; Baker 2000) that the topology of a native protein influences the rate with which it is formed. This implies that the transition state for protein folding resembles the native state of a protein. The influence of topology on folding rate is expressed in the contact order (Alm and Baker 1999), which reflects the average sequence separation of residues in contact in the native state. The larger the contact order, the slower a protein will fold.

Contact order can be used to predict the part of a protein that forms first during kinetic folding (Grantcharova et al. 2001). Parts of a protein with a low local contact order, like e.g.  $\alpha$ -helices, will form more rapidly than parts of a protein with a high contact order, like e.g. parallel  $\beta$ -sheets. These predictions can in some cases be verified by comparison with experimentally determined Phi-values (Fersht et al. 1992). The Phi-value expresses to what extent native-like structure is formed around a specific residue in the transition state for folding, as determined from the effect a mutation has on the folding rate.

Although the simple contact order model has successfully captured the overall features of protein folding, it is shown to be an oversimplification of reality. Especially in larger proteins or proteins that display a certain degree of symmetry there are often several parallel routes to the native state. These parallel routes may be equally consistent with the topological constraints on the transition state for folding and their population is determined by the magnitude of the local stabilities of the structured protein regions formed in the corresponding transition states for folding. As a consequence, a mutation or a solvent condition that specifically stabilises part of a protein may significantly alter the observed folding mechanism. Incorporation of the free energy of the structures that can be formed on parallel routes into the contact order model considerably improves the ability of the model to predict structures of transition states (Grantcharova et al. 2001).

A good approach to experimentally decipher the influence of the topology of a protein on its folding mechanism is by comparing the experimental results obtained on the folding of proteins which share the same topology but which have no sequence homology. This has been done for the SH3 domains (Martinez and Serrano 1999; Riddle et al. 1999), acylphosphatase and procarboxypeptidase A2 (Villegas et al. 1998; Chiti et al. 1999), and immunoglobulin-like  $\beta$ -sandwich proteins (Clarke et al. 1999). Mutational analysis showed in these cases that indeed the transition state for folding is generally conserved among proteins with the same topology. The question remains however, whether the appearance of folding intermediates and their specific role in kinetic folding is conserved among topologically related proteins that have no sequence homology. This aspect of protein folding has hardly been experimentally investigated so far (Clarke et al. 1999; Gunasekaran et al. 2001). The subject is difficult to resolve, because most protein folding intermediates form rapidly, often within the dead time of stopped-flow techniques, and are not substantially populated at equilibrium.

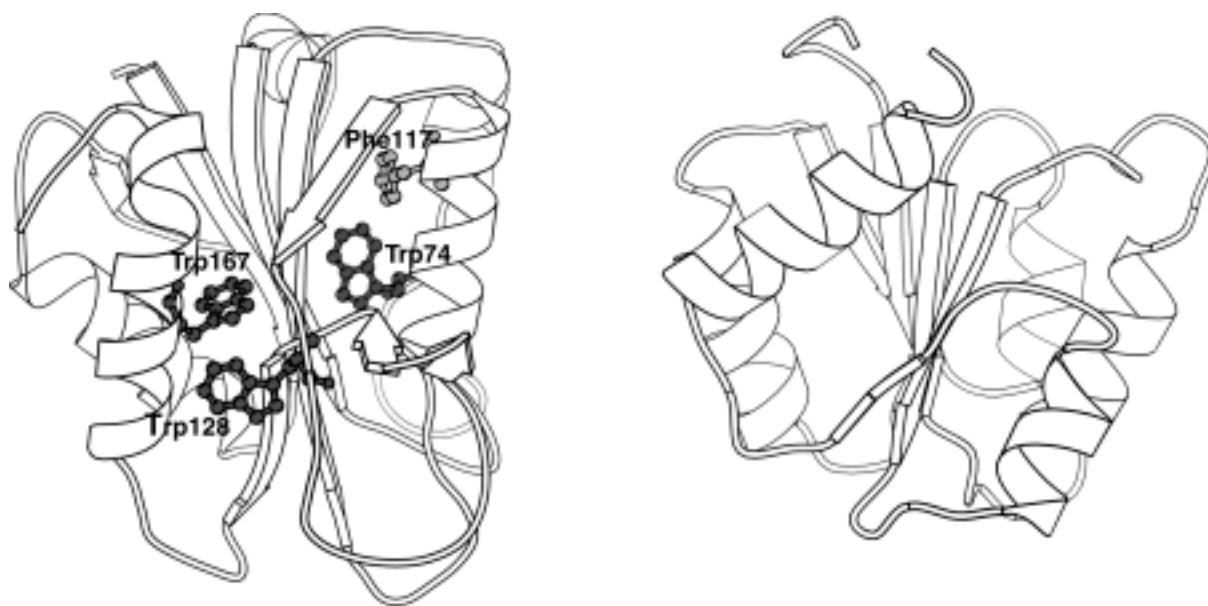
The study of the role of intermediates during folding is facilitated when these intermediates populate both kinetically and at equilibrium. The relatively large, 179-residue

apoflavodoxin from *Azotobacter vinelandii* populates an intermediate with molten-globule like characteristics during denaturant- and thermally-induced equilibrium unfolding (van Mierlo et al. 1998; van Mierlo et al. 2000). In addition, *A. vinelandii* apoflavodoxin is shown to kinetically fold via two intermediates (Chapter 2 of this thesis):



$I_1$  is an off-pathway intermediate that populates heavily during refolding and also populates during denaturant-induced equilibrium unfolding of apoflavodoxin.  $I_2$  is a high-energy folding intermediate on the direct folding route between unfolded and native apoflavodoxin.

Flavodoxins consist of a single structural domain and adopt the so-called doubly-wound or  $\alpha$ - $\beta$  parallel topology. The doubly-wound topology is a rather popular fold: it belongs to the five most common observed folds, together with the TIM-barrel, Rossmann, thiamin-binding and P-loop hydrolase folds (Gerstein 1997). In contrast to most protein folds, this doubly-wound topology is shared by many (i.e. nine) protein superfamilies (Brenner et al. 1997). These nine superfamilies exhibit little or no sequence similarity and comprise a broad range of unrelated proteins with different functions like catalases, chemotactic proteins, lipases, esterases, and flavodoxins. They are all characterised by a five-stranded parallel  $\beta$ -sheet surrounded by  $\alpha$ -helices at either side of the sheet (Figure 1).



**Figure 1.** Left: Molscript cartoon drawing (Kraulis 1991) of the X-ray structure of *Azotobacter chroococcum* flavodoxin (Thorneley et al. 1994), the sequence of which is 95 % identical to *A. vinelandii* flavodoxin. The sidechains of the three tryptophan residues and of Phe117 of *A. vinelandii* flavodoxin are shown in ball-and-stick representation. Right: Molscript cartoon drawing of *E. coli* CheY (3chy.pdb), which shares the  $\alpha$ - $\beta$  parallel topology with flavodoxin.

The population of on- and off-pathway intermediates during kinetic and equilibrium (un)folding of apoflavodoxin from *A. vinelandii* and the comparison with folding data obtained for other  $\alpha$ - $\beta$  parallel proteins allows the investigation of the role protein topology has on the presence of intermediates during protein folding. Such a comparison leads to a better understanding of the fundamental rules describing the folding of proteins with a doubly-wound topology.

### **Comparison of the folding of *A. vinelandii* apoflavodoxin with the folding of other doubly-wound proteins**

#### *a. A. vinelandii apoflavodoxin folding results*

The folding kinetics of *A. vinelandii* apoflavodoxin have been followed by stopped-flow experiments monitored by fluorescence intensity and anisotropy (Chapter 2 of this thesis). Single jump and interrupted refolding experiments show that the refolding kinetics involve four processes yielding native molecules. Interrupted unfolding experiments show that the two slowest folding processes are due to Xaa-Pro peptide bond isomerisation in unfolded apoflavodoxin.

The denaturant dependence of the folding kinetics (i.e. the chevron plot) is complex. Under strongly unfolding conditions ( $> 2.5$  M GuHCl), single exponential kinetics are observed. The slope of the chevron plot changes between 3 and 5 M denaturant. This, together with the absence of an additional unfolding process reveals the presence of two consecutive transition states on a linear pathway that surround a high-energy on-pathway intermediate. Under refolding conditions, two processes are observed for the folding of apoflavodoxin molecules with native Xaa-Pro peptide bond conformations, which implies the population of an intermediate. The slowest of these two processes becomes faster with increasing denaturant concentration, meaning that an unfolding step is rate-limiting for folding of the majority of apoflavodoxin molecules. This, together with the absence of a lag in the formation of native molecules, means that the intermediate that populates during refolding is off-pathway.

A molten globule-like intermediate  $I_1$  populates during denaturant-induced equilibrium unfolding of apoflavodoxin. Its kinetic behaviour shows that it is not the high-energy on-pathway intermediate  $I_2$  detected during apoflavodoxin unfolding. The experimental data obtained on apoflavodoxin folding are consistent with the linear four-state folding mechanism:



with the off-pathway intermediate  $I_1$  being the one that also populates during denaturant-induced equilibrium unfolding of apoflavodoxin.  $I_1$  has to unfold before native apoflavodoxin can be formed. All folding *A. vinelandii* apoflavodoxin molecules pass through  $I_2$  before reaching the native state.

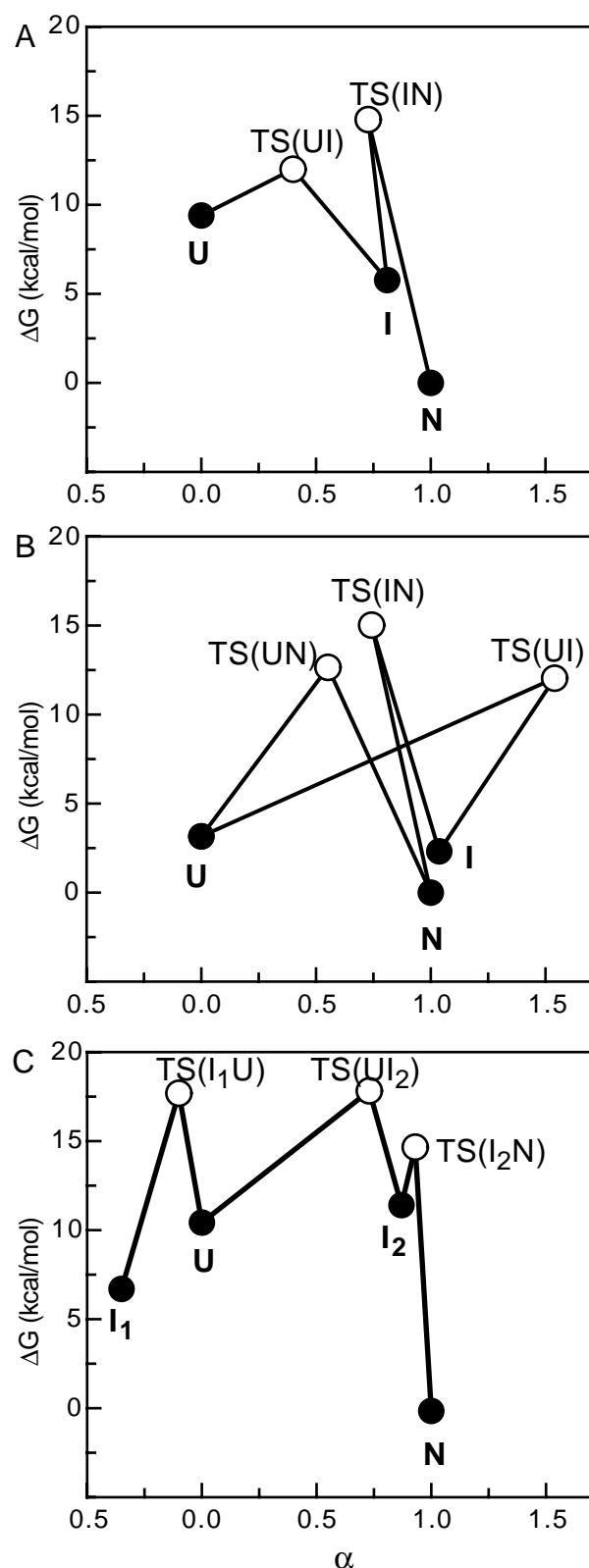
### b. Comparison of the folding of *A. vinelandii* apoflavodoxin with CheY folding

The  $\alpha$ - $\beta$  parallel protein the folding of which has been most intensely studied so far is CheY (129 amino acids, Figure 1). CheY is both sequentially and functionally unrelated to flavodoxin. A structural alignment of 184 backbone atoms of CheY and of *A. vinelandii* apoflavodoxin with an rmsd of 1.62 Å shows 4 % sequence identity and 20 % similarity according to the PAM250 matrix (Pearson 1990). The chevron plot recorded for CheY(un)folding displays a curvature on its folding side, indicative of the presence of an intermediate (Lopez-Hernandez and Serrano 1996; Lopez-Hernandez et al. 1997). The CheY kinetic folding data are described by a three-state folding mechanism that involves an on-pathway intermediate I:  $U \rightleftharpoons I \rightleftharpoons N$ . However, according to Lopez-Hernandez et al. (1997) it can not be excluded that the observed intermediate is off-pathway because the first step in CheY folding is too fast to be resolved.

Quantitative analysis of kinetic folding data results in rate constants for folding and unfolding, and kinetic  $m$ -values. A kinetic  $m$ -value, i.e. the denaturant concentration dependence of a specific folding or unfolding rate constant, informs about the difference in denaturant accessible surface area between a specific folding species and the transition state involved (Myers et al. 1995). In case of an unfolding process, the denaturant accessibility increases upon protein unfolding and the corresponding  $m$ -value is positive. In case of a folding process, the denaturant accessibility decreases upon folding and the corresponding  $m$ -value is thus negative.

The kinetic  $m$ -values obtained for CheY folding contain indirect information about the role of the observed intermediate during kinetic folding of CheY. The trajectory the folding molecules follow in terms of free energy and solvent accessibility in the mechanism for CheY folding proposed by Lopez-Hernandez et al. (1997) is shown in Figure 2A. As discussed, the mechanism used for the analysis of the experimental CheY folding data contains an on-pathway intermediate (Lopez-Hernandez et al. 1997). The use of this mechanism leads to  $m$ -values for the I to N and the N to I transition that are both positive (the reported values are both negative, due to the opposite definition of the  $m$ -value by Lopez-Hernandez et al. (1997) compared to the one used here). This implies that the transition state between the intermediate and the native state is less compact than the intermediate, and that the on-pathway intermediate of CheY has to unfold partially before productive folding to the native state occurs. Although the latter is not impossible, it is unlikely. Instead, just as in case of *A. vinelandii* apoflavodoxin folding, we think that the observed curvature on the folding side of the chevron plot of CheY is better explained by an off-pathway intermediate.

The folding behaviour of a number of CheY mutants has been studied. Inspection of Figure 1D in Lopez-Hernandez et al. (1997) clearly shows that the kinetic refolding limb of the chevron plot of the Hel2 mutant of CheY, a mutant in which helix two has been stabilised, has a positive slope at low denaturant concentrations. A positive slope (i.e. a positive  $m$ -value) implies that an unfolding process is rate-limiting in the formation of native CheY molecules. In our opinion, this positive slope confirms the involvement of an off-pathway folding intermediate during CheY folding. In addition, the observation made for CheY that stabilisation of any of its five helices causes slower refolding at low denaturant



**Figure 2.** Schematic representation of the folding trajectories of CheY (A) and of apoflavodoxin from *Anabaena* (B) and from *A. vinelandii* (C). The free energies and denaturant accessibilities of the unfolded (U), intermediate (I, I<sub>1</sub>, I<sub>2</sub>), and native (N) states and of the corresponding transition states of the three proteins are shown on the vertical and horizontal axis, respectively. The denaturant accessibility is expressed as the  $\alpha$ -value, i.e. the ratio of the  $m$ -value of a state and the  $m$ -value of the native protein. Unfolded protein has an  $\alpha$ -value of 0 and native protein an  $\alpha$ -value of 1. I represents the CheY or the *Anabaena* apoflavodoxin folding intermediate, I<sub>1</sub> and I<sub>2</sub> are the *A. vinelandii* apoflavodoxin folding intermediates presented in Scheme 1. TS(xy) is the transition state between species x and y. The heights of the barriers  $\Delta G^\ddagger$  are calculated from the reported folding and unfolding rates according to  $\Delta G^\ddagger = -RT \ln(k/k_0)$  using a value for  $k_0$  of  $10^8$  (Krieger et al. 2003). Folding molecules follow the lines connecting two points. (A) Folding trajectory of CheY assuming that the protein folds according to the linear folding scheme: Unfolded  $\rightleftharpoons$  Intermediate  $\rightleftharpoons$  Native as is proposed by Lopez-Hernandez and Serrano (1996) and Lopez-Hernandez et al. (1997). Note that the transition state between native CheY and its folding intermediate, TS(IN), is 10 % less solvent accessible than the intermediate. The values for the free energy and denaturant accessibility of TS(UI) are chosen arbitrarily but in such a manner that they are consistent with the reported formation of I being too rapid to be observed. The data shown are extracted from (Lopez-Hernandez and Serrano 1996; Lopez-Hernandez et al. 1997) for the pseudo-wildtype protein (i.e. the F14N mutant of CheY). (B) Folding trajectory of *Anabaena* apoflavodoxin assuming that the protein folds according to the triangular folding scheme 2 (see main text), as is proposed by Fernandez-Recio et al. (2001). Note that the transition state between U and I, TS(UI), is 50 % less solvent accessible than native *Anabaena* apoflavodoxin. The  $\alpha$ -values show that the transition between I and N involves an unfolding process. The free energy- and  $\alpha$ -values presented are calculated based on the results of Fernández-Recio et al. (2001). (C) Folding trajectory of *A. vinelandii* apoflavodoxin assuming that the protein folds according to the linear four-state folding scheme I<sub>1</sub>  $\rightleftharpoons$  U  $\rightleftharpoons$  I<sub>2</sub>  $\rightleftharpoons$  N as proposed in Chapter 2 of this thesis. The off-pathway intermediate I<sub>1</sub> is represented on the left-hand side of the unfolded

state, whereas both the on-pathway intermediate I<sub>2</sub> and native apoflavodoxin reside on the right-hand side of the unfolded state. The free energy- and  $\alpha$ -values are taken from Chapter 2. Note that in contrast to the proposed folding schemes for *Anabaena* apoflavodoxin and CheY folding, the denaturant accessibility of the folding *A. vinelandii* apoflavodoxin molecules gradually decreases on the productive folding route to their native state. concentrations compared to native CheY supports the conclusion that this intermediate, in which these helices are structured, is off-pathway.



Furthermore, inspection of Figure 2 in Munoz et al. (1994) reveals a curvature in the unfolding limb of the chevron plot of wild-type CheY folding, similar to as is observed in the kinetic *A. vinelandii* apoflavodoxin unfolding data. In more recent studies of CheY folding, this curvature is no longer observed, which is probably due to the change of buffer from 5 mM sodium phosphate to 50 mM PIPES. The curvature in the unfolding limb of the chevron plot of CheY is most likely caused by the presence of a high-energy on-pathway intermediate on the direct (un)folding route of CheY.

In conclusion, the CheY folding data published suggest that similar kinetic on- and off-pathway intermediates observed during *A. vinelandii* apoflavodoxin folding are also involved during kinetic CheY folding.

*c. Comparison of the folding of A. vinelandii apoflavodoxin with Anabaena apoflavodoxin folding*

Besides the presentation of the kinetic folding of *A. vinelandii* apoflavodoxin in Chapter 2 of this thesis, recently a rather limited number of kinetic folding experiments on an apoflavodoxin extracted from *Anabaena* have been reported (Fernandez-Recio et al. 2001). In the latter study, a kinetic mechanism for *Anabaena* apoflavodoxin folding is proposed that involves only one intermediate that is off the direct route linking unfolded and native apoflavodoxin. The unfolding of this intermediate is rate-limiting at low urea concentrations, as is clear from the increase of one of the two observed *Anabaena* apoflavodoxin refolding rates upon increasing the denaturant concentration.

The kinetic *Anabaena* apoflavodoxin folding data are according to Fernandez-Recio et al. (2001) best described by a three-state triangular model:



However, the folding intermediate of *Anabaena* apoflavodoxin as presented in Scheme 2 needs to have rather unusual properties, as recognised by Fernandez-Recio et al. Kinetic  $m$ -value analysis shows that its conformation seems to be as compact as the native state. In addition, the kinetic  $m$ -values indicate that interconversion between the intermediate and the native state of *Anabaena* apoflavodoxin involves a partial unfolding of the intermediate, as both  $m_{IN}$  and  $m_{NI}$  are positive (i.e. 0.48 and 0.42 kcal/mol·M<sup>-1</sup>, respectively (Fernandez-Recio et al. 2001)).

Even more remarkable, the unfolding of the *Anabaena* apoflavodoxin folding intermediate to the unfolded state is associated with a negative  $m$ -value ( $m_{IU} = -0.88 \pm 0.04$  kcal/mol·M<sup>-1</sup> (Fernandez-Recio et al. 2001)). This negative  $m$ -value implies a peculiar, super-compact transition state to occur between unfolded *Anabaena* apoflavodoxin and its folding intermediate. This transition state is approximately 50 % more compact than both the intermediate and the native state of the protein. Apoflavodoxin structures that are more compact than the native state might theoretically exist. Native apoflavodoxins contain a

relatively open, flexible flavin binding site (Steensma and van Mierlo 1998), which potentially could close resulting in a transition state that is more compact than the native state is. However, the presence of an apoflavodoxin conformation that needs to be 50 % more compact than native apoflavodoxin to be able to explain the negative  $m$ -value discussed is highly unlikely.

The trajectory the folding apoflavodoxin molecules follow in terms of free energy and solvent accessibility, assuming the proposed triangular mechanism for *Anabaena* apoflavodoxin folding to be valid, is shown in Figure 2B. For comparison, in Figure 2C the analogous trajectory for *A. vinelandii* apoflavodoxin folding is shown (data are taken from Chapter 2 of this thesis). In case of *A. vinelandii* apoflavodoxin folding the denaturant accessibility  $\alpha$  of the protein gradually decreases upon following the productive folding that leads from unfolded to native protein. Instead, if *Anabaena* apoflavodoxin folding would occur according to Scheme 2, large unrealistic changes in solvent accessibility of the folding protein are observed on the different folding routes individual *Anabaena* apoflavodoxin molecules can follow (Figure 2B).

In conclusion, the kinetic folding data published show that an off-pathway intermediate plays a significant role during *Anabaena* apoflavodoxin folding, just as is the case for *A. vinelandii* apoflavodoxin folding. However, the triangular *Anabaena* apoflavodoxin folding mechanism (Scheme 2) proposed by Fernandez-Recio et al. leads to, as discussed, some highly unlikely features of the protein states involved. Apparently, Scheme 2 does not correctly describe the kinetic folding of *Anabaena* apoflavodoxin. Further details of the kinetic mechanism of *Anabaena* apoflavodoxin folding might be revealed in the future by interrupted refolding and unfolding experiments similar to those presented in Chapter 2 for *A. vinelandii* apoflavodoxin.

#### *d. Available simulation data on the folding of an $\alpha$ - $\beta$ parallel protein*

The folding of CheY has been studied theoretically using a simple model for protein folding (Clementi et al. 2000). The model consists of a Gō-like potential, in which the only forces used arise from contacts which are present in the native state, whereas the energetic frustration is drastically reduced by not including residue-specific parameters. As the latter contacts are highly conserved among topologically related proteins, the simulations by Clementi et al. are expected to inform about the influence topology has on folding. The simulation roughly reproduces the structure of the transition state for folding of CheY as determined experimentally by means of mutational analysis (Lopez-Hernandez and Serrano 1996; Lopez-Hernandez et al. 1997). Furthermore, the simulations predict two possible intermediates: a short-lived on-pathway intermediate and an intermediate which is a misfolded trap (Clementi et al. 2000). The latter intermediate has all five helices which are present in native CheY well structured, but lacks its central parallel  $\beta$ -sheet. During the simulation, this misfolded species has to unfold, at least partially, before the native state is reached. Since non-native interactions are not possible in the model by Clementi et al. (2000), the misfolded species is interpreted to be the result of topological constraints.

In conclusion, since the constraints in the theoretical model of Clementi et al. (2000) are derived from contacts in the native state of CheY, application of this model to flavodoxin or any other protein that shares the doubly wound topology should produce the involvement of similar folding intermediates as theoretically predicted for CheY. The kinetic results obtained in Chapter 2 are the first experimental identification of both folding intermediates predicted theoretically to occur during the folding of an  $\alpha$ - $\beta$  parallel protein.

### Structural characteristics of the off-pathway intermediate observed during the folding of $\alpha$ - $\beta$ doubly wound proteins

In this section, the structural characteristics of the off-pathway folding intermediate of *A. vinelandii* apoflavodoxin are compared with those of the off-pathway folding intermediate observed during CheY folding. In case topology determines the formation of intermediates, the structure of these intermediates is expected also to depend on topology.

The off-pathway intermediate  $I_1$  of *A. vinelandii* apoflavodoxin (Scheme 1) populates during denaturant-induced equilibrium unfolding of the protein. A comparison of the fluorescence and CD data in the transition zone of unfolding leads to the conclusion that  $I_1$  has conserved secondary structure (approximately 65 % native  $\alpha$ -helicity, see Chapter 2) but that it lacks the characteristic tertiary structure of native apoflavodoxin.  $I_1$  has molten-globule like features. This can explain the relatively low sensitivity of this state to denaturant ( $m_{UI_1} = -1.83$  kcal/mol·M<sup>-1</sup> versus  $m_{UN} = -6.23$  kcal/mol·M<sup>-1</sup>, Table 1 of Chapter 2). These features of  $I_1$  are consistent with the observation by Clementi et al. (2000), Lopez-Hernandez and Serrano (1996) and Lopez-Hernandez et al. (1997) that the off-pathway intermediate in CheY folding has all native helices formed but lacks most of the long range interactions observed in the native state of the protein.

An off-pathway intermediate appears to be present in the folding of all three  $\alpha$ - $\beta$  doubly wound proteins the folding of which has been experimentally or theoretically investigated (i.e. CheY from *E. coli* and apoflavodoxin from *A. vinelandii* and *Anabaena*). Its presence is most likely determined by the topology of the native protein. Part of the interactions stabilising the off-pathway intermediate will thus be native-like. Some interactions that stabilise the intermediate might be non-native and can be sequence-specific. The latter may explain the observed differences between the stabilities of the off-pathway intermediates found in the folding of different doubly wound proteins. The off-pathway folding intermediate of *A. vinelandii* apoflavodoxin populates during denaturant-induced equilibrium unfolding experiments, whereas the corresponding intermediates of CheY and of *Anabaena* apoflavodoxin do not, or hardly, populate at equilibrium.

The off-pathway folding intermediate of CheY contains mainly helices as ordered structural elements, and ellipticity values indicate this to be the case for the corresponding *A. vinelandii* apoflavodoxin folding intermediate as well. The helical nature of both off-pathway intermediates explains why they are formed so easily during CheY and apoflavodoxin folding. Helices are formed much more rapidly than sheets, especially when parallel  $\beta$ -sheets are involved as is the case for apoflavodoxin and CheY. This is due to the highly local

character of the interactions in helices, whereas the residues that have to be brought into contact to form a parallel  $\beta$ -sheet are separated by many residues from one another (Bieri and Kiefhaber 1999). Apparently, rapid formation of  $\alpha$ -helices followed by some docking prevents the formation of the parallel  $\beta$ -sheet of apoflavodoxin and of CheY, and as a result the off-pathway folding intermediate has to unfold to yield native protein.

In case of CheY folding, phi-value analysis shows that helices 4 and 5, which are folded in the off-pathway intermediate, are unfolded in the subsequent rate limiting transition state. In addition, in this transition state  $\beta$ -strands 1 to 3 are formed (Lopez-Hernandez and Serrano 1996). Whether or not a similar situation occurs during *A. vinelandii* apoflavodoxin folding remains to be studied. However, the fluorescence anisotropy data presented in Chapter 2 allow for some speculation. The native state of apoflavodoxin has a low anisotropy due to specific side-chain interactions as discussed in Chapter 2. The off-pathway apoflavodoxin folding intermediate, which also populates during equilibrium unfolding of the protein, has a fluorescence anisotropy which is higher than that of the native protein (Figure 2 of Chapter 2). Consequently, the tryptophans that give rise to the fluorescence signal must be immobilised in the intermediate but the native interaction between Trp 128 and Trp 167 (Figure 1) is not formed in the intermediate. These tryptophan residues are located in helix 5 and  $\beta$ -strand 5a, respectively. Helix 5 and  $\beta$ -strand 5a may thus be formed in the off-pathway folding intermediate of *A. vinelandii* apoflavodoxin but they are definitely not natively oriented towards one another, as is also the case in the off-pathway intermediate of CheY (Lopez-Hernandez and Serrano 1996; Lopez-Hernandez et al. 1997; Clementi et al. 2000).

In conclusion, the structural characteristics of the off-pathway *A. vinelandii* apoflavodoxin folding intermediate and those of the CheY off-pathway intermediate are similar. This is to be expected when the conformation of off-pathway intermediates is determined to a substantial extent by the  $\alpha$ - $\beta$  parallel protein topology.

### **Structural characteristics of the on-pathway folding intermediate observed during the folding of *A. vinelandii* apoflavodoxin**

The on-pathway intermediate observed during the folding of *A. vinelandii* apoflavodoxin never significantly populates, neither kinetically not at equilibrium. Consequently, it is not possible to extract the precise conformational properties of this intermediate from the apoflavodoxin folding data presented in Chapter 2 of this thesis. The on-pathway intermediate has a denaturant accessibility that is comparable to those of both transition states that surround it on the direct folding route of *A. vinelandii* apoflavodoxin (Figure 2B). Consequently, the on-pathway folding intermediate probably structurally resembles these two transition states. The most denaturant-accessible one of these transition states (i.e. TS(UI<sub>2</sub>)) is the one that limits the folding rate under native conditions (Chapter 2). The conformation of this transition state most likely resembles the conformation of the rate-limiting transition state of CheY because, as discussed, the structure of transition states is conserved among structurally homologous proteins.

In case of CheY folding, phi-value analysis shows that helices 1 to 3 and  $\beta$ -strands 1 to 3 are formed in the rate-limiting transition state, whereas the C-terminal part of the protein (helices 4 and 5 and strands 4 and 5) are unstructured (Lopez-Hernandez and Serrano 1996; Lopez-Hernandez et al. 1997). The conformation of this transition state is roughly reproduced by folding simulations (Clementi et al. 2000). If the on-pathway intermediate of *A. vinelandii* apoflavodoxin indeed resembles the structure of the CheY transition state, its N-terminal half is structured, whereas its C-terminal half is unstructured.

## Conclusion

The on- and off-pathway intermediate experimentally observed during *A. vinelandii* apoflavodoxin folding have remarkable similarities with the two folding intermediates that are observed during the computational folding simulation of an  $\alpha$ - $\beta$  parallel protein. Both other  $\alpha$ - $\beta$  parallel proteins the folding of which has been studied experimentally, i.e. CheY and apoflavodoxin from *Anabaena*, populate at least one intermediate during kinetic folding. Based on the curved appearance of the folding limb of the chevron plots of CheY and *Anabaena* apoflavodoxin folding and based on the analysis of the corresponding *m*-values, both folding intermediates must be off-pathway. These intermediates need to unfold, just like the off-pathway intermediate of *A. vinelandii* apoflavodoxin, before productive folding occurs. The limited structural information that is available about the *A. vinelandii* apoflavodoxin off-pathway folding intermediate shows that its conformational properties are consistent with those of the misfolded, off-pathway intermediate observed during CheY folding. The conformation of the latter intermediate is known in more detail via mutational analysis studies and computational folding simulations. The curved unfolding limb of the chevron plot of CheY obtained with 5 mM sodium phosphate as buffer suggests that also an on-pathway intermediate is involved in CheY folding. Apparently, protein topology determines the appearance and kinetic role (i.e. being on- or off-pathway) of protein folding intermediates during the folding of proteins that have an  $\alpha$ - $\beta$  parallel topology. This implies that the overall shape of the free energy landscape of protein folding depends on protein topology.

## References

- Alm, E. and D. Baker (1999). "Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures." Proc Natl Acad Sci USA **96**(20): 11305-10.
- Baker, D. (2000). "A surprising simplicity to protein folding." Nature **405**(6782): 39-42.
- Bieri, O. and T. Kiefhaber (1999). "Elementary steps in protein folding." Biol Chem **380**(7-8): 923-9.
- Brenner, S. E., C. Chothia and T. J. Hubbard (1997). "Population statistics of protein structures: lessons from structural classifications." Curr Opin Struct Biol **7**(3): 369-76.

- Chiti, F., N. Taddei, P. M. White, M. Bucciantini, F. Magherini, M. Stefani and C. M. Dobson (1999). "Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding." Nat Struct Biol **6**(11): 1005-9.
- Clarke, J., E. Cota, S. B. Fowler and S. J. Hamill (1999). "Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway." Structure Fold Des **7**(9): 1145-53.
- Clementi, C., H. Nymeyer and J. N. Onuchic (2000). "Topological and energetic 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**(5): 937-53.
- Fernandez-Recio, J., C. G. Genzor and J. Sancho (2001). "Apoflavodoxin Folding Mechanism: An alpha/beta Protein with an Essentially Off-Pathway Intermediate." Biochemistry **40**(50): 15234-15245.
- Fersht, A. R., A. Matouschek and L. Serrano (1992). "The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding." J Mol Biol **224**(3): 771-82.
- Gerstein, M. (1997). "A structural census of genomes: comparing bacterial, eukaryotic, and archaeal genomes in terms of protein structure." J Mol Biol **274**(4): 562-76.
- Goldberg, D. P. (1999). "Finding the right fold." Nat Struct Biol **6**(11): 987-990.
- Grantcharova, V., E. J. Alm, D. Baker and A. L. Horwich (2001). "Mechanisms of protein folding." Curr Opin Struct Biol **11**(1): 70-82.
- Gunasekaran, K., S. J. Eyles, A. T. Hagler and L. M. Gierasch (2001). "Keeping it in the family: folding studies of related proteins." Curr Opin Struct Biol **11**(1): 83-93.
- Kraulis, P. J. (1991). "MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures." J Appl Cryst **24**: 946-950.
- Krieger, F., B. Fierz, O. Bieri, M. Drewello and T. Kiefhaber (2003). "Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding." J Mol Biol **332**(1): 265-274.
- Lopez-Hernandez, E., P. Cronet, L. Serrano and V. Munoz (1997). "Folding kinetics of CheY mutants with enhanced native alpha-helix propensities." J Mol Biol **266**(3): 610-20.
- Lopez-Hernandez, E. and L. Serrano (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**(1): 43-55.
- Martinez, J. C. and L. Serrano (1999). "The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved." Nat Struct Biol **6**(11): 1010-6.
- Munoz, V., E. M. Lopez, M. Jager and L. Serrano (1994). "Kinetic characterization of the chemotactic protein from Escherichia coli, CheY. Kinetic analysis of the inverse hydrophobic effect." Biochemistry **33**(19): 5858-66.
- Myers, J. K., C. N. Pace and J. M. Scholtz (1995). "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding." Protein Sci **4**(10): 2138-48.

- Pearson, W. R. (1990). "Rapid and sensitive sequence comparison with FASTP and FASTA." Methods Enzymol **183**: 63-98.
- Riddle, D. S., V. P. Grantcharova, J. V. Santiago, E. Alm, I. Ruczinski and D. Baker (1999). "Experiment and theory highlight role of native state topology in SH3 folding." Nat Struct Biol **6**(11): 1016-24.
- Steensma, E. and C. P. van Mierlo (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**(3): 653-66.
- Thorneley, R. N. F., G. A. Ashby, M. H. Drummond, R. R. Eady, D. L. Hughes, G. Ford, P. M. Harrison, A. Shaw, R. L. Robson, J. Kazlauskaitė and H. A. O. Hill (1994). Flavodoxin and nitrogen fixation: Structure, electrochemistry and posttranslational modification by coenzyme A. Flavins and flavoproteins 1993. K. Yagi. Berlin, Walter de Gruyter & Co.: 343-354.
- van Mierlo, C. P. M., J. M. van den Oever and E. Steensma (2000). "Apoflavodoxin (un)folding followed at the residue level by NMR." Protein Sci **9**(1): 145-57.
- van Mierlo, C. P. M., W. M. van Dongen, F. Vergeldt, W. J. H. van Berkel and E. Steensma (1998). "The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate." Protein Sci **7**(11): 2331-44.
- Villegas, V., J. C. Martinez, F. X. Aviles and L. Serrano (1998). "Structure of the transition state in the folding process of human procarboxypeptidase A2 activation domain." J Mol Biol **283**(5): 1027-36.





# 4 Partially unfolded forms of *Azotobacter vinelandii* apoflavodoxin identified by native state hydrogen exchange and NMR spectroscopy are off the productive folding route

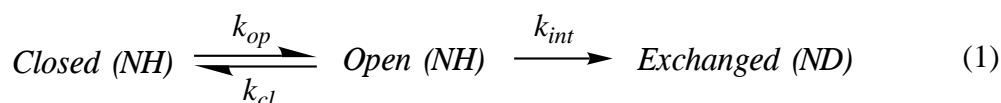
*Yves J.M. Bollen, Monique B. Kamphuis and Carlo P.M. van Mierlo*

Hydrogen deuterium (H/D) exchange detected by NMR spectroscopy shows that under native conditions the stabilities against local unfolding of all residues within *Azotobacter vinelandii* apoflavodoxin are lower than the one determined for global unfolding of the protein. Only at GuDCI concentrations exceeding 0.5 M, a few of the backbone amide protons of apoflavodoxin exchange via global protein unfolding. In contrast to the observations made on many other proteins, the global stability of the 179-residue apoflavodoxin can not be derived from the H/D exchange rates of the slowest exchanging backbone amide protons. The use of native state H/D exchange leads to the identification of five clusters of residues in apoflavodoxin the backbone amide protons of which exchange via five separate co-operative (sub)global unfolding events. Four partially unfolded forms (PUFs) of apoflavodoxin are proposed in which some non-native interactions apparently exist. The rates of interconversion of these PUFs with native apoflavodoxin are determined via the pD-dependence of the corresponding amide proton exchange rates. Stopped-flow kinetic folding rates show that these interconversion rates are inconsistent with any of the identified PUFs being on the direct productive folding route between globally unfolded and native apoflavodoxin. PUF1 and PUF2 are on an unfolding route starting from native apoflavodoxin that does not lead to globally unfolded apoflavodoxin. PUF3 and PUF4 are on a non-productive folding route that starts with globally unfolded apoflavodoxin. A common free energy barrier separates both PUF3 and PUF4 from globally unfolded apoflavodoxin. This barrier is proposed to be the same as the one that separates a known off-pathway apoflavodoxin folding intermediate ( $I_1$ ), which has molten globule characteristics, from the productive apoflavodoxin folding route. Both PUF3 and PUF4, as well as  $I_1$ , need to unfold before productive folding of apoflavodoxin can occur. The H/D exchange data obtained imply that apparently the molten globule  $I_1$  is similarly protected against H/D exchange than native apoflavodoxin is. Native state H/D exchange and classical denaturant-induced equilibrium and kinetic folding experiments together provide detailed information about the free energy landscape for protein folding, as is shown here for *A. vinelandii* apoflavodoxin folding.

## Introduction

Hydrogen deuterium (H/D) exchange detected by NMR spectroscopy is a powerful technique that can give detailed information about the stability and dynamics of a protein at the level of single amino acids. Backbone amide protons, provided they are water-accessible, exchange gradually with the protons of water or with deuterons in case D<sub>2</sub>O is used as a solvent. Amide protons involved in hydrogen bonding networks or buried from the solvent exchange more slowly than freely accessible amide protons since the protein has to unfold, partially or completely, to facilitate exchange of these buried or bonded protons.

Quantitative interpretation of hydrogen deuterium exchange is possible using a simple two-state model (Hvidt and Nielsen 1966):



In this model, the open or exchange-competent form and the closed or exchange-incompetent form of a protein at the site of a particular amide proton are in equilibrium with one another, and interconvert with rate constants for opening ( $k_{op}$ ) and closing ( $k_{cl}$ ). This interconversion might reflect local opening of the structure or a (sub)global unfolding process. From the open state exchange takes place with the intrinsic rate constant  $k_{int}$  that depends on amino acid sequence, pD and temperature and can be calculated using values derived from work on model peptides (Bai et al. 1993). The timecourse of exchange can be monitored at the residue level by NMR, since amide protons give rise to a <sup>1</sup>H NMR signal, and replacement of the proton by a deuteron leads to disappearance of this signal.

Under conditions favouring the closed state (i.e.  $k_{op} \ll k_{cl}$ ), the observed exchange rate constant  $k_{ex}$  is given by:

$$k_{ex} = \frac{k_{op}k_{int}}{k_{cl} + k_{int}} \quad (2)$$

Depending on the ratio of  $k_{cl}$  and  $k_{int}$ , two limiting cases may be reached. If  $k_{cl} \ll k_{int}$ , equation 2 reduces to

$$k_{ex} = k_{op} \quad (3)$$

Under these conditions, referred to as EX1, the measured exchange rate constant of a certain amide proton informs about the rate constant for local conversion between the closed and the open state of its micro-environment. If  $k_{cl} \gg k_{int}$ , equation 2 reduces to

$$k_{ex} = \frac{k_{op}k_{int}}{k_{cl}} = K_{op}k_{int} \quad (4)$$

In this limiting case, referred to as EX2, the ratio of the measured exchange rate constant  $k_{ex}$  and  $k_{int}$  provides the equilibrium constant for local opening of the protein structure  $K_{op}$ , which is the ratio of  $k_{op}$  and  $k_{cl}$  and is a measure for the local stability  $\Delta G_{op}$ :

$$\Delta G_{op} = -RT \ln(K_{op}) = -RT \ln(k_{op}/k_{cl}) = -RT \ln(k_{ex}/k_{int}) \quad (5)$$

Under EX2 conditions, it has been shown for several proteins that the  $\Delta G_{op}$  of the most stable residues corresponds to the free energy for global protein unfolding  $\Delta G_{NU}$  as determined by macroscopic spectroscopic probes, like fluorescence and circular dichroism spectroscopy. This identity in stability is observed when the corresponding experiments are performed under identical solvent conditions. A correction for the difference in free energies of the unfolded state in the H/D exchange experiment compared to the equilibrium unfolding experiment due to differences in Xaa-Pro peptide bond isomerisations needs to be applied (Huyghues-Despointes et al. 1999).

Native state hydrogen exchange experiments performed under EX2 conditions in the presence of small amounts of denaturant allow the identification of clusters of residues that reach their respective exchange-competent state by a concerted sub-global or global protein unfolding process (Bai et al. 1995; Chamberlain et al. 1996). In this approach, the free energy for a (sub)global opening process is considered to depend linearly on the denaturant concentration:

$$\Delta G_{op}(D) = \Delta G_{op}(0) - m[D] \quad (6)$$

with  $\Delta G_{op}(0)$  the local stability of an amide proton in the absence of denaturant and with  $[D]$  the concentration of denaturant. The  $m$ -value in equation 6 correlates to the amount of protein surface that becomes accessible to the denaturant in the respective opening process (Myers et al. 1995).

Three types of residues can be identified based on their  $\Delta G_{op}(0)$  and  $m$ -value. Amide protons of residues of type 1 are located in the stable core of a protein and exchange only from the globally unfolded state. These residues have  $\Delta G_{op}(0)$ -values that match the global protein stability against unfolding, and the corresponding  $m$ -values match the  $m$ -value associated with global protein unfolding. Under EX1 conditions, all residues of type 1 will have the same rate constant for amide proton exchange, which equals the global unfolding (i.e. native to globally unfolded) rate constant of the protein.

The second type of residues are protected against exchange in the native state and become water-exposed in a partially unfolded form (PUF) of the protein. The amide proton exchange behaviour of these residues is dominated by a common opening process that leads to exchange. As a consequence, under EX2 conditions amide protons of residues of type 2 that exchange from the same PUF all have the same free energy for opening  $\Delta G_{op}(0)$ , which equals the free energy for subglobal unfolding of native protein to the specific PUF,  $\Delta G_{N-PUF}(0)$ .  $\Delta G_{N-PUF}(0)$  is smaller than the  $\Delta G$  associated with global protein unfolding. The

corresponding  $m$ -values (equation 6) are smaller than the  $m$ -value associated with global protein unfolding since a smaller amount of protein surface becomes accessible to denaturant in the PUF compared to the globally unfolded state. Under EX1 conditions, type 2 residues that exchange from the same PUF all have the same amide proton exchange rate constants, which equal the subglobal opening rate constant, i.e. the rate constant for conversion from the native state to the PUF. In case more than one PUF exists for a specific protein, each PUF will be characterised by its own  $\Delta G_{N-PUF}(0)$ ,  $m$ -value and opening rate constant (Bai et al. 1995). Under EX2 conditions, type 2 residues thus can be assigned to become solvent exposed in a specific PUF according to the similarity in  $\Delta G_{op}(0)$  and  $m$ -values. They can also be assigned to a specific PUF according to the similarities of their exchange rates (i.e. opening rates) under EX1 conditions (Sivaraman et al. 2001).

The backbone amide protons of the third type of residues exchange through local fluctuations of the protein structure. They all have differing  $\Delta G_{op}(0)$  values that are lower than the  $\Delta G$  associated with global protein unfolding, and the corresponding  $m$ -values (equation 6) are equal to or close to zero, because only a small amount of protein surface becomes accessible to denaturant in such a local opening process. Under EX1 conditions, residues of type 3 have different amide proton exchange rate constants, because the corresponding rate constants for local opening of the protein structure differ. Above a specific denaturant concentration, subglobal or global protein unfolding will occur more often than local opening of the protein structure. This is because local opening processes are insensitive to denaturant concentration whereas (sub)global protein unfolding processes are promoted by denaturant. The (sub)global unfolding process will then start to dominate the amide proton exchange behaviour of the residue involved. As a consequence, the denaturant concentration dependent  $\Delta G_{op}$  isotherm of a type 3 residue will curve from virtually horizontal ( $m \sim 0$ ) at low concentration denaturant towards the isotherm that characterises the corresponding (sub)global protein unfolding event. The amide proton exchange behaviour of these type 3 residues is analysed according to a two-process model (Woodward and Hilton 1980; Woodward et al. 1982) in which exchange either occurs from a locally opened structure or from a (sub)globally unfolded structure, according to equation 7:

$$\Delta G_{op} = -RT \ln(K_{loc} + K_{glob}); \quad K_{loc} = \frac{[open]}{[closed]}; \quad K_{glob} = \frac{[unfolded]}{[closed]} \quad (7)$$

In equation 7  $[open]$  is the fraction of molecules that are in the locally opened, exchange competent state at the site of the residue of interest,  $[closed]$  is the fraction of molecules that are in the closed, exchange incompetent conformation at the site of interest, and  $[unfolded]$  is the fraction of molecules that are in the (sub)globally unfolded form, respectively.

As discussed, native state H/D exchange allows the characterisation of partially unfolded forms of a protein (Bai et al. 1995). These PUFs are generally interpreted to be high-energy folding intermediates that are often undetectable by other techniques, e.g. because they occur kinetically after the highest energy transition state for folding and do not populate at equilibrium (Chu et al. 2002; Englander et al. 2002). PUFs are believed to fold

one after the other in a strictly hierarchical order determined by the free energies of the PUFs, where the folding of one PUF facilitates the formation of the next (Chu et al. 2002; Englander et al. 2002). However, little evidence exists that links PUFs to equilibrium or kinetic folding intermediates.

The combination of native state H/D exchange under both EX1 and EX2 conditions provides unique information about the residues involved in specific co-operative (un)folding processes (Arrington and Robertson 2000; Canet et al. 2002; Hoang et al. 2002). Since amide proton exchange is base-catalysed above pD 4, the intrinsic exchange rate constant  $k_{int}$  increases tenfold by every increase of the pD by 1 unit above pD 4. As a consequence, the mechanism for amide proton exchange for a specific protein can be altered from EX2 to EX1 by increasing the pD sufficiently. Once residues have been identified to exchange via a specific (sub)global unfolding process under EX2 conditions in absence of denaturant, the exchange rate constant of these residues under EX1 conditions informs about the kinetic role of a specific PUF. Under these circumstances, the rate constant of conversion of the native state to the PUF can be determined (Yan et al. 2002). Comparison of the latter rate constants with the more conventionally determined kinetic folding data of a protein can give clues about the position of a PUF in the kinetic protein folding scheme.

Here, native state H/D exchange measurements are used to characterise at the residue level the (sub)global (un)folding events occurring during the (un)folding of *Azotobacter vinelandii* apoflavodoxin. Flavodoxins are small monomeric proteins involved in electron transport. They adopt an  $\alpha$ - $\beta$  parallel topology, and contain a non-covalently bound FMN cofactor (Mayhew and Tollin 1992). Apoflavodoxin from *Azotobacter vinelandii* (i.e. flavodoxin in the absence of the FMN cofactor) is structurally identical to flavodoxin except for some dynamic disorder in the flavin-binding region (Steensma et al. 1998; Steensma and van Mierlo 1998). The protein remains folded up to a concentration of 0.9 M GuDCI (van Mierlo et al. 1998; Chapter 2 of this thesis) and up to this denaturant concentration native-state H/D exchange measurements are feasible.

*A. vinelandii* apoflavodoxin populates a relatively stable folding intermediate in equilibrium unfolding studies (van Mierlo et al. 1998; van Mierlo et al. 2000). The change in denaturant accessibility upon unfolding of the equilibrium folding intermediate ( $I_1$ ) is 29 % of that of native apoflavodoxin and  $I_1$  is off-pathway: it needs to unfold to enable the production of native apoflavodoxin (Chapter 2 of this thesis). In addition to this intermediate, another folding intermediate ( $I_2$ ) is detected in kinetic apoflavodoxin (un)folding.  $I_2$  is on the direct route between unfolded and native apoflavodoxin and is high-energetic (Chapter 2 of this thesis); it does not significantly populate in both kinetic and equilibrium (un)folding experiments.  $I_2$  has approximately 80-90 % of the compactness of native apoflavodoxin and all apoflavodoxin molecules pass through this intermediate before reaching the native state (Chapter 2 of this thesis). The experimental data obtained on apoflavodoxin folding are consistent with the linear four state folding scheme presented in equation 8. The rate constants shown in equation 8 are the values extrapolated to 0 M GuHCl, in units of  $s^{-1}$ .



The rate constants for departure from  $I_2$ , either to N or to U, are not absolute. The kinetic data obtained only allow the determination of their ratio. Since  $I_2$  is shown to never populate significantly,  $k_{I_2N}$  must be large and is fixed to  $10^5 \text{ s}^{-1}$  (Chapter 2 of this thesis).

Native state H/D exchange data of *A. vinelandii* apoflavodoxin at various GuDCI concentrations are reported here. The data allow the determination of several clusters of amino acids the amide protons of which exchange via separate co-operative (sub)global unfolding events. Partially unfolded forms of apoflavodoxin are thus identified. The rate constants for opening and closing of these partially unfolded forms of apoflavodoxin could be determined via the measurement of the exchange rate constants of the amide protons of characteristic residues of these PUFs at high pD values. This informs about the position of these partially unfolded forms of apoflavodoxin within the kinetic scheme for apoflavodoxin folding.

## Materials and methods

### Materials

Guanidinium chloride (ultrapure) and potassium pyrophosphate were from Sigma (Bornem, Belgium).  $^{15}\text{N}$  Ammonium chloride was purchased from Campro Scientific, Veenendaal, The Netherlands. Uniformly deuterated guanidinium chloride (GuDCI) was obtained by a series of lyophilisation steps from  $\text{D}_2\text{O}$  (Cambridge Isotope Laboratories, Andover, UK).

### Protein purification

Uniformly  $^{15}\text{N}$ -labelled recombinant C69A *A. vinelandii* flavodoxin was obtained as described previously (Steensma et al. 1998; van Mierlo et al. 1998). Monomeric apoflavodoxin was subsequently prepared by a trichloroacetic acid preparation (Edmondson and Tollin 1971; van Mierlo et al. 1998) followed by size exclusion chromatography (Chapter 2 of this thesis).

### pD dependent H/D exchange

H/D Exchange experiments were started by dissolving lyophilised samples of uniformly  $^{15}\text{N}$ -labelled apoflavodoxin to a final protein concentration of  $\pm 1 \text{ mM}$  in  $\text{D}_2\text{O}$ , which contains  $100 \text{ mM}$  potassium pyrophosphate at various pD values and  $100 \mu\text{M}$  DSS. The samples were immediately transferred into the NMR machine, and a series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded at  $25^\circ\text{C}$ . A similarly prepared sample without containing protein was used to shim the magnetic field, tune the probe, and to determine the length of the  $90^\circ$  proton pulse prior to the start of the H/D exchange experiment. The dead-time of the exchange experiment, i.e. the difference in time between the mixing of  $\text{D}_2\text{O}$  with lyophilised

protein and the recording of the first HSQC spectrum, was approximately 5 minutes. The pD of the sample was measured using a glass electrode after the exchange experiment was completed. Experiments were done at pD values of 6.28, 6.80, and 8.23 (uncorrected pH-meter readings).

### **Native state H/D exchange**

In case of the native state H/D exchange experiments, lyophilised samples of uniformly  $^{15}\text{N}$ -labelled apoflavodoxin were dissolved to a final protein concentration of  $\pm 1$  mM in  $\text{D}_2\text{O}$ , which contains 100 mM potassium pyrophosphate and 100  $\mu\text{M}$  DSS and GuDCI in the concentration range of 0 to 750 mM. Potassium chloride was added in various amounts to obtain a salt concentration (i.e. GuDCI + KCl) of 750 mM in all the NMR samples. The pD of the samples was measured after exchange was complete, and varied between 5.7 and 5.9 (uncorrected pH meter readings). The further approach followed to measure H/D exchange was identical to the one described for the pD dependent H/D exchange experiments. Again, the dead-time of this exchange experiment was approximately 5 minutes.

### **NMR spectroscopy and data analysis**

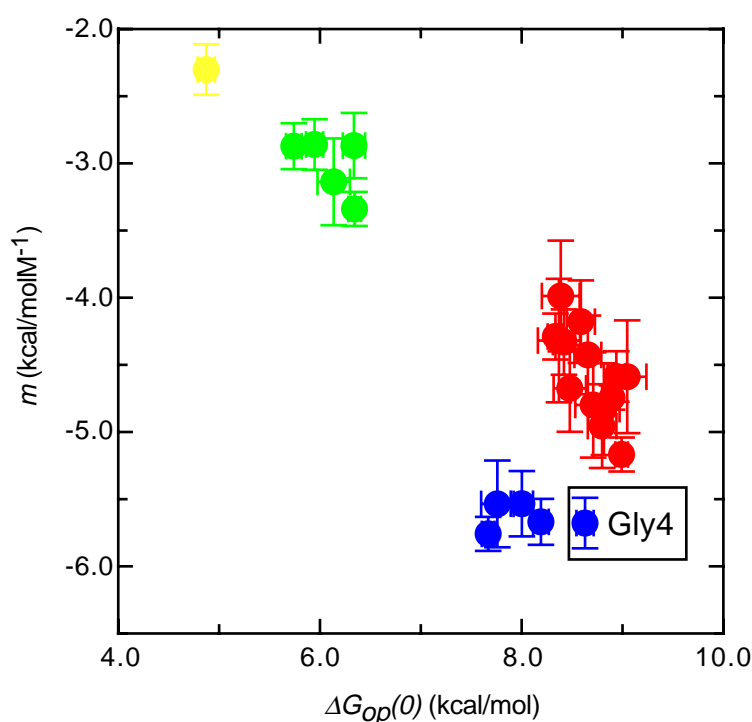
Gradient-enhanced  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (Palmer III et al. 1991; Kay et al. 1992) were recorded at 25 °C on a Bruker AMX500 spectrometer operating at a proton frequency of 500.13 MHz as described previously (van Mierlo et al. 2000). In all exchange experiments, first 9 HSQC spectra were recorded with 1024 complex data points in the direct  $^1\text{H}$  dimension and 96 complex data points in the indirect  $^{15}\text{N}$  dimension with two scans per increment. Each HSQC spectrum took 8 minutes and 40 seconds to acquire. As a result, a reasonably high sampling rate constant in the beginning of the exchange experiment could be obtained. Subsequently, a series of spectra with identical settings but with an increased number of complex data points (i.e. 256) in the indirect dimension was recorded. Each HSQC spectrum now lasted 22 minutes and 46 seconds before completion and has a higher spectral resolution than the former HSQC spectra.

The NMR data were processed on a Silicon Graphics Octane workstation using the program Felix 2D (MSI, San Diego, CA, USA). In the direct  $^1\text{H}$  dimension, data were multiplied by a Gaussian function and zero-filled to 2048 points prior to Fourier transform. In the indirect  $^{15}\text{N}$  dimension data were zero-filled once (to 192 points if 96 complex points were recorded, or to 512 points in case 256 complex points were recorded) and multiplied by a squared cosine bell prior to Fourier transform. Spectra were referenced to internal DSS. Peak heights were automatically determined using the macros and scripts provided by A. G. Palmer III (Palmer III), and a single exponential function was fitted to the data using the program CurveFit (Palmer III) to obtain the exchange rate constants  $k_{\text{ex}}$ . Assignments of the NMR crosspeaks were adapted from (Steensma and van Mierlo 1998) by either a pD titration or a KCl titration to the conditions used in the present study.

## Results and Discussion

### Identification of five co-operatively (un)folding clusters within *A. vinelandii* apoflavodoxin

The local free energy values for opening,  $\Delta G_{op}$ , of *A. vinelandii* apoflavodoxin at different GuDCI concentrations are listed in Table 1 for those residues the amide protons of which exchange sufficiently slowly to allow for an accurate determination of  $\Delta G_{op}$  at more than one GuDCI concentration. Residues that display a linear non-zero dependence of  $\Delta G_{op}$  on GuDCI concentration over the complete concentration range used (i.e. 0 - 0.75 M) are highlighted in Table 1. A linear function (equation 6) is fitted to this dependence, and the corresponding slopes ( $m$ -values) are plotted against  $\Delta G_{op}$  at zero molar GuDCI ( $\Delta G_{op}(0)$ ) in Figure 1. Residues are grouped based on similar  $\Delta G_{op}(0)$ - and  $m$ -values. Four clusters are identified within apoflavodoxin, each with specific  $\Delta G_{op}(0)$ - and  $m$ -values, indicating that a specific (sub)global unfolding process leads to amide proton exchange for all residues within one cluster. This is in contrast with the observation made by Marqusee and co-workers for



**Figure 1.** Free energy difference for local opening of native *A. vinelandii* apoflavodoxin at the residue-level at 0 M GuDCI (i.e.  $\Delta G_{op}(0)$ ) versus the corresponding dependence of  $\Delta G_{op}$  on concentration denaturant (i.e. the  $m$ -value), determined from native state H/D exchange experiments. Equation 6 is fitted to the denaturant-dependent  $\Delta G_{op}$  data of residues for which this dependence is fully linear in the denaturant-range studied (i.e. 0 - 0.75 M GuDCI). The resulting  $\Delta G_{op}(0)$ - and corresponding  $m$ -values are shown together with their standard fitting errors. The amino acid residues group in four clusters: Cluster 1 (yellow, Val 141), Cluster 2 (green, residues Phe146, Val147, Gly148, Leu149, Ala150), Cluster 3 (blue, residues Ala4, Leu5, Phe6, Ile21, Lys22), Cluster 4 (red, residues Ile51, Leu52, Gly53, Val91, Ala92, Leu93, Phe94, Trp167, Leu168, Ala169, Gln170, Ile171, Ala172). The data are obtained at 25 °C in 100 mM potassium pyrophosphate. The pD of the samples varied between 5.7 and 5.9.



**Table 1.** Calculated free energy difference ( $\Delta G_{op}$ ) for local opening of *A. vinelandii* apoflavodoxin at the residue-level as a function of GuDCI concentration, and rates for local opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of apoflavodoxin in the absence of denaturant. <sup>a</sup>

residue	cluster	$\Delta G_{op}(0)$	$\Delta G_{op}(0.15)$	$\Delta G_{op}(0.30)$	$\Delta G_{op}(0.45)$	$\Delta G_{op}(0.60)$	$\Delta G_{op}(0.75)$
3	3	6.33	5.90	5.62	5.08	4.76	3.76
<b>4</b>	<b>3</b>	<b>8.66</b>	<b>7.81</b>	<b>6.95</b>	<b>5.86</b>	<b>5.31</b>	<b>4.41</b>
<b>5</b>	<b>3</b>	<b>8.18</b>	<b>7.25</b>	<b>6.64</b>	<b>5.59</b>	<b>4.87</b>	<b>3.86</b>
<b>6</b>	<b>3</b>	<b>7.90</b>	<b>7.22</b>	<b>6.55</b>	<b>5.33</b>	<b>4.72</b>	<b>3.84</b>
8	3	5.87	5.84	5.90	5.77	5.10	4.15
17	3	5.22	5.06	4.93	4.45	4.12	3.49
18	3	6.76	6.50	6.27	5.11	4.68	3.87
19	3	6.90	7.02	6.27	5.35	4.64	3.89
<b>21</b>	<b>3</b>	<b>7.12</b>	<b>6.91</b>	<b>5.92</b>	<b>4.85</b>	<b>4.28</b>	<b>3.41</b>
<b>22</b>	<b>3</b>	<b>7.49</b>	<b>6.93</b>	<b>6.12</b>	<b>5.21</b>	<b>4.47</b>	<b>3.60</b>
24	3	4.66	4.43	4.54	4.21	4.13	3.79
30	3	5.47	5.27	5.60	4.88	4.60	4.22
46	4	3.99	4.33	3.98	3.65	3.68	3.45
47	4	5.20	4.96	5.06	4.83	4.93	4.63
49	4	6.90	6.63	6.59	6.27	6.07	5.39
50	4	8.74	8.55	7.39	6.91	6.41	5.31
<b>51</b>	<b>4</b>	<b>8.27</b>	<b>7.69</b>	<b>7.21</b>	<b>6.20</b>	<b>5.98</b>	<b>5.00</b>
<b>52</b>	<b>4</b>	<b>8.25</b>	<b>7.65</b>	<b>7.60</b>	<b>6.34</b>	<b>5.93</b>	<b>4.99</b>
<b>53</b>	<b>4</b>	<b>8.84</b>	<b>8.41</b>	<b>7.49</b>		<b>5.69</b>	<b>5.21</b>
81	4	7.32	6.82	6.54	5.89	5.60	4.63
82	3	4.89	4.70	4.62	4.27	4.05	3.52
84	3	5.33	5.00	5.01	4.91	4.80	4.20
87	4	4.73	4.58	4.49	4.46	4.30	4.17
89	4	7.47	7.06	6.73	5.97	5.61	4.75
<b>91</b>	<b>4</b>	<b>8.52</b>	<b>7.61</b>	<b>7.35</b>	<b>6.05</b>	<b>5.93</b>	<b>4.89</b>
<b>92</b>	<b>4</b>	<b>8.93</b>	<b>8.03</b>	<b>7.70</b>	<b>6.56</b>	<b>6.21</b>	<b>5.27</b>
<b>93</b>	<b>4</b>	<b>8.65</b>	<b>7.93</b>	<b>7.50</b>	<b>6.35</b>	<b>6.07</b>	<b>4.96</b>
<b>94</b>	<b>4</b>	<b>8.57</b>	<b>7.88</b>	<b>7.73</b>	<b>6.42</b>	<b>6.13</b>	<b>5.24</b>
95	4	7.98	7.46	7.18	6.20	5.87	5.04
110	5	5.69	5.60	5.41	5.29	5.15	4.86
111	5	7.50	7.14	6.84	6.36	6.17	5.54
114	5	8.90	8.66	8.57	7.75	6.86	6.17
115	5	7.59	7.34	7.34	6.73	6.61	5.95
116	5	7.78	7.63	7.66	7.19	7.02	6.16
117	5	9.28	8.85	8.32	7.58	6.96	5.83
119	4		4.80	4.09	3.67	3.73	3.77
120	4	6.96	6.71	6.53	6.05	5.89	4.95
123	4	7.24	6.88	6.65	6.20	5.89	5.05
125	4	7.05	6.68	6.69	5.81	5.50	4.62
128	3	5.73	5.40	5.13	4.74	4.38	3.47
140	1	4.65	4.36	4.16	3.62	3.54	3.18
<b>141</b>	<b>1</b>	<b>4.90</b>	<b>4.45</b>	<b>4.22</b>	<b>3.84</b>	<b>3.50</b>	<b>3.13</b>
142	1	4.35	3.98	3.92	3.49	3.24	2.93
145	2	4.59	4.42	4.32	4.07	3.84	3.69
<b>146</b>	<b>2</b>	<b>5.98</b>	<b>5.48</b>	<b>5.09</b>	<b>4.58</b>	<b>4.37</b>	<b>3.75</b>

Table 1 continued

residue	cluster	$\Delta G_{op}(0)$	$\Delta G_{op}(0.15)$	$\Delta G_{op}(0.30)$	$\Delta G_{op}(0.45)$	$\Delta G_{op}(0.60)$	$\Delta G_{op}(0.75)$
<b>147</b>	<b>2</b>	<b>5.66</b>	<b>5.44</b>	<b>4.78</b>	<b>4.49</b>	<b>4.09</b>	<b>3.52</b>
<b>148</b>	<b>2</b>	<b>6.34</b>	<b>5.89</b>	<b>5.49</b>	<b>5.01</b>	<b>4.72</b>	<b>4.13</b>
<b>149</b>	<b>2</b>	<b>6.18</b>	<b>5.89</b>	<b>5.55</b>	<b>4.92</b>	<b>4.14</b>	<b>3.86</b>
<b>150</b>	<b>2</b>	<b>6.09</b>	<b>5.73</b>	<b>5.25</b>	<b>4.64</b>	<b>4.25</b>	<b>3.80</b>
151	2	5.75	5.42	5.10	4.40	4.12	3.71
163	4	5.10	5.01	4.74	4.52	4.30	
164	4	6.93	6.37	5.96	5.46	5.17	4.40
166	4	6.67	6.31	6.13	5.72	5.39	4.78
<b>167</b>	<b>4</b>	<b>8.42</b>	<b>7.48</b>	<b>7.32</b>	<b>6.32</b>	<b>5.98</b>	<b>4.98</b>
<b>168</b>	<b>4</b>	<b>9.04</b>	<b>8.02</b>	<b>6.96</b>	<b>6.36</b>	<b>6.17</b>	<b>5.07</b>
<b>169</b>	<b>4</b>	<b>8.89</b>	<b>8.28</b>	<b>8.11</b>	<b>6.79</b>	<b>6.46</b>	<b>5.43</b>
<b>170</b>	<b>4</b>	<b>8.47</b>	<b>7.92</b>	<b>7.64</b>	<b>6.55</b>	<b>6.22</b>	<b>5.32</b>
<b>171</b>	<b>4</b>	<b>8.65</b>	<b>7.93</b>	<b>7.98</b>	<b>6.41</b>	<b>6.01</b>	<b>5.05</b>
<b>172</b>	<b>4</b>	<b>8.24</b>	<b>7.71</b>	<b>7.57</b>	<b>6.50</b>	<b>6.11</b>	<b>5.23</b>
175	4	6.56	6.13	5.90	5.25	4.85	4.23

<sup>a</sup> Free energy values are in kcal/mol. Experiments were done at 25 °C in 100 mM potassium pyrophosphate, at the specific GuHCl concentrations indicated, i.e.  $\Delta G_{op}(0.15)$  is determined in a native-state H/D exchange experiment at 0.15 M GuHCl, etceteras. The pD of the different samples varied between 5.7 and 5.9.

Residues are grouped into clusters that exchange via the same subglobal unfolding event at 0.75 M GuHCl. The residues for which the subglobal unfolding reaction is the dominant mechanism for exchange of their amide proton down to 0 M GuHCl are indicated in bold. In case of the other residues, a local opening reaction determines the amide proton exchange rate at low GuHCl concentration (see main text).

*Escherichia coli* Ribonuclease HI, where  $m$ -values plotted against  $\Delta G_{op}(0)$ -values do not cluster, but spread rather broadly. The latter is interpreted to be an indication for the involvement of a continuum of states instead of a limited number of states in the folding of Ribonuclease HI (Parker and Marqusee 1999).

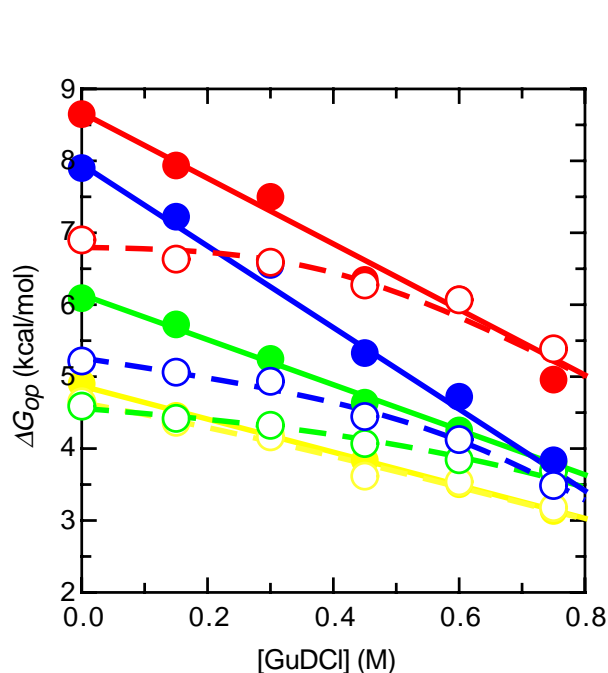
The backbone amide protons of the residues not shown in Figure 1 mostly display small  $m$ -values at low GuDCl concentration. At higher GuDCl concentrations (up to 0.75 M GuDCl) the denaturant-dependence of their  $\Delta G_{op}$ -values curves towards the isotherm corresponding to one of the four clusters identified in Figure 1. Consequently, at low GuDCl concentration a local opening process is the dominant mechanism for exchange of these residues. The amide proton exchange data of these residues have been analysed according to a two-process model with equation 7 and the results are listed in Table 1. These residues are assigned to belong to one of the specific subglobally unfolding clusters identified when equation 7 can be fitted to their curving  $\Delta G_{op}$  data with  $K_{glob}$  fixed based on the values for  $\Delta G_{op}$  and  $m$  of that specific cluster. Most of these locally unfolding residues display an  $m$ -value that deviates significantly from zero, and no correlation with the value of the local  $\Delta G_{op}(0)$  is observed. This indicates that the corresponding local opening events involve several residues but have a low co-operativity. All these residues are positioned at the periphery of elements of secondary structure. In contrast, the residues shown in Figure 1 that exchange only via (sub)global unfolding events are in the centre of elements of secondary structure of apoflavodoxin.

In Figure 2, the differences observed in  $\Delta G_{op}$ -dependence on denaturant concentration are shown for two typical residues per individual co-operatively unfolding cluster identified within apoflavodoxin. Cluster 1 within apoflavodoxin, which is coloured yellow in Figure 2 (characterised by  $\Delta G_{op}(0) = 4.87 \pm 0.03$  kcal/mol, and  $m = -2.3 \pm 0.1$  kcal/mol M<sup>-1</sup>) has only one residue, Val141, the  $\Delta G_{op}$  of which depends linearly on the GuDCI concentration up to 0.75 M. Its neighbour, Ala140, has a somewhat lower  $\Delta G_{op}$  at zero molar GuDCI, and its isotherm curves smoothly towards the isotherm corresponding to Val141. Also Val142 is a member of the same cluster. Cluster 1 is shown in yellow in the tertiary structure of apoflavodoxin (Figure 3). The three residues discussed form one of the two strands of a small beta structure in an inserted loop typical for longchain flavodoxins that connects beta strand 5A and 5B.

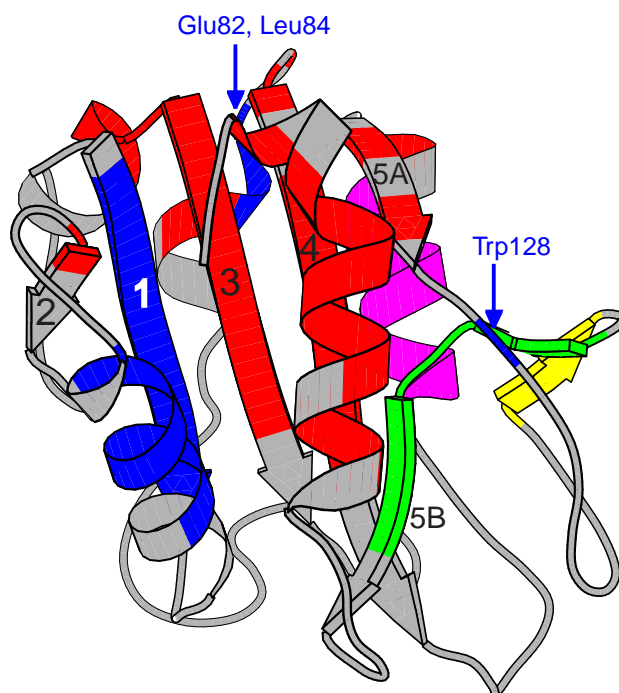
Cluster 2 within apoflavodoxin consists of the other strand of the small beta structure in the loop characteristic for longchain flavodoxins together with beta strand 5B (coloured green in Figures 1 to 3,  $\Delta G_{op}(0) = 6.14 \pm 0.05$  kcal/mol,  $m = -3.1 \pm 0.1$  kcal/mol M<sup>-1</sup>).

Cluster 3 within apoflavodoxin (coloured blue in Figures 1 to 3,  $\Delta G_{op}(0) = 7.95 \pm 0.09$  kcal/mol,  $m = -5.7 \pm 0.2$  kcal/mol M<sup>-1</sup>) involves mainly a large part of the first 24 N-terminal residues of apoflavodoxin. Interestingly, at least two residues (i.e. Trp128 and Glu82) that are far from the N terminus both sequentially and spatially (Figure 3), show a curvature of their  $\Delta G_{op}$  isotherm towards the linear one characterising the co-operative unfolding of Cluster 3 (Figure 4). The isotherm corresponding to Leu84 displays a similar curvature, and although the isotherm ends at a somewhat higher  $\Delta G_{op}$  value at 0.75 M GuDCI than the average for this cluster, Leu84 most likely belongs to Cluster 3 (Figure 4). This suggests that the subglobal unfolding of the N-terminal part of apoflavodoxin also affects other distant parts of the protein structure. Of the residues belonging to Cluster 3, Gly4 has a significantly higher  $\Delta G_{op}(0)$  (Figure 1). Since the corresponding  $m$ -value is identical within error to the average  $m$ -value of residues belonging to Cluster 3, Gly4 is considered to be a member of this cluster as well. Gly4 seems to have residual protection against exchange in the exchange-competent state of Cluster 3; its  $\Delta G_{op}(0)$  and  $m$ -value are not used for the determination of the average  $\Delta G_{op}(0)$  and  $m$ -value of Cluster 3.

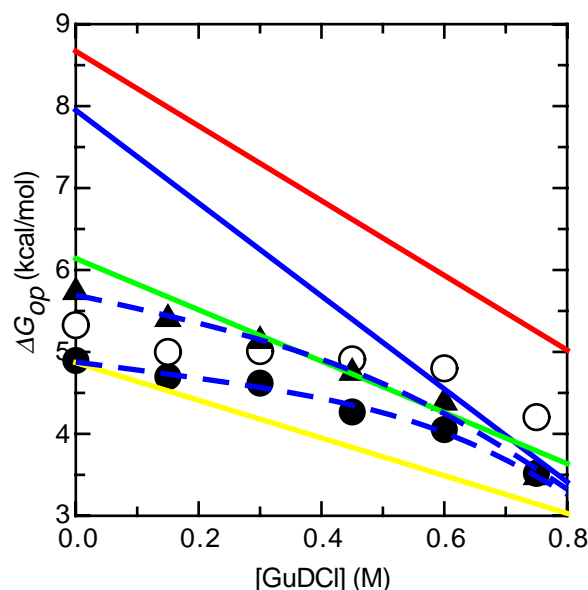
Cluster 4 within apoflavodoxin ( $\Delta G_{op}(0) = 8.67 \pm 0.05$  kcal/mol,  $m = -4.6 \pm 0.1$  kcal/mol M<sup>-1</sup>) comprises the largest number of residues of the protein, which are positioned in most of the remaining secondary structural elements (coloured red in Figure 3).



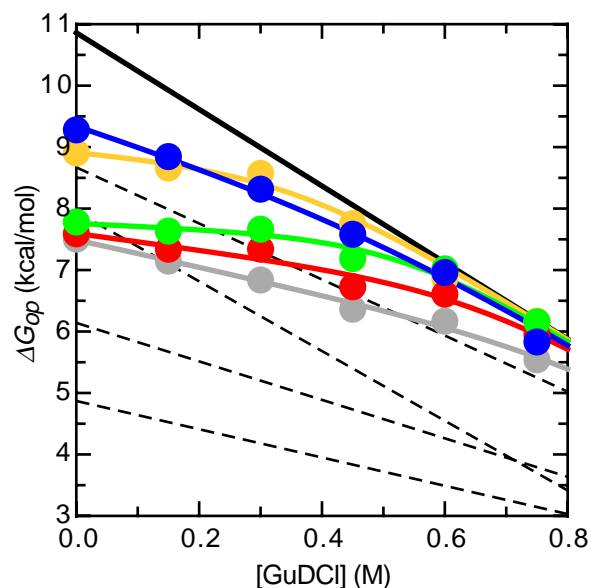
**Figure 2.** Typical examples of the GuDCl dependence of  $\Delta G_{op}$  of the backbone amide protons of residues belonging to each of the four clusters identified by native state H/D exchange within *A. vinelandii* apoflavodoxin. The colour-coding is as in Figure 1. Each solid line shows the characteristic linear isotherm of a specific cluster shown in Figure 1, which is defined by the average  $\Delta G_{op}(0)$  and  $m$ -value of that specific cluster. Data of residues for which the denaturant-dependence of  $\Delta G_{op}$  is linear in the native-state region (i.e. 0 - 0.75 M GuDCl) are depicted by filled circles. Instead, data of residues for which  $\Delta G_{op}$  joins the latter isotherm at denaturant concentrations larger than 0 M are depicted by open circles and the corresponding isotherm, resulting of a fit of a two-process model (equation 7) to the data, is dashed. Cluster 1 (yellow) is represented by Val141 (●) and Ala140 (○), Cluster 2 (green) by Ala150 (●) and Lys145 (○), Cluster 3 (blue) by Phe6 (●) and Val17 (○), and Cluster 4 (red) by Leu93 (●) and Phe49 (○). The data are obtained at 25 °C in 100 mM potassium pyrophosphate. The pD of the samples varied between 5.7 and 5.9.



**Figure 3.** Molscript (Kraulis 1991) cartoon drawing of the X-ray structure of *A. chroococcum* flavodoxin (Thorneley et al. 1994), the sequence of which is 95 % identical to *A. vinelandii* flavodoxin. The secondary structure depicted is as determined for *A. vinelandii* apoflavodoxin by NMR spectroscopy (Steensma and van Mierlo 1998). The clusters of co-operatively unfolding residues within *A. vinelandii* apoflavodoxin as determined by native state H/D exchange are coloured yellow (Cluster 1), green (Cluster 2), blue (Cluster 3), red (Cluster 4), and purple (Cluster 5), respectively. The  $\beta$ -sheets are numbered based on the primary sequence of the protein.



**Figure 4.** The dependence of  $\Delta G_{op}$  of the backbone amides of Glu82 (●), Leu84 (○), and Trp128 (▲) of *A. vinelandii* apoflavodoxin on the GuDCl concentration. The linear isotherms describing the co-operative unfolding of Clusters 1 to 4 are shown in the same colours as used in Figures 1 to 3. The dashed lines are best fits of a two-process model (equation 7) to the Glu82 and the Trp128 data joining the isotherm of Cluster 3. The data are obtained at 25 °C in 100 mM potassium pyrophosphate. The pD of the samples varied between 5.7 and 5.9.



**Figure 5.** GuDCl-dependence of the  $\Delta G_{op}$  of the backbone amide protons of amino acid residues that are not categorised to belong to the four in Figure 1 identified co-operatively unfolding clusters of residues within *A. vinelandii* apoflavodoxin (Gly111: grey, Tyr114: orange, Ser115: red, Phe116: green, Phe117: blue). The isotherm describing the global unfolding of apoflavodoxin as determined via denaturant-induced equilibrium unfolding measurements, which is corrected for Xaa-Pro peptide bond isomerisations, is shown in black (see main text). The best fit of a two-process model (equation 7) to the denaturant-dependence of the  $\Delta G_{op}$  data of the individual residues are shown as coloured lines. The two-process model assumes that at high concentrations denaturant the corresponding isotherms join the global unfolding isotherm in black. For comparison, the linear  $\Delta G_{op}$  isotherms corresponding to the four identified co-operative unfolding clusters within flavodoxin are shown by dashed lines. Note that the  $\Delta G_{op}$ -values of Gly111, Tyr114, Ser115, Phe116 and Phe117 are larger than the  $\Delta G_{op}$ -values of the latter isotherms at various GuDCl concentrations. The data are obtained at 25 °C in 100 mM potassium pyrophosphate. The pD of the samples varied between 5.7 and 5.9.

Interestingly, the remaining, non-categorised residues of apoflavodoxin have curved  $\Delta G_{op}$  isotherms that display  $\Delta G_{op}$  values that are above the isotherms of the four identified clusters at various GuDCl concentrations (Figure 5). These residues all reside in alpha helix 4 of apoflavodoxin (Figure 3) and the corresponding isotherms curve towards a common isotherm at high concentrations denaturant (shown in black in Figure 5). This common isotherm shows the denaturant-dependence of the free energy difference for global unfolding of apoflavodoxin ( $\Delta G_{NU}$ ) as determined from equilibrium (un)folding experiments.  $\Delta G_{NU}$  is corrected for Xaa-Pro peptide bond isomerisations that occur in equilibrium unfolded apoflavodoxin, but do not occur during the short time the unfolded state is visited in the native state H/D exchange experiment. The global stability of apoflavodoxin ( $\Delta G_{NU}$ ) as

derived from the global analysis of denaturant-induced equilibrium unfolding curves obtained by using various spectroscopic probes is  $10.45 \pm 0.52$  kcal/mol, with  $m_{NU} = -6.23 \pm 0.23$  kcal/mol M<sup>-1</sup> (Chapter 2 of this thesis). To correct for Xaa-Pro peptide bond isomerisation events in equilibrium unfolded apoflavodoxin, 0.42 kcal/mol needs to be added to  $\Delta G_{NU}$  (Reimer et al. 1998). Clearly, the amide protons of the discussed residues of alpha helix 4 of apoflavodoxin exchange via global unfolding of apoflavodoxin at high GuDCI concentrations. Together they form Cluster 5 within apoflavodoxin (with  $\Delta G_{op}(0) = \Delta G_{NU} + 0.42$  kcal/mol =  $10.87 \pm 0.52$  kcal/mol,  $m = -6.23 \pm 0.23$  kcal/mol M<sup>-1</sup>).

At low GuDCI concentrations (< 0.75 M), local opening processes dominate the amide proton exchange process for residues belonging to Cluster 5 within apoflavodoxin. As a consequence, in contrary to what is observed for many other proteins (Huyghues-Despointes et al. 1999), the  $\Delta G_{op}(0)$  values of the most stable residues of *A. vinelandii* apoflavodoxin are not identical to the global stability ( $\Delta G_{NU}$ ) of the protein. In the absence of denaturant, none of the identified clusters within *A. vinelandii* apoflavodoxin exchange their amide protons from the globally unfolded state of the protein.

### **Partially unfolded forms of *A. vinelandii* apoflavodoxin**

Native state H/D exchange data allow the identification of five clusters of residues within apoflavodoxin. Residues in a particular cluster share a common co-operative subglobal unfolding process that brings these residues into a structure in which they are water accessible. The latter structures are thus partially unfolded forms (PUFs) of a protein (Bai et al. 1995). It is believed that PUFs are snapshots of the direct (un)folding route of proteins (Englander et al. 2002). The order in which PUFs appear on this route is determined by their stability difference with respect to the native state. The PUF with the lowest stability difference with respect to the native state is formed first starting from the native state. The PUF with the second lowest stability difference with respect to the native state is subsequently formed from the first PUF, etceteras. This order is thought to be caused by the co-operative nature of protein folding (Englander et al. 2002). Assuming such sequential unfolding to occur for native apoflavodoxin, structures of the partially unfolded forms of apoflavodoxin have been created (Figure 6). Starting from native apoflavodoxin, in PUF1 residues of Cluster 1 are unfolded, subsequently PUF2 is formed in which residues of Cluster 2 are unfolded in addition to those of Cluster 1, and so on.

### **Are the partially unfolded forms of *A. vinelandii* apoflavodoxin partially misfolded?**

An important assumption underlying the idea that PUFs represent a sequential order in the (un)folding of a protein is that the parts of the PUFs that are folded are all native-like. However, some observations on apoflavodoxin argue against this assumption. The first observation is that in PUF1 one of the two strands of a little  $\beta$ -sheet is unfolded, whereas the other strand remains folded (Figures 3 and 6). It is unlikely however, that residues in the remaining folded strand are observed to be shielded from water without non-native interactions with other folded parts of the protein being involved. The second observation is that in PUF2 the long loop typical for long-chain flavodoxins and  $\beta$ -strand 5B are unfolded,

whereas the backbone amide of a single residue in this part of the protein (i.e. Trp128) remains protected against exchange (Figures 3 and 6). Most likely, Trp128 has non-native interactions with the part of the protein that remains folded in PUF2. The third observation is that in PUF3  $\beta$ -strand 2, which is at the periphery of apoflavodoxin, should be natively folded according to the Englander model, whereas its neighbour  $\beta$ -strand 1, which belongs to the core of apoflavodoxin, is unfolded (Figures 3 and 6). It seems unlikely however, for  $\beta$ -strand 2 to be in a native conformation in PUF3 while  $\beta$ -strand 1 is unstructured (Figure 6).

In conclusion, although PUFs are useful to illustrate which structural elements of a protein unfold co-operatively, some care is required in the interpretation of their exact conformation, as non-native interactions apparently play a role in stabilising PUFs.

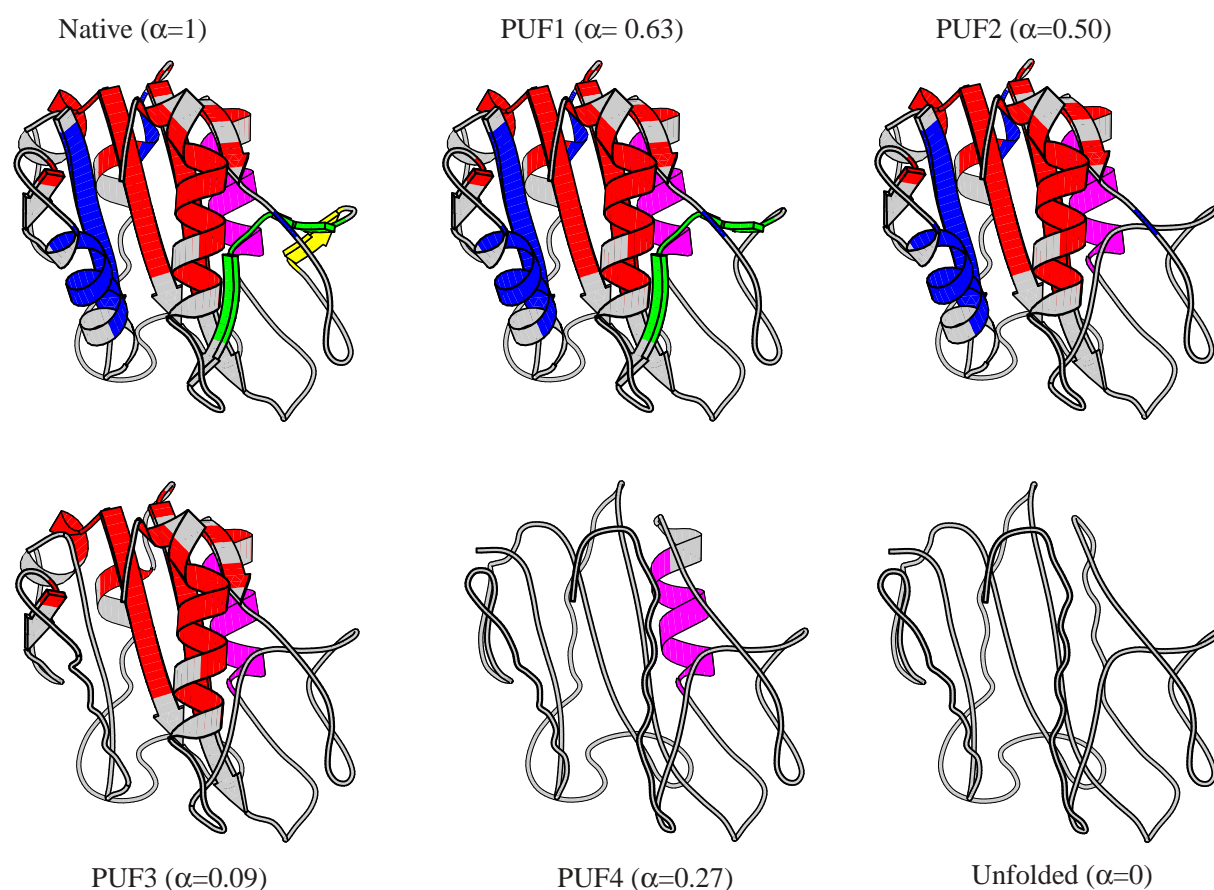
### Comparison of the solvent- and denaturant-accessibilities of the PUFs of *A. vinelandii* apoflavodoxin

The number of residues the amide protons of which exchange from a PUF informs directly about the PUF's overall water accessibility. All residues with a  $\Delta G_{op}$  that is larger than the free energy difference between native apoflavodoxin and a specific PUF, i.e.  $\Delta G_{N-PUF}$ , are protected against amide proton exchange in this specific PUF and are thus inaccessible to water. Residues with  $\Delta G_{op}$  values that equal  $\Delta G_{N-PUF}$  are water accessible in this specific PUF as their amide proton exchange behaviour defines this specific PUF. Are residues that have  $\Delta G_{op}$  values that are lower than a specific  $\Delta G_{N-PUF}$  water accessible in that specific PUF? This question can not be resolved, since the amide protons of the corresponding residues can exchange via local opening processes in native apoflavodoxin or from PUFs that populate more often than the specific PUF involved. For example, residues of Cluster 2 may be folded in PUF4 but remain undetected to be so by H/D exchange. The contribution of H/D exchange from PUF4 to the overall exchange rate of these amide protons is negligible compared to the much more rapid H/D exchange of these amide protons from PUF2, provided the difference in free energy between PUF2 and PUF4 is sufficiently large.

The  $m$ -value of a PUF informs about the PUF's denaturant accessibility (Myers et al. 1995). It represents the dependence of the free energy difference between the PUF and the native state (i.e.  $\Delta G_{N-PUF}$ ) on the denaturant concentration. The denaturant accessibility of a PUF is expressed in the  $\alpha$  factor, which is the normalised  $m$ -value of a certain PUF, i.e.  $(m_{NU} - m_{PUF})/m_{NU}$ , with  $m_{NU}$  being the  $m$ -value associated with global unfolding of apoflavodoxin (Chapter 2 of this thesis). In case  $\alpha = 0$ , the corresponding PUF is as denaturant accessible as fully unfolded apoflavodoxin, and in case  $\alpha = 1$  it is as denaturant accessible as native apoflavodoxin. The  $m$ -values of the five co-operatively unfolding clusters of apoflavodoxin are extracted from the denaturant-dependence of  $\Delta G_{op}$  as measured in the native-state H/D exchange experiments (Figures 2, 4 and 5) and subsequently the corresponding  $\alpha$ -factors are calculated. The  $\alpha$ -value of each PUF of apoflavodoxin is shown in Figure 6. The comparison of the conformations of the PUFs in Figure 6 with the corresponding  $\alpha$ -values shows a striking difference between the water accessibility of the PUFs (i.e. the less structured the more water accessible) and their denaturant accessibility. Whereas PUF1, PUF2, and PUF3

seem to be more accessible to denaturant than to water, PUF4 appears to be more accessible to water than to denaturant.

PUF1 has an  $\alpha$  of 0.63, which appears to be too small as only one of the two strands of the small beta structure in the loop connecting  $\beta$ -strands 5A and 5B in native apoflavodoxin is detected to be unfolded and D<sub>2</sub>O accessible in this PUF. PUF2 has a lower  $\alpha$ -value than PUF1 (i.e.  $\alpha = 0.50$  compared to  $\alpha = 0.63$ ). This is to be expected as PUF2 is less folded than PUF1 (Figure 6). However, the  $\alpha$ -value associated with PUF2 is relatively small regarding that more than 80 % of the observed amide protons are shielded from water in PUF2. Remarkably, PUF3 has an  $\alpha$ -value of only 0.09 (Figure 6). It suggests that 91 % of PUF3 of apoflavodoxin is as accessible for denaturant as the globally unfolded state of apoflavodoxin is. However, this contradicts with 3 out of 5  $\beta$ -strands and 4 out of 5  $\alpha$ -helices of native apoflavodoxin having residues that are protected against amide proton exchange



**Figure 6.** Partially unfolded forms (PUFs) of *A. vinelandii* apoflavodoxin created on basis of the native state H/D exchange data of the protein. The Molscript (Kraulis 1991) cartoon drawings of the PUFs of *A. vinelandii* apoflavodoxin are presented in increasing order of stability difference with respect to native apoflavodoxin (i.e. PUF1 to PUF4). Going from native apoflavodoxin to its unfolded state, clusters of residues are supposed to co-operatively unfold one by one in the order of this increasing stability difference. Elements of secondary structure that are structured in a specific PUF are drawn in a cartoon fashion, whereas elements of secondary structure that are unfolded and water accessible in a specific PUF are drawn as grey coils. Secondary structure elements are assumed to unfold co-operatively unless data opposing this exist. Residues protected against amide proton exchange in a specific PUF are coloured yellow (Cluster 1), green (Cluster 2), blue (Cluster 3), red (Cluster 4), and purple (Cluster 5). The  $\alpha$ -values (i.e. the normalised denaturant accessibility of a protein structure;  $\alpha=1$  means native-like accessibility;  $\alpha=0$  means as accessible as the unfolded state) of the individual PUFs are given in brackets (see main text).



with D<sub>2</sub>O in PUF3. At least 32 residues in total, i.e. 18 % of the sequence of apoflavodoxin, are protected against amide proton exchange in PUF3.

Apparently, PUF1, PUF2, and PUF3 are more accessible to denaturant than to water. One possible explanation for this is that loop regions in native apoflavodoxin are water-accessible but not denaturant-accessible. Most loops in apoflavodoxin have low local stabilities and high associated amide proton exchange rate constants and thus are not observed in the native state H/D exchange experiments. These loop regions may participate in the same subglobal unfolding step that characterises a specific PUF and become denaturant-accessible in this specific PUF. This could cause the low  $\alpha$ -values detected. Another explanation for the apparent discrepancy between  $\alpha$ -values and conformations of PUF1, PUF2 and PUF3 may be the conformational disruption of stable parts of apoflavodoxin in the subglobal unfolding step leading to one of these PUFs, which makes these parts denaturant accessible in the respective PUFs. These apoflavodoxin parts, however, need to remain protected against amide proton exchange in the respective PUF (as they are not detected to belong to the unfolded, exchange competent part of the specific PUF) until a more drastic disruption of the partially unfolded apoflavodoxin molecule occurs.

PUF4 is associated with an  $\alpha$ -value of 0.27, which is somewhat high regarding that only six residues of a single  $\alpha$ -helix are detected to be protected against amide proton exchange in PUF4. While the other three PUFs appear to be more accessible to denaturant than to water, PUF4 displays the opposite behaviour. Perhaps additional residues compared to the ones shown in Figure 6 are also protected against amide proton exchange in PUF4. These additional residues may remain undetected in PUF4 because the amide protons of these residues exchange more rapidly than amide protons belonging to Cluster 4. This can be caused by exchange taking place from another PUF with a lower  $\Delta G_{N-PUF}$  than the one that characterises PUF4. Another source for the  $\Delta G_{op}$ -values for these additional residues being lower than  $\Delta G_{N-PUF4}$ , can be H/D exchange from locally open states of native apoflavodoxin.

### **Why is the stable folding intermediate of *A. vinelandii* apoflavodoxin not observed by native state H/D exchange?**

The off-pathway folding intermediate  $I_1$  (equation 8) populates heavily during apoflavodoxin kinetic refolding, and also populates in apoflavodoxin equilibrium unfolding studies (Chapter 2 of this thesis). Next to native apoflavodoxin,  $I_1$  is the most populated species up to 1.50 M GuDCl. Remarkably, none of the four apoflavodoxin PUFs identified has a stability and  $m$ -value that matches those of  $I_1$  (i.e.  $\Delta G_{NI_1} = 6.70 \pm 0.17$  kcal/mol;  $m_{NI_1} = -4.40 \pm 0.11$  kcal/molM<sup>-1</sup> (Chapter 2 of this thesis)). The majority of the residues of apoflavodoxin that reside in secondary structure elements exchange either from PUF3, PUF4, or U. Although  $I_1$  populates more often than PUF3, PUF4 and globally unfolded apoflavodoxin do (i.e.  $\Delta G_{NI_1} < \Delta G_{N-PUF3} < \Delta G_{N-PUF4}$ ),  $I_1$  is not detected as a folding intermediate by amide proton H/D exchange. Apparently, most observable amide protons of the secondary structure elements of apoflavodoxin are similarly protected against H/D exchange in the molten globule-like folding intermediate  $I_1$  as they are in native

apoflavodoxin. This is the case, despite that the change in denaturant accessibility upon unfolding of I<sub>1</sub> is only 29 % of that of native apoflavodoxin (Chapter 2 of this thesis).

The exchange behaviour observed for the molten globule of apoflavodoxin is consistent with the one observed for the molten globule of human carbonic anhydrase: both the molten globule and the native state of human carbonic anhydrase are similarly protected against H/D exchange for all of the observable amide protons (Kjellsson et al. 2003). Water seems not to efficiently penetrate the interior of the molten globule. However, backbone amide protons of a molten globule exchange more rapidly than those of the native state of a protein as a molten globule globally unfolds more often than native protein due to its low stability compared to the native state. Denisov et al. (Denisov et al. 1999) conclude from water <sup>17</sup>O relaxation dispersion studies that molten globule proteins preserve most of the native internal hydration sites and have native-like surface hydration. Our observations together with those of Kjellsson et al. (Kjellsson et al. 2003) and Denisov et al. (Denisov et al. 1999) substantiate the conclusion that molten globules are similarly protected against H/D exchange than the native state is.

### **Rates for structural opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of *A. vinelandii* apoflavodoxin at the amino acid level determined from pD dependent exchange rates**

Equation 2 describes the pD-dependence of the amide proton exchange rate constant  $k_{ex}$ . The logarithm of  $k_{ex}$  depends linearly on pD in the EX2 pD regime with a slope of 1, and curves at higher pD values towards a horizontal line where  $\log(k_{ex})$  equals  $\log(k_{op})$  (i.e. the EX1 pD regime). Equation 2, together with  $k_{int}$  values calculated at different pD values according to (Bai et al. 1993), is used to fit the pD dependent backbone amide H/D exchange rate constants of apoflavodoxin, as illustrated for Ile51 in Figure 7A. In this manner, values for the rates of local opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of apoflavodoxin at the level of individual amino acid residues are determined (Table 2).

In case of EX2 behaviour over the entire pD range studied and  $\log(k_{ex})$  does not curve as a function of pD, no absolute value for the opening and closing rate constants can be determined for the residues involved. However, the data do allow for an estimation of the lower limits of  $k_{op}$  and  $k_{cl}$ . The opening rate constant for these residues must be larger than the exchange rate constant measured at the highest pD used, otherwise a curvature would have been observed in the pD-dependent  $\log(k_{ex})$  data. As under EX2 conditions the ratio of  $k_{op}$  and  $k_{cl}$  is defined by  $\Delta G_{op}$ , the lower limit for  $k_{op}$  automatically determines the lower limit for  $k_{cl}$  as well.

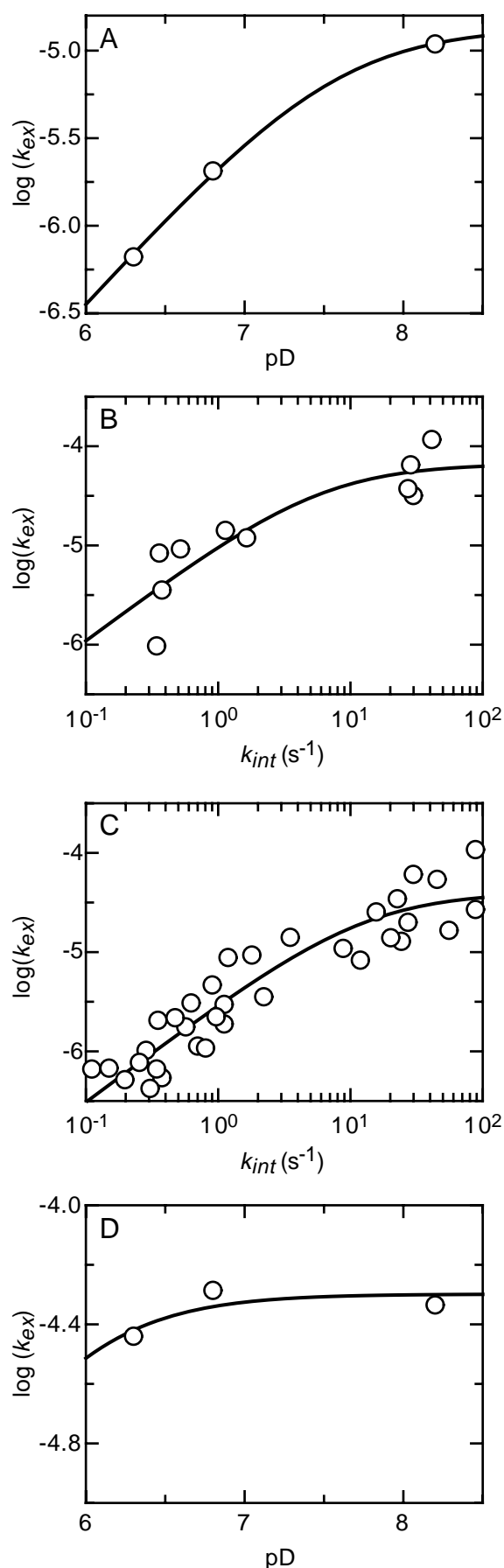
#### *a. Rates for local opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of apoflavodoxin*

In case of amide protons that exchange via local opening processes in absence of denaturant, the  $k_{op}$  and  $k_{cl}$  values determined inform about the local dynamics within the protein at the site of the residue of interest. A large number of these residues (i.e. Val17, Met30, Ala45, Tyr47, Phe49, Leu50, Glu82, Leu84,

**Table 2. Calculated rate constants for local opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of apoflavodoxin in the absence of denaturant.**

residue	cluster	$k_{op}$	$k_{cl}$
<b>5</b>	<b>3</b>	<b><math>(7 \pm 4) 10^{-5}</math></b>	<b><math>6 \pm 4</math></b>
<b>6</b>	<b>3</b>	<b><math>(7 \pm 4) 10^{-5}</math></b>	<b><math>6 \pm 4</math></b>
17	3	$> 6 10^{-3}$	$> 9$
18	3	$(7 \pm 3) 10^{-4}$	$22 \pm 14$
19	3	$(7 \pm 3) 10^{-4}$	$22 \pm 14$
<b>21</b>	<b>3</b>	<b><math>(7 \pm 4) 10^{-5}</math></b>	<b><math>6 \pm 4</math></b>
<b>22</b>	<b>3</b>	<b><math>(7 \pm 4) 10^{-5}</math></b>	<b><math>6 \pm 4</math></b>
25		$(5 \pm 1) 10^{-5}$	$0.3 \pm 0.3$
30	3	$> 3 10^{-3}$	$> 50$
47	4	$> 2.6 10^{-3}$	$> 8$
49	4	$> 1 10^{-3}$	$> 70$
50	4	$> 7.5 10^{-5}$	$> 180$
<b>51</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>52</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>53</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
81	4	$(9 \pm 4) 10^{-5}$	$10 \pm 5$
82	3	$> 8 10^{-4}$	$> 2.7$
84	3	$> 8 10^{-4}$	$> 5.7$
87	4		
89	4	$> 3 10^{-4}$	$> 36$
<b>91</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>92</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>93</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>94</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
114	5	$> 3 10^{-5}$	$> 38$
115	5	$> 8 10^{-4}$	$> 355$
116	5	$> 2 10^{-4}$	$> 175$
120	4	$(5 \pm 2) 10^{-4}$	$54 \pm 27$
123	4	$(5 \pm 2) 10^{-4}$	$54 \pm 27$
125	4	$(5 \pm 2) 10^{-4}$	$54 \pm 27$
164	4	$> 3 10^{-4}$	$> 4.5$
<b>167</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>168</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>169</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>170</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>171</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>
<b>172</b>	<b>4</b>	<b><math>(4 \pm 1) 10^{-5}</math></b>	<b><math>13 \pm 4</math></b>

Rates are in  $s^{-1}$ . For backbone amide protons that exchange too rapidly to be followed by NMR spectroscopy at high pD values, no  $k_{op}$ - and  $k_{cl}$ -values are reported as they could not be determined. The residues for which a subglobal unfolding reaction is the dominant mechanism for exchange of their amide proton down to 0 M GuHCl are indicated in bold. Experiments were done at 25 °C in 100 mM potassium pyrophosphate at pD 6.28, 6.80, and 8.23.



Lys89, Tyr114, Ser115, Phe116, Val164) display EX2 behaviour over the entire pD range studied. This means that the corresponding local dynamics are fast compared to the intrinsic exchange rate constants, which are at the highest pD value studied approximately between 10 and 100 s<sup>-1</sup>.

Interestingly, some residues that exchange via local unfolding in the absence of denaturant (i.e. Ala18, Lys19, Ile81, Arg120, Lys123, Val125) have a curved pD dependence of the logarithm of the exchange rate constants of their amide protons. In these cases, rate constants for local opening and closing can be determined (Table 2). The local closing rate constants for these residues must be relatively small, since for EX1 exchange behaviour to occur the closing rate constants

**Figure 7.** Dependence of the H/D exchange rate constants of A. *vinelandii* apoflavodoxin on pD. (A) The H/D exchange rate constants of the backbone amide proton of Ile51 curve from EX2 to EX1 behaviour upon increasing the pD. (B) The exchange rate constants of the backbone amide protons of residues of the subglobally unfolding Cluster 3 that have a linear dependence of  $\Delta G_{op}$  on the concentration of GuDCI (i.e. residues Ala4, Leu5, Phe6, Ile21, Lys22) display a pD-dependent curvature from EX2 to EX1 behaviour in the pD range studied. The exchange rate constants determined by NMR spectroscopy ( $k_{ex}$ ) are plotted against their intrinsic exchange rate constants ( $k_{int}$ ) to make the curves of all residues of Cluster 3 coincide. (C) The backbone amide proton exchange rate constants of residues of the subglobally unfolding Cluster 4 that have a linear dependence of  $\Delta G_{op}$  on the concentration of GuDCI (i.e. residues Ile51, Leu52, Gly53, Val91, Ala92, Leu93, Phe94, Trp167, Leu168, Ala169, Gln170, Ile171, Ala172) display a pD-dependent curvature from EX2 to EX1 behaviour in the pD range studied. Solid lines in all panels are the best fit of equation 2 to the data shown. The  $k_{op}$  and  $k_{cl}$  values obtained from the fits are given in the main text and in Table 2. (D) Phe25 is the only residue in apoflavodoxin the backbone amide proton of which displays pure EX1 behaviour down to a pD-value of approximately 6.5. This behaviour is characteristic for exchange occurring from the off-pathway folding intermediate I<sub>1</sub> of A. *vinelandii* apoflavodoxin (see main text).

must be smaller than the intrinsic exchange rate constants (approximately between 10 and 100 s<sup>-1</sup> at pD 8.23). In addition, these residues have quite small opening rate constants ranging between 9 10<sup>-5</sup> and 7 10<sup>-4</sup> s<sup>-1</sup>. Apparently, at these sites in apoflavodoxin slow local motions of the protein structure occur. Residues Ala18 and Lys19 are characterised by common opening and closing rate constants, as is also the case for Arg120, Lys123 and Val125 (Table 2), which illustrates that some co-operativity is involved in the corresponding local opening processes.

*b. Co-operative subglobal unfolding and corresponding folding rates within apoflavodoxin*

Each PUF of apoflavodoxin is characterised by a specific unfolding rate with which it is formed starting from native protein, and by a specific folding rate with which it folds back to native apoflavodoxin. The rate constants with which a specific PUF and native apoflavodoxin interconvert can be determined from the opening and closing rate constants of the residues the amide protons of which exchange from this specific PUF in absence of denaturant. These residues are shown in bold in Table 1. Equation 2 is globally fitted to the pD-dependent H/D exchange rate constants of all residues of a specific cluster that are shown in bold in Table 2, while each residue has its own individual  $k_{int}$  (Figure 7 C and D). The thus extracted rate constants for subglobal opening and closing are listed in bold in Table 2.

Unfortunately, backbone amide protons belonging to the cores of Clusters 1 and 2 of apoflavodoxin (i.e. Val141, Phe146, Val147, Gly148, Leu149, Ala150) which can be detected at pD 6.28 exchange too rapidly to be followed at higher pD values. As the backbone amide proton exchange of these residues apparently increases drastically with increasing pD, EX2 behaviour is likely. Consequently, the closing rate constants of these residues need to be fast compared to the intrinsic exchange rate constants, which are at the highest pD value studied approximately between 10 and 100 s<sup>-1</sup>.

In case of Cluster 3 of apoflavodoxin, the pD dependence of the backbone amide proton exchange could be determined for four of its core residues (i.e. Leu5, Phe6, Ile21, and Lys22). The fit of equation 2 to the corresponding data leads to  $k_{op} = (7 \pm 3) 10^{-5} \text{ s}^{-1}$  and  $k_{cl} = 6 \pm 4 \text{ s}^{-1}$  for Cluster 3 (Figure 7C). Native apoflavodoxin and PUF3 thus interconvert with the latter rate constants.

Similarly, in case of residues belonging to Cluster 4 of apoflavodoxin, the following opening and closing rate constants are determined:  $k_{op} = (4 \pm 1) 10^{-5} \text{ s}^{-1}$  and  $k_{cl} = 13 \pm 4 \text{ s}^{-1}$  (Figure 7D), which are the rate constants that characterise the interconversion between native apoflavodoxin and PUF4.

As discussed, the amide protons of residues that belong to Cluster 5 exchange only from the globally unfolded state of apoflavodoxin in the presence of GuDCl concentrations exceeding 0.5 M. Consequently, the global apoflavodoxin folding and unfolding rate constants can not be determined from the pD-dependent H/D exchange results presented here, as they are obtained in absence of denaturant.

### Positioning of the PUFs within the kinetic scheme for *A. vinelandii* apoflavodoxin folding

In the preceding section, rate constants for interconversion between native apoflavodoxin and its PUFs are determined. Comparison of these rate constants with the apoflavodoxin folding and unfolding rate constants extracted from stopped-flow denaturant-induced kinetic folding experiments can identify whether PUFs coincide with the kinetically important folding intermediates  $I_1$  and  $I_2$  (equation 8).

#### *a. Are the identified PUFs on the direct apoflavodoxin folding route?*

Unfolded apoflavodoxin (U) forms native apoflavodoxin (N) with a rate constant of  $7.3 \cdot 10^4 \text{ s}^{-1}$ , and N goes to U with an unfolding rate constant of  $0.013 \text{ s}^{-1}$ , respectively. This has been calculated by using the experimentally determined rate constants shown in equation 8, which are extracted from both kinetic and equilibrium (un)folding experiments (Chapter 2 of this thesis), and by applying equation 9:

$$k_{NU} = \frac{k_{NI_2} k_{I_2U}}{k_{I_2U} + k_{I_2N}} \quad k_{UN} = \frac{k_{UI_2} k_{I_2N}}{k_{I_2U} + k_{I_2N}} \quad (9)$$

Any folding intermediate that is positioned on the direct apoflavodoxin folding route must form native protein with a rate constant that is the same or higher than  $k_{UN}$ . Consequently, backbone amide protons that exchange from U, from  $I_2$  or from any other potential apoflavodoxin folding intermediate on the productive folding route between N and U are expected to display EX2 exchange behaviour within the entire pD range studied (i.e. between 6.3 and 8.2). In case of the residues of Cluster 1 and 2 of apoflavodoxin, the backbone amide proton exchange is too rapid to be followed above pD 6.28. As discussed, EX2 behaviour seems likely for the amide protons involved, and thus both PUF1 and PUF2 could be on the productive folding route between U and N. However, PUF1 ( $\alpha = 0.63$ ) and PUF2 ( $\alpha = 0.50$ ) are too accessible for denaturant to coincide with the kinetic apoflavodoxin folding intermediate  $I_2$  (the  $\alpha$  of which falls somewhere between 0.8 and 0.9 (Chapter 2 of this thesis)). None of the backbone amide protons that belong to Clusters 3 and 4 of apoflavodoxin display pure EX2 H/D exchange behaviour over the entire pD range studied. Thus PUF3 and PUF4 of apoflavodoxin are not on the productive folding route between U and N.

#### *b. Comparison of the rate constants with which PUFs and $I_1$ interconvert with native apoflavodoxin*

The off-pathway intermediate  $I_1$  forms native apoflavodoxin (N) with a rate constant of  $0.73 \text{ s}^{-1}$ , and N goes to  $I_1$  with an unfolding rate constant of  $7.4 \cdot 10^{-5} \text{ s}^{-1}$ , respectively. This has been calculated by again using the experimentally determined folding rate constants shown in equation 8 (Chapter 2 of this thesis) and by applying equation 10:

$$k_{N I_1} = \frac{k_{N I_2} k_{I_2 U} k_{U I_1}}{(k_{I_2 U} + k_{I_2 N})(k_{U I_1} + k_{U I_2})} \quad k_{I_1 N} = \frac{k_{I_1 U} k_{U I_2} k_{I_2 N}}{(k_{I_2 U} + k_{I_2 N})(k_{U I_1} + k_{U I_2})} \quad (10)$$

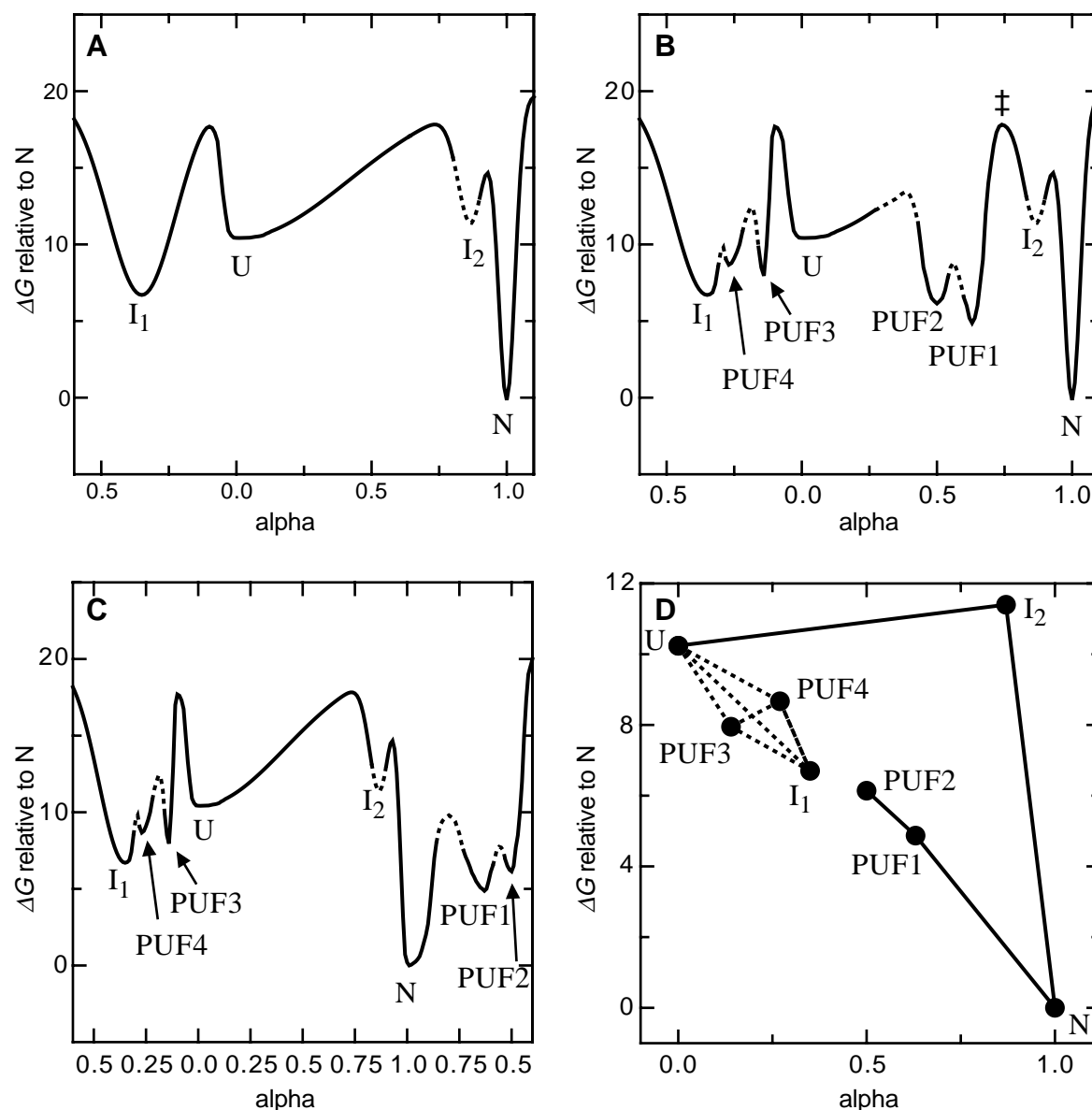
The  $I_1$  to N folding (i.e. closing) rate constant is such that amide protons that can exchange from  $I_1$  will mainly do so according to an EX1 mechanism in the pD range studied here. However, such exchange behaviour is observed for only a single amino acid residue of apoflavodoxin: Phe25 (Figure 7B). Thus the pD dependent H/D exchange data confirm our conclusion based on the denaturant concentration dependent H/D exchange data that most amide protons are similarly protected against H/D exchange in the molten globule  $I_1$  as they are in native apoflavodoxin. Fitting of the pD dependent amide proton exchange rate constant of Phe25 leads to  $k_{op} = (5.0 \pm 0.7) 10^{-5} \text{ s}^{-1}$  and  $k_{cl} = 0.3 \pm 0.3 \text{ s}^{-1}$ , values that roughly correspond to the N to  $I_1$  unfolding rate constant of  $(7.4 \pm 0.7) 10^{-5} \text{ s}^{-1}$  and to the  $I_1$  to N folding rate constant of  $0.733 \pm 0.002 \text{ s}^{-1}$ .

Remarkably, the opening rate constants of Cluster 3  $((7 \pm 3) 10^{-5} \text{ s}^{-1})$  and Cluster 4  $((4 \pm 1) 10^{-5} \text{ s}^{-1})$  are within error identical to the macroscopic N to  $I_1$  unfolding rate constant of  $(7.4 \pm 0.7) 10^{-5} \text{ s}^{-1}$ . However, the rate constants for closing of Cluster 3  $(6 \pm 4 \text{ s}^{-1})$  and Cluster 4  $(13 \pm 4 \text{ s}^{-1})$  are higher than the macroscopically determined  $I_1$  to N folding rate constant ( $k_{I_1 N} = 0.733 \pm 0.002 \text{ s}^{-1}$ ). Consequently, the free energy difference of Cluster 4 with respect to native apoflavodoxin calculated from H/D exchange rates ( $\Delta G_{op} = -RT \ln (k_{op}/k_{cl})$ , equation 5) should be larger than the one of Cluster 3, both of which should be larger than the free energy difference between N and  $I_1$  as determined from denaturant-induced equilibrium (un)folding experiments (i.e.  $\Delta G_{N-PUF4} > \Delta G_{N-PUF3} > \Delta G_{N I_1}$ ). The latter is indeed observed. Apparently, the apoflavodoxin folding intermediate  $I_1$ , detected by classical denaturant-induced equilibrium and kinetic (un)folding studies, and Clusters 3 and 4 of apoflavodoxin detected by native state H/D exchange share a single rate limiting step when these species are formed from the native state.

As discussed, both PUF3 and PUF4 have opening and closing rate constants that are inconsistent with them being on a productive folding route between fully unfolded and native apoflavodoxin. Just as equilibrium intermediate  $I_1$ , PUF3 and PUF4 need to be off the productive folding route of apoflavodoxin, and unfold to enable production of native apoflavodoxin.

### Free energy landscape for *A. vinelandii* apoflavodoxin folding

Protein folding can be described by using a free energy landscape model (Bryngelson et al. 1995; Dill and Chan 1997). In this model, an unfolded protein molecule descends along a funnel describing its free energy until it reaches the state with the lowest free energy, which is the native state of a protein. The denaturant-induced equilibrium and kinetic apoflavodoxin (un)folding results have been used to construct a model of the free energy landscape for apoflavodoxin folding, as shown in Figure 8A (Chapter 2 of this thesis). In this model, the reaction co-ordinate is the alpha value (i.e. the ratio of the m-value of a folding species and the m-value of native apoflavodoxin).



**Figure 8.** (A) Schematic representation of the free energy landscape for folding of *A. vinelandii* apoflavodoxin derived from the results obtained from denaturant-induced equilibrium and kinetic (un)folding studies and as presented by (Chapter 2 of this thesis). The horizontal axis represents the reaction co-ordinate, the vertical axis the free energy difference between a particular species and native apoflavodoxin. The horizontal axis scales to denaturant accessibility because it is expressed as the  $\alpha$ -value, i.e. the ratio of the  $m$ -value of a folding species and the  $m$ -value of native apoflavodoxin. U ( $\alpha = 0$ ) and N ( $\alpha = 1$ ) represent unfolded and native apoflavodoxin, respectively; I<sub>1</sub> ( $\alpha = 0.29$ ) and I<sub>2</sub> ( $\alpha$  lies somewhere between 0.8 and 0.9) are the two folding intermediates presented in equation 8. The off-pathway intermediate I<sub>1</sub> is represented on the left-hand side of the unfolded state, whereas both the on-pathway intermediate I<sub>2</sub> and native apoflavodoxin reside on the right-hand side of the unfolded state. The heights of the barriers  $\Delta G_{op}^\ddagger$  between the individual species are calculated from the opening (i.e. unfolding) rate constants  $k_{op}$  according to  $\Delta G_{op}^\ddagger = -RT \ln(k_{op}/k_0)$  using a value for  $k_0$  of  $10^8$  (Krieger et al. 2003). The depth of the minimum in the free energy landscape in which I<sub>2</sub> resides is unknown, and is therefore represented by a dashed line. (B) Schematic representation of the free energy landscape for apoflavodoxin folding as shown in (A), but now the partially unfolded forms of apoflavodoxin detected by native state H/D exchange are included. PUF3 ( $\alpha = 0.09$ ) and PUF4 ( $\alpha = 0.27$ ) have within error the same  $\Delta G_{op}^\ddagger$  as I<sub>1</sub>, according to their opening rate constants as determined from pD dependent H/D exchange data. The



heights of the free energy barriers between these three species are unknown, and are therefore represented by dashed lines. PUF1 ( $\alpha = 0.63$ ) and PUF2 ( $\alpha = 0.50$ ) are assumed to be on the productive folding route. The height and position of the transition states that are kinetically important are determined accurately and are shown by continuous lines. The other barriers are created such that they do not influence the observed stopped-flow folding and unfolding kinetics; otherwise their position and height is chosen arbitrarily as is shown by dashed lines. Note that the energy barrier between PUF1 and the transition state marked with ‡ is very large (12.95 kcal/mol) and causes the PUF1 to N folding rate constant to be  $0.029\text{ s}^{-1}$ . This would be the rate-limiting step for folding, and would cause EX1 amide proton exchange behaviour of all residues the amides of which exchange from a partially or entirely unfolded apoflavodoxin molecule. This is not observed to be the case. (C) Schematic representation of the free energy landscape for apoflavodoxin folding in which PUF1 and PUF2 are off the direct folding route between U and N and reside on a separate unfolding excursion from N. As in panel B, all free energy barriers that are not present in panel A are chosen such that they do not influence the observed stopped-flow (un)folding kinetics. Note that on the right-hand side of N ( $\alpha = 1$ ) the  $\alpha$ -values decrease. (D) Schematic representation of the apoflavodoxin folding trajectory in which PUF1 and PUF2 are positioned on an unfolding excursion starting from N, as is also the case in panel C. Note that the denaturant accessibility axis is differently presented as compared to panels A to C; it now starts at  $\alpha=0$  and ends at  $\alpha=1$ . Observed transitions between two species are indicated by connecting them with solid lines. The productive folding route from U to N is via the high-energy intermediate  $I_2$  (Chapter 2 of this thesis). The non-productive folding route starting from U and the non-productive unfolding route starting from N end up in species  $I_1$  and PUF2, respectively. Both species have similar free energies and denaturant accessibilities. The dashed lines reflect that no information is available about how U, PUF3, PUF4 and  $I_1$  interconvert with one another.

The presented H/D exchange data can be incorporated into the latter free energy landscape model. The off-pathway intermediate  $I_1$  has to unfold before native apoflavodoxin can be formed and the unfolding of  $I_1$  to U is the rate-limiting step in the production of native apoflavodoxin. As discussed, PUF3 and PUF4 can not reside on the productive folding route between native and unfolded apoflavodoxin. Additionally, the rate constant with which native apoflavodoxin forms  $I_1$  as extracted from stopped-flow experiments (Chapter 2 of this thesis) and the opening rate constants of Clusters 3 and 4 extracted from native state H/D exchange experiments coincide within error. The most simple explanation for this observation is that in the proposed two-dimensional free energy landscape model PUF3 ( $\alpha = 0.09$ ), PUF4 ( $\alpha = 0.27$ ) and  $I_1$  ( $\alpha = 0.29$ ) reside in a free energy region that is separated by a high free energy barrier from unfolded apoflavodoxin (Figure 8B). Both PUF3 and PUF4, as well as  $I_1$ , need to unfold before the productive folding route to native apoflavodoxin can be followed. No information is available about how these three species transform into one another. The heights of the free energy barriers that separate them are set arbitrarily in Figure 8B.

As discussed, the closing rate constants of PUF1 and PUF2 are consistent with these two PUFs being folding intermediates on the productive folding route. However, as denaturant accessibilities show, PUF1 ( $\alpha = 0.63$ ) and PUF2 ( $\alpha = 0.50$ ) do not coincide with the apoflavodoxin kinetic folding intermediate  $I_2$  ( $\alpha$  between 0.8 and 0.9 (Chapter 2 of this thesis)). Figure 8B shows the free energy landscape for apoflavodoxin folding with PUF1 and PUF2 positioned on the productive folding route in such a way to accord for their free energies and denaturant accessibilities. Both PUF1 and PUF2 are significantly more denaturant accessible than the rate-limiting transition state for folding ( $\alpha = 0.73$  (Chapter 2 of this thesis)), marked by ‡ in Figure 8B. As a consequence, the relatively stable PUF1 is on

the unfolded side of the transition state ‡. The energy barrier between PUF1 and ‡ is then 12.95 kcal/mol, which theoretically leads to a PUF1 to N folding rate constant of 0.029 s<sup>-1</sup>. This process would be the rate-limiting step in apoflavodoxin folding, and it is so slow that almost all residues that exchange from a PUF of apoflavodoxin or from unfolded apoflavodoxin would display pure EX1 exchange behaviour at the pD values studied. This is in contradiction with all apoflavodoxin kinetic and H/D exchange data discussed. Thus, PUF1 and PUF2 cannot be on the productive folding route of apoflavodoxin.

PUF1 and PUF2 must be excursions from the native state that are separated from the main (un)folding route that links native and globally unfolded apoflavodoxin. Both PUF1 and PUF2 are not necessarily positioned on the same unfolding route, but could also be two separate excursions from the native state. However, for reasons of simplicity, PUF1 and PUF2 are positioned on a single unfolding trajectory in Figure 8C. This unfolding trajectory is not the one that leads to unfolded apoflavodoxin as is represented in Figure 8C by  $\alpha$ -values that decrease again on the right-hand side of N. The energy barriers that separate PUF1 and PUF2 from native apoflavodoxin and from each other are chosen arbitrarily in such a way that they are sufficiently low not to influence the observed stopped-flow (un)folding kinetics.

The schematic energy landscape presented in Figure 8C is consistent with the experimental apoflavodoxin equilibrium and kinetic folding data (Chapter 2 of this thesis) as well as with the H/D exchange data presented here, although the linear two-dimensional model is an oversimplification of reality. Considered from a conformational point of view (Figure 6) it makes sense that PUF2 develops from PUF1, although there is no concrete evidence that supports this assumption. The conformations of PUF3 and PUF4 (Figure 6) however, are in conflict with the two-dimensional model, which suggests that on the non-productive folding route to I<sub>1</sub> unfolded apoflavodoxin first forms PUF3, then PUF4 and finally I<sub>1</sub>. This would mean that unfolded apoflavodoxin would fold to the relatively structured PUF3, then unfold to form PUF4 and then needs to fold to form the molten-globule like species I<sub>1</sub>. However, no information is available on how these three species interconvert with one another. Furthermore, the conformations of PUF3 and PUF4 presented in Figure 6 are probably incomplete or may be even incorrect, as is indicated by the discrepancies between their solvent- and denaturant accessibilities discussed in a previous section. As PUF3 and PUF4 are on a non-productive folding route, they may be misfolded species that contain non-natively structured parts.

In contrast to common belief, none of the identified PUFs of *A. vinelandii* apoflavodoxin appears to be positioned on the productive folding route between unfolded and native protein. The corresponding apoflavodoxin folding trajectories in terms of free energies and denaturant accessibilities are shown in Figure 8D. Note that the off-pathway folding intermediate I<sub>1</sub> and PUF2 are quite close to one another in the latter representation. Figure 8D suggests that a direct folding route between U and N via I<sub>1</sub> and all PUFs identified should be possible. However, as the apoflavodoxin equilibrium and kinetic data show, such a route is not significantly populated (Chapter 2 of this thesis). The latter must be caused by a high energy barrier between I<sub>1</sub> and PUF2. As a consequence, apoflavodoxin folds via the high-energy intermediate I<sub>2</sub>.

In conclusion, native state H/D exchange experiments and classical denaturant-induced equilibrium and kinetic folding experiments are complementary. Together they provide a detailed picture of the free energy landscape for protein folding, as is shown here for *A. vinelandii* apoflavodoxin folding. None of the PUFs detected for the 179-residue apoflavodoxin appear to be on the productive folding route, which starts with unfolded protein and ends with native apoflavodoxin. The accompanying energy landscape for apoflavodoxin folding is complex and includes numerous minima and maxima. Complex free energy landscapes for protein folding may be typical for the folding of large proteins (> 100 amino acids).

## References

- Arrington, C. B. and A. D. Robertson (2000). "Microsecond to minute dynamics revealed by EX1-type hydrogen exchange at nearly every backbone hydrogen bond in a native protein." *J Mol Biol* **296**(5): 1307-17.
- Bai, Y., J. S. Milne, L. Mayne and S. W. Englander (1993). "Primary structure effects on peptide group hydrogen exchange." *Proteins* **17**(1): 75-86.
- Bai, Y., T. R. Sosnick, L. Mayne and S. W. Englander (1995). "Protein folding intermediates: native-state hydrogen exchange." *Science* **269**(5221): 192-7.
- Bryngelson, J. D., J. N. Onuchic, N. D. Socci and P. G. Wolynes (1995). "Funnels, pathways, and the energy landscape of protein folding: a synthesis." *Proteins* **21**(3): 167-95.
- Canet, D., A. M. Last, P. Tito, M. Sunde, A. Spencer, D. B. Archer, C. Redfield, C. V. Robinson and C. M. Dobson (2002). "Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme." *Nat Struct Biol* **9**(4): 308-15.
- Chamberlain, A. K., T. M. Handel and S. Marqusee (1996). "Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH." *Nat Struct Biol* **3**(9): 782-7.
- Chu, R., W. Pei, J. Takei and Y. Bai (2002). "Relationship between the native-state hydrogen exchange and folding pathways of a four-helix bundle protein." *Biochemistry* **41**(25): 7998-8003.
- Denisov, V. P., B. H. Jonsson and B. Halle (1999). "Hydration of denatured and molten globule proteins." *Nat Struct Biol* **6**(3): 253-60.
- Dill, K. A. and H. S. Chan (1997). "From Levinthal to pathways to funnels." *Nat Struct Biol* **4**(1): 10-9.
- Edmondson, D. E. and G. Tollin (1971). "Chemical and physical characterization of the Shethna flavoprotein and apoprotein and kinetics and thermodynamics of flavin analog binding to the apoprotein." *Biochemistry* **10**(1): 124-32.
- Englander, S. W., L. Mayne and J. N. Rumbley (2002). "Submolecular cooperativity produces multi-state protein unfolding and refolding." *Biophysical chemistry* **101-102**(1): 57-65.

- Hoang, L., S. Bedard, M. M. Krishna, Y. Lin and S. W. Englander (2002). "Cytochrome c folding pathway: kinetic native-state hydrogen exchange." Proc Natl Acad Sci U S A **99**(19): 12173-8.
- Huyghues-Despointes, B. M., J. M. Scholtz and C. N. Pace (1999). "Protein conformational stabilities can be determined from hydrogen exchange rates." Nat Struct Biol **6**(10): 910-2.
- Hvidt, A. and S. O. Nielsen (1966). "Hydrogen exchange in proteins." Adv Protein Chem **21**: 287-386.
- Kay, L. E., P. Keifer and T. Saarinen (1992). "Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity." J Am Chem Soc **114**: 10663-10665.
- Kjellsson, A., I. Sethson and B. H. Jonsson (2003). "Hydrogen exchange in a large 29 kD protein and characterization of molten globule aggregation by NMR." Biochemistry **42**(2): 363-74.
- Kraulis, P. J. (1991). "MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures." J Appl Cryst **24**: 946-950.
- Krieger, F., B. Fierz, O. Bieri, M. Drewello and T. Kiefhaber (2003). "Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding." J Mol Biol **332**(1): 265-274.
- Mayhew, S. G. and G. Tollin (1992). General properties of flavodoxins. Chemistry and biochemistry of flavoenzymes. F. Müller. Boca Raton, CRC Press. **3**: 389-426.
- Myers, J. K., C. N. Pace and J. M. Scholtz (1995). "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding." Protein Sci **4**(10): 2138-48.
- Palmer III, A. G. <http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/nmr2modelfree.html>.
- Palmer III, A. G., J. Cavanagh, P. E. Wright and M. Rance (1991). "Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy." Journal of Magnetic Resonance **93**: 151-170.
- Parker, M. J. and S. Marqusee (1999). "The cooperativity of burst phase reactions explored." J Mol Biol **293**(5): 1195-210.
- Reimer, U., G. Scherer, M. Drewello, S. Kruber, M. Schutkowski and G. Fischer (1998). "Side-chain effects on peptidyl-prolyl cis/trans isomerisation." J Mol Biol **279**(2): 449-60.
- Sivaraman, T., C. B. Arrington and A. D. Robertson (2001). "Kinetics of unfolding and folding from amide hydrogen exchange in native ubiquitin." Nat Struct Biol **8**(4): 331-3.
- Steensma, E., M. J. Nijman, Y. J. Bollen, P. A. de Jager, W. A. van den Berg, W. M. van Dongen and C. P. van Mierlo (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**(2): 306-17.

- Steensma, E. and C. P. van Mierlo (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**(3): 653-66.
- Thorneley, R. N. F., G. A. Ashby, M. H. Drummond, R. R. Eady, D. L. Hughes, G. Ford, P. M. Harrison, A. Shaw, R. L. Robson, J. Kazlauskaitė and H. A. O. Hill (1994). Flavodoxin and nitrogen fixation: Structure, electrochemistry and posttranslational modification by coenzyme A. *Flavins and flavoproteins 1993*. K. Yagi. Berlin, Walter de Gruyter & Co.: 343-354.
- van Mierlo, C. P., J. M. van den Oever and E. Steensma (2000). "Apoflavodoxin (un)folding followed at the residue level by NMR." *Protein Sci* **9**(1): 145-57.
- van Mierlo, C. P., W. M. van Dongen, F. Vergeldt, W. J. van Berkel and E. Steensma (1998). "The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate." *Protein Sci* **7**(11): 2331-44.
- Woodward, C., I. Simon and E. Tuchsén (1982). "Hydrogen exchange and the dynamic structure of proteins." *Mol Cell Biochem* **48**(3): 135-60.
- Woodward, C. K. and B. D. Hilton (1980). "Hydrogen isotope exchange kinetics of single protons in bovine pancreatic trypsin inhibitor." *Biophys J* **32**(1): 561-75.
- Yan, S., S. D. Kennedy and S. Koide (2002). "Thermodynamic and kinetic exploration of the energy landscape of *Borrelia burgdorferi* OspA by native-state hydrogen exchange." *J Mol Biol* **323**(2): 363-75.



# 5 Last in, first out: the role of cofactor binding in *Azotobacter vinelandii* flavodoxin folding

Yves J.M. Bollen, Willem J.H. van Berkel and Carlo P.M. van Mierlo

Despite that many proteins require the binding of a ligand to be functional, the (kinetic) role of ligand-binding during folding is poorly understood. Here, the influence of the presence of the non-covalently bound flavin mononucleotide (FMN) cofactor on the global stability and on the kinetic folding of *Azotobacter vinelandii* holoflavodoxin are reported. The denaturant-induced equilibrium (un)folding data of holoflavodoxin in the presence and absence of excess FMN are excellently described by a model in which only native apoflavodoxin binds to FMN. As an intermediate  $I_1$  populates during apoflavodoxin equilibrium (un)folding, the holoflavodoxin equilibrium (un)folding model consists of four species: unfolded apoflavodoxin, the apoflavodoxin folding intermediate  $I_1$ , native apoflavodoxin and holoflavodoxin molecules. The folding intermediate  $I_1$  has molten globule-like characteristics. Cofactor binding to apoflavodoxin is shown to affect the protein stability in a theoretically predictable manner. FMN binding to native apoflavodoxin occurs with two kinetically observable rate constants at all denaturant and protein concentrations studied. These two rate constants arise from two conformationally differing apoflavodoxin species, which most likely exist due to the binding of inorganic phosphate to the FMN phosphate binding site of a fraction of the *A. vinelandii* apoflavodoxin molecules. Excess FMN does not accelerate flavodoxin folding, and FMN does not act as a nucleation site for flavodoxin folding. During kinetic folding of holoflavodoxin formation of native apoflavodoxin precedes ligand binding. Even under strongly denaturing conditions, global unfolding of holoflavodoxin occurs only after release of its FMN. The model that describes *A. vinelandii* apoflavodoxin kinetic folding, which includes the stable off-pathway intermediate  $I_1$  and a high-energy on-pathway intermediate  $I_2$ , can now be extended to describe kinetic holoflavodoxin folding:  $I_1 + \text{FMN} \rightleftharpoons \text{unfolded apoflavodoxin} + \text{FMN} \rightleftharpoons I_2 + \text{FMN} \rightleftharpoons \text{native apoflavodoxin} + \text{FMN} \rightleftharpoons \text{holoflavodoxin}$ .

## Introduction

Despite that many proteins require the binding of a non-covalently bound ligand to be functional, the role of ligand-binding during folding is poorly understood. In principle, a ligand could bind to an unfolded protein and reduce the conformational freedom of the polypeptide involved and thereby reduce the conformational space sampled during protein folding. Such bound ligands are potentially able to speed up the folding process by acting as a nucleation site. In some cases, ligands are indeed shown to stay bound to the unfolded protein (Bertini et al. 1997; Robinson et al. 1997). However, in other cases, ligands may not interact with non-native protein states and only become incorporated in the protein during the final stages of protein folding. As kinetic folding studies of proteins in the presence of their ligand are sparse, the kinetic role of ligand binding during protein folding remains unclear.

Flavoproteins offer a good opportunity to study the role of ligand binding during protein folding. The flavin cofactor is generally non-covalently bound and can be reversibly removed by a number of methods (Hefti et al. 2003). Recombination of the obtained apoprotein with flavin can be studied due to the clear change in the fluorescence emission spectrum of the flavin that generally accompanies flavin binding by an apoprotein.

Here, the influence of the presence of the flavin mononucleotide (FMN) cofactor on the folding of *Azotobacter vinelandii* flavodoxin is investigated. Flavodoxins are small monomeric proteins that function as low-potential one-electron carriers, which is made possible by the use of the bound FMN cofactor. Both the denaturant-induced equilibrium and kinetic folding of *A. vinelandii* apoflavodoxin, i.e. flavodoxin in absence of its cofactor, have recently been studied in great detail (van Mierlo et al. 1998; van Mierlo et al. 2000; Chapters 2 and 4 of this thesis). During denaturant-induced equilibrium unfolding experiments, apoflavodoxin populates a relatively stable folding intermediate  $I_1$  with molten globule-like characteristics between 1 and 3 M guanidinium chloride (GuHCl), as expressed in equation 1:



with U unfolded and N native apoflavodoxin. Equation 1 is an equilibrium folding scheme, in which the species involved are organised according to their stability. This scheme does not contain kinetic information about apoflavodoxin folding.

During kinetic *A. vinelandii* apoflavodoxin folding, intermediate  $I_1$  acts as a trap. The intermediate is off the direct folding route between unfolded and native apoflavodoxin, and has to unfold before native apoflavodoxin can be formed. Some folding apoflavodoxin molecules manage to circumvent this trap and fold *via* a direct and rapid route to the native state, on which a second folding intermediate  $I_2$  is located. This second intermediate is highly unstable and never populates to a significant extent, and is thus not observed during denaturant-induced equilibrium unfolding of apoflavodoxin. As a result, apoflavodoxin kinetic folding can be described by the model shown in equation 2 (Chapter 2 of this thesis):





In this study we report on (i) the kinetics of FMN binding to native *A. vinelandii* apoflavodoxin, (ii) the kinetics of FMN binding during *A. vinelandii* holoflavodoxin folding and (iii) the release of FMN from *A. vinelandii* holoflavodoxin. In addition, the influence of FMN binding on the global stability of *A. vinelandii* holoflavodoxin is investigated. By doing so, the role of FMN during *A. vinelandii* flavodoxin folding is deciphered.

## Materials and methods

### Materials

Guanidinium chloride (GuHCl, ultrapure) and potassium pyrophosphate were from Sigma (Bornem, Belgium). In holoflavodoxin equilibrium unfolding experiments, FMN purchased from Sigma was used without further purification. In all other experiments, FMN obtained during the preparation of apoflavodoxin (see below) and purified by reverse-phase HPLC was used.

### Protein expression and purification

The single cysteine residue 69 in wild-type *A. vinelandii* (strain ATCC 478) flavodoxin II was replaced by an alanine to avoid covalent dimerisation of apoflavodoxin. The mutant protein was shown to be similar to wild-type flavodoxin regarding both redox potential of the holoprotein and global stability of the apoprotein (Steensma et al. 1996; van Mierlo et al. 1998). Recombinant *A. vinelandii* C69A holoflavodoxin was obtained and purified as described previously (Steensma et al. 1998; van Mierlo et al. 1998). Apoflavodoxin was subsequently prepared by trichloroacetic acid preparation (Edmondson and Tollin 1971; van Mierlo et al. 1998) followed by gel filtration on a Superdex 200 prep grade column (Pharmacia, Uppsala, Sweden) to remove apoflavodoxin molecules in an oligomeric state (Chapter 2 of this thesis).

### Dissociation constant of the apoflavodoxin-FMN complex

The dissociation constant of the apoflavodoxin-FMN complex was determined using the quenching of FMN fluorescence upon binding to the apoprotein (Mayhew and Wassink 1980). A solution of 1.5 ml containing 210 nM FMN (based on the extinction coefficient at 445 nm of  $12.2 \text{ mM}^{-1}\text{cm}^{-1}$  (Lostao et al. 1997)) in 100 mM potassium pyrophosphate pH 6.0 was titrated with aliquots of 4.12  $\mu\text{M}$  apoflavodoxin in the same buffer. The apoflavodoxin concentration was determined spectrophotometrically using an extinction coefficient at 280 nm of  $29 \text{ mM}^{-1}\text{cm}^{-1}$  (Barman and Tollin 1972). After each addition of protein, the system was allowed to equilibrate for 5 minutes. Subsequently, the FMN fluorescence intensity was determined using a Cary eclipse fluorimeter equipped with a peltier accessory operating at 25 °C (Varian, Palo Alto, CA, USA). Excitation was at 445 nm with a slit of 5 nm; emission was recorded during 1 second at 525 nm with a slit of 10 nm. The dissociation constant of the apoflavodoxin-FMN complex was determined by fitting the fluorescence emission to equation 3, a slightly modified equation compared to the one described by Lostao et al. (1997).

$$F = dF_{end} + F_{\delta} \left( dC_F - \frac{(C_A + K_D + dC_F) - \sqrt{(C_A + K_D + dC_F)^2 - 4C_A dC_F}}{2} \right) \quad (3)$$

where  $F$  is the observed fluorescence intensity after each addition,  $d$  the dilution factor (initial volume/total volume),  $F_{end}$  the remaining fluorescence intensity after the titration (resulting from both fluorescence of holoflavodoxin and of traces of modified FMN that is unable to bind to apoflavodoxin),  $F_{\delta}$  the difference in molar emission intensity between holoflavodoxin and unbound FMN,  $C_F$  the initial concentration of FMN,  $C_A$  the total protein concentration after each addition (i.e. apo + holo), and  $K_D$  the dissociation constant of the apoflavodoxin-FMN complex.

### FMN binding kinetics

Kinetic FMN binding experiments were performed on a BioLogic (Claix, France) SFM-4 stopped-flow machine. Solutions were thermostated at 25 °C using a circulating waterbath. FMN and apoflavodoxin, both in 100 mM potassium pyrophosphate pH 6.0, were mixed in different ratios. Pseudo-first order kinetics were obtained by assuring at least a tenfold excess of apoflavodoxin relative to FMN. The final FMN concentration was 0.10  $\mu$ M in all cases, except for the two highest protein concentrations (i.e. 10 and 12.5  $\mu$ M apoflavodoxin) where an FMN concentration of 1.0  $\mu$ M was used. Binding of FMN to apoflavodoxin was monitored via the accompanying quenching of the FMN fluorescence. The excitation wavelength was 446 nm using an 8 nm slit, emission was recorded above 475 nm. Exponential equations were fitted to the kinetic traces using ProFit (Quantumsoft, Zürich, Switzerland).

FMN binding kinetics were also measured as a function of GuHCl concentration. In this case, FMN, apoflavodoxin and GuHCl, all in 100 mM potassium pyrophosphate pH 6.0, were mixed to a final FMN concentration of 10  $\mu$ M, a final protein concentration of 1.0  $\mu$ M, and GuHCl concentrations varying between 0.20 and 1.0 M. Above 1.0 M GuHCl, no binding experiments were done as at these denaturant concentrations apoflavodoxin populates non-native states to a significant extent (i.e. for more than 2 %), which could potentially complicate the observed binding kinetics. Solutions were thermostated at 25 °C using a circulating waterbath. The kinetic traces were fitted to exponential equations using ProFit (Quantumsoft, Zürich, Switzerland).

### Denaturant-induced equilibrium unfolding of holoflavodoxin

The global stability of holoflavodoxin from *A. vinelandii* was determined via its equilibrium unfolding in GuHCl in 100 mM potassium pyrophosphate pH 6.0 at 25 °C. To obtain a virtually constant FMN concentration throughout the experiment, the equilibrium unfolding was performed in the presence of 100  $\mu$ M excess FMN. Before measurements, all samples were equilibrated for 24 hours at 25 °C in the dark. The protein concentration was 4.0  $\mu$ M.

Steady-state far-UV CD measurements were performed on a Jasco J715 spectropolarimeter (Tokyo, Japan) equipped with a PTC-348WI peltier temperature control system. GuHCl unfolding samples were measured in a 1 mm quartz cuvette (Starna, Hainault, England) at 222 and 255 nm, and averaged over 3 minutes per wavelength at a temperature of 25 °C. The ellipticity at 255 nm was subtracted from the 222 nm ellipticity as a baseline value. During all experiments the cell chamber was purged with dry nitrogen gas at a flow rate of 5 l/min.

Tryptophan fluorescence intensity was determined using a Cary eclipse fluorimeter equipped with a peltier accessory operating at 25 °C (Varian, Palo Alto, CA, USA). Excitation was at 280 nm using a slit of 5 nm; emission was measured for 5 seconds at 350 nm with a slit of 2.5 nm. As the sample has a high optical density due to the presence of excess FMN, the fluorescence emission was recorded using a 0.5 ml cuvette (Helma 115F-QS, Rijswijk, The Netherlands), which was positioned with its long side (1 cm length) perpendicular to the excitation beam, resulting in an excitation path length of 2 mm.

The resulting GuHCl-induced equilibrium unfolding curves as observed by fluorescence- and CD-spectroscopy were fitted to the following set of equations:



$$K_D = [Apo][FMN]/[Holo] \quad (5)$$

$$K_1 = [Apo]/[I_1] \quad K_2 = [I_1]/[U] \quad (6)$$

$$K_i(D) = K_i(0) \exp(-m_i[D]/RT) \quad (7)$$

$$Y_{obs} = \sum (\alpha_j + \beta_j[D])[j] \quad (8)$$

with *Holo* being holo flavodoxin, *Apo* being native apoflavodoxin, *I<sub>1</sub>* the folding intermediate observed during apoflavodoxin equilibrium unfolding (van Mierlo et al. 1998; Chapter 2 of this thesis), and *U* being the unfolded protein. *K<sub>D</sub>* is the dissociation constant of the apoflavodoxin-FMN complex, *K<sub>1</sub>* and *K<sub>2</sub>* describe the equilibrium between *Apo* and *I<sub>1</sub>*, and between *I<sub>1</sub>* and *U*, respectively. The dependence of the equilibrium constants *K<sub>D</sub>*, *K<sub>1</sub>* and *K<sub>2</sub>* on the GuHCl concentration is given by equation 7, in which *[D]* represents the denaturant concentration, *K<sub>i</sub>(0)* is the equilibrium constant (i.e. *K<sub>D</sub>*, *K<sub>1</sub>* or *K<sub>2</sub>*) in absence of denaturant, *m* is a constant of proportionality that describes the denaturant dependence of the equilibrium constant *K*, *R* is the gas constant, and *T* is the absolute temperature. The observed spectroscopic signal *Y* (i.e. fluorescence or CD intensity) is described by equation 8 as the summation of contributions of all four components *j* (i.e. *Holo*, *Apo*, *I<sub>1</sub>* and *U*). The molar spectroscopic property of species *j* in absence of denaturant is *α<sub>j</sub>*, *β<sub>j</sub>* describes the dependence of *α<sub>j</sub>* on the concentration denaturant, and *[j]* is the fractional population of species *j*. The kinetic apoflavodoxin folding intermediate *I<sub>2</sub>* (equation 2) is not incorporated in equation 4 as

it is shown not to populate during denaturant-induced equilibrium unfolding of apoflavodoxin (Chapter 2 of this thesis).

Each equilibrium constant in equation 5 and 6 is related to a free energy difference  $\Delta G$ .

$$\Delta G = -RT \ln(K) \quad (9)$$

The equilibrium between apo- and holoflavodoxin depends on the concentration of free FMN in solution. Consequently, the apparent stability of holoflavodoxin against unfolding  $\Delta G_{unf}^{app}$  depends on the concentration of free FMN, as is expressed in equation 10 (Creighton 1993):

$$\Delta G_{unf}^{app} = \Delta G_{apo} + RT \ln \left( 1 + \frac{[FMN]}{K_D} \right) \quad (10)$$

in which  $\Delta G_{apo}$  is the global stability of apoflavodoxin.

The holoflavodoxin denaturant-induced equilibrium unfolding data were analysed globally together with those of apoflavodoxin taken from (Chapter 2 of this thesis). The dissociation constant  $K_D$  of holoflavodoxin in absence of denaturant is fixed to  $3.4 \cdot 10^{-10} \text{ M}^{-1}$ , the value determined in the present study. The denaturant-induced equilibrium unfolding data of apoflavodoxin are described by equation 4 upon removal of holoflavodoxin and FMN, and by equations 6 to 8. The stabilities of the folding intermediate  $I_1$  and of native apoflavodoxin, as well as the corresponding  $m$ -values, are taken from (Chapter 2 of this thesis). Both unfolded flavodoxin and the flavodoxin folding intermediate  $I_1$  are thus assumed not to interact significantly with FMN. The spectroscopic properties of each of the four protein species in equation 4 are treated as global parameters, which means that the fluorescence and CD signal of native apoflavodoxin, folding intermediate  $I_1$  and of unfolded apoflavodoxin are assumed to be identical in the apoflavodoxin and the holoflavodoxin denaturant-induced unfolding experiments. The dependence of the spectroscopic properties of a specific species (i.e.  $\beta$ -values in equation 8) can be determined accurately only when the species populates for 100 % over a significant range of denaturant concentrations. This is the case for native holoflavodoxin in the holoflavodoxin equilibrium unfolding experiment, for native apoflavodoxin in the apoflavodoxin unfolding experiment, and for unfolded flavodoxin in both unfolding experiments. However, the equilibrium folding intermediate  $I_1$  does not populate for 100 %, and thus the denaturant-dependence of its CD and fluorescence intensity (i.e.  $\beta$ -values) can not be determined. Therefore, the  $\beta$ -values of  $I_1$  are fixed to zero (Chapter 2 of this thesis). The stabilities of native apoflavodoxin and of the folding intermediate  $I_1$ , their corresponding  $m$ -values, and the dissociation constant for FMN release from holoflavodoxin are known (either determined here or in Chapter 2 of this thesis). The only free parameters during the global fit of the denaturant-induced equilibrium (un)folding data of holoflavodoxin are thus the four  $\alpha$ -values, the corresponding  $\beta$ -values (except in case of

$I_1$ ), and the denaturant concentration dependence ( $m$ -value) of the dissociation constant for FMN release from holo flavodoxin.

### Denaturant-dependence of flavodoxin folding kinetics

Folding kinetics of holo flavodoxin were measured by monitoring the FMN binding to folding apo flavodoxin molecules. The following solutions were mixed in a BioLogic SFM-4 stopped-flow machine: (i) FMN, (ii) unfolded apo flavodoxin in 3.0 M GuHCl, and (iii) a GuHCl solution of varying concentration, all three in 100 mM potassium pyrophosphate pH 6.0. The final FMN concentration was 10  $\mu$ M, the final protein concentration was 1.0  $\mu$ M and GuHCl concentrations ranged between 0.3 and 1.2 M. To determine the fluorescence intensity of unbound FMN in the denaturant range studied, a blank experiment was done that is identical to the one described, except that the apo flavodoxin solution in 3.0 M GuHCl is replaced by a solution of 3.0 M GuHCl without apo flavodoxin. Solutions were thermostated at 25 °C using a circulating waterbath. Binding of FMN to apo flavodoxin was monitored via the accompanying quenching of FMN fluorescence, using an excitation wavelength of 446 nm with an excitation slit of 8 nm, and recording emission above 475 nm. Kinetic traces were fitted to exponential equations using the program ProFit.

The influence of Xaa-Pro peptide bond isomerisation on the observed holo flavodoxin folding kinetics was examined by repeating the experiment described above with freshly unfolded apo flavodoxin, which has all Xaa-Pro peptide bonds in the native conformation. This was achieved by using a double-jump stopped flow experiment in which native apo flavodoxin was first unfolded for a period of 600 ms in 3.0 M GuHCl. The resulting freshly unfolded apo flavodoxin solution was subsequently immediately mixed with 100 mM potassium pyrophosphate pH 6.0 containing FMN, to final concentrations of 10  $\mu$ M FMN, 1.0  $\mu$ M apo flavodoxin, and 0.50 M GuHCl.

### Denaturant-dependence of holo flavodoxin unfolding kinetics

The denaturant-dependence of the unfolding kinetics of holo flavodoxin were measured by mixing holo flavodoxin and GuHCl-solutions, both in 100 mM potassium pyrophosphate pH 6.0, in a BioLogic SFM-4 stopped-flow machine to a final protein concentration of 1.0  $\mu$ M and final GuHCl concentrations that range between 3.5 and 5.7 M. Unfolding of holo flavodoxin was monitored by the release of FMN, which results in a strong increase of the FMN fluorescence intensity. Excitation was done at 446 nm using a 8 nm slit, and emission was recorded above 475 nm.

### Models used that describe FMN binding to apo flavodoxin

Two rate constants for FMN binding to apo flavodoxin are observed at all concentrations denaturant used (see Results and Discussion). In general when two rate constants are observed in a kinetic process, three species play a role (Szabo 1969). Three FMN binding models consisting of three species were tested for their capacity to reproduce the GuHCl concentration dependence of both FMN binding rate constants and corresponding amplitudes, the GuHCl concentration dependence of the holo flavodoxin unfolding rate

constant and the equilibrium dissociation constant of flavodoxin in absence of GuHCl. Rate constants for FMN binding were assumed to depend exponentially on the GuHCl concentration, similar to as is observed for the rate constants for protein folding (Tanford 1968):

$$k = k(0) \exp(m[D]/RT) \quad (11)$$

where  $k$  is the observed rate constant for binding in the presence of denaturant,  $k(0)$  the observed binding rate constant in the absence of denaturant,  $m$  a constant of proportionality and  $[D]$  the denaturant concentration. Binding rate constants can be treated as first order reaction constants because all kinetic holoflavodoxin folding experiments are done with a sufficient excess amount of FMN to have pseudo-first order reaction conditions.

In FMN binding model 1, an equilibrium between two apoflavodoxin species is assumed to exist in the absence of FMN. Each species has its own FMN binding rate constant, and FMN binding leads to conformationally identical holoflavodoxin molecules:



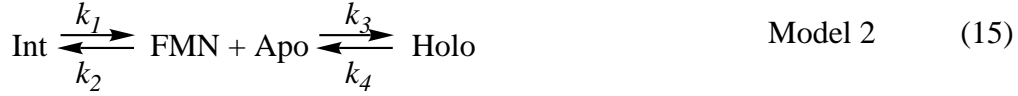
A single rate constant for FMN release from holoflavodoxin is observed (see Results and Discussion). As a consequence, both apoflavodoxin species have a different equilibrium FMN binding constants. The equilibrium constant  $K_{apo1,2}$  that describes the equilibrium between the two apoflavodoxin forms is assumed to depend exponentially on the denaturant concentration, and can be converted into a difference in free energy  $\Delta G_{apo1,2}$  between the two apoflavodoxin species:

$$K_{apo1,2} = \frac{[Apo2]}{[Apo1]} = K_{apo1,2}(0) \exp(m[D]) \quad \Delta G_{apo1,2} = -RT \ln(K_{apo1,2}) \quad (13)$$

In model 1, the GuHCl concentration dependence of each of the two observed rate constants for FMN binding as well as the single rate constant for FMN release are fitted to equation 11; the amplitudes corresponding to both FMN binding processes,  $Amp1$  and  $Amp2$ , are fitted to the fraction of apoflavodoxin molecules that exists as apoflavodoxin species 1 or 2 as determined by the equilibrium constant  $K_{apo1,2}$  (equation 13):

$$\begin{aligned}
 Amp1 &= \frac{[Apo1]}{[Apo1] + [Apo2]} = \frac{1}{1 + K_{apo1,2}} \\
 Amp2 &= \frac{[Apo2]}{[Apo1] + [Apo2]} = \frac{K_{apo1,2}}{1 + K_{apo1,2}}
 \end{aligned} \quad (14)$$

In FMN binding model 2, apoflavodoxin can bind FMN in a native or a non-native manner. The FMN molecules that are bound in a non-native manner can be released and can subsequently be incorporated in flavodoxin in a native manner, as expressed in equation 15:



with Int being a transient intermediate in the binding process in which FMN is bound to flavodoxin in a non-native manner. The two observed binding rate constants (named  $\lambda_1$  and  $\lambda_2$ ) are a complex function of the four microscopic rate constants ( $k_1 - k_4$ ) as expressed in equation 16, which has been derived in general for first order reactions (Matsen and Franklin 1950; Szabo 1969) and is often used to describe three-state protein folding reactions (Ikai and Tanford 1973; Hagerman and Baldwin 1976).

$$\lambda_{1,2} = \frac{k_1 + k_2 + k_3 + k_4 \pm \sqrt{(k_1 + k_2 + k_3 + k_4)^2 - 4(k_1k_3 + k_1k_4 + k_2k_4)}}{2} \quad (16)$$

The normalised amplitudes corresponding to  $\lambda_1$  and  $\lambda_2$  depend on the four microscopic rate constants as well, as is expressed in equation 17:

$$A_1 = \frac{Q_{rel}k_2(\lambda_1 - k_4) + k_3(\lambda_1 - k_1)}{\lambda_1(\lambda_1 - \lambda_2)} \quad A_2 = \frac{-Q_{rel}k_2(\lambda_2 - k_4) + k_3(\lambda_2 - k_1)}{\lambda_2(\lambda_1 - \lambda_2)} \quad (17)$$

where  $Q_{rel}$  is the efficiency of quenching of FMN fluorescence by the intermediate relative to the quenching of FMN fluorescence by holoflavodoxin:

$$Q_{rel} = \frac{F_{Int} - F_{FMN}}{F_{Holo} - F_{FMN}} \quad (18)$$

with  $F_x$  the fluorescence intensity of species  $x$ . The rate constant for flavin release as monitored by the rate constant of unfolding of native holoflavodoxin at GuHCl concentrations above 3.5 M is assumed to equal  $k_4$  in model 2, as at these GuHCl concentrations FMN binding to apoflavodoxin can be neglected because apoflavodoxin unfolds rapidly under these circumstances.

In FMN binding model 3 two parallel routes for FMN binding exist: one direct route from apoflavodoxin to holoflavodoxin, and one route that involves an FMN-containing intermediate, as shown in equation 19:



In model 3, the two observed rate constants  $\lambda_1$  and  $\lambda_2$  are defined by equation 20:

$$\lambda_{1,2} = \frac{\Sigma k \pm \sqrt{(\Sigma k)^2 - 4 \Sigma k k}}{2} \quad (20)$$

with  $\Sigma k = k_1 + k_2 + k_3 + k_4 + k_5 + k_6$ , and  $\Sigma k k = k_1(k_3 + k_4 + k_6) + k_2(k_4 + k_5 + k_6) + k_3(k_5 + k_6) + k_4 k_5$  (Matsen and Franklin 1950; Szabo 1969; Ikai and Tanford 1973). For reasons of symmetry,  $k_1 k_3 k_6 = k_2 k_4 k_5$ , thus each rate constant can be expressed into the five others. The corresponding normalised amplitudes are given by equations 21 and 22:

$$A_1 = \frac{Q_{rel}(k_2(\lambda_1 - k_4 - k_6) - k_3 k_6) + k_3(\lambda_1 - k_1 - k_5) - k_2 k_5}{\lambda_1(\lambda_1 - \lambda_2)} \quad (21)$$

$$A_2 = \frac{-Q_{rel}(k_2(\lambda_2 - k_4 - k_6) - k_3 k_6) + k_3(\lambda_2 - k_1 - k_5) - k_2 k_5}{\lambda_2(\lambda_1 - \lambda_2)} \quad (22)$$

The rate constant for flavin release as monitored by the rate constant of unfolding of holoflavodoxin at GuHCl concentrations above 3.5 M is fitted to  $\lambda_2$  with  $k_2$  and  $k_3$  fixed to zero in model 3, as at these GuHCl concentrations FMN binding to apoflavodoxin can be neglected because apoflavodoxin unfolds rapidly under these conditions.

## Results and discussion

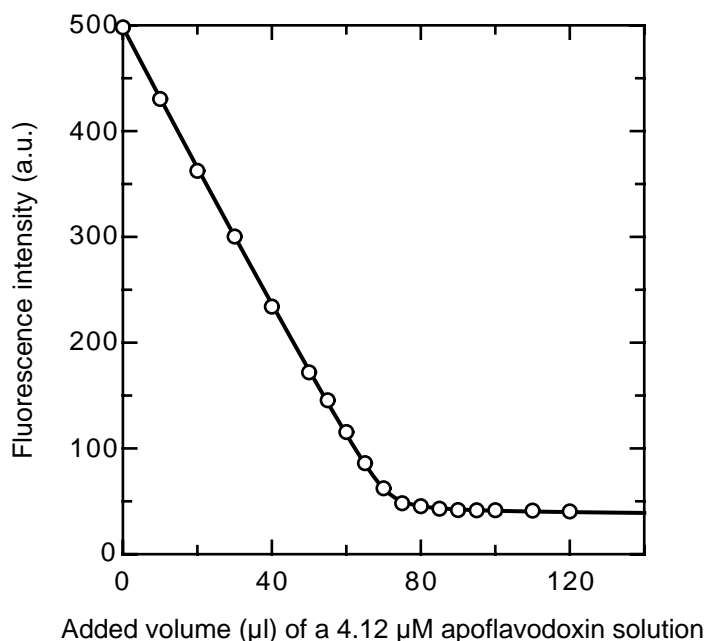
### Dissociation constant of the *A. vinelandii* apoflavodoxin–FMN complex

The FMN dissociation constant of *A. vinelandii* holoflavodoxin in 100 mM potassium pyrophosphate at pH 6.0 has been determined by a procedure in which aliquots of apoflavodoxin are titrated to an FMN solution (see Materials and Methods). In this procedure, use is made of the quenching of FMN fluorescence upon its binding to apoflavodoxin (Mayhew 1971; Lostao et al. 1997). The corresponding binding curve is shown in Figure 1 and the fitted dissociation constant turns out to be  $(3.4 \pm 0.6) \cdot 10^{-10} \text{ M}^{-1}$ . This value is identical within error to the reported value of  $(4.4 \pm 0.9) \cdot 10^{-10} \text{ M}^{-1}$  for *A. vinelandii* apoflavodoxin in 50 mM sodium phosphate at pH 7.0 (Pueyo et al. 1996).

### Determination of the global stability of *A. vinelandii* holoflavodoxin

The global stability of holoflavodoxin is determined by its equilibrium unfolding in GuHCl. Both holoflavodoxin and apoflavodoxin equilibrium unfolding data sets shown in Figure 2 are obtained by measuring tryptophan fluorescence at 350 nm and circular dichroism (CD) at 222 nm according to the procedure described in Materials and Methods. Holoflavodoxin GuHCl-induced equilibrium unfolding is not a two-state process, as can be inferred from the presented (un)folding data (Figure 2). Separate fitting of a two-state (un)folding model to the holoflavodoxin equilibrium unfolding data observed by either CD or

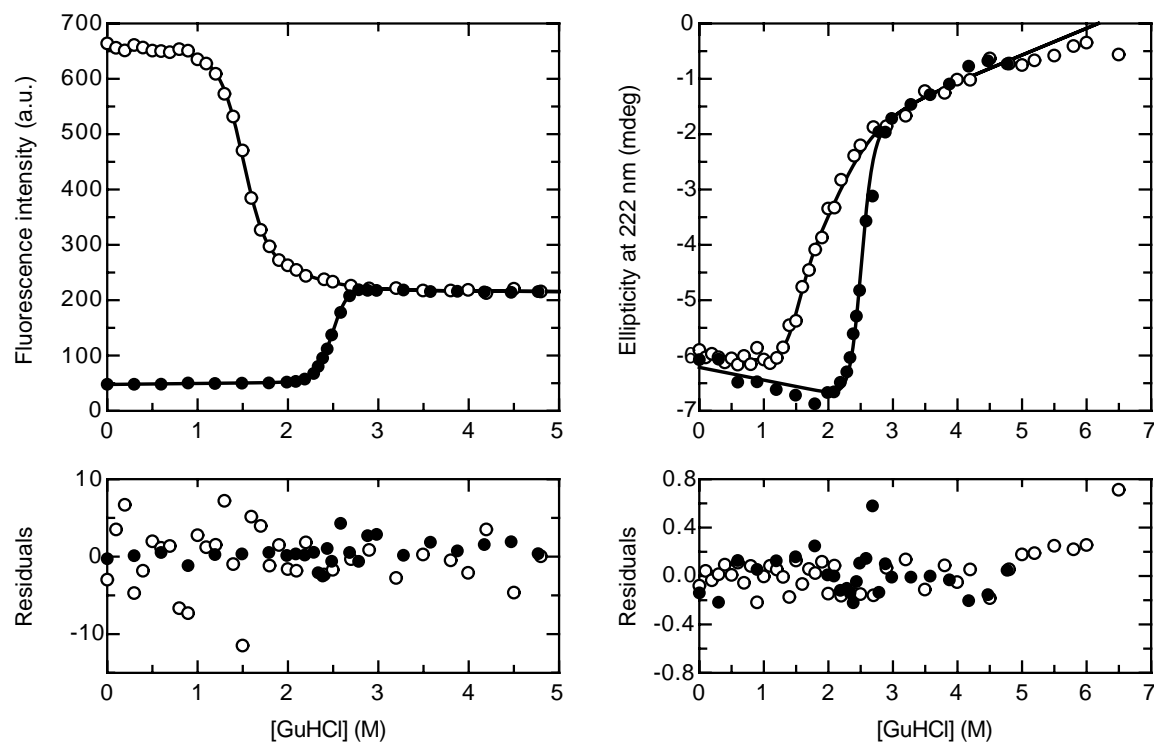




**Figure 1.** Determination of the dissociation constant of the *A. vinelandii* apoflavodoxin-FMN complex using the quenching of FMN fluorescence upon its binding to apoflavodoxin. A solution of 210 nM FMN is titrated with aliquots of a 4.12  $\mu\text{M}$  apoflavodoxin solution. Equation 3 is fitted to the resulting fluorescence intensity data as described in Materials and Methods. The dissociation constant is determined to be  $(3.4 \pm 0.6) \cdot 10^{-10}$  M. The protein is in 100 mM potassium pyrophosphate, pH 6.0 at 25  $^{\circ}\text{C}$ .

fluorescence leads to significantly differing apparent global stabilities for holoflavodoxin (i.e.  $16.52 \pm 0.31$  kcal/mol extracted from the fluorescence data, and  $11.65 \pm 0.78$  kcal/mol extracted from the CD data). The transition observed in the CD data is less steep than the one observed in the fluorescence data ( $m = -4.62 \pm 0.31$  kcal/molM $^{-1}$  extracted from the CD data versus  $-6.65 \pm 0.21$  kcal/molM $^{-1}$  extracted from the fluorescence data). Consequently, the normalised holoflavodoxin unfolding curves, expressed as fraction folded molecules, do not coincide (data not shown) and at least one other species than native and unfolded holoflavodoxin must also populate during denaturant-induced holoflavodoxin equilibrium unfolding.

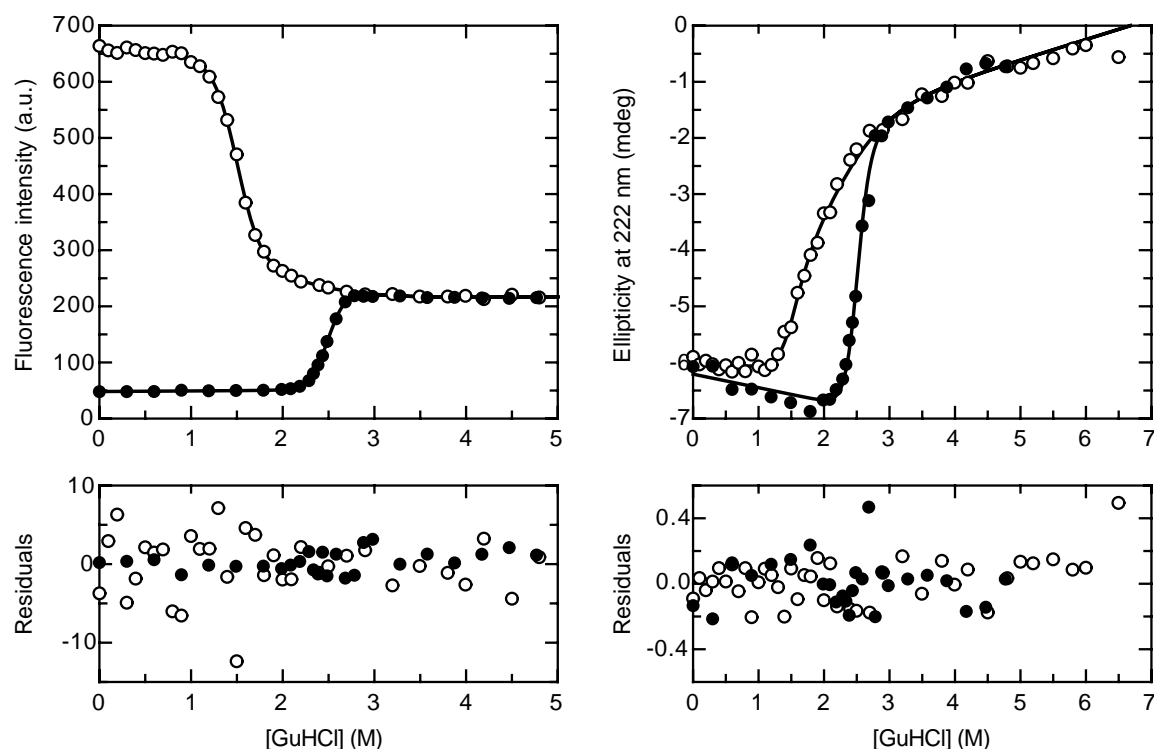
*A. vinelandii* apoflavodoxin populates one stable intermediate during denaturant-induced equilibrium unfolding (van Mierlo et al. 1998; van Mierlo et al. 2000; Chapter 2 of this thesis). Consequently, the equilibrium unfolding data of holo- and apoflavodoxin are analysed simultaneously. The four-state (un)folding model described by equation 4, which incorporates the presence of a folding intermediate, is globally fitted to the GuHCl-induced equilibrium (un)folding data of holoflavodoxin as well as of apoflavodoxin (Figure 2). In case of the analysis of the apoflavodoxin unfolding data, the four-state model described by equation 4 is reduced to the three-state model described by equation 1 by removing holoflavodoxin from equation 4 and setting the FMN concentration to zero. In the four-state model that describes holoflavodoxin (un)folding, only native apoflavodoxin is assumed to interact significantly with FMN. The equilibrium (un)folding of holoflavodoxin is measured in the presence of a 25-fold excess of FMN (i.e. 100  $\mu\text{M}$ ), to ensure a virtually constant FMN concentration at all denaturant concentrations. The dissociation constant of the



**Figure 2.** GuHCl-induced equilibrium unfolding data of *A. vinelandii* holoflavodoxin (●) and of apoflavodoxin (O) (top panels). The unfolding is monitored by tryptophan fluorescence at 350 nm (left panel) and by circular dichroism at 222 nm (right panel). The solid lines are the results of a global fit of a four-state (un)folding model ( $Holo \rightleftharpoons Apo+FMN \rightleftharpoons I_1+FMN \rightleftharpoons U+FMN$ , equation 4) to the data, as described in Materials and Methods. The global stability values of native apoflavodoxin and of the apoflavodoxin folding intermediate  $I_1$  both at 0 M GuHCl, and the corresponding  $m$ -values are fixed to those determined previously (Chapter 2 of this Thesis). The (un)folding is done at 25 °C in 100 mM potassium pyrophosphate pH 6.0 and the protein concentration is 4.0  $\mu$ M. Holoflavodoxin (un)folding data are measured in the presence of 100  $\mu$ M excess FMN to ensure a virtually constant FMN concentration at all denaturant concentrations. Residuals of the global fit of the four-state (un)folding model to the experimental unfolding data of holoflavodoxin (●) and of apoflavodoxin (O) are also shown (bottom panels).

apoflavodoxin–FMN complex in absence of denaturant is fixed to  $3.4 \cdot 10^{-10}$  M. The linear dependence of this dissociation constant on GuHCl concentration is allowed to float freely during the fit. The global stabilities of apoflavodoxin and of the flavodoxin folding intermediate  $I_1$ , both in absence of denaturant, and the corresponding  $m$ -values are taken from the extensive study of apoflavodoxin equilibrium unfolding presented in Chapter 2 of this thesis and are set fixed. Figure 2 shows that the holoflavodoxin and apoflavodoxin equilibrium unfolding data are well described by the four-state (un)folding model discussed. The results of the fit are summarised in Table 1.

In case the global stabilities of apoflavodoxin and of the apoflavodoxin folding intermediate  $I_1$ , both in absence of denaturant, and the corresponding  $m$ -values are allowed to float freely during the simultaneous fit of the four-state (un)folding model to the apo- and holoflavodoxin (un)folding data, the residuals improve slightly compared to the former fit (Figure 3 and Table 1). Now, the global stability of the apoflavodoxin folding intermediate in absence of denaturant decreases slightly compared to the value found by a three-state fit to only the apoflavodoxin (un)folding data presented in Chapter 2 of this thesis. As the stability



**Figure 3.** GuHCl-induced equilibrium unfolding data of *A. vinelandii* holoflavodoxin (●) and of apoflavodoxin (○) (top panels); the data are taken from Figure 2. The same four-state (un)folding model (equation 4) as used in Figure 2 is fitted to the data, but now the global stability values of native apoflavodoxin and of its folding intermediate  $I_1$ , both in absence of denaturant, and the corresponding  $m$ -values are allowed to float freely during the fit. Residuals of the global fit of the four-state (un)folding model to the experimental unfolding data of holoflavodoxin (●) and of apoflavodoxin (○) are also shown (bottom panels).

of native apoflavodoxin relative to the intermediate does not alter significantly, the global stability of native apoflavodoxin relative to the unfolded state now is decreased compared to the value previously found (Chapter 2 of this thesis). This lower apoflavodoxin global stability (i.e.  $9.02 \pm 0.52$  kcal/mol instead of the reported value of  $10.44 \pm 0.52$  kcal/mol, Table 1) is in good agreement with the stability value of native apoflavodoxin derived from kinetic apoflavodoxin folding and unfolding experiments (i.e.  $9.17 \pm 0.01$  kcal/mol)(Chapter 2 of this thesis). The populations at equilibrium of holoflavodoxin, native apoflavodoxin, the flavodoxin folding intermediate  $I_1$  and of unfolded flavodoxin, all as a function of the GuHCl concentration, are shown in Figure 4.

The excellent fit of the four-state (un)folding model (equation 4) to the holoflavodoxin equilibrium (un)folding data implies that the derived global stability of *A. vinelandii* holoflavodoxin equals the sum of the global stability of apoflavodoxin and the free energy associated with FMN-binding to native apoflavodoxin. This must be the case as equations 4 to 8 are fully based on the latter assumption. Our conclusion contradicts with the one of Wittung-Stafshede and co-workers (Apiyo et al. 2000), who state that non-covalent cofactor binding has no effect on the denaturant-induced equilibrium unfolding of *Desulfovibrio desulfuricans* holoflavodoxin. This latter statement however, is based on

GuHCl-induced equilibrium unfolding curves that consist of only 9 to 13 data points in which the native and unfolded baselines are not sampled sufficiently. The extensive (un)folding data presented here on *A. vinelandii* holoflavodoxin clearly show that non-covalent cofactor binding to *A. vinelandii* apoflavodoxin affects the protein stability in a theoretically predictable manner. The stability of *A. vinelandii* holoflavodoxin at a free cofactor concentration of 1 M FMN, as calculated using equation 10, is reported in Table 1.

**Table 1. Thermodynamic parameters obtained from the denaturant-induced equilibrium unfolding of *A. vinelandii* holoflavodoxin<sup>a</sup>**

		fitted with fixed values <sup>b</sup>	fitted freely <sup>c</sup>
$\Delta G_{UI}$	(kcal/mol)	3.74 <sup>d</sup>	2.45 ± 0.47
$m_{UI}$	(kcal/molM <sup>-1</sup> )	-1.83 <sup>d</sup>	-1.29 ± 0.17
$\Delta G_{IA}$	(kcal/mol)	6.70 <sup>d</sup>	6.57 ± 0.24
$m_{IA}$	(kcal/molM <sup>-1</sup> )	-4.40 <sup>d</sup>	-4.25 ± 0.15
$\Delta G_{UA}$	(kcal/mol)	10.44	9.02 ± 0.52
$m_{UA}$	(kcal/molM <sup>-1</sup> )	-6.23	-5.54 ± 0.23
$\Delta G_D$	(kcal/mol)	12.86 <sup>e</sup>	12.86 <sup>e</sup>
$m_D$	(kcal/molM <sup>-1</sup> )	-0.881 ± 0.007	-0.986 ± 0.065
$\Delta G_{unf}^{app}$	(kcal/mol)	23.30	21.88 ± 0.53
$m_{unf}^{app}$	(kcal/molM <sup>-1</sup> )	-7.11 ± 0.01	-6.53 ± 0.23

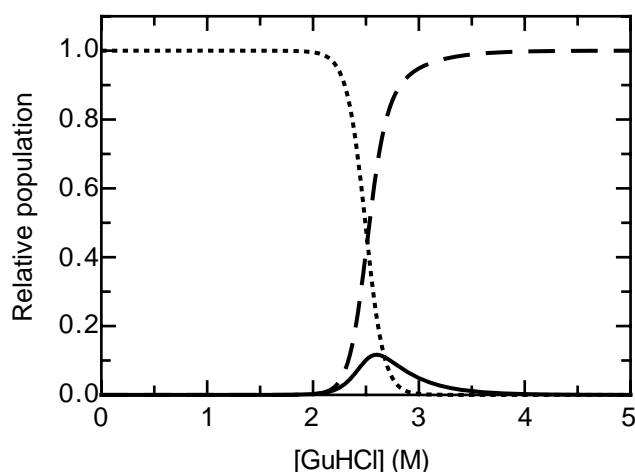
<sup>a</sup> A four-state model for equilibrium unfolding (*Holo* ⇌ *Apo*+FMN ⇌ *I<sub>I</sub>*+FMN ⇌ *U*+FMN, equations 4 to 8) is fitted to the holoflavodoxin (un)folding data. The free energy differences in absence of denaturant ( $\Delta G$ ) and corresponding denaturant-concentration dependencies ( $m$ ) are given for the equilibria between the apoflavodoxin folding intermediate *I<sub>I</sub>* and unfolded flavodoxin *U* ( $\Delta G_{UI}$  and  $m_{UI}$ ), for the equilibrium between native apoflavodoxin and intermediate *I<sub>I</sub>* ( $\Delta G_{IA}$  and  $m_{IA}$ ), and for the equilibrium between native and unfolded apoflavodoxin ( $\Delta G_{UA}$  and  $m_{UA}$ ), respectively.  $\Delta G_D$  is the free energy that is gained upon FMN binding to apoflavodoxin in the theoretical presence of 1 M FMN (calculated using  $\Delta G_D = -RT \ln K_D$ ).  $\Delta G_{unf}^{app}$  is the apparent stability of holoflavodoxin against unfolding in the presence of 1 M FMN (calculated according to equation 10), and  $m_{unf}^{app}$  is the corresponding  $m$ -value. The errors given are standard fitting errors.

<sup>b</sup> Data were fitted with  $\Delta G_{UI}$ ,  $m_{UI}$ ,  $\Delta G_{IA}$ , and  $m_{IA}$  fixed to the previously determined values (Chapter 2 of this thesis).

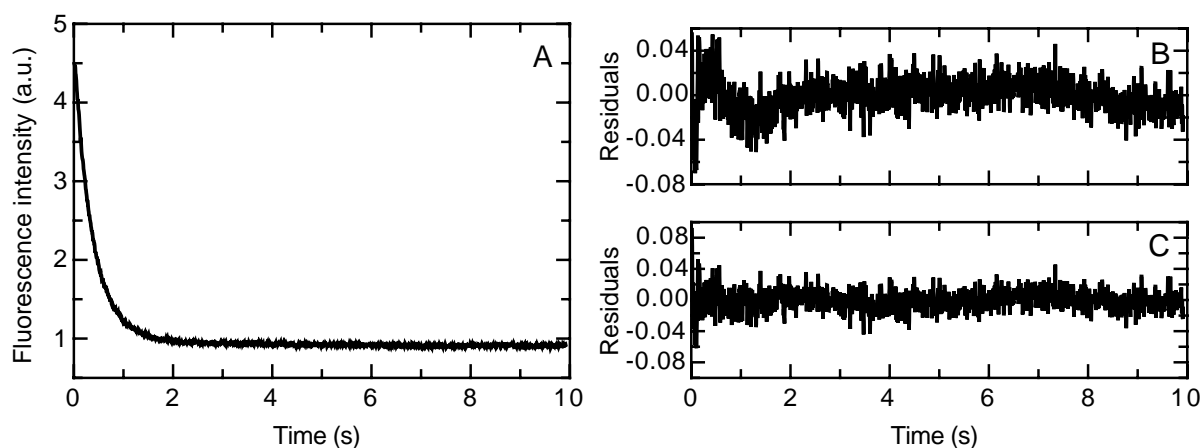
<sup>c</sup> Data were fitted with  $\Delta G_{UI}$ ,  $m_{UI}$ ,  $\Delta G_{IA}$ , and  $m_{IA}$  floating freely.

<sup>d</sup> Taken from Chapter 2 of this thesis.

<sup>e</sup> The value for  $K_D$  used is 3.4 · 10<sup>-10</sup> M determined in the present study.



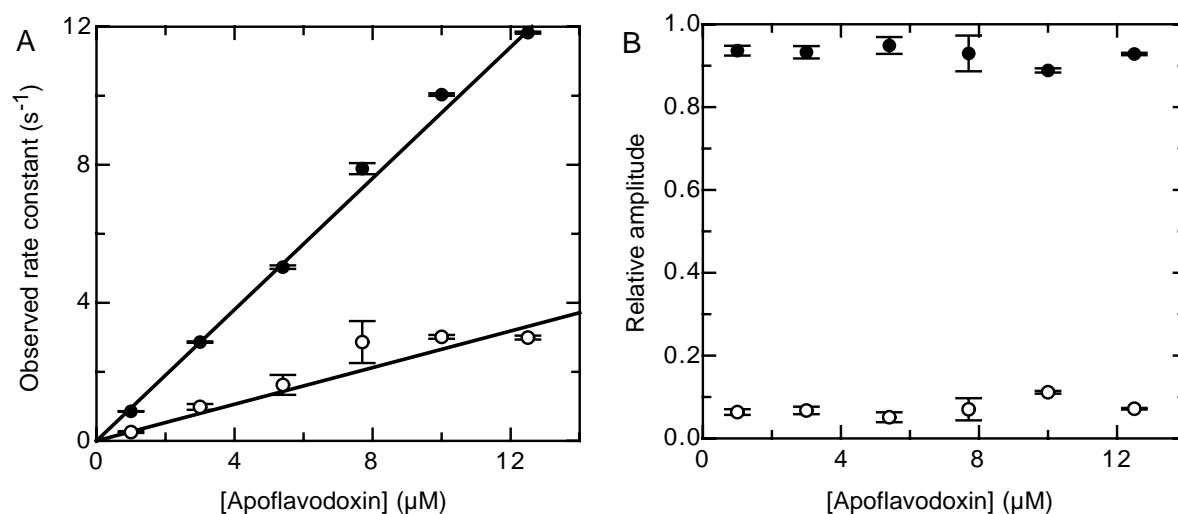
**Figure 4.** Equilibrium population of *A. vinelandii* holoflavodoxin (dotted line), unfolded flavodoxin (dashed line) and of the flavodoxin folding intermediate  $I_1$  (solid line) as a function of the concentration GuHCl in the presence of 100  $\mu\text{M}$  excess FMN. The populations are derived from the results of the four-state global fit to the experimental flavodoxin (un)folding data shown in Figure 2 and summarised in Table 1. The results of the four-state fit in which the stabilities of native apoflavodoxin and of its folding intermediate  $I_1$  and corresponding denaturant-dependencies are fixed to those previously determined (Chapter 2 of this thesis) are used to construct this graph. The folding intermediate  $I_1$  populates maximally for 11.7 % at 2.60 M GuHCl, whereas native apoflavodoxin does not significantly populate at all (it populates maximally for 0.01 % at 2.33 M GuHCl and is thus not shown).



**Figure 5.** Two rate constants describe FMN binding to native *A. vinelandii* apoflavodoxin. (A) Kinetics of FMN binding to native apoflavodoxin as monitored by the quenching of the FMN fluorescence emission above 475 nm in a stopped-flow instrument. (B) Residuals of a fit of a single exponential equation to the data. (C) Residuals of a fit of a sum of two exponential equations to the data with the fitted rate constants being  $2.86 \pm 0.02$  and  $0.98 \pm 0.08 \text{ s}^{-1}$ , respectively. The final flavodoxin concentration is 3  $\mu\text{M}$  and the final FMN concentration is 0.10  $\mu\text{M}$ , both in 100 mM potassium pyrophosphate pH 6.0, at 25  $^{\circ}\text{C}$ .

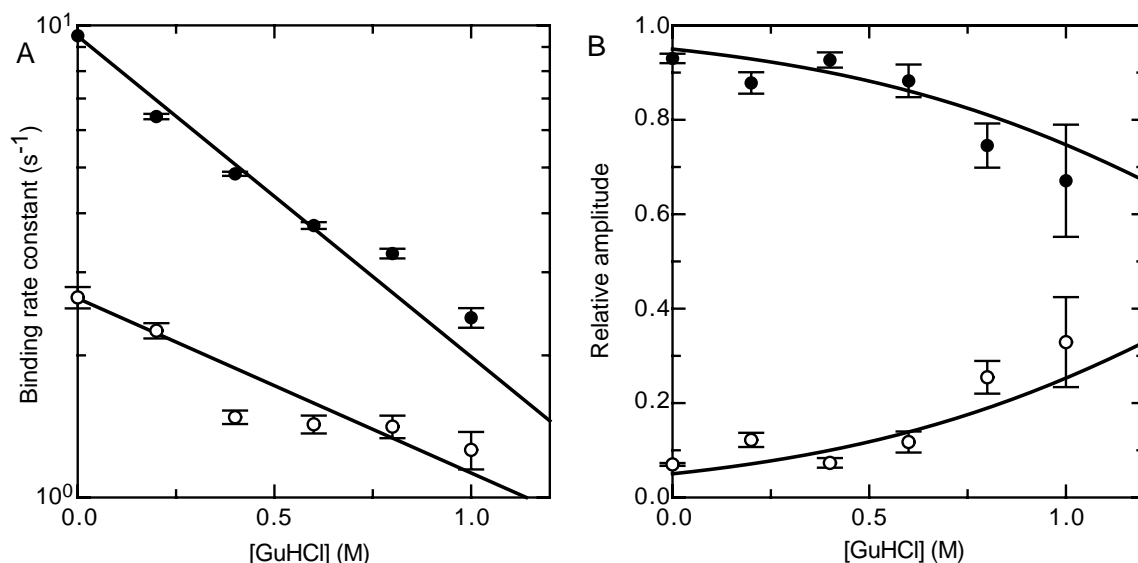
### The kinetics of FMN binding to *A. vinelandii* apoflavodoxin in the presence and absence of denaturant involves two rate constants

The kinetics of FMN binding to apoflavodoxin are determined at different protein concentrations. Pseudo-first order kinetics are obtained by using at least a tenfold excess of apoflavodoxin relative to FMN. In Figure 5A, a fluorescence intensity trace of FMN mixed with native apoflavodoxin is shown. Analysis of the residuals of a fit of a single exponential equation (Figure 5B) and of a sum of two exponential equations (Figure 5C) to the fluorescence trace reveals that two exponentials are required to properly describe the data. Two rate constants for FMN binding to apoflavodoxin are observed at all protein concentrations used, with both rate constants linearly depending on the protein concentration (Figure 6A). The larger binding rate constant is associated with a large amplitude (93 % of the total signal on average), whereas the slower binding process contributes for approximately 7 % to the observed FMN-binding kinetics (Figure 6B). The small amplitude associated with the slower binding process causes the observed scatter in the corresponding binding rate constants (Figure 6A). Second order rate constants for FMN binding to *A. vinelandii* apoflavodoxin of  $0.951 \pm 0.002$  and  $0.265 \pm 0.003 \mu\text{M}^{-1}\text{s}^{-1}$ , respectively, are derived from the slopes of the protein-concentration dependence of both FMN binding rate constants (Figure 6A).



**Figure 6.** Observed rate constants for FMN binding to *A. vinelandii* apoflavodoxin as a function of the final protein concentration in 100 mM potassium pyrophosphate pH 6.0, at 25 °C. The final FMN concentration is 0.10  $\mu\text{M}$  in all cases, except for the two highest protein concentrations used (i.e. 10 and 12.5  $\mu\text{M}$ ), where it is 1.0  $\mu\text{M}$ . (A) The protein-concentration dependence of both observed rate constants is fitted to a straight line that crosses the origin, the corresponding slopes are the pseudo-first order rate constants. The faster of the two observed FMN binding rate constants (●) fits to a pseudo-first order rate constant of  $0.951 \pm 0.002 \mu\text{M}^{-1}\text{s}^{-1}$ , whereas the slower one (○) fits to a pseudo-first order rate constant of  $0.265 \pm 0.003 \mu\text{M}^{-1}\text{s}^{-1}$ . (B) Relative amplitudes associated with the faster (●) and the slower (○) observed rate constants for FMN binding to apoflavodoxin as a function of the protein concentration. The error bars give the standard fitting errors of the rate constants and of the relative amplitudes, respectively.

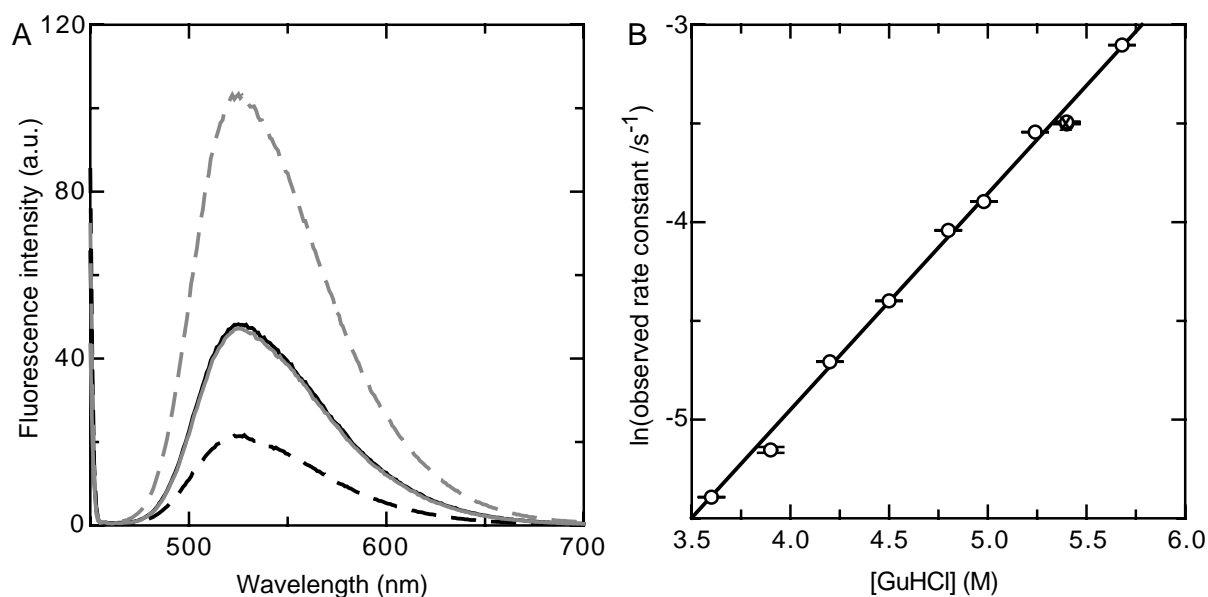
FMN binding to *A. vinelandii* apoflavodoxin is also studied as a function of the GuHCl concentration. This can inform about the co-operativity of the binding process. Figure 7A shows that, analogous to the protein concentration dependent FMN-binding experiments, two distinct rate constants for FMN binding are observed at all denaturant concentrations at which apoflavodoxin is native (i.e. up to 1.0 M GuHCl). Both FMN binding rate constants decrease with increasing denaturant concentration. In protein folding, folding rate constants decrease and unfolding rate constants increase upon increasing the denaturant concentration. The magnitude of the slope (i.e.  $m$ -value) scales to the difference in solvent accessible surface area between the reactant(s) and the corresponding transition state (Tanford 1968; Myers et al. 1995). In analogy with this, both transition states associated with FMN binding to *A. vinelandii* apoflavodoxin must be less denaturant accessible than apoflavodoxin and free FMN taken together. Figure 7B shows that at GuHCl concentrations above 0.5 M the amplitude associated with the slower FMN binding process increases, whereas the amplitude associated with the faster process decreases. The denaturant concentration dependence of both FMN binding rate constants and corresponding amplitudes can be used to discriminate between possible mechanisms for FMN binding to apoflavodoxin, as discussed below.



**Figure 7.** (A) Logarithm of the fast (●) and slow (○) observed rate constants for FMN binding to *A. vinelandii* apoflavodoxin as a function of the GuHCl concentration, detected via the quenching of the FMN fluorescence emission above 475 nm. (B) Relative amplitudes associated with the fast (●) and slow (○) observed rate constants for FMN binding to *A. vinelandii* apoflavodoxin as a function of GuHCl concentration. The rate constants and amplitudes shown at 0 M GuHCl are those obtained from extrapolating the data in Figure 6. Solid lines are the results of a fit of model 1 (equation 12) to the data. Model 1 includes two apoflavodoxin species, each with a different FMN binding rate constant. Final conditions are 10  $\mu$ M FMN and 1.0  $\mu$ M flavodoxin in 100 mM potassium pyrophosphate pH 6.0, at 25 °C. The error bars show the standard errors of the binding rate constants and of the relative amplitudes.

### The rate of FMN release determines the *A. vinelandii* holoflavodoxin unfolding rate

The rate of FMN release from *A. vinelandii* holoflavodoxin is determined by investigating the holoflavodoxin unfolding kinetics. Release of FMN results in a strong increase of its fluorescence intensity (Figure 8A). To proof that FMN-release is the rate-limiting step in *A. vinelandii* holoflavodoxin unfolding, as is suggested previously based on H/D exchange data (Steensma et al. 1998; van Mierlo and Steensma 1999), the denaturant-concentration dependence of the holoflavodoxin unfolding rate constant is measured (Figure 8B). At each GuHCl concentration studied, only one unfolding rate constant is observed (i.e. 3.5 to 5.7 M GuHCl). Clearly, the natural logarithm of the holoflavodoxin unfolding rate constant depends linearly on the GuHCl concentration. The linear extrapolation of the holoflavodoxin unfolding rate constants presented in Figure 8B to zero molar GuHCl leads to a holoflavodoxin unfolding rate constant in water of  $(8.88 \pm 0.02) \cdot 10^{-5} \text{ s}^{-1}$ , and the corresponding  $m$ -value is  $0.6459 \pm 0.0003 \text{ kcal/mol} \cdot \text{M}^{-1}$ . In case FMN release is indeed the rate-limiting step in *A. vinelandii* holoflavodoxin unfolding, the holoflavodoxin unfolding rate constant determined should equal the rate constant for FMN release from holoflavodoxin. Combination of this holoflavodoxin unfolding rate constant with the two second-order apoflavodoxin-FMN binding rate constants determined in absence of denaturant results in values for  $K_D$  of  $(9.34 \pm 0.03) \cdot 10^{-11} \text{ M}$  and  $(3.35 \pm 0.04) \cdot 10^{-10} \text{ M}$ , respectively. These



**Figure 8.** (A) Fluorescence emission spectra of free FMN (grey lines) and of FMN in the presence of *A. vinelandii* apoflavodoxin (black lines), which show the effect of the absence (dashed lines) and the presence (solid lines) of 4 M GuHCl. In the absence of denaturant, FMN binds to apoflavodoxin and as a result its fluorescence intensity is strongly quenched. Apoflavodoxin is unfolded in 4 M GuHCl, and as a result the fluorescence emission spectrum is indistinguishable from the one of free FMN recorded under the same conditions. Apoflavodoxin (when present) and FMN concentrations are 1  $\mu\text{M}$  in 100 mM potassium pyrophosphate pH 6.0 at 25 °C. Excitation is at 446 nm and a 5 nm slit is used. (B) Natural logarithm of the single observed rate constant for *A. vinelandii* holoflavodoxin unfolding (O) as a function of the GuHCl concentration. At the X sign, two data points overlap. The solid line is the result of a linear fit to all data shown. The unfolding rate constant extrapolated to water is  $(8.88 \pm 0.02) \cdot 10^{-5} \text{ s}^{-1}$ , and the corresponding  $m$ -value is  $0.6459 \pm 0.0003 \text{ kcal/mol} \cdot \text{M}^{-1}$ . The error bars show the standard errors of the rate constants. Final conditions are 1.0  $\mu\text{M}$  flavodoxin in 100 mM potassium pyrophosphate pH 6.0 at 25 °C.



values are in good agreement with the directly determined value for  $K_D$  of  $(3.4 \pm 0.4) 10^{-10}$  M (Figure 1). Thus, FMN release must indeed be the rate-limiting step in *A. vinelandii* holoflavodoxin global unfolding, as is also the case for unfolding of *Desulfovibrio vulgaris* holoflavodoxin (Nuallain and Mayhew 2002). The data shown in Figure 8B in fact represent the GuHCl concentration-dependence of the rate constant for FMN release from *A. vinelandii* holoflavodoxin. These data can be used to discriminate between possible mechanisms for FMN binding to apoflavodoxin, as discussed below.

### Kinetic models that describe FMN binding to *A. vinelandii* apoflavodoxin

Several possible origins exist for the observed two rate constants for FMN binding to *A. vinelandii* apoflavodoxin, but some of them can be excluded. As both rate constants increase with increasing protein concentration (Figure 6), each of these rate constants represents a second order association process. This excludes binding models that describe the two observed rate constants for FMN binding by employing two consecutive processes, like for example native holoflavodoxin being formed via a structural adaptation of an apoflavodoxin-FMN complex (i.e.  $\text{Apo} + \text{FMN} \rightleftharpoons \text{X} \rightleftharpoons \text{Holo}$ ). NMR experiments show no indications for structural heterogeneity to exist in *A. vinelandii* holoflavodoxin (Steensma et al. 1998). In addition, a single rate constant is observed for holoflavodoxin unfolding at various GuHCl concentrations. Consequently, the ensemble of holoflavodoxin molecules can be considered to be conformationally identical, and therefore all possible FMN-binding models need to end with a single holoflavodoxin conformation.

Three FMN-binding models have been constructed (see Materials and Methods) that can potentially explain the observed two rate constants for FMN binding to *A. vinelandii* apoflavodoxin and that fulfil the above requirements. Each of these binding models has been tested on its ability to describe the observed denaturant-concentration dependence of the FMN binding rate constants and associated amplitudes, the denaturant-concentration dependence of the rate constant of FMN release from holoflavodoxin, and the equilibrium dissociation constant of the apoflavodoxin-FMN complex.

The first FMN binding model tested (equation 12) assumes an equilibrium between two conformationally differing apoflavodoxin species, each with its own FMN binding rate constant:



This is possible as the flavin binding site of *A. vinelandii* apoflavodoxin is shown to be flexible (Steensma and van Mierlo 1998), which could reflect a relatively slow dynamic equilibrium between two protein substates. One of the apoflavodoxin species binds FMN rapidly, whereas the other one binds FMN more slowly. The denaturant-dependence of the relative amplitudes of the two observed binding processes (Figure 7) can in this model be explained by denaturant-induced changes in the relative population of both apoflavodoxin

species. The model describes the observed GuHCl concentration dependence of the binding rate constants and corresponding amplitudes rather well (Figure 7). The results of the fit are summarised in Table 2. The two apoflavodoxin species involved have a difference in free energy of 1.7 kcal/mol in absence of denaturant (determined according to equation 13), and this free energy difference has a GuHCl concentration dependence (i.e.  $m$ -value) of  $-1.09$  kcal/molM<sup>-1</sup>. Combination of the two apoflavodoxin-FMN binding rates of model 1 with the holoflavodoxin unfolding rate (i.e. the rate of FMN release from holoflavodoxin) of  $8.88 \cdot 10^{-5}$  s<sup>-1</sup> (Figure 8B), all rates in absence of denaturant, results in equilibrium FMN dissociation constants of the two apoflavodoxin species ( $K_D(1)$  and  $K_D(2)$ , Table 2) that are in good agreement with the  $K_D$  measured according to the classical method (i.e.  $3.4 \cdot 10^{-10}$  M, Figure 1).

In the second FMN binding model (equation 15), FMN can bind to apoflavodoxin in two different manners, with the first one leading to native holoflavodoxin and the second one leading to a protein molecule with incorrectly bound FMN. Incorrectly bound FMN needs to be released before subsequent native binding can occur. The mathematical description of FMN binding model 2 is analogous to a three-state kinetic protein folding model that involves an off-pathway intermediate (Matsen and Franklin 1950; Szabo 1969; Ikai and

**Table 2. Parameters extracted from the global fit of a model including two apoflavodoxin species (model 1, equation 12), each with a differing FMN binding rate constant, to the experimentally obtained GuHCl-dependence of the FMN binding rate constants and associated amplitudes of *A. vinelandii* apoflavodoxin (the corresponding data are presented in Figure 7).**

$K_{apo1,2}$	$18.9 \pm 2.6$	$m_{apo1,2}$	$-1.09 \pm 0.16$ kcal/molM <sup>-1</sup>
$k_{off}$	$8.88 \pm 0.02$ s <sup>-1</sup>	$m_{off}$	$0.6459 \pm 0.0003$ kcal/molM <sup>-1</sup>
$k_{bind}(1)$	$0.949 \pm 0.002$ μM <sup>-1</sup> s <sup>-1</sup>	$k_{bind}(2)$	$0.264 \pm 0.003$ μM <sup>-1</sup> s <sup>-1</sup>
$m_{bind}(1)$	$-0.922 \pm 0.009$ kcal/molM <sup>-1</sup>	$m_{bind}(2)$	$-0.50 \pm 0.03$ kcal/molM <sup>-1</sup>
$K_D(1)$	$(9.36 \pm 0.03) \cdot 10^{-11}$ M	$K_D(2)$	$(3.36 \pm 0.04) \cdot 10^{-10}$ M
$m_D(1)$	$-1.568 \pm 0.009$ kcal/molM <sup>-1</sup>	$m_D(2)$	$-1.14 \pm 0.03$ kcal/molM <sup>-1</sup>

The equilibrium constant  $K_{apo1,2}$  describes the equilibrium between the two apoflavodoxin species in absence of denaturant,  $k_{off}$  is the single observed rate constant for release of FMN from holoflavodoxin,  $k_{bind}(1)$  and  $k_{bind}(2)$  are the second-order rate constants for FMN binding to each of the two apoflavodoxin species,  $K_D(1)$  and  $K_D(2)$  are the equilibrium dissociation constants of each of the two apoflavodoxin-FMN complexes (i.e.  $k_{off}/k_{bind}$ ). For each rate constant and equilibrium constant the corresponding denaturant-concentration dependencies (i.e.  $m$ -values) are shown. For further details of the fitting procedure see Materials and Methods. The errors given are standard fitting errors.

Tanford 1973; Hagerman and Baldwin 1976):



This model is unable to describe the experimental data with physically meaningful parameters (fit results not shown). Firstly, the intermediate formed in model 2 is significantly less denaturant accessible than native holoflavodoxin, as can be inferred from the extracted *m*-values, that scale to denaturant accessibility (Myers et al. 1995). Secondly, the fitted fluorescence intensity of the protein molecule with incorrectly bound FMN is strongly negative, which is impossible.

The third and final FMN binding model used assumes that the two observed FMN binding rate constants originate from two parallel mechanisms for FMN binding to *A. vinelandii* apoflavodoxin that kinetically partition (equation 19):



However, FMN binding model 3 only fits to the data by ascribing some highly unlikely properties to the intermediate with bound FMN (fit results not shown). Firstly, the intermediate formed in this binding model is much less denaturant accessible than native holoflavodoxin. Secondly, the transition state associated with the conversion of the intermediate to native holoflavodoxin is even less denaturant accessible than this intermediate.

The above considerations suggest that only model 1, which includes a heterogeneous population of apoflavodoxin molecules consisting of two species, is able to describe the FMN binding data of *A. vinelandii* apoflavodoxin. The question remains what causes the existence of these two conformationally differing *A. vinelandii* apoflavodoxin species.

In case of *D. vulgaris* apoflavodoxin, monophasic FMN binding kinetics are observed in the absence of inorganic phosphate in the buffer, whereas in the presence of phosphate FMN binding is biphasic (Murray and Swenson 2003). Based on these observations and additional studies (Murray et al. 2003) it is proposed that in absence of free phosphate the 5'-phosphate of FMN binds first to apoflavodoxin, followed by binding of the isoalloxazine ring. Inorganic phosphate is proposed to bind to the phosphate binding site of a fraction of the *D. vulgaris* apoflavodoxin molecules and leads to a pre-formed isoalloxazine binding site (Murray and Swenson 2003). The latter requires the isoalloxazine ring of FMN to bind first to apoflavodoxin in these molecules. Consequently, biphasic FMN binding kinetics are observed for *D. vulgaris* apoflavodoxin in the presence of phosphate.

The studies of Murray and Swenson prompted us to determine the amount of free inorganic phosphate in our buffer (i.e. 100 mM potassium pyrophosphate). <sup>31</sup>P NMR spectroscopy shows that approximately 2 % of the <sup>31</sup>P nuclei present in this buffer is present

as phosphate (data not shown). Thus approximately 4 mM inorganic phosphate is present in the FMN binding studies of *A. vinelandii* apoflavodoxin. Most likely, the heterogeneity in *A. vinelandii* apoflavodoxin is caused by the binding of inorganic phosphate to the phosphate binding site of a fraction of the *A. vinelandii* apoflavodoxin molecules, similar to as observed for *D. vulgaris* apoflavodoxin (Murray and Swenson 2003).

In conclusion, the kinetics of FMN binding to *A. vinelandii* apoflavodoxin in the presence and absence of denaturant involves two rate constants, which are caused by the presence of two conformationally differing apoflavodoxin molecules. The binding of inorganic phosphate to a fraction of the apoflavodoxin molecules most likely causes the existence of these two apoflavodoxin species.

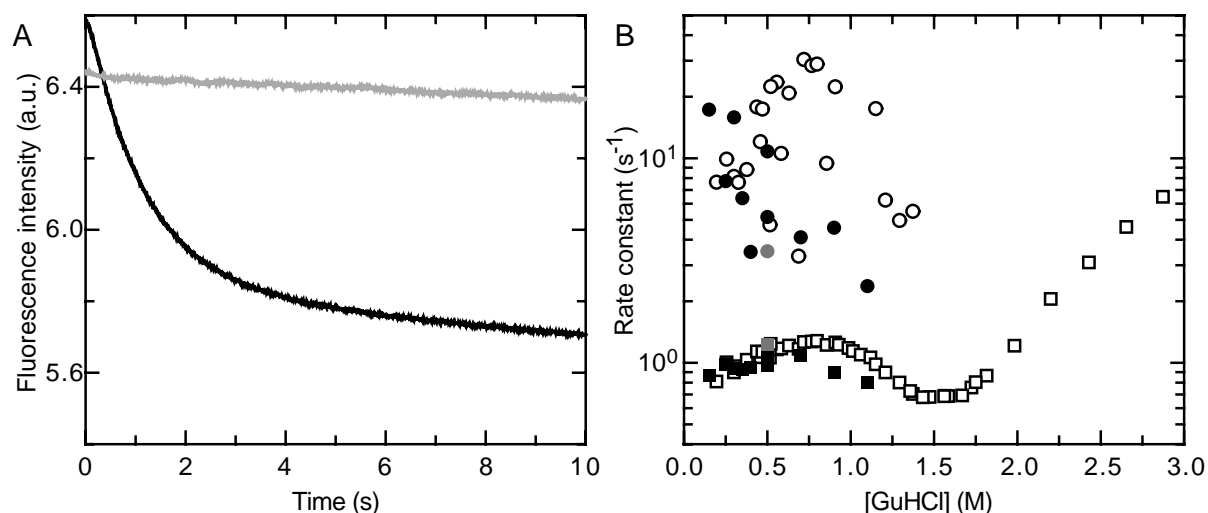
### The kinetics of *A. vinelandii* holoflavodoxin folding

The kinetics of the folding of apoflavodoxin from *A. vinelandii* in absence of FMN has been studied in great detail (Chapter 2 of this thesis). Single jump and interrupted refolding experiments show that the refolding kinetics of *A. vinelandii* apoflavodoxin are complex and involve four processes, with rate constants  $\lambda_1$  to  $\lambda_4$  in order of decreasing value. All four processes yield native molecules. The two slowest folding processes (with rate constants  $\lambda_3$  and  $\lambda_4$ ) are due to Xaa-Pro peptide bond isomerisation in the unfolded state. Under refolding conditions, two processes are observed for the folding of apoflavodoxin molecules with all Xaa-Pro peptide bonds in the native conformation, which implies the population of a folding intermediate  $I_1$ . The slowest of these two processes (i.e.  $\lambda_2$ ) becomes faster with increasing denaturant concentration, meaning that an unfolding step is rate-limiting for folding of the majority of apoflavodoxin molecules. This, together with the absence of a lag in the formation of native molecules, means that the intermediate  $I_1$  that populates during refolding is off-pathway. A second intermediate ( $I_2$ ) is observed during apoflavodoxin unfolding. It is on the direct route between unfolded and native apoflavodoxin and is high-energetic. All folding apoflavodoxin molecules pass through this high-energy intermediate before reaching the native state. The experimental data obtained on apoflavodoxin folding are consistent with the linear four-state folding mechanism (Chapter 2 of this thesis):



In this study, the influence of the presence of FMN on the folding kinetics of *A. vinelandii* flavodoxin is investigated. Unfolded apoflavodoxin in 3.0 M GuHCl is mixed into buffer containing FMN and small amounts of denaturant, and the subsequent binding of FMN to the folding apoflavodoxin molecules is monitored, as shown in Figure 9A. A sum of three exponential equations can be fitted to the time-dependent fluorescence signal obtained, which leads to the identification of three observable folding rate constants.

The slowest folding rate constant observed in the holoflavodoxin folding experiment is equal to  $\lambda_3$  observed during apoflavodoxin kinetic folding, which originates from Xaa-Pro peptide bond isomerisations (a rate constant of approximately  $0.3 \text{ s}^{-1}$  in absence of denaturant,



**Figure 9.** (A) Time-dependent fluorescence signal resulting from FMN binding to folding *A. vinelandii* apoflavodoxin molecules (black line) detected via the quenching of the FMN fluorescence above 475 nm. The trace is obtained by mixing unfolded apoflavodoxin (in 3.0 M GuHCl) with buffer containing FMN to final conditions of 10  $\mu$ M FMN, 1.0  $\mu$ M flavodoxin, 0.50 M GuHCl, 100 mM potassium pyrophosphate pH 6.0 at 25 °C. Under these conditions apoflavodoxin is in its native state. The grey line shows the trace obtained when repeating the same experiment in absence of protein. The decrease in fluorescence emission observed in the latter trace is caused by the photobleaching of FMN. (B) Chevron plot of the two largest observed rate constants for *A. vinelandii* holo flavodoxin folding as observed by changes in FMN fluorescence above 475 nm (●:  $\lambda_1$ , ■:  $\lambda_2$ ). For comparison, the corresponding *A. vinelandii* apoflavodoxin folding rate constants as observed by changes in tryptophan fluorescence intensity are shown as well (open symbols). The two lowest folding rate constants for both apo- and holo flavodoxin folding (i.e.  $\lambda_3$  and  $\lambda_4$ ) are not shown as they do not inform about the folding mechanism of flavodoxin, because they originate from Xaa-Pro peptide bond isomerisations. Up to 0.7 M GuHCl, the folding rate constants observed for both holo- and apoflavodoxin are identical within error. The  $\lambda_1$  and  $\lambda_2$  rate constants for refolding of freshly unfolded apoflavodoxin (made by unfolding apoflavodoxin for a period of 600 ms in 3.0 M GuHCl) in the presence of FMN at 0.5 M GuHCl are shown in grey. The rate constants are plotted on a logarithmic scale. Final conditions are 1.0  $\mu$ M flavodoxin in 100 mM potassium pyrophosphate pH 6.0 at 25 °C. In case of holo flavodoxin, 10  $\mu$ M FMN is present as well. Apoflavodoxin folding data are taken from Chapter 2 of this thesis.

data not shown). The slowest apoflavodoxin folding rate constant  $\lambda_4$  of 0.03 s<sup>-1</sup> in absence of denaturant, which also originates from Xaa-Pro peptide bond isomerisations, is not sufficiently sampled in the holo flavodoxin kinetic folding traces, as these traces are recorded for a period of only 10 seconds. As both  $\lambda_3$  and  $\lambda_4$  originate from Xaa-Pro peptide bond isomerisations, they do not inform about the folding mechanism of both apo- and holo flavodoxin.

The dependence of the two fastest rate constants for both apo- and holo flavodoxin folding (i.e.  $\lambda_1$  and  $\lambda_2$ ) on the GuHCl concentration is shown in Figure 9B. During apoflavodoxin kinetic folding, the processes with these rate constants correspond to the folding of molecules with all Xaa-Pro peptide bonds in the correct, native conformation (Chapter 2 of this thesis). Up to 0.7 M GuHCl, the rate constants for holo flavodoxin folding as observed via FMN binding are identical within error to those observed for apoflavodoxin folding. Above 0.7 M GuHCl, the FMN binding rate constants deviate from the

apoflavodoxin folding rate constants. At these concentrations denaturant both folding and unfolding processes contribute to the observed kinetics, and as holoflavodoxin unfolds slower than apoflavodoxin, the resulting observed rate constants for holoflavodoxin folding are lower than those of apoflavodoxin folding. Clearly, the presence of excess FMN does not accelerate flavodoxin folding.

The largest folding rate observed during apoflavodoxin folding (i.e.  $\lambda_1$ ) contains contributions from the formation of folding intermediates with non-native Xaa-Pro peptide bonds (Chapter 2 of this thesis). In case these intermediates are able to bind FMN, this could influence the observed holoflavodoxin folding kinetics as monitored by FMN binding (Figure 9B). However, freshly unfolded apoflavodoxin (made by unfolding apoflavodoxin for 600 ms in 3.0 M GuHCl), in which the vast majority of the Xaa-Pro peptide bond isomers is correct, gives upon folding in presence of FMN at 0.5 M GuHCl rate constants for FMN binding that are identical within error to the rate constants for folding of equilibrium unfolded apo- and holoflavodoxin (Figure 9B). Consequently, apoflavodoxin folding intermediates with non-native Xaa-Pro peptide bonds do not bind FMN.

In Figure 9A, the time-dependent fluorescence intensity of unbound FMN in a kinetic folding experiment is shown, recorded under identical circumstances as those used for holoflavodoxin folding, except that now no protein is present. The initial fluorescence intensity of FMN during kinetic holoflavodoxin folding is higher than observed in the blank experiment in which flavodoxin is absent. This higher intensity is possibly explained by a weak aspecific association of FMN with the folding apoflavodoxin molecules, which would result in FMN being less accessible to GuHCl compared to unbound FMN. We observe that the fluorescence intensity of unbound FMN is quenched by the presence of GuHCl (24 % quenching at 0.5 M GuHCl, data not shown).

The possible association of FMN with unfolded apoflavodoxin is investigated under circumstances at which this state is significantly populated. However, no indications for the association of FMN with apoflavodoxin unfolded in 4.0 M GuHCl are observed. Firstly, at this denaturant concentration the fluorescence emission spectra of FMN in the presence and in the absence of apoflavodoxin are identical (Figure 8A). Secondly,  $^1\text{H}$ -NMR spectra of holoflavodoxin unfolded in 4 M GuHCl show sharp resonances of FMN protons at proton frequencies that coincide within the spectral resolution of 0.01 ppm with those of FMN in 4 M GuHCl in absence of flavodoxin (data not shown). Thirdly, the observed flavodoxin folding kinetics are not affected by the presence of excess FMN (Figure 9B). Fourthly, the holoflavodoxin denaturant-induced equilibrium unfolding data (Figures 2 and 3) are excellently described by assuming that only native apoflavodoxin binds FMN. In summary, if any association of FMN with unfolded apoflavodoxin molecules exists, it must be weak.

The possible association of FMN with the apoflavodoxin folding intermediate  $I_1$  can, if it exists, only be weak, because at 2.6 M GuHCl, where  $I_1$  is maximally populated for 11.7 % (see Figure 4) the fluorescence emission spectra of FMN in the presence and in the absence of apoflavodoxin are identical (data not shown). In addition, up to 0.7 M GuHCl the observed flavodoxin folding kinetics are not affected (Figure 9B):  $\lambda_2$  increases with increasing denaturant concentration for both apo- and holoflavodoxin. This increase reflects

the unfolding of the off-pathway apoflavodoxin folding intermediate  $I_1$  (Chapter 2 of this thesis). As the presence of FMN leads to no alteration in the latter phenomenon, binding of FMN to  $I_1$  seems unlikely. The latter is again supported by the holoflavodoxin equilibrium unfolding data being well described by assuming that only native apoflavodoxin binds FMN to a significant extent. In summary, the off-pathway intermediate  $I_1$  does not bind FMN with an appreciable binding constant. In addition, the high-energy on-pathway intermediate  $I_2$  (equation 2) exists for such a short time (Chapter 2 of this thesis) that FMN binding to this intermediate must also be negligible.

The data presented show that FMN has no appreciable interaction with non-native apoflavodoxin molecules. The data also show that the presence of excess FMN does not accelerate *A. vinelandii* flavodoxin folding. As under the conditions investigated the formation of native apoflavodoxin molecules is the rate-limiting step in holoflavodoxin folding, FMN binding to native apoflavodoxin must be the final step in *A. vinelandii* holoflavodoxin folding.

In contrast to our observations, Wittung-Stafshede and co-workers propose that the presence of FMN speeds up the folding of *D. desulfuricans* flavodoxin (Apiyo and Wittung-Stafshede 2002). However, in the latter study only a limited number of folding and unfolding rate constants as a function of concentration denaturant are reported for both apoflavodoxin and holoflavodoxin kinetic folding. In addition, in case of the *D. desulfuricans* holoflavodoxin kinetic folding study, FMN and apoflavodoxin are present in equal amounts, resulting in an FMN binding rate constant that must decrease during the course of the folding experiment. No corrections were made for this phenomenon.

In summary, the folding data of *A. vinelandii* holoflavodoxin show that first apoflavodoxin folds to its native state according to the mechanism described in Chapter 2 of this thesis and that subsequently FMN binds to native apoflavodoxin, thereby forming holoflavodoxin. Excess FMN does not accelerate holoflavodoxin kinetic folding and FMN does not act as a nucleation site for flavodoxin folding. No indications exist that unfolded apoflavodoxin, the molten globule-like apoflavodoxin off-pathway folding intermediate  $I_1$ , or the high-energy on-pathway intermediate  $I_2$  interact to a significant extent with FMN. Whereas FMN binding is the final step in *A. vinelandii* holoflavodoxin folding, holoflavodoxin is so stable that FMN needs to be released before global unfolding of the protein can occur. The global unfolding of *A. vinelandii* holoflavodoxin is a rare event, occurring approximately once per 3 hours in the absence of denaturant.

## References

- Apiyo, D., J. Guidry and P. Wittung-Stafshede (2000). "No cofactor effect on equilibrium unfolding of *Desulfovibrio desulfuricans* flavodoxin." *Biochim Biophys Acta* 1479(1-2): 214-24.
- Apiyo, D. and P. Wittung-Stafshede (2002). "Presence of the cofactor speeds up folding of *Desulfovibrio desulfuricans* flavodoxin." *Protein Sci* 11(5): 1129-35.

- Barman, B. G. and G. Tollin (1972). "Flavine-protein interactions in flavoenzymes. Temperature-jump and stopped-flow studies of flavine analog binding to the apoprotein of *Azotobacter* flavodoxin." Biochemistry 11(25): 4746-54.
- Bertini, I., J. A. Cowan, C. Luchinat, K. Natarajan and M. Piccioli (1997). "Characterization of a partially unfolded high potential iron protein." Biochemistry 36(31): 9332-9.
- Creighton, T. E. (1993). Proteins, structures and molecular properties. New York, W. H. Freeman and Company.
- Edmondson, D. E. and G. Tollin (1971). "Chemical and physical characterization of the Shethna flavoprotein and apoprotein and kinetics and thermodynamics of flavin analog binding to the apoprotein." Biochemistry 10(1): 124-32.
- Hagerman, P. J. and R. L. Baldwin (1976). "A quantitative treatment of the kinetics of the folding transition of ribonuclease A." Biochemistry 15(7): 1462-73.
- Hefti, M. H., J. Vervoort and W. J. H. van Berkel (2003). "Deflavination and reconstitution of flavoproteins." Eur J Biochem 270(21): 4227-42.
- Ikai, A. and C. Tanford (1973). "Kinetics of unfolding and refolding of proteins. I. Mathematical analysis." J Mol Biol 73(2): 145-63.
- Lostao, A., C. Gomez-Moreno, S. G. Mayhew and J. Sancho (1997). "Differential stabilization of the three FMN redox forms by tyrosine 94 and tryptophan 57 in flavodoxin from *Anabaena* and its influence on the redox potentials." Biochemistry 36(47): 14334-44.
- Matsen, F. A. and J. L. Franklin (1950). "A general theory of coupled sets of first order reactions." J Am Chem Soc 72(53): 3337-3341.
- Mayhew, S. G. (1971). "Studies on flavin binding in flavodoxins." Biochim Biophys Acta 235(2): 289-302.
- Mayhew, S. G. and J. H. Wassink (1980). "Determination of FMN and FAD by fluorescence titration with apoflavodoxin." Methods Enzymol 66: 217-20.
- Murray, T. A., M. P. Foster and R. P. Swenson (2003). "Mechanism of flavin mononucleotide cofactor binding to the *Desulfovibrio vulgaris* flavodoxin. 2. Evidence for cooperative conformational changes involving tryptophan 60 in the interaction between the phosphate- and ring-binding subsites." Biochemistry 42(8): 2317-27.
- Murray, T. A. and R. P. Swenson (2003). "Mechanism of flavin mononucleotide cofactor binding to the *Desulfovibrio vulgaris* flavodoxin. 1. Kinetic evidence for cooperative effects associated with the binding of inorganic phosphate and the 5'-phosphate moiety of the cofactor." Biochemistry 42(8): 2307-16.
- Myers, J. K., C. N. Pace and J. M. Scholtz (1995). "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding." Protein Sci 4(10): 2138-48.
- Nuallain, B. O. and S. G. Mayhew (2002). "A comparison of the urea-induced unfolding of apoflavodoxin and flavodoxin from *Desulfovibrio vulgaris*." Eur J Biochem 269(1): 212-23.



- Pueyo, J. J., G. P. Curley and S. G. Mayhew (1996). "Kinetics and thermodynamics of the binding of riboflavin, riboflavin 5'-phosphate and riboflavin 3',5'-bisphosphate by apoflavodoxins." Biochem J 313(Pt 3): 855-61.
- Robinson, C. R., Y. Liu, J. A. Thomson, J. M. Sturtevant and S. G. Sligar (1997). "Energetics of heme binding to native and denatured states of cytochrome b562." Biochemistry 36(51): 16141-6.
- Steensma, E., H. A. Heering, W. R. Hagen and C. P. M. Van Mierlo (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(1-2): 167-72.
- Steensma, E., M. J. Nijman, Y. J. Bollen, P. A. de Jager, W. A. van den Berg, W. M. van Dongen and C. P. M. van Mierlo (1998). "Apparent local stability of the secondary structure of *Azotobacter vinelandii* holo flavodoxin II as probed by hydrogen exchange: implications for redox potential regulation and flavodoxin folding." Protein Sci 7(2): 306-17.
- Steensma, E. and C. P. M. van Mierlo (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(3): 653-66.
- Szabo, Z. G. (1969). Kinetic Characterization of complex reaction systems. Comprehensive chemical kinetics. C. H. Bamford and C. F. H. Tipper. Amsterdam, Elsevier publishing company. 2: 1-80.
- Tanford, C. (1968). "Protein denaturation." Adv Protein Chem 23: 121-282.
- van Mierlo, C. P. M. and E. Steensma (1999). "Stabilization centers differ between structurally homologous proteins as shown by NMR spectroscopy." Journal of molecular catalysis B: Enzymatic 7: 147-156.
- van Mierlo, C. P. M., J. M. van den Oever and E. Steensma (2000). "Apoflavodoxin (un)folding followed at the residue level by NMR." Protein Sci 9(1): 145-57.
- van Mierlo, C. P. M., W. M. van Dongen, F. Vergeldt, W. J. H. van Berkel and E. Steensma (1998). "The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate." Protein Sci 7(11): 2331-44.



# 6 The influence of cofactor binding on local unfolding of *Azotobacter vinelandii* flavodoxin as determined by H/D exchange and NMR spectroscopy

Yves J.M. Bollen and Carlo P.M. van Mierlo

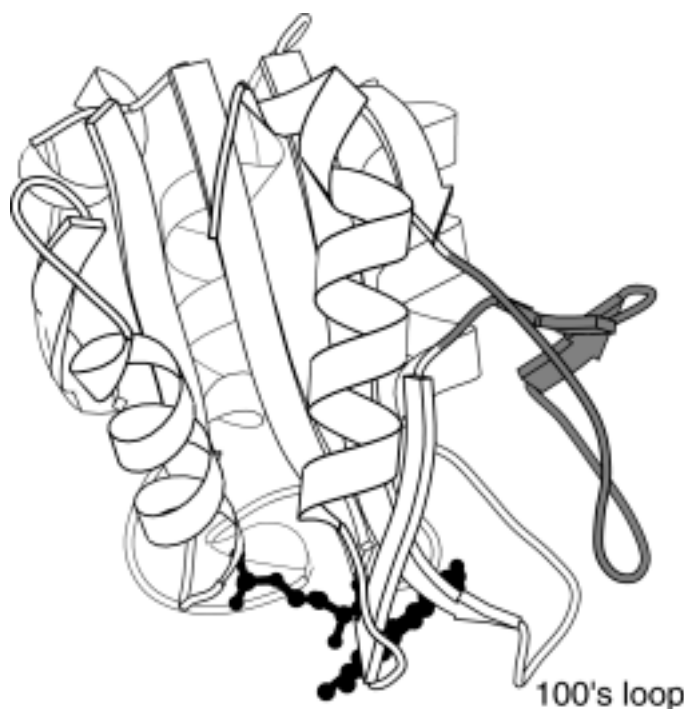
Native state H/D exchange combined with NMR spectroscopy is used to probe the influence of FMN binding on the stability of *Azotobacter vinelandii* flavodoxin against local, subglobal and global unfolding. The global stability of *A. vinelandii* holoflavodoxin can not be determined from the H/D exchange data. This is caused by the unavoidable occasional FMN release from the holoprotein, which causes none of its backbone amide protons to exchange solely from the globally unfolded state of the protein. In absence of denaturant, H/D exchange occurs via local or subglobal unfolding processes, just as is the case for apoflavodoxin. No areas are detected in flavodoxin where FMN binding results in an increase of the local dynamics. Almost the entire flavodoxin backbone is substantially more rigid in holoflavodoxin than in apoflavodoxin. Part of the 23-residue loop that is inserted into  $\beta$ -strand 5 of flavodoxin, which is characteristic for the class of long chain flavodoxins, is in strong contrast to apoflavodoxin inaccessible to water in holoflavodoxin. Similarly, the N(3)H proton of the FMN cofactor is inaccessible to water as long as FMN is bound and its exchange with solvent deuterons only takes place upon release of FMN. At least three out of the four partially unfolded forms that apoflavodoxin occasionally adopts under native conditions are inaccessible to holoflavodoxin. Holoflavodoxin can form these partially unfolded conformations only when FMN is released. Taking all observations together, a schematic representation of the free energy landscape for folding of *A. vinelandii* holoflavodoxin is presented.

## Introduction

Flavodoxins are small monomeric proteins involved in electron transport. They adopt an  $\alpha$ - $\beta$  parallel topology, and contain a non-covalently bound flavin mononucleotide (FMN) cofactor (Figure 1) (Mayhew and Tollin 1992). Apoflavodoxin from *Azotobacter vinelandii* (i.e. flavodoxin in the absence of the FMN cofactor) is structurally identical to flavodoxin except for some dynamic disorder in the flavin-binding region (Steensma et al. 1998; Steensma and van Mierlo 1998). The folding of apoflavodoxin from *A. vinelandii* has been studied in great detail. Apoflavodoxin is shown to populate an intermediate during equilibrium unfolding (van Mierlo et al. 1998; van Mierlo et al. 2000), which is kinetically off the productive folding route (Chapter 2 of this thesis). The formation of a second intermediate is shown to be an obligatory step on the productive folding route of the protein. This second intermediate is highly unstable and never populates significantly at equilibrium or during kinetic folding of apoflavodoxin (Chapter 2 of this thesis). Apoflavodoxin folding can be summarised as follows:



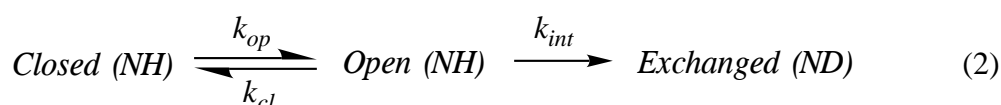
in which  $I_1$  is the off-pathway intermediate that populates at equilibrium, U is unfolded apoflavodoxin,  $I_2$  is the unstable on-pathway intermediate and N is native apoflavodoxin.



**Figure 1.** Molscript cartoon drawing (Kraulis 1991) of the X-ray structure of *Azotobacter chroococcum* flavodoxin (Thorneley et al. 1994), the sequence of which is 95 % identical to *A. vinelandii* flavodoxin. The secondary structure depicted is as determined for *A. vinelandii* holoflavodoxin by NMR spectroscopy (Steensma et al. 1998). The FMN cofactor is shown in black in ball-and-stick representation. The 23-residue insertion into  $\beta$ -strand 5 that is characteristic for long-chain flavodoxins is depicted in grey.

To characterise the local dynamics of *A. vinelandii* apoflavodoxin, native-state hydrogen deuterium (H/D) exchange experiments have been reported (Chapter 4 of this thesis). H/D exchange detected by NMR spectroscopy is a powerful technique that can give detailed information about the stability and dynamics of a protein at the level of single amino acids. Backbone amide protons, provided they are water-accessible, exchange gradually with the protons of water or with deuterons in case D<sub>2</sub>O is used as a solvent. Amide protons involved in hydrogen bonding networks or buried from the solvent exchange more slowly than freely accessible amide protons since the protein has to unfold, partially or completely, to facilitate exchange of these buried or bonded protons.

Quantitative interpretation of hydrogen deuterium exchange is possible using a simple two-state model (Hvidt and Nielsen 1966):



In this model, the open or exchange-competent form and the closed or exchange-incompetent form of a protein at the site of a particular amide proton are in equilibrium with one another, and interconvert with rate constants for opening ( $k_{op}$ ) and closing ( $k_{cl}$ ). This interconversion reflects either local opening of the structure or a (sub)global unfolding process. From the open state exchange takes place with the intrinsic rate constant  $k_{int}$  that depends on amino acid sequence, pD and temperature and can be calculated using values derived from work on model peptides (Bai et al. 1993). The time course of exchange can be monitored at the residue level by NMR, since amide protons give rise to a <sup>1</sup>H-NMR signal, and replacement of the proton by a deuteron leads to disappearance of this signal.

Under conditions favouring the closed state (i.e.  $k_{op} \ll k_{cl}$ ), the observed exchange rate  $k_{ex}$  is given by:

$$k_{ex} = \frac{k_{op}k_{int}}{k_{cl} + k_{int}} \quad (3)$$

Depending on the ratio of  $k_{cl}$  and  $k_{int}$ , two limiting cases may be reached. If  $k_{cl} \ll k_{int}$ , equation 3 reduces to

$$k_{ex} = k_{op} \quad (4)$$

Under these conditions, referred to as EX1, the measured exchange rate of a certain amide proton informs about the local conversion rate between the closed and the open state of its micro-environment. If  $k_{cl} \gg k_{int}$ , equation 3 reduces to

$$k_{ex} = \frac{k_{op}k_{int}}{k_{cl}} = K_{op}k_{int} \quad (5)$$

In this limiting case, referred to as EX2, the ratio of the measured exchange rate  $k_{ex}$  and  $k_{int}$  provides the equilibrium constant for local opening of the protein structure  $K_{op}$ , which is the ratio of  $k_{op}$  and  $k_{cl}$  and is a measure for the local stability  $\Delta G_{op}$ :

$$\Delta G_{op} = -RT \ln(K_{op}) = -RT \ln(k_{op}/k_{cl}) = -RT \ln(k_{ex}/k_{int}) \quad (6)$$

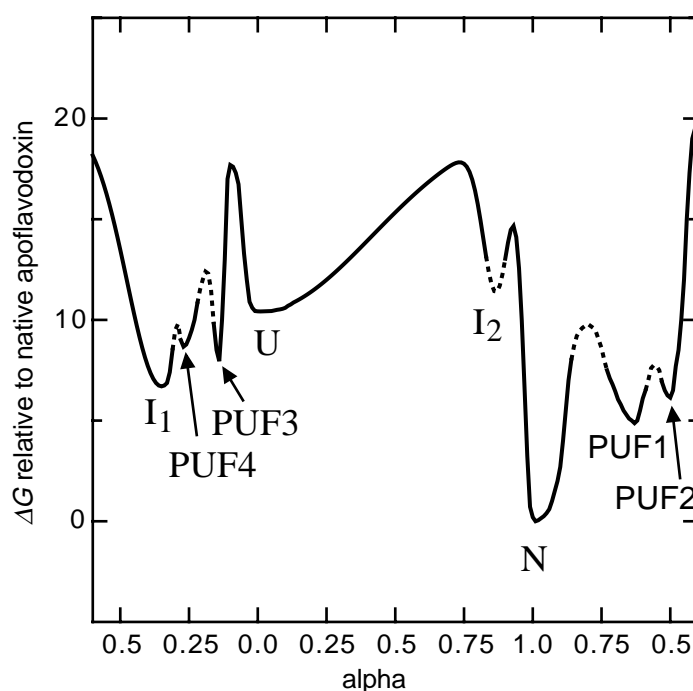
An amide proton can exchange with solvent deuterons via local opening of the surrounding protein structure, via subglobal (i.e. partial) unfolding of the protein or via global protein unfolding. Native state hydrogen exchange experiments performed under EX2 conditions in the presence and absence of small amounts of denaturant allow the identification of clusters of residues that reach their respective exchange-competent state by a concerted sub-global or global protein unfolding process (Bai et al. 1995; Chamberlain et al. 1996). In case an amide proton exchanges via local opening processes its  $\Delta G_{op}$  is insensitive to denaturant because only a small amount of protein surface becomes accessible to denaturant in such a local fluctuation of the protein structure. In case an amide proton exchanges via (sub)global unfolding processes its  $\Delta G_{op}$  decreases with increasing denaturant concentration. Amide protons that exchange via the same (sub)global unfolding process are characterised by the same denaturant concentration-dependence of  $\Delta G_{op}$  (see chapter 4 of this thesis for details).

Once residues have been identified to exchange via a specific (sub)global unfolding process under EX2 conditions in absence of denaturant, the exchange rate of these residues under EX1 conditions informs about the rate of that specific unfolding process (Yan et al. 2002). Since amide proton exchange is base-catalysed above pD 4, the intrinsic exchange rate  $k_{int}$  increases tenfold by every increase of the pD by 1 unit above pD 4. As a consequence, the mechanism for amide proton exchange for a specific protein can be altered from EX2 to EX1 by increasing the pD sufficiently.

In absence of denaturant, most amide protons of *A. vinelandii* apoflavodoxin exchange via local structural opening processes (Chapter 4 of this thesis). Of the amide protons that exchange via (sub)global protein unfolding, only a single residue (Phe25) that is shielded from the solvent in the native state becomes solvent accessible in the equilibrium intermediate  $I_1$ . The latter is remarkable as  $I_1$  is the most stable apoflavodoxin species besides the native state (Chapter 4 of this thesis). Four clusters of residues have been identified in apoflavodoxin that become water accessible in a specific partially unfolded form (PUF) of the protein. These residues are categorised according to the denaturant-concentration dependence of their  $\Delta G_{op}$ : backbone amide protons that become water accessible in a particular PUF all have the same denaturant concentration-dependence of their  $\Delta G_{op}$ . The identification of four of such clusters within apoflavodoxin leads to the characterisation of four apoflavodoxin PUFs. A fifth cluster of residues is identified within *A. vinelandii* apoflavodoxin the amide protons of which exchange from unfolded apoflavodoxin in presence of GuHCl concentrations exceeding 0.5 M. However, in absence of denaturant these amide protons all exchange via local opening processes (Chapter 4 of this thesis). Under

native conditions not one backbone amide proton in apoflavodoxin exchanges only from the globally unfolded state. The separate clusters of co-operatively unfolding residues as determined by native state H/D exchange of *A. vinelandii* apoflavodoxin are shown in Figure 3 of Chapter 4 at page 76. Starting from native apoflavodoxin, in PUF1 residues of Cluster 1 are unfolded, subsequently PUF2 is formed in which residues of Cluster 2 are unfolded in addition to those of Cluster 1, and so on.

The rates of interconversion of the apoflavodoxin PUFs with native apoflavodoxin are determined via the pD-dependence of the corresponding amide proton exchange rates. Comparison of these rates with stopped-flow kinetic folding rate constants shows that none of the identified PUFs are on the direct productive folding route between globally unfolded and native apoflavodoxin (Chapter 2 of this thesis). PUF3 and PUF4 are on a non-productive



**Figure 2.** Schematic representation of the free energy landscape for folding of *A. vinelandii* apoflavodoxin derived from the results obtained from denaturant-induced equilibrium and kinetic (un)folding studies as presented in Chapter 2 of this thesis, and obtained from native state H/D exchange studies (Chapter 4 of this Thesis). The horizontal axis represents the reaction co-ordinate, the vertical axis the free energy. The horizontal axis scales to the denaturant accessibility of the protein; it is expressed as the  $\alpha$ -value, with native apoflavodoxin having an  $\alpha$ -value of 1 and unfolded apoflavodoxin an  $\alpha$ -value of 0. U and N represent unfolded and native apoflavodoxin, respectively;  $I_1$  and  $I_2$  are the two folding intermediates presented in equation 1. The off-pathway intermediate  $I_1$  is represented on the left-hand side of the unfolded state, whereas both the on-pathway intermediate  $I_2$  and native apoflavodoxin reside on the right-hand side of the unfolded state. PUF3 and PUF4 have within error the same free energy barrier for opening as  $I_1$ , according to their opening rates as determined from pD dependent H/D exchange data. The heights of the free energy barriers between these latter three species are unknown, and are therefore represented by dashed lines. PUF1 and PUF2 are on a separate unfolding route starting from N. Note that on the right-hand side of N ( $\alpha = 1$ ) the  $\alpha$ -values decrease again. The heights of the barriers  $\Delta G_{op}^\ddagger$  were calculated from the opening (i.e. unfolding) rates according to  $\Delta G_{op}^\ddagger = -RT \ln(k_{op}/k_0)$  using a value for  $k_0$  of  $10^8$  (Krieger et al. 2003). All free energy barriers the heights of which are not known from kinetic (un)folding studies are chosen such that they do not influence the (un)folding kinetics, and are represented by dashed lines. The depth of the minimum in the free energy landscape in which  $I_2$  resides is unknown, and is therefore represented by a dashed line.

folding route that starts with unfolded apoflavodoxin. A common free energy barrier separates both PUF3 and PUF4 from unfolded apoflavodoxin. This barrier is proposed to be the same as the one that separates the off-pathway apoflavodoxin folding intermediate  $I_1$  from the native state. Both PUF3 and PUF4, as well as  $I_1$ , need to unfold before productive folding of apoflavodoxin can occur. PUF1 and PUF2 are unfolding excursions from native apoflavodoxin that are not on the unfolding route that leads to unfolded apoflavodoxin. A simple two-dimensional model of the free energy landscape for apoflavodoxin folding that is consistent with the experimental data on apoflavodoxin equilibrium and kinetic folding and unfolding as well as with the H/D exchange behaviour of apoflavodoxin's backbone amide protons is shown in Figure 2 (see also Chapter 4 of this thesis).

In case of *A. vinelandii* holoflavodoxin it has been shown that the equilibrium stability of the protein equals the sum of the stability of apoflavodoxin and the free energy that is gained upon binding of FMN to apoflavodoxin (Chapter 5 of this thesis). During kinetic folding of flavodoxin, binding of FMN only occurs after native apoflavodoxin has been formed. The presence of FMN does not accelerate the folding process. Release of FMN is the rate-limiting process during global unfolding of flavodoxin. As long as FMN stays bound to flavodoxin, the protein remains native even under strong denaturing conditions. However, as soon as apoflavodoxin is formed under these conditions, it unfolds rapidly. The kinetic folding scheme for holoflavodoxin (Chapter 5 of this thesis) is given by:



Currently, little is known about the influence ligand binding has on the internal dynamics of proteins. In several proteins, ligand binding induces large-scale structural changes (see (Luque et al. 2002) for a recent review). The propagation of binding interactions to distal sites in the protein is proposed to enable allosteric regulation and signal transduction (Luque et al. 2002). In most cases, the backbone conformation of the protein becomes less flexible upon ligand binding. Remarkably, in case of mouse major urinary protein the protein backbone is observed to become more flexible upon binding of a small hydrophobic ligand (Zidek et al. 1999). Clearly, more experimental data are required to further insight into the role ligands have on the internal dynamics of proteins.

Here, native state H/D exchange is used to probe the influence FMN binding has on the local stabilities of the residues within *A. vinelandii* flavodoxin. Measurements are reported at different pD values, which allows the determination of local opening and closing rates within holoflavodoxin. Furthermore, the holoflavodoxin data presented here are combined with the recently obtained insight into the folding kinetics of apoflavodoxin and into the subglobal unfolding processes that take place within apoflavodoxin. As a result, a detailed picture of the influence of ligand binding on the folding and stability of *A. vinelandii* flavodoxin emerges.



## Materials and Methods

### Materials

Potassium pyrophosphate was from Sigma (Bornem, Belgium).  $^{15}\text{N}$  Ammonium chloride was purchased from Campro Scientific, Veenendaal, The Netherlands.

### Protein purification

Uniformly  $^{15}\text{N}$ -labelled recombinant C69A *A. vinelandii* holoflavodoxin was obtained as described previously (Steensma et al. 1998; van Mierlo et al. 1998).

### pD dependent H/D exchange of *A. vinelandii* holoflavodoxin

H/D Exchange experiments were started by dissolving lyophilised samples of uniformly  $^{15}\text{N}$ -labelled holoflavodoxin in  $\text{D}_2\text{O}$  to a final protein concentration of 2.5 to 3 mM. The samples were immediately transferred into the NMR machine, and a series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded at 25 °C. A similarly prepared sample without containing protein was used to shim the magnetic field, tune the probe, and to determine the length of the 90° proton pulse prior to the start of the H/D exchange experiment. The dead-time of the exchange experiment, i.e. the difference in time between the mixing of  $\text{D}_2\text{O}$  with lyophilised protein and the recording of the first HSQC spectrum, was approximately 5 minutes. The samples contained 100 mM potassium pyrophosphate at various pD values and 100  $\mu\text{M}$  DSS. The pD of the sample was measured using a glass electrode after the exchange experiment was completed. Experiments were done at pD values of 6.19, 6.26, 7.73 and 8.82 (uncorrected pH-meter readings). The experiments at pD 6.26 and 8.82 lasted for 11 months. The experiments at pD 6.19 and 7.73 were stopped after 2 months due to the observation of unwanted proteolysis taking place. In all cases, a few intense  $^1\text{H}$ - $^{15}\text{N}$  NMR cross peaks were still present in the final HSQC spectra acquired. As exchange of these protons was far from complete, this implies that the exchange rates of the most slowly exchanging amide protons could not be determined accurately.

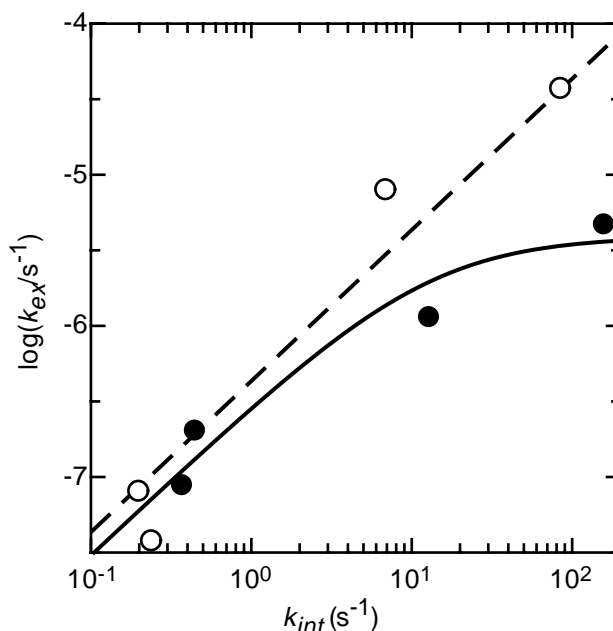
### NMR spectroscopy and data processing

Gradient-enhanced  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (Palmer III et al. 1991; Kay et al. 1992) were recorded at 25 °C on a Bruker AMX500 spectrometer operating at a proton frequency of 500.13 MHz as described previously (van Mierlo et al. 2000). In all exchange experiments, first 9 HSQC spectra were recorded with 1024 complex data points in the direct  $^1\text{H}$  dimension and 96 complex data points in the indirect  $^{15}\text{N}$  dimension with two scans per increment. Each HSQC spectrum took 8 minutes and 40 seconds to acquire. As a result, a reasonably high sampling rate in the beginning of the exchange experiment could be obtained. Subsequently, a series of spectra with identical settings but with an increased number of complex data points (i.e. 256) in the indirect dimension was recorded. Each HSQC spectrum now lasted 22 minutes and 46 seconds before completion and has a higher spectral resolution than the former HSQC spectra.

The NMR data were processed on a Silicon Graphics Octane workstation using the program Felix 2D (MSI, San Diego, CA, USA). In the direct  $^1\text{H}$  dimension, data were multiplied by a Gaussian function and zerofilled to 2048 points prior to Fourier transform. In the indirect  $^{15}\text{N}$  dimension data were zerofilled once (to 192 points if 96 complex points were recorded, or to 512 points in case 256 complex points were recorded) and multiplied by a squared cosine bell prior to Fourier transform. Spectra were referenced to internal DSS. Peak heights were automatically determined using the macros and scripts provided by A. G. Palmer III (Palmer III), and a single exponential function was fitted to the data using the program CurveFit (Palmer III) to obtain the exchange rate constants  $k_{ex}$ . Assignments of most of the NMR cross peaks were adapted from (Steensma et al. 1998) to the pD and temperature values used in the present study.

## Results and discussion

The H/D exchange rates of those backbone amide protons of *A. vinelandii* holoflavodoxin that exchange sufficiently slowly to be followed by NMR spectroscopy are determined at four different pD values. Equation 5 is fitted to the pD-dependence of the H/D exchange rate constant ( $k_{ex}$ ) of those amide protons for which EX2 exchange behaviour was observed over the entire pD range studied (i.e. pD 6.19 to 8.82), as is shown for Leu34 in Figure 3. Stabilities against local structural opening of holoflavodoxin ( $\Delta G_{op}$ ) are calculated using the  $K_{op}$  values obtained from the fit results and equation 6. In case the pD dependent



**Figure 3.** Dependence of the detected backbone amide H/D exchange rates ( $k_{ex}$ ) of Leu34 (O) and Thr56 (●) of *A. vinelandii* holoflavodoxin on the intrinsic exchange rate ( $k_{int}$ ). In case  $k_{ex}$  increases tenfold with an increase of pD by one unit, as seen for Leu34, the stability against local unfolding of the micro-environment of a specific backbone amide proton is determined by fitting equation 5 to the pD-dependent  $k_{ex}$  data. The result of the latter fit to the Leu34 data is shown as a dashed line. In case the dependence of  $k_{ex}$  on  $k_{int}$  curves, as is the case here for the Thr56 data, rate constants for local opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of the micro-environment around this specific backbone amide proton are determined by fitting equation 3 to the data (solid line). The data are obtained from 2.5 to 3 mM holoflavodoxin in 100 mM potassium pyrophosphate at 25 °C.

$\log(k_{ex})$  data of a specific amide proton deviate from a straight line with a slope of unity, as shown for Thr56 in Figure 3, equation 3 is fitted to these data. The latter fit gives rate constants for local opening ( $k_{op}$ ) and closing ( $k_{cl}$ ) of the micro-environment around this specific amide proton and the corresponding  $\Delta G_{op}$  is determined from the ratio of  $k_{op}$  and  $k_{cl}$ , according to equation 6. The  $k_{op}$ -,  $k_{cl}$ - and  $\Delta G_{op}$ -values determined for the backbone amide protons of *A. vinelandii* holoflavodoxin are listed in Table 1. In addition, when available, the corresponding values of apoflavodoxin are also reported in Table 1 as well as the local stability difference at the residue-level between apo- and holoflavodoxin ( $\Delta\Delta G$ ).

**Table 1. Dependence of the backbone amide proton H/D exchange rates ( $k_{ex}$ ) of *A. vinelandii* holoflavodoxin on pD, fitted local opening and closing rates ( $k_{op}$  and  $k_{cl}$ , respectively), and stabilities ( $\Delta G_{op}$ ) against local unfolding of the protein.<sup>a</sup>**

Residue number	$k_{ex}(6.19)$	$k_{ex}(6.26)$	$k_{ex}(7.73)$	$k_{ex}(8.82)$
3	8.29E-07	1.91E-06		
5		4.07E-09	6.56E-07	2.90E-06
6		1.97E-09		1.15E-08
7				3.46E-08
8		7.87E-09	8.89E-07	2.79E-06
13	8.13E-04	4.37E-04	1.98E-04	
14	4.60E-06	4.33E-06	3.19E-05	2.11E-04
15		5.44E-07	4.17E-05	1.56E-04
16	4.70E-07	7.16E-07	2.42E-04	
17		2.56E-08	1.02E-05	4.65E-05
19		1.45E-08	2.85E-06	1.80E-05
21		4.29E-08	1.22E-05	8.65E-05
22	6.85E-08	7.68E-08	3.26E-05	1.97E-04
23	1.80E-06	4.88E-06		
24	1.34E-04	4.12E-04		
25	1.94E-06	4.64E-06		
29	1.52E-03			
30	1.34E-05	3.93E-05		
31	1.56E-04	4.50E-04		
33	8.48E-04	9.87E-04		
34	8.11E-08	3.81E-08	8.03E-06	3.75E-05
37	2.11E-05	5.89E-05		
38	7.67E-04	1.85E-03		
39	8.29E-05	2.58E-04		
45	1.06E-05	3.27E-05		
46	4.49E-04			
47	4.89E-05	2.75E-04		
49	1.70E-06	4.14E-06		
50		1.69E-09		1.08E-08
51		2.23E-09		8.87E-09
52		7.11E-10		6.29E-09
53		1.53E-09		1.04E-08
54		2.03E-09		2.64E-08
56	8.90E-08	2.04E-07	1.15E-06	4.73E-06
57	6.85E-04	6.57E-04		
58	4.76E-04	1.01E-03		
60	6.15E-04	8.64E-04		
61		3.98E-04		
64	1.54E-03			
65	8.32E-04	1.27E-03		

Table 1 continued

Residue nr	$k_{op}$	$k_{cl}$	$\Delta G_{op}$	$\Delta G_{apo}$	$\Delta\Delta G$
3			7.07 $\pm$ 0.32	6.42	0.65
5			10.30 $\pm$ 0.50	6.82	3.48
6	1.16E-08	1.57	11.04	7.54	3.50
7			13.29		
8			11.13 $\pm$ 0.46	5.36	5.77
13					
14	1.67E-04 $\pm$ 1.12E-04	41.75 $\pm$ 37.28	7.33 $\pm$ 1.12		
15			8.68 $\pm$ 0.37		
16			8.09 $\pm$ 0.86		
17			8.63 $\pm$ 0.76	4.54	4.09
19			9.91 $\pm$ 0.56	6.44	3.47
21			8.57 $\pm$ 0.69	6.29	2.29
22			8.54 $\pm$ 0.68	6.45	2.09
23			7.50 $\pm$ 0.41		
24			5.09 $\pm$ 0.47	4.43	0.66
25			7.36 $\pm$ 0.34	5.98	1.38
29			3.36		
30			6.44 $\pm$ 0.44	5.67	0.77
31			5.40 $\pm$ 0.44		
33			3.77 $\pm$ 0.02		
34			8.65 $\pm$ 0.41		
37			6.41 $\pm$ 0.42		
38			4.40 $\pm$ 0.35		
39			4.44 $\pm$ 0.47		
45			6.39 $\pm$ 0.47	6.04	0.35
46			4.46	4.64	-0.18
47			5.14 $\pm$ 0.77	5.53	-0.39
49			7.40 $\pm$ 0.35	6.86	0.54
50	7.59E-09 $\pm$ 2.01E-09	0.84 $\pm$ 0.44	10.93 $\pm$ 0.59	7.31	3.61
51	7.59E-09 $\pm$ 2.01E-09	0.84 $\pm$ 0.44	10.93 $\pm$ 0.59	7.09	3.84
52	7.59E-09 $\pm$ 2.01E-09	0.84 $\pm$ 0.44	10.93 $\pm$ 0.59	7.26	3.67
53	7.59E-09 $\pm$ 2.01E-09	0.84 $\pm$ 0.44	10.93 $\pm$ 0.59	7.84	3.09
54					
56	3.90E-06 $\pm$ 2.09E-06	12.79 $\pm$ 9.36	8.85 $\pm$ 0.91		
57			3.68 $\pm$ 0.11		
58			4.26 $\pm$ 0.28		
60			4.31 $\pm$ 0.08		
61			4.10		
64			3.69		
65			3.39 $\pm$ 0.12		

**Table 1 continued**

Residue number	$k_{ex}(6.19)$	$k_{ex}(6.26)$	$k_{ex}(7.73)$	$k_{ex}(8.82)$
73	2.52E-04	1.97E-04		
74	1.90E-04	2.04E-04		
75		2.43E-04		
76		1.73E-05		
77	5.70E-07	7.53E-07	1.59E-04	
80	5.20E-06	6.63E-06		
81		9.14E-09	2.79E-06	1.36E-05
82	4.97E-06	1.03E-05		
84	7.97E-06	2.38E-05		
89	1.04E-06	2.32E-06		
91		6.74E-09	2.20E-06	9.28E-06
92				6.50E-09
93		6.93E-09	9.72E-07	4.42E-06
94		1.46E-09		8.76E-09
95		2.18E-09		2.12E-07
96				1.67E-07
98	1.18E-07	2.49E-07	3.33E-07	3.31E-07
99		2.25E-08	3.35E-06	2.15E-05
100			2.63E-06	1.19E-05
101	2.70E-06	2.32E-06	1.03E-04	
102		1.25E-07	2.88E-07	4.65E-07
105	3.77E-05	8.01E-05		
106		3.91E-08	9.35E-06	5.45E-05
108				5.68E-07
110		2.15E-08	4.43E-06	2.20E-05
111		3.18E-08	6.97E-06	4.52E-05
112	1.46E-04	2.76E-04		
113	1.56E-07	3.50E-08		
114				6.35E-08
118		2.74E-08	1.01E-05	6.24E-05
119	2.95E-04	3.85E-04		
120	1.41E-06	2.83E-06		
121	5.24E-04			
123	1.45E-06	3.49E-06		
124	5.47E-04	1.83E-03		
125	3.75E-07	5.16E-07	1.80E-04	
128	4.92E-07	1.02E-06	4.52E-04	
131	9.23E-04	1.71E-03		
133		2.56E-04		
134	1.28E-04	3.78E-04		
135	8.37E-04	1.81E-03		
137	6.04E-05	1.53E-04		
139	2.67E-04	2.14E-04		
141	4.23E-07	7.86E-07	4.15E-04	

Table 1 continued

Residue nr	$k_{op}$	$k_{cl}$	$\Delta G_{op}$	$\Delta G_{apo}$	$\Delta\Delta G$
73			5.14 ± 0.22		
74			4.77 ± 0.06		
75			4.01		
76			5.51		
77			7.35 ± 0.68		
80			6.61 ± 0.03		
81			9.29 ± 0.69	6.81	2.47
82			5.87 ± 0.27	4.93	0.94
84			5.93 ± 0.45	5.38	0.55
89			7.95 ± 0.31	7.12	0.83
91			9.62 ± 0.70	7.96	1.66
92	7.59E-09 ± 2.01E-09	0.84 ± 0.44	10.93 ± 0.59	7.87	3.06
93			9.64 ± 0.43	7.49	2.15
94				7.76	
95			12.54 ± 0.65	6.51	6.02
96			12.05		
98	3.43E-07 ± 8.20E-08	0.63 ± 0.42	8.51 ± 0.71		
99			9.65 ± 0.48		
100			9.12 ± 0.50		
101			7.62 ± 0.13		
102	3.84E-07 ± 8.46E-08	1.52 ± 0.85	8.96 ± 0.60		
105			6.13 ± 0.28		
106			9.39 ± 0.62		
108			11.2		
110			8.97 ± 0.57	5.19	3.78
111			9.53 ± 0.60	6.78	2.75
112			4.44 ± 0.22		
113			8.58 ± 0.84	5.8	2.78
114			12.53	7.99	4.54
118			9.37 ± 0.76	8.34	1.03
119			4.36 ± 0.04		
120			7.49 ± 0.26	7.06	0.43
121			4.95		
123			7.50 ± 0.35	7.07	0.43
124			3.13 ± 0.51		
125			6.75 ± 0.85	6.69	0.07
128			7.50 ± 1.01	4.69	2.81
131			3.69 ± 0.22		
133			4.68		
134			4.23 ± 0.45		
135			3.31 ± 0.30		
137			5.64 ± 0.37		
139			4.52 ± 0.20		
141			6.79 ± 1.05	4.04	2.76

**Table 1 continued**

Residue number	$k_{ex}(6.19)$	$k_{ex}(6.26)$	$k_{ex}(7.73)$	$k_{ex}(8.82)$
142	3.11E-06	8.76E-06		
143	7.45E-08			
145		4.01E-04		
146		1.43E-08	5.63E-06	3.13E-05
147		6.92E-09	2.34E-06	1.16E-05
148		3.47E-09	1.46E-07	2.89E-07
150		3.04E-09	4.00E-07	2.91E-07
151		3.96E-09		
153	2.59E-04	9.93E-04		
154		7.54E-04		
155	4.90E-05	1.31E-04		
163	3.85E-06	1.09E-05		
167			1.31E-07	5.12E-07
168				1.19E-07
169	1.47E-05	4.63E-08	2.06E-05	9.20E-05
170	1.83E-07	2.87E-07	1.66E-04	
171		6.27E-09	4.26E-06	2.53E-05
172		2.57E-07		
174	1.08E-03			
175	1.26E-06	2.60E-06		
177	5.12E-04	1.57E-03		



**Table 1 continued**

Residue nr	$k_{op}$	$k_{cl}$	$\Delta G_{op}$	$\Delta G_{apo}$	$\Delta\Delta G$
142			5.92 $\pm$ 0.43	3.83	2.09
143			8.92		
145			4.73	4.43	0.30
146			9.56 $\pm$ 0.77	4.77	4.79
147			9.37 $\pm$ 0.72	4.76	4.61
148	4.04E-07 $\pm$ 2.46E-07	106.9 $\pm$ 89.45	11.44 $\pm$ 1.03	4.89	6.55
150			11.18 $\pm$ 0.82	4.80	6.38
151			10.56	4.76	5.80
153			3.33 $\pm$ 0.58		
154			3.49		
155			5.87 $\pm$ 0.40		
163			6.84 $\pm$ 0.43		
167			11.05 $\pm$ 0.57	7.94	3.12
168			11.87	7.58	4.29
169			8.30 $\pm$ 0.51	7.70	0.60
170			8.37 $\pm$ 1.05	7.57	0.80
171			9.26 $\pm$ 0.95	7.40	1.87
172			7.77 $\pm$ 1.38	8.61	-0.03
174			2.84		
175			7.16 $\pm$ 0.27		
177			3.47 $\pm$ 0.47	3.20	0.27

<sup>a</sup> The backbone amide proton H/D exchange rates  $k_{ex}$  are determined on samples containing 2.5 to 3 mM holoflavodoxin in 100 mM potassium pyrophosphate at 25 °C at the pD value indicated between brackets (i.e. kex(6.26) are the kex values determined in a H/D exchange experiment at pD 6.26). Rate constants for exchange ( $k_{ex}$ ), local opening ( $k_{op}$ ) and local closing ( $k_{cl}$ ) are in s<sup>-1</sup>. The stabilities at the residue level against local opening ( $\Delta G_{op}$ ) as well as the corresponding local stabilities in apoflavodoxin ( $\Delta G_{apo}$ ) and the differences in local stabilities between holo- and apoflavodoxin ( $\Delta\Delta G$ ) are in kcal/mol. The standard fitting errors are shown. In cases where too few data points are available to meaningfully determine a standard fitting error, the latter error is not shown.

### The global stability of *A. vinelandii* holoflavodoxin can not be determined from pD-dependent H/D exchange data

As a result of the influence of the non-covalently bound FMN, the stability against global unfolding of holoflavodoxin is higher than the one characterising apoflavodoxin. The magnitude of the stability difference between holo- and apoflavodoxin depends on the free FMN concentration in the holoflavodoxin sample. This is the result of FMN shifting the equilibrium between apo- and holoflavodoxin towards the holo form, as is expressed in equation 8:

$$\Delta G_{holo} = \Delta G_{apo} + RT \ln \left( 1 + \frac{[FMN]}{K_D} \right) \quad (8)$$

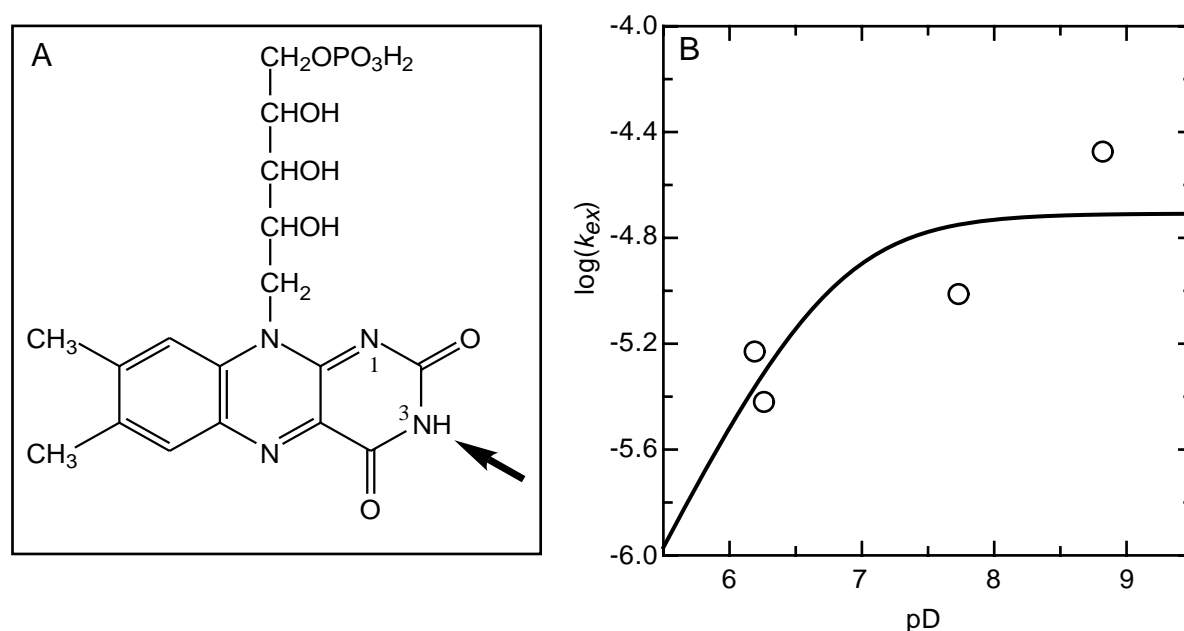
in which  $\Delta G_{holo}$  is the stability against global unfolding of holoflavodoxin,  $\Delta G_{apo}$  is the stability against global unfolding of apoflavodoxin (i.e.  $10.45 \pm 0.52$  kcal/mol, Chapter 2 of this thesis),  $R$  the gas constant,  $T$  the absolute temperature,  $[FMN]$  the concentration of free FMN in solution, and  $K_D$  the dissociation constant of the apoflavodoxin-FMN complex (i.e.  $[FMN][Apo]/[Holo] = (3.4 \pm 0.4) \cdot 10^{-10}$  M, see Chapter 5 of this thesis). In the study presented, no free FMN is added to the NMR samples. As a consequence, the concentration of free FMN in solution is determined by the total protein concentration and by the magnitude of the dissociation constant. In a typical H/D exchange experiment, the total flavodoxin concentration is 2.75 mM, which results in a free FMN concentration of  $0.97 \pm 0.01$   $\mu$ M. Under this circumstance the global stability of holoflavodoxin according to equation 8 is 15.15 kcal/mol, which is 4.7 kcal/mol higher than the global stability of apoflavodoxin.

In many proteins, the local stability of the most slowly exchanging amide protons equals the global stability of the protein (Huyghues-Despointes et al. 1999). However, in case of *A. vinelandii* apoflavodoxin under native conditions not a single amide proton is detected to exchange solely from the unfolded state of the protein (Chapter 4 of this thesis). All amide protons of apoflavodoxin exchange via local or subglobal unfolding processes. Only when global unfolding is promoted by the addition of denaturant ( $> 0.5$  M GuHCl), some amide protons exchange mainly from unfolded apoflavodoxin (Chapter 4). Consequently, the global stability of apoflavodoxin in absence of denaturant can not be determined from H/D exchange data.

Even if sites would exist in holoflavodoxin where local and subglobal unfolding processes do not contribute to H/D exchange, the local stabilities determined at these sites will not match the stability against global unfolding of holoflavodoxin. The unavoidable occasional release of FMN from holoflavodoxin causes H/D exchange of backbone amide protons at these sites to be the case via local or subglobal unfolding processes, just as occurs in apoflavodoxin. As a result, the global stability of *A. vinelandii* holoflavodoxin can not be determined from H/D exchange data. Indeed, none of the experimentally determined  $\Delta G_{op}$  values of holoflavodoxin corresponds within error to the calculated global stability of 15.15 kcal/mol (Table 1).

### The N(3)H of FMN is inaccessible to water when FMN is bound to *A. vinelandii* flavodoxin

The N(3)H proton of oxidised FMN (Figure 4A) can exchange with water. The corresponding H/D exchange rate informs about the water accessibility of the FMN cofactor when it is bound to *A. vinelandii* flavodoxin. Unfortunately, the intrinsic exchange rate of the N(3)H proton is not known. Consequently, the N(3)H exchange rates can not be transferred into free energy values. However, in case the N(3)H proton exchanges under EX1 conditions the corresponding rate equals the rate for opening of the local micro-environment around N(3)H. This opening rate can be determined without any knowledge about the intrinsic exchange rate of N(3)H by using equation 4.



**Figure 4.** The N(3)H of FMN in *A. vinelandii* holoflavodoxin is indicated by an arrow in the FMN structure, and can exchange with water (A). Fitting of equation 3 to the dependence of the H/D exchange rate constant of this amide proton on pD (B) leads to  $k_{op} = (2 \pm 1) 10^{-5} \text{ s}^{-1}$ , and  $k_{cl}/k_{int} = 3 \pm 2 \text{ s}^{-1}$ . The data are obtained from 2.5 to 3 mM holoflavodoxin in 100 mM potassium pyrophosphate at 25 °C.

Figure 4B shows the H/D exchange behaviour of N(3)H at different pD values. The exchange rate of N(3)H does not increase tenfold with every increase of pD by one unit in the pD range studied. Instead, above pD 7 the exchange rate levels off upon increasing pD. This switch from EX2 to EX1 amide proton exchange behaviour allows the determination of the opening rate ( $k_{op}$ ) of the process that makes the N(3)H proton of FMN accessible to the solvent. Fitting of equation 3, with  $k_{int}$  fixed arbitrarily to  $1 \text{ s}^{-1}$ , to the data of Figure 4B leads to  $k_{op} = (2 \pm 1) 10^{-5} \text{ s}^{-1}$ . Fixing  $k_{int}$  to a different value would lead to a different  $k_{cl}$ -value, but instead the value found for  $k_{op}$  is unaffected as it does not depend on  $k_{int}$ .

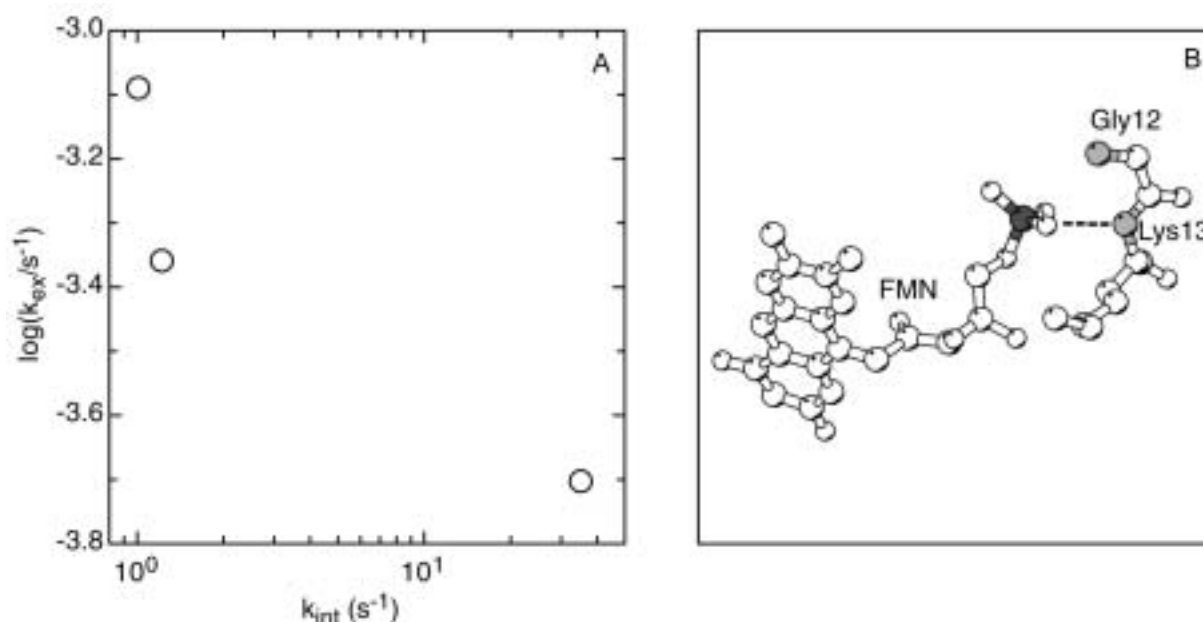
Based on holoflavodoxin unfolding rates at guanidinium chloride concentrations exceeding 3.5 M, with a protein concentration of 1  $\mu\text{M}$ , the rate constant for flavin release

from *A. vinelandii* holoflavodoxin is  $8.88 \pm 0.02 \cdot 10^{-5} \text{ s}^{-1}$  (Chapter 5 of this thesis). The latter holoflavodoxin concentration is more than 3 orders of magnitude lower than the protein concentrations used in the present H/D exchange experiments. The rate constant for flavin release of  $8.88 \pm 0.02 \cdot 10^{-5} \text{ s}^{-1}$  and the  $k_{op}$  value for N(3)H exchange of  $(2 \pm 1) \cdot 10^{-5} \text{ s}^{-1}$  are in good agreement regarding the large differences in sample conditions between the two experiments.

In summary, we conclude that the N(3)H proton of the FMN cofactor is inaccessible to water when FMN is bound to *A. vinelandii* flavodoxin. Exchange of the N(3)H proton with solvent deuterons only takes place once FMN is released from holoflavodoxin. The solvent inaccessibility of the nonpolar environment around N(3) could at least in part establish the low redox potential of *A. vinelandii* holoflavodoxin (i.e. -485 mV, (Steensma et al. 1996)).

### The backbone amide H/D exchange rate of Lys13 decreases with increasing pD

The backbone amide proton H/D exchange rate of Lys13 is the only exchange rate of *A. vinelandii* holoflavodoxin that decreases with increasing pD (Figure 5A). This behaviour can not be explained by H/D exchange theory. The backbone amide proton of Lys13 is hydrogen bonded to the phosphate moiety of the FMN cofactor (Figure 5B). Probably the observed decrease of the H/D exchange rate of this amide proton with increasing pD is caused by a deprotonation of the phosphate group of FMN. The pK of this phosphate group (mono-anion to di-anion) of FMN within *A. vinelandii* holoflavodoxin is reported to be 5.5 (MacKnight et al. 1973). In D<sub>2</sub>O this pK-value may be slightly different. Apparently, the deprotonation of the FMN phosphate results in a stronger hydrogen bond between Lys13 and the now di-anionic phosphate group, which leads to a higher stability against H/D exchange.



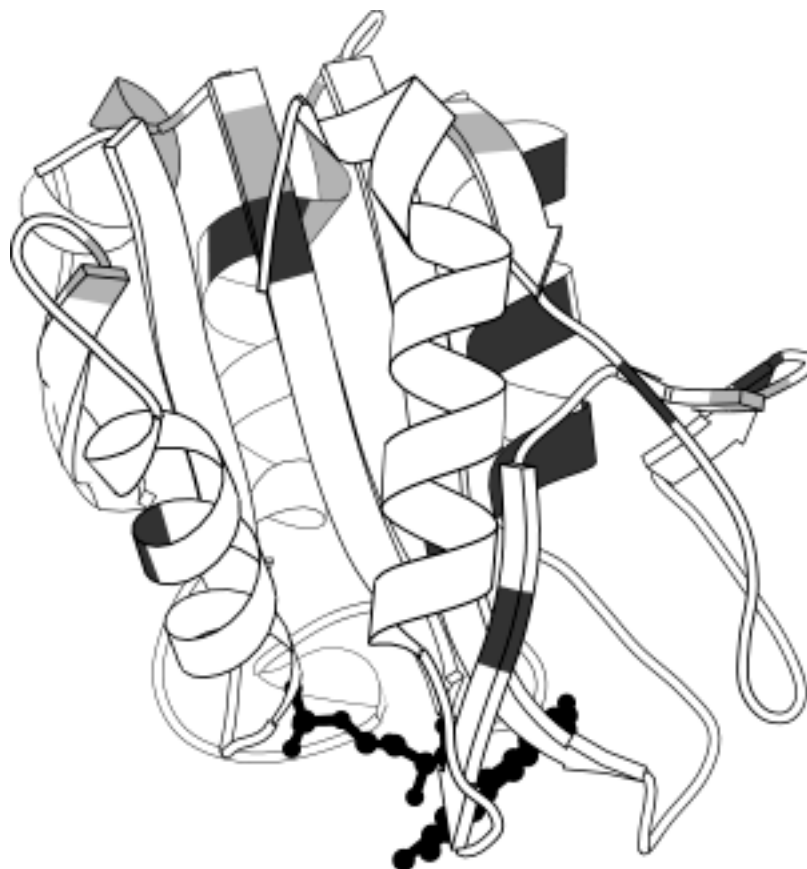
**Figure 5.** (A) Dependence of the backbone amide proton H/D exchange rate,  $k_{ex}$ , of Lys13 of *A. vinelandii* holoflavodoxin on its intrinsic exchange rate ( $k_{int}$ ). Remarkably, Lys13 is the only residue of holoflavodoxin the  $k_{ex}$  of its backbone amide proton decreases with increasing  $k_{int}$  ( $k_{int}$  is increased by increasing the pD of the solution). The data are obtained from 2.5 to 3 mM holoflavodoxin in 100 mM potassium pyrophosphate at 25 °C. (B) The backbone amide of Lys13 in *A. vinelandii* holoflavodoxin is hydrogen bonded to the phosphate group of the FMN cofactor, as indicated by the dashed line. The graph is generated using Molscrip (Kraulis 1991).

This stronger hydrogen bond can not contribute significantly to the apoflavodoxin-FMN binding energy. If the latter would be the case, besides the amide proton of Lys13 many amide protons would display a decrease of  $k_{ex}$  with increasing pD, which is clearly not observed (Table 1).

**Cofactor binding influences the dynamics of almost all residues the amide protons of which exchange via local processes in *A. vinelandii* flavodoxin**

In holoflavodoxin, (sub)global protein unfolding is expected to occur less frequently than in apoflavodoxin due to the increased global stability of holoflavodoxin compared to apoflavodoxin. In this section, focus lies on the amide protons in holoflavodoxin that exchange via local opening processes in apoflavodoxin (Chapter 4 of this thesis). These amide protons most likely exchange mainly via local opening processes in holoflavodoxin as well. If this is the case, the differences in local stabilities of these backbone amide protons between holo- and apoflavodoxin inform about the extent to which FMN binding influences the dynamics of the protein the local opening processes in flavodoxin.

Figure 6 shows the free energy differences between apo- and holoflavodoxin for those residues the backbone amide protons of which exchange via local opening processes in



**Figure 6.** Molscript cartoon drawing (Kraulis 1991) of *A. vinelandii* holoflavodoxin as presented in Figure 1. Residues of *A. vinelandii* flavodoxin the backbone amide protons of which exchange via local opening processes in apoflavodoxin are grey shaded. Light grey is used when the stability against local unfolding increases by less than 1 kcal/mol upon FMN binding, and dark grey is used when this value increases by more than 1 kcal/mol upon FMN binding.

apoflavodoxin. None of the residues of holoflavodoxin highlighted in Figure 6 have micro-environments with significantly lower stabilities against local opening than as observed in apoflavodoxin. FMN binding thus does not lead to an increase in the local dynamics within *A. vinelandii* flavodoxin. In case of some residues, the difference in local stability between holo- and apoflavodoxin is small (i.e. < 1 kcal/mol more stable in holoflavodoxin) and FMN binding only marginally diminishes the local dynamics at these sites. These sites are all at the periphery of the flavodoxin structure and distant from the FMN binding site (Figure 6). The stabilities against local unfolding of the remaining backbone amide protons that exchange via local opening processes all increase substantially (i.e. by more than 1 kcal/mol) upon binding of FMN to apoflavodoxin.

### **The influence of FMN binding on subglobal unfolding processes in *A. vinelandii* flavodoxin**

Four PUFs that can be accessed by *A. vinelandii* apoflavodoxin have been identified by H/D exchange measurements. Determination of the H/D exchange rates of the corresponding residues in holoflavodoxin can reveal whether these PUFs are also formed by holoflavodoxin. By doing so, the influence of FMN binding on the free energy landscape for flavodoxin folding is deciphered.

#### *a. The formation of PUF1 from native holoflavodoxin*

Residues the amide protons of which become water accessible in PUF1 of apoflavodoxin are coloured yellow in Figure 3 of Chapter 4 (i.e. Ala140, Val141, Val142). In absence of denaturant, only the backbone amide proton exchange behaviour of Val141 characterises PUF1 (Chapter 4 of this thesis). The backbone amide protons of Ala140 and Val142 exchange via local structural openings and the addition of denaturant causes these amide protons to exchange from PUF1 (see Chapter 4 of this thesis for details). The stability against local opening of Val141 is 2.76 kcal/mol higher in holoflavodoxin compared to in apoflavodoxin. This stability increase is substantially lower than the increase in global flavodoxin stability of 4.7 kcal/mol upon FMN binding. Holoflavodoxin can unfold to PUF1 just as apoflavodoxin does, although it apparently happens less frequently. However, the backbone amide proton of Val 141 may not exchange via subglobal unfolding of holoflavodoxin, but instead via a local conformational fluctuation. The stability against local opening of the protein structure surrounding Val141 may be such that it is larger than the free energy difference between native apoflavodoxin and PUF1 ( $\Delta G_{apo-PUF1}$ ), but smaller than the free energy difference between native holoflavodoxin and PUF1 ( $\Delta G_{holo-PUF1}$ ). If this is the case,  $\Delta G_{holo-PUF1}$  is larger than the local stability of Val141 in holoflavodoxin. Whether in holoflavodoxin local or subglobal unfolding events determine the amide proton exchange of residues of PUF1 of apoflavodoxin will become clear by the future study of the denaturant concentration dependence of the corresponding local stabilities in holoflavodoxin.

#### *b. The formation of PUF2 from native holoflavodoxin*

Residues the amide protons of which become water accessible in PUF2 of apoflavodoxin are coloured green in Figure 3 of Chapter 4 (i.e. Lys145, Phe146, Val147,

Gly148, Leu149, Ala150, Leu151). In case of apoflavodoxin in the absence of denaturant, five of these residues, i.e. Phe146, Val147, Gly148, Leu149 and Ala150, display the characteristic backbone amide proton exchange behaviour from PUF2. The backbone amide protons of the remaining residues of PUF2 exchange mainly via local opening processes, unless subglobal unfolding of apoflavodoxin to PUF2 is sufficiently promoted by the addition of denaturant.

The amide protons of residues 146 to 150 exchange very slowly in holoflavodoxin. The intensities of the corresponding cross peaks in the HSQC spectra decrease only 10 to 30 % in a 11 month period of exchange in D<sub>2</sub>O at pD 6.26. Consequently, the exchange rates of these amide protons can not be determined accurately. The stability against subglobal unfolding of the part of holoflavodoxin surrounding these residues is at least 4.6 kcal/mol higher than it is in apoflavodoxin. This stability increase corresponds to the global stability difference between holoflavodoxin and apoflavodoxin under the experimental conditions used, which is due to FMN binding to apoflavodoxin. PUF2 must thus be inaccessible to holoflavodoxin. Flavodoxin can form the partially unfolded structure PUF2 only once FMN is released.

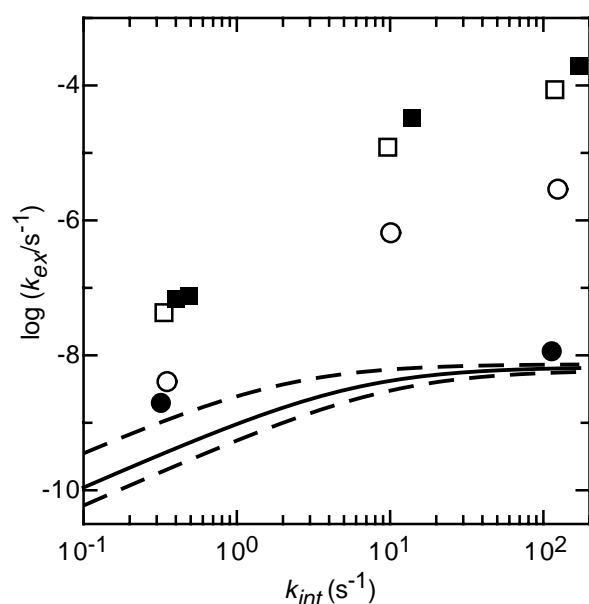
Some of the residues the amide protons of which become solvent accessible in PUF2 are part of a 23-residue loop that is inserted into  $\beta$ -strand 5 of *A. vinelandii* flavodoxin. This insertion is characteristic for the class of long-chain flavodoxins. In case of *A. vinelandii* flavodoxin the loop comprises residues Gly126 to Gly148. The function of this 23-residue insertion is unknown. The insertion is proposed to enhance the stability of holoflavodoxin by enabling the 100's loop (i.e. amino acids 100 to 106) to form a rigid hydrogen bonding network with the FMN cofactor (Steensma et al. 1998). The insertion contains a small  $\beta$ -sheet (Figure 1). In PUF1 of apoflavodoxin, one strand of this  $\beta$ -sheet is unfolded. In PUF2, the remainder of this  $\beta$ -sheet unfolds, together with the second half of  $\beta$ -strand 5 of the central  $\beta$ -sheet of apoflavodoxin (Figure 3 of Chapter 4).

The residues of the 23-residue insertion that are not part of the small  $\beta$ -sheet structure discussed and do not belong to any of the PUFs identified in apoflavodoxin, i.e. residues 129 to 139, have no regular secondary structure (Figure 3 of Chapter 4). In apoflavodoxin their backbone amide protons exchange too rapidly to be detected by NMR spectroscopy, which implies that these residues reside in a highly unstable region of the protein. In holoflavodoxin however, the backbone amide exchange of these residues shows that now the local stability against unfolding of this region varies between 4 and 5 kcal/mol. This matches well with the stability gain due to FMN binding to apoflavodoxin of 4.7 kcal/mol. Apparently, the entire loop formed by residues 129 to 139 is water inaccessible in holoflavodoxin. The corresponding backbone amide protons exchange only upon release of the FMN cofactor. The only exception to the latter forms the backbone amide proton of Phe135, which has a local stability of only 3.31 kcal/mol in holoflavodoxin. Possibly Phe135 is locally dynamic in holoflavodoxin causing the low local stability observed. The hydrogen exchange results obtained indicate that the loop formed by the residues 129 to 139 plays an important role in the co-operative FMN binding to *A. vinelandii* apoflavodoxin.

### c. The formation of PUF3 from native holoflavodoxin

Residues the amide protons of which become water accessible in PUF3 of apoflavodoxin are coloured blue in Figure 3 of Chapter 4 (i.e. Ile3, Gly4, Leu5, Phe6, Gly8, Val17, Ala18, Lys19, Ile21, Lys22, Arg24, Met30, Glu82, Gly83, Trp128). PUF3 is a partially folded apoflavodoxin species that resides on a non-productive folding route starting from unfolded apoflavodoxin (Figure 2). Native flavodoxin has to unfold globally before this non-productive folding route can be entered. Consequently, due to the influence of FMN binding, the difference in free energy between native holoflavodoxin and PUF3 is expected to be 4.7 kcal/mol higher than the free energy difference between native apoflavodoxin and PUF3.

In apoflavodoxin, five amide protons, i.e. those of Gly4, Leu5, Phe6, Ile21, and Lys22, display the H/D exchange behaviour characteristic for exchange from PUF3 in absence of denaturant. As the cross peak associated with the backbone amide of Gly4



**Figure 7.** Dependence of the H/D exchange rates,  $k_{ex}$ , of the backbone amide protons of Leu5 (O), Phe6 (●), Ile21 (□) and of Lys22 (■) of *A. vinelandii* holoflavodoxin on the intrinsic exchange rate,  $k_{int}$ . The theoretically expected dependence (as described by equation 3) of  $k_{ex}$  on  $k_{int}$  for amide protons that exchange from PUF3 is plotted as a solid line, with the dashed lines indicating the standard deviation. The data are obtained from 2.5 to 3 mM holoflavodoxin in 100 mM potassium pyrophosphate at 25 °C.

overlaps with another cross peak in the HSQC spectrum of holoflavodoxin, the amide proton exchange rate of Gly4 can not be determined in holoflavodoxin. The backbone amide protons of Ile21 and Lys22 have H/D exchange rates that show that the corresponding local stabilities increase relative to apoflavodoxin by only 2.3 and 2.1 kcal/mol, respectively. Consequently, the amide protons of Ile21 and Lys22 must exchange via local opening processes that do not reflect the transition of native holoflavodoxin to PUF3. H/D exchange shows that the stabilities of the remaining residues that exchange from PUF3 in apoflavodoxin in absence of denaturant (i.e. residues Leu5 and Phe6) are 3.5 kcal/mol higher in holoflavodoxin compared to in apoflavodoxin. This increase is reasonably close to the expected stability difference of 4.7 kcal/mol and thus Leu5 and Phe6 seem to exchange from PUF3 in holoflavodoxin.

The dependence of the exchange rates of the amide protons of Leu5 and Phe6 on pD can give additional information about whether these amide protons indeed exchange from PUF3 in holoflavodoxin. Figure 7 shows  $k_{ex}$  as a function of the intrinsic exchange rate  $k_{int}$  for



the backbone amide protons of Leu5, Phe6, Ile21 and Lys22. The rate constants for interconversion between native apoflavodoxin and PUF3 are taken from the pD dependent H/D exchange study of apoflavodoxin (Chapter 4 of this thesis). Use of these rate constants together with those for FMN binding to apoflavodoxin and for FMN release from holoflavodoxin enables the calculation of the expected rate constants for interconversion between PUF3 and native holoflavodoxin, according to equation 9:

$$k_{Holo-PUF3} = \frac{k_{Holo-Apo} k_{Apo-PUF3}}{k_{Apo-PUF3} + k_{Apo-Holo}} \quad k_{PUF3-Holo} = \frac{k_{PUF3-Apo} k_{Apo-Holo}}{k_{Apo-PUF3} + k_{Apo-Holo}} \quad (9)$$

with  $k_{X-Y}$  being the rate constant with which species X is converted into species Y. The values used are  $k_{Holo-Apo} = (8.88 \pm 0.02) 10^{-5} \text{ s}^{-1}$ ,  $k_{Apo-Holo} = 0.92 \pm 0.01 \text{ s}^{-1}$ ,  $k_{PUF3-Apo} = 6 \pm 4 \text{ s}^{-1}$ ,  $k_{Apo-PUF3} = (7.4 \pm 0.7) 10^{-5} \text{ s}^{-1}$ . The values calculated are  $k_{PUF3-Holo} = 6 \pm 4 \text{ s}^{-1}$ , and  $k_{Holo-PUF3} = (6.7 \pm 0.7) 10^{-9} \text{ s}^{-1}$  (see Chapter 4 of this thesis).

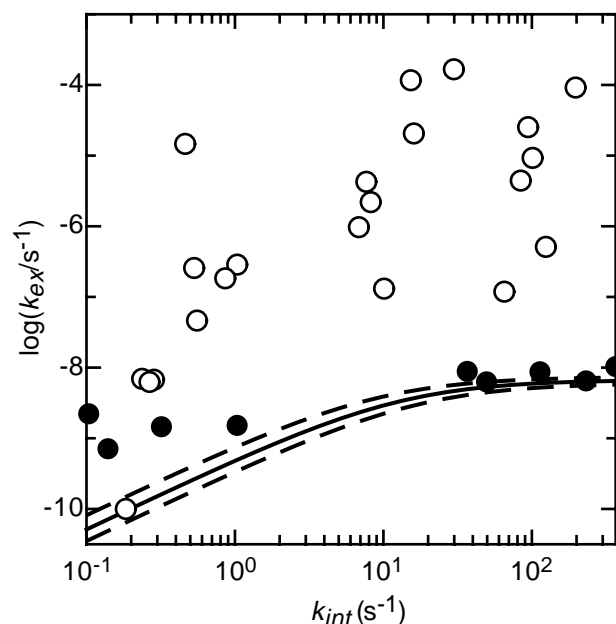
The theoretically expected dependence of the exchange rate of a backbone amide proton that exchanges with deuterons from PUF3 on pD is plotted as a solid curve in Figure 7 (the dashed lines indicate the standard deviation). Use is made of equation 3 and the rate constants determined above. The curve is calculated assuming that PUF3, PUF4 and I<sub>1</sub> can only be formed from unfolded apoflavodoxin. Only the exchange behaviour of the backbone amide proton of Phe6 comes close to the latter curve. Its exchange from PUF3 is a bit faster than theoretically expected. Note however, that the exchange of the backbone amide proton of Phe6 is extremely slow. During the 2-month lasting exchange experiments at pD 6.19 and pD 7.73 no decrease in intensity of the NMR cross peak corresponding to Phe6 is observed at all. At pD 6.26 the backbone amide proton exchange rate of Phe6 is  $1.97 10^{-9} \text{ s}^{-1}$  (Table 1), which corresponds to a lifetime of 16 years. The height of the HSQC cross peak corresponding to the backbone amide of Phe6 diminishes only by a few percent during the 11-month lasting H/D exchange experiment at pD 6.26. This most likely causes the overestimation of the corresponding exchange rate constant.

In conclusion, the backbone amide proton of Phe6 seems to exchange from PUF3 in holoflavodoxin. The other amide protons that exchange from PUF3 in apoflavodoxin, i.e. those of Leu5, Ile21 and Lys22, must exchange in holoflavodoxin via local opening processes. The pD-dependent H/D exchange data of the backbone amide proton of Phe6 show that PUF3 is only accessible to native holoflavodoxin upon release of the FMN cofactor followed by global protein unfolding. Subsequently the non-productive folding route that leads to PUF3 is followed.

#### d. The formation of PUF4 from native holoflavodoxin

Residues the amide protons of which become water accessible in PUF4 in apoflavodoxin are coloured red in Figure 3 of Chapter 4 (i.e. Gln46, Tyr47, Phe49, Leu50, Ile51, Leu52, Gly53, Ile81, Ser87, Lys89, Val91, Ala92, Leu93, Phe94, Gly95, Asp119, Arg120, Lys123, Val125, Arg163, Val164, Ala166, Trp167, Leu168, Ala169, Gln170, Ile171, Ala172, Phe175). PUF4, like PUF3, is a partially folded apoflavodoxin species that

resides on a non-productive folding route that starts from unfolded apoflavodoxin, and is positioned close to  $I_1$  in the apoflavodoxin folding scheme (equation 1, Figure 2).



**Figure 8.** Dependence of the H/D exchange rates,  $k_{ex}$ , of the backbone amide protons of Val91, Leu93, Trp167, Leu168, Ala169, Gln170, Ile171, Ala172 (all depicted as O), of Ile51, Leu52, Gly53, Ala92, and of Phe94 (all depicted as ●) of *A. vinelandii* holoflavodoxin on the intrinsic exchange rate,  $k_{int}$ . The theoretically expected dependence (as described by equation 3) of  $k_{ex}$  on  $k_{int}$  for amide protons that exchange from PUF4 is plotted as a solid line, with the dashed lines indicating the standard deviation. The data are obtained from 2.5 to 3 mM holoflavodoxin in 100 mM potassium pyrophosphate at 25 °C.

proton exchange rates (Figure 8). The rate constants for interconversion between native holoflavodoxin and PUF4 are calculated according to equation 9 in which PUF3 is substituted by PUF4. The values used for  $k_{PUF4-Apo}$  and  $k_{Apo-PUF4}$  are  $13 \pm 4 \text{ s}^{-1}$  and  $(7.4 \pm 0.7) 10^{-5} \text{ s}^{-1}$ , respectively (Chapter 4 of this thesis). This results in  $k_{PUF4-Holo} = 13 \pm 4 \text{ s}^{-1}$  and  $k_{Holo-PUF4} = (6.7 \pm 0.7) 10^{-9} \text{ s}^{-1}$ . Both  $k_{Holo-PUF4}$  and  $k_{Holo-PUF3}$  have identical values. This is caused by the same energy barrier being rate limiting for the formation of  $I_1$ , PUF3 as well as of PUF4 from both native apo- and holoflavodoxin (Figure 2). The theoretically expected dependence of the amide proton exchange rate of a backbone amide proton that exchanges with deuterons from PUF4 on pD is plotted in Figure 8 as a solid curve (the dashed lines indicate the standard deviation).

The dependence of the  $k_{ex}$ -values on  $k_{int}$  of the amide protons of five residues of holoflavodoxin, i.e. Ile51, Leu52, Gly53, Ala92 and Phe94, is similar to the one described by the isotherm for H/D exchange from PUF4. At high pD (i.e. at high  $k_{int}$  values) the amide proton exchange rates of these residues match with the opening rate calculated for the

Consequently, due to the influence of FMN binding, the difference in free energy between native holoflavodoxin and PUF4 is expected to be 4.7 kcal/mol higher than the free energy difference between native apoflavodoxin and PUF4.

In apoflavodoxin, the amide protons that characterise the H/D exchange from PUF4 are of Ile51, Leu52, Gly53, Val91, Ala92, Leu93, Phe94, Trp167, Leu168, Ala169, Gln170, Ile171, and of Ala172. In holoflavodoxin the corresponding  $\Delta G_{op}$ -values are 0 to 3.84 kcal/mol higher than in apoflavodoxin (Table 1). Of the amide protons mentioned, those that have the largest difference in  $\Delta G_{op}$  between apo- and holoflavodoxin might exchange from PUF4 in holoflavodoxin.

Further insight into whether some of the amide protons that exchange from PUF4 in apoflavodoxin also exchange from PUF4 in holoflavodoxin can be gained by investigating the pD dependence of the corresponding amide

transition of native holoflavodoxin into PUF4. At low pD the exchange rates of these five amide protons are higher than theoretically expected (Figure 8). This overestimation is the result of the rate constants in question (i.e.  $k_{ex}$  varies between  $10^{-9}$  and  $10^{-10}$  s $^{-1}$ ) being determined close to the detection limit of the 11-month H/D exchange experiment at pD 6.26, as is also the case for Phe6 (Figure 7). In case of the H/D exchange experiments done at pD 6.19 and 7.73, which lasted for a period of two months, no exchange at all is observed for the backbone amide protons of Ile51, Leu52, Gly53, Ala92 and Phe94. Fitting of equation 3 to the pD-dependent H/D exchange data of the backbone amide protons of Ile51, Leu52, Gly53, Ala92 and Phe94 shown in Figure 8 results in  $k_{op} = (7 \pm 2) 10^{-9}$  s $^{-1}$  and  $k_{cl} = 0.8 \pm 0.4$  s $^{-1}$ .

In conclusion, five backbone amide protons of holoflavodoxin, i.e. those of Ile51, Leu52, Gly53, Ala92 and Phe94 have pD-dependent H/D exchange rates that are consistent with PUF4 being only accessible to native holoflavodoxin after release of its FMN cofactor.

### The free energy landscape for *A. vinelandii* holoflavodoxin folding

Combining the results obtained to date about *A. vinelandii* flavodoxin folding allows the construction of a model of the free energy landscape for holoflavodoxin folding.

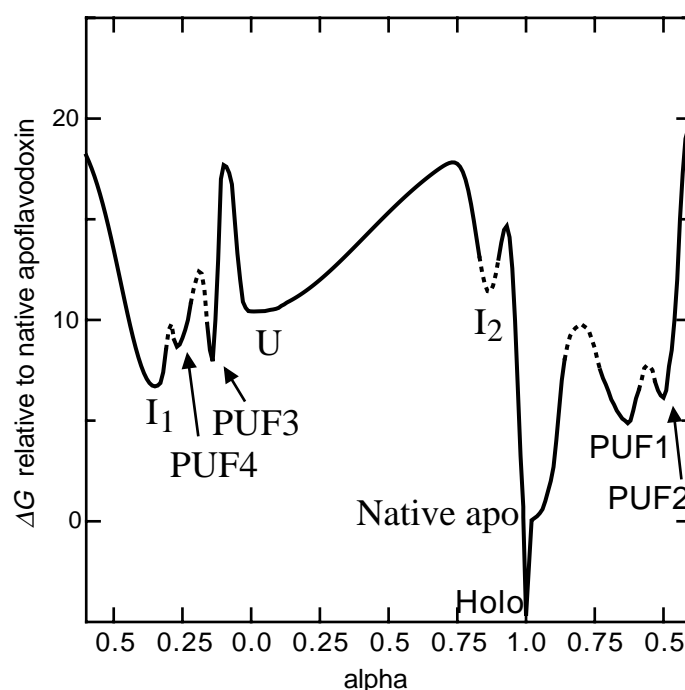
The off-pathway folding intermediate  $I_1$  that populates during both apo- and holoflavodoxin denaturant-induced equilibrium (un)folding does not bind FMN. Neither does unfolded apoflavodoxin (equation 7, and Chapter 5 of this thesis). The results presented here show that PUF2 can not be formed by holoflavodoxin. Whether PUF1 is accessible to holoflavodoxin can not be deciphered on basis of the available data. However, as FMN binding is shown to significantly stabilise most of the flavodoxin structure we propose that PUF1 is also inaccessible to holoflavodoxin. Both PUF3 and PUF4 can not be formed by holoflavodoxin as long as the FMN molecule remains bound. This implies that both PUF3 and PUF4 of apoflavodoxin do not bind FMN. As a consequence, the presence of FMN does not alter the part of the apoflavodoxin folding free energy landscape in which U, PUF3, PUF4 and  $I_1$  reside (Figure 2). In addition, as discussed, PUF3 and PUF4 are formed from holoflavodoxin with a rate constant that can be predicted by taking into account that FMN is released first from the protein. This implies that native apoflavodoxin is the actual species from which PUF3 and PUF4 are formed with rates known from the subglobal unfolding of apoflavodoxin. The formation of PUF3 and PUF4 starting with holoflavodoxin implies that the productive part of the free energy landscape of apoflavodoxin folding (i.e.  $U \rightleftharpoons I_2 \rightleftharpoons N$ ) must have been ascended (Figure 2). As the corresponding rates can be precisely predicted on basis of the apoflavodoxin folding results, the free energy landscape for apoflavodoxin folding is not affected by the presence of FMN with one exception. Binding of FMN to native apoflavodoxin leads to a free energy of the resulting complex that is much lower than that of apoflavodoxin itself.

Taking all observations together, the two-dimensional model of the apoflavodoxin folding landscape (Figure 2) can be easily adjusted to incorporate FMN binding. FMN binds solely to native apoflavodoxin resulting in the formation of holoflavodoxin. Holoflavodoxin from *A. vinelandii* is structurally identical to apoflavodoxin except that in case of holoflavodoxin the dynamic disorder in the flavin-binding region does not exist (Steensma et

al. 1998; Steensma and van Mierlo 1998). In addition, holoflavodoxin is much more stable against unfolding than apoflavodoxin is. FMN binding to apoflavodoxin thus leads to a lowering of the minimum in the free energy landscape in which the native state of the protein resides, as is shown in Figure 9. For reasons of simplicity, the relatively small difference in denaturant accessibility between apo- and holoflavodoxin is not accounted for in the free energy landscape for holoflavodoxin folding.

### Concluding remarks

H/D exchange combined with NMR spectroscopy shows that ligand binding affects the dynamics of almost the entire backbone of *A. vinelandii* flavodoxin. Most local opening processes, even those that involve amino acids distant from the FMN binding site, occur less frequently in holoflavodoxin compared to in apoflavodoxin. Subglobal and global unfolding of holoflavodoxin does not occur as long as FMN is bound to flavodoxin. The entire flavodoxin molecule is more rigid when FMN is bound.



**Figure 9.** Schematic representation of the free energy landscape for folding of *A. vinelandii* flavodoxin. The model for apoflavodoxin folding as presented in Figure 2 is now extended to incorporate holoflavodoxin as the product of FMN binding to solely native apoflavodoxin. Both flavodoxin species are structurally identical except that apoflavodoxin has some dynamic disorder in the flavin-binding region (Steensma and van Mierlo 1998). Due to the stabilising effect of FMN incorporation, the minimum in the free energy landscape in which holoflavodoxin resides is deeper than the minimum in which apoflavodoxin resides. No experimental indications exist for the presence of FMN influencing any other parts of the free energy landscape for flavodoxin folding.

## References

- Bai, Y., J. S. Milne, L. Mayne and S. W. Englander (1993). "Primary structure effects on peptide group hydrogen exchange." *Proteins* **17**(1): 75-86.
- Bai, Y., T. R. Sosnick, L. Mayne and S. W. Englander (1995). "Protein folding intermediates: native-state hydrogen exchange." *Science* **269**(5221): 192-7.
- Chamberlain, A. K., T. M. Handel and S. Marqusee (1996). "Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH." *Nat Struct Biol* **3**(9): 782-7.
- Huyghues-Despointes, B. M., J. M. Scholtz and C. N. Pace (1999). "Protein conformational stabilities can be determined from hydrogen exchange rates." *Nat Struct Biol* **6**(10): 910-2.
- Hvidt, A. and S. O. Nielsen (1966). "Hydrogen exchange in proteins." *Adv Protein Chem* **21**: 287-386.
- Kay, L. E., P. Keifer and T. Saarinen (1992). "Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity." *J Am Chem Soc* **114**: 10663-10665.
- Kraulis, P. J. (1991). "MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures." *J Appl Cryst* **24**: 946-950.
- Krieger, F., B. Fierz, O. Bieri, M. Drewello and T. Kiefhaber (2003). "Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding." *J Mol Biol* **332**(1): 265-274.
- Luque, I., S. A. Leavitt and E. Freire (2002). "The linkage between protein folding and functional cooperativity: two sides of the same coin?" *Annu Rev Biophys Biomol Struct* **31**: 235-56.
- MacKnight, M. L., J. M. Gillard and G. Tollin (1973). "Flavine-protein interactions in flavoenzymes. pH dependence of the binding of flavine mononucleotide and riboflavine to *Azotobacter flavodoxin*." *Biochemistry* **12**(21): 4200-6.
- Mayhew, S. G. and G. Tollin (1992). General properties of flavodoxins. *Chemistry and biochemistry of flavoenzymes*. F. Müller. Boca Raton, CRC Press. **3**: 389-426.
- Palmer III, A. G. <http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/nmr2modelfree.html>.
- Palmer III, A. G., J. Cavanagh, P. E. Wright and M. Rance (1991). "Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy." *Journal of Magnetic Resonance* **93**: 151-170.
- Steensma, E., H. A. Heering, W. R. Hagen and C. P. M. Van Mierlo (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**(1-2): 167-72.
- Steensma, E., M. J. Nijman, Y. J. Bollen, P. A. de Jager, W. A. van den Berg, W. M. van Dongen and C. P. M. van Mierlo (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**(2): 306-17.
- Steensma, E. and C. P. M. van Mierlo (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**(3): 653-66.
- Thorneley, R. N. F., G. A. Ashby, M. H. Drummond, R. R. Eady, D. L. Hughes, G. Ford, P. M. Harrison, A. Shaw, R. L. Robson, J. Kazlauskaitė and H. A. O. Hill (1994). Flavodoxin and nitrogen fixation: Structure, electrochemistry and posttranslational modification by coenzyme A. Flavins and flavoproteins 1993. K. Yagi. Berlin, Walter de Gruyter & Co.: 343-354.
- van Mierlo, C. P. M., J. M. van den Oever and E. Steensma (2000). “Apoflavodoxin (un)folding followed at the residue level by NMR.” Protein Sci **9**(1): 145-57.
- van Mierlo, C. P. M., W. M. van Dongen, F. Vergeldt, W. J. van Berkel and E. Steensma (1998). “The equilibrium unfolding of *Azotobacter vinelandii* apoflavodoxin II occurs via a relatively stable folding intermediate.” Protein Sci **7**(11): 2331-44.
- Yan, S., S. D. Kennedy and S. Koide (2002). “Thermodynamic and kinetic exploration of the energy landscape of *Borrelia burgdorferi* OspA by native-state hydrogen exchange.” J Mol Biol **323**(2): 363-75.
- Zidek, L., M. V. Novotny and M. J. Stone (1999). “Increased protein backbone conformational entropy upon hydrophobic ligand binding.” Nat Struct Biol **6**(12): 1118-21.

## Summary

The research described in this thesis has been carried out to obtain a better understanding of the fundamental rules describing protein folding. Protein folding is the process in which a linear chain of amino acids contracts to a compact state in which it is active. Flavodoxin from *Azotobacter vinelandii* is chosen as the representative of the group of  $\alpha$ - $\beta$  parallel proteins. Flavodoxins are small monomeric proteins that contain a non-covalently bound FMN cofactor. The  $\alpha$ - $\beta$  parallel topology is characterised by a five-stranded parallel  $\beta$ -sheet surrounded by  $\alpha$ -helices at either side of the sheet. The doubly-wound topology is a rather popular fold: it belongs to the five most common observed folds, together with the TIM-barrel, Rossmann, thiamin-binding and P-loop hydrolase folds. In contrast to most protein folds, this topology is shared by many (i.e. nine) protein superfamilies. These nine superfamilies exhibit little or no sequence similarity and comprise a broad range of unrelated proteins with different functions like catalases, chemotactic proteins, lipases, esterases, and flavodoxins. By studying the folding behaviour of *A. vinelandii* flavodoxin insight can be gained into how this large group of proteins folds.

First, the equilibrium (un)folding of apoflavodoxin from *A. vinelandii* (i.e. flavodoxin in the absence of the FMN cofactor) is investigated. Apoflavodoxin is structurally identical to holoflavodoxin except for some dynamic disorder in the flavin-binding region. A molten globule-like intermediate is shown to populate during denaturant-induced equilibrium unfolding of apoflavodoxin (Chapter 2).

Subsequently, the folding and unfolding kinetics of the 179-residue *A. vinelandii* apoflavodoxin have been followed by stopped-flow experiments monitored by fluorescence intensity and anisotropy (Chapter 2). The denaturant concentration dependence of the folding kinetics is complex. Under strongly unfolding conditions, the kinetics can be described by a single rate constant. When this unfolding rate constant is plotted against the denaturant concentration, a change in the slope is observed. This, together with the absence of an additional unfolding process reveals the presence of two consecutive transition states on a linear pathway that surround a high-energy on-pathway intermediate.

Under refolding conditions, two folding processes are observed. The slowest of these two processes is the one that is populated most, and it becomes faster with increasing denaturant concentration. This means that an unfolding step is rate-limiting for folding of the majority of apoflavodoxin molecules. This, together with the absence of a lag in the formation of native molecules, means that the intermediate that populates during refolding is off-pathway.

The experimental data obtained on apoflavodoxin folding are consistent with the linear four-state folding mechanism  $I_1 \rightleftharpoons \text{unfolded apoflavodoxin} \rightleftharpoons I_2 \rightleftharpoons \text{native apoflavodoxin}$ . The off-pathway intermediate  $I_1$  is the one that populates during refolding and that also populates during denaturant-induced equilibrium unfolding of apoflavodoxin.  $I_2$  is the unstable intermediate that is observed during kinetic unfolding.

The presence of such on-pathway and off-pathway intermediates in the folding kinetics of proteins with an  $\alpha$ - $\beta$  parallel topology is predicted from simulations of Gō-like

protein models. In addition, two kinetic folding intermediates, one on-pathway and the other off-pathway, seem to be present under specific experimental conditions during the folding of all proteins with an  $\alpha$ - $\beta$  parallel topology that have been investigated. The appearance of folding intermediates in this class of proteins is apparently governed by protein topology (Chapter 3).

Next, the local dynamics of apoflavodoxin have been studied by hydrogen deuterium exchange detected by heteronuclear NMR spectroscopy (Chapter 4). The use of native state hydrogen deuterium exchange detected by NMR spectroscopy leads to the identification of four partially unfolded forms (PUFs) of apoflavodoxin in which some non-native interactions apparently play a role. The rates of interconversion of these PUFs with native apoflavodoxin are determined. These rates are inconsistent with the PUFs being on a direct folding route between native and globally unfolded apoflavodoxin. PUF1 and PUF2 are on an unfolding route starting from native apoflavodoxin that does not lead to the globally unfolded state of the protein. PUF3 and PUF4 are on a non-productive folding route starting from globally unfolded apoflavodoxin. A common free energy barrier separates both PUF3 and PUF4 from unfolded apoflavodoxin. This barrier has the same height as the one determined from stopped-flow kinetic folding studies that separates the known off-pathway apoflavodoxin folding intermediate  $I_1$  from the productive folding route. Therefore a single energy barrier is proposed to separate both PUF3 and PUF4 as well as  $I_1$  from the productive folding route. All three species thus need to unfold before productive folding of apoflavodoxin can occur (Chapter 4).

The influence of the presence of the non-covalently bound flavin mononucleotide (FMN) cofactor on the global stability and on the kinetic folding of *A. vinelandii* holoflavodoxin (i.e. flavodoxin in presence of the FMN co-factor) are reported in Chapter 5. The denaturant-induced equilibrium (un)folding data of flavodoxin in the presence and absence of FMN are excellently described by a model in which only native apoflavodoxin binds to FMN. As the intermediate  $I_1$  populates during apoflavodoxin equilibrium (un)folding, the holoflavodoxin equilibrium (un)folding model consists of four species: unfolded apoflavodoxin, the apoflavodoxin folding intermediate  $I_1$ , native apoflavodoxin and holoflavodoxin molecules. Cofactor binding to apoflavodoxin is shown to affect the protein stability in a theoretically predictable manner.

Despite that many proteins require the binding of a ligand to be functional, the kinetic role of ligand-binding during folding is poorly understood. FMN binding to native apoflavodoxin occurs with two kinetically observable rate constants at all denaturant and protein concentrations studied, as is shown in Chapter 5. These two rate constants arise from two conformationally differing apoflavodoxin species, which most likely exist due to the binding of inorganic phosphate to the FMN phosphate binding site of a fraction of the *A. vinelandii* apoflavodoxin molecules.

In Chapter 5 it is also shown that excess FMN does not accelerate flavodoxin folding, and FMN does not act as a nucleation site for flavodoxin folding. During kinetic folding of holoflavodoxin formation of native apoflavodoxin precedes ligand binding. Even under strongly denaturing conditions, global unfolding of holoflavodoxin occurs only after release



of its FMN. The model that describes *A. vinelandii* apoflavodoxin kinetic folding, which includes the stable off-pathway intermediate  $I_1$  and a high-energy on-pathway intermediate  $I_2$ , can now be extended to describe kinetic holoflavodoxin folding:  $I_1 + \text{FMN} \rightleftharpoons \text{unfolded apoflavodoxin} + \text{FMN} \rightleftharpoons I_2 + \text{FMN} \rightleftharpoons \text{native apoflavodoxin} + \text{FMN} \rightleftharpoons \text{holoflavodoxin}$  (Chapter 5).

Finally, in Chapter 6 native state H/D exchange combined with NMR spectroscopy is used to probe the influence of FMN binding on the stability of *A. vinelandii* flavodoxin against local, subglobal and global unfolding. Almost the entire flavodoxin backbone is substantially more rigid in holoflavodoxin than in apoflavodoxin. No areas are detected in flavodoxin where FMN binding results in an increase of the local dynamics. Occasional release of FMN from holoflavodoxin results in the population of apoflavodoxin. Until FMN is rebound, these apoflavodoxin molecules behave as described in Chapter 4. Consequently, they will adopt the previously described partially unfolded forms (PUFs). At least three out of the four partially unfolded forms that apoflavodoxin occasionally adopts under native conditions are inaccessible to holoflavodoxin. Holoflavodoxin can form these partially unfolded conformations only when FMN is released.

All observations described in this thesis are used to create a schematic free energy landscape of folding of *A. vinelandii* flavodoxin. This schematic energy landscape provides insight into how a protein molecule that adopts the  $\alpha$ - $\beta$  parallel topology surfs from its unfolded state to its characteristic folded state in which it is active.

## Outlook

As described in Chapter 3 of this thesis, the appearance of both on- and off-pathway intermediates during the folding of *A. vinelandii* apoflavodoxin appears to be governed by its  $\alpha$ - $\beta$  parallel topology. Folding kinetics of other  $\alpha$ - $\beta$  parallel proteins than the ones mentioned in this thesis need to be determined to verify this hypothesis.

An interesting question is why intermediate  $I_1$  that *A. vinelandii* apoflavodoxin populates during its denaturant-induced equilibrium (un)folding is off the direct folding route. This question may be resolved by studying the structure of intermediate  $I_1$  using among others multidimensional NMR experiments. In addition, the investigation of possible residual structure within unfolded apoflavodoxin can inform about the origin of the kinetic partitioning of folding apoflavodoxin molecules into two routes, one leading to native apoflavodoxin, the other one leading to the molten globule-like intermediate  $I_1$ .

The kinetic model for *A. vinelandii* apoflavodoxin folding presented in this thesis implies that apoflavodoxin molecules once they have formed the intermediate  $I_1$  need to unfold before folding to the native state can proceed. Studying the folding behaviour of single *A. vinelandii* apoflavodoxin molecules using sensitive fluorescence techniques can reveal to what extent an apoflavodoxin molecule has to unfold in the latter process. To date, only the folding kinetics of small proteins that fold in one step have been studied by single molecule detection techniques. Studying the folding of individual apoflavodoxin molecules can reveal the general dynamics involved in the partitioning of individual protein molecules into two separate folding trajectories.

Finally, it will be highly interesting to study the folding behaviour of proteins in their natural environment. As pointed out in the first chapter of this thesis, the high concentration of biomacromolecules in cells is bound to influence the latter behaviour. Therefore, studying the influence macromolecular crowding agents have on folding flavodoxin molecules will be of great interest. In this thesis, a solid, and strongly necessary, basis is laid for the future perspective of the *in vivo* investigation of flavodoxin folding in the living cell.

## Samenvatting

Het doel van het onderzoek beschreven in dit proefschrift is een beter inzicht te verkrijgen in de fundamentele regels die eiwitvouwing beschrijven. Eiwitvouwing is het proces waarmee een lineaire keten van aminozuren zich samentrekt tot een toestand waarin het eiwit compact, gestructureerd en actief is. Het onderzoek is uitgevoerd aan flavodoxine van *Azotobacter vinelandii*, dat is gekozen als vertegenwoordiger van de familie van  $\alpha$ - $\beta$  parallelle eiwitten. Flavodoxines zijn kleine monomere eiwitten die een niet-covalent gebonden flavine mononucleotide (FMN) co-factor bevatten. De  $\alpha$ - $\beta$  parallelle topologie van het eiwit wordt gekenmerkt door een centrale parallelle  $\beta$ -sheet bestaande uit vijf strengen, die aan beide zijden omgeven is door  $\alpha$ -helices. Deze  $\alpha$ - $\beta$  parallelle topologie behoort tot de vijf meest geobserveerde topologieën, samen met de TIM-barrel, Rossman, thiamine-bindende en de P-loop hydrolase topologieën. In tegenstelling tot de meeste andere topologieën wordt de  $\alpha$ - $\beta$  parallelle topologie gedeeld door meerdere, namelijk negen, eiwit superfamilies. Deze negen superfamilies hebben weinig of geen homologie in hun aminozuursequenties en omvatten een breed scala aan eiwitten met verschillende functies, zoals catalases, chemotactische eiwitten, lipases, esterasen en flavodoxines. Door de vouwing van *A. vinelandii* flavodoxine te bestuderen kan hopelijk inzicht verkregen worden in hoe deze grote groep van eiwitten zich opvouwt tot hun actieve toestand.

Ten eerste is de evenwichtsontvouwing van flavodoxine in afwezigheid van de FMN co-factor (apoflavodoxine) bestudeerd. Apoflavodoxine is structureel identiek aan flavodoxine, behalve dat de flavine bindingsplaats dynamischer is in apoflavodoxine. Tijdens dit onderzoek is aangetoond dat apoflavodoxine een *molten globule*-achtige intermediair ( $I_1$ ) vormt tijdens denaturant-geïnduceerde evenwichtsontvouwing (Hoofdstuk 2).

Vervolgens is de vouwingskinetiek van apoflavodoxine bestudeerd met behulp van *stopped-flow* experimenten, waarbij het vouwingsproces in de tijd gevolgd wordt aan de hand van fluorescentie intensiteit en anisotropie (Hoofdstuk 2). De denaturant concentratieafhankelijkheid van de vouwingskinetiek van apoflavodoxine is complex. Onder sterk denaturerende condities kan de kinetiek beschreven worden met één ontvouwingssnelheid. Wanneer de logaritme van de ontvouwingssnelheid wordt geplot tegen de denaturant concentratie, wordt een verandering in de helling waargenomen. Deze verandering wordt niet vergezeld door een additioneel ontvouwingsproces. Dit duidt op de aanwezigheid van twee opeenvolgende transitietoestanden op een lineaire (ont)vouwingsroute. Deze twee transitietoestanden omsluiten een hoogenergetische intermediair die op de directe (ont)vouwingsroute van het eiwit ligt.

Tijdens hervouwing van ontvouwen apoflavodoxine worden twee vouwingsgerelateerde processen geobserveerd. Het langzaamste vouwingsproces wordt gevolgd door de meeste apoflavodoxine moleculen en wordt sneller met toenemende denaturant concentratie. Dit kan enkel verklaard worden door de aanwezigheid van een ontvouwingsproces dat snelheidsbepalend is tijdens de vorming van natieve apoflavodoxine moleculen. Additionele experimenten laten zien dat geen vertraging optreedt in de vorming van natieve apoflavodoxine moleculen. Hieruit is geconcludeerd dat een intermediair wordt

gevormd tijdens de hervouwing van apoflavodoxine die niet op de productieve vouwingsroute ligt (Hoofdstuk 2).

De experimenteel verkregen apoflavodoxine vouwingsdata zijn consistent met het volgende lineaire viertoestanden vouwingsmechanisme:  $I_1 \Leftrightarrow$  ontvouwen apoflavodoxine  $\Leftrightarrow I_2 \Leftrightarrow$  natief apoflavodoxine. De intermediair  $I_1$  die niet op de productieve vouwingsroute ligt is de intermediair die populeert tijdens het hervouwen van ontvouwen apoflavodoxine en tijdens evenwichtsontvouwing van het eiwit. Intermediair  $I_2$  is de hoogenergetische intermediair die tijdens kinetische ontvouwing van apoflavodoxine wordt waargenomen.

De aanwezigheid van twee eiwitvouwingsintermediären gedurende de vouwingskinetiek van eiwitten met een  $\alpha$ - $\beta$  parallelle topologie was voorspeld middels computersimulaties die gebruik maken van Gö-achtige eiwit modellen. Onder bepaalde experimentele condities lijken zulke vouwingsintermediären voor te komen tijdens de vouwing van alle eiwitten met een  $\alpha$ - $\beta$  parallelle topologie die tot nu toe experimenteel zijn bestudeerd. Blijkbaar bepaalt de topologie in deze klasse van eiwitten de aanwezigheid van vouwingsintermediären (Hoofdstuk 3).

Vervolgens is de lokale dynamica van apoflavodoxine bestudeerd met behulp van proton deuterium (H/D) uitwisselingsexperimenten en heteronucleaire NMR (Nuclear Magnetic Resonance = kernspin resonantie) spectroscopie (Hoofdstuk 4). H/D uitwisselingsexperimenten in aanwezigheid van kleine hoeveelheden denaturant leiden tot de identificatie van vier gedeeltelijk ontvouwen vormen ("Partially Unfolded Forms", PUFs) van apoflavodoxine. De interconversiesnelheden van deze PUFs met natief apoflavodoxine zijn bepaald en hieruit bleek dat deze snelheden inconsistent zijn met een positie van de PUFs op de directe (ont)vouwingsroute tussen natief en globaal ontvouwen apoflavodoxine. PUF1 en PUF2 blijken te liggen op een ontvouwingsroute die met natief apoflavodoxine begint maar niet leidt tot de globaal ontvouwen vorm van het eiwit. PUF3 en PUF4 liggen op een niet-productieve vouwingsroute die begint bij globaal ontvouwen apoflavodoxine en waarbij een gemeenschappelijke energiebarrière zowel PUF3 als PUF4 belemmert om naar de globaal ontvouwen toestand van apoflavodoxine terug te keren. Deze energiebarrière heeft dezelfde hoogte als de energiebarrière die  $I_1$  scheidt van de productieve vouwingsroute zoals bepaald met behulp van *stopped-flow* kinetiek (Hoofdstuk 2). Verondersteld wordt daarom dat één barrière zowel PUF3, PUF4 als  $I_1$  scheidt van de productieve vouwingsroute. Zowel PUF3 en PUF4 als  $I_1$  moeten dus ontvouwen voordat natief apoflavodoxine gevormd kan worden (Hoofdstuk 4).

In hoofdstuk 5 van dit proefschrift wordt verslag gedaan van de invloed van de niet-covalent gebonden FMN co-factor op de globale stabiliteit en op de vouwingskinetiek van *A. vinelandii* holoflavodoxine (dat is flavodoxine in aanwezigheid van de FMN co-factor). De denaturant-geïnduceerde evenwichtsontvouwingsdata van flavodoxine in aan- en afwezigheid van FMN worden uitstekend beschreven door een model waarin alleen natief apoflavodoxine FMN kan binden. Aangezien intermediair  $I_1$  populeert tijdens de evenwichtsontvouwing van apoflavodoxine moet het model dat de evenwichtsontvouwing van holoflavodoxine beschrijft uit vier toestanden bestaan: ontvouwen apoflavodoxine, de apoflavodoxine

vouwingsintermediair  $I_1$ , natief apoflavodoxine en holoflavodoxine. ( $I_2$  wordt niet geobserveerd tijdens evenwichtsontvouwing omdat deze te onstabiel is.)

Ondanks dat voor vele eiwitten het binden van een ligand een voorwaarde is voor functionaliteit, is weinig bekend over de kinetische rol van ligandbinding tijdens eiwitvouwing. Natief apoflavodoxine blijkt FMN te binden met twee observeerbare snelheden. Dit wordt veroorzaakt doordat een deel van de apoflavodoxine moleculen anorganisch fosfaat bindt op de plaats waar de fosfaatgroep van FMN hoort te binden. Het anorganisch fosfaat moet loslaten voordat FMN kan binden (Hoofdstuk 5).

In hoofdstuk 5 wordt tevens uitgelegd dat een overmaat FMN het vouwingsproces niet versnelt, en dat FMN niet functioneert als een condensatiekern voor de vouwing van flavodoxine. Tijdens de kinetische vouwing van holoflavodoxine gaat de vorming van natief apoflavodoxine vooraf aan FMN binding. Zelfs onder sterk denaturerende condities treedt globale ontvouwing van holoflavodoxine alleen op nadat het FMN is vrijgekomen. Het eerder genoemde model dat de vouwingskinetiek van apoflavodoxine beschrijft middels de stabiele intermediair  $I_1$  die niet op de productieve vouwingsroute ligt, en de onstabiele intermediair  $I_2$  die op de directe vouwingsroute van ontvouwen naar natief apoflavodoxine ligt, kan nu worden uitgebreid om holoflavodoxine vouwing te beschrijven:  $I_1 + \text{FMN} \rightleftharpoons \text{ontvouwen apoflavodoxine} + \text{FMN} \rightleftharpoons I_2 + \text{FMN} \rightleftharpoons \text{natief apoflavodoxine} + \text{FMN} \rightleftharpoons \text{holoflavodoxine}$  (Hoofdstuk 5).

Tenslotte wordt in hoofdstuk 6 H/D uitwisseling en NMR spectroscopie gebruikt om de invloed van FMN-binding op de stabiliteit van flavodoxine tegen lokale en subglobale ontvouwing te bestuderen. Bijna de gehele hoofdketen van flavodoxine is substantieel stijver in holoflavodoxine dan in apoflavodoxine. In holoflavodoxine kunnen geen regio's worden gevonden waar FMN binding leidt tot een verhoging van de lokale beweeglijkheid. Incidenteel laat holoflavodoxine de FMN co-factor los, wat resulteert in apoflavodoxine. Dit apoflavodoxine gedraagt zich als beschreven in hoofdstuk 4, en vormt de reeds beschreven gedeeltelijk ontvouwen vormen (PUFs). Ten minste drie van de vier PUFs die apoflavodoxine incidenteel aanneemt onder natieve condities zijn ontoegankelijk voor holoflavodoxine. Holoflavodoxine kan deze structuren alleen vormen als het zijn FMN loslaat.

Alle waarnemingen die in dit proefschrift beschreven zijn, zijn gebruikt om een schematisch vrij energie landschap te maken voor de vouwing van *A. vinelandii* flavodoxine. Dit energie landschap biedt inzicht in hoe een eiwitmolecuul met een  $\alpha$ - $\beta$  parallelle topologie surft van zijn ontvouwen toestand naar de karakteristieke gevouwen toestand waarin het eiwit actief is.

## Vooruitblik

In hoofdstuk 3 van dit proefschrift staat beschreven dat de aanwezigheid tijdens de vouwing van *A. vinelandii* apoflavodoxine van zowel een intermediair op de directe vouwingsroute als een intermediair die zich niet op die route bevindt bepaald lijkt te zijn door de topologie van het eiwit. Om deze hypothese te verifiëren moet de vouwingskinetiek van meer  $\alpha$ - $\beta$  parallelle eiwitten worden bestudeerd.

Een interessante vraag is waarom intermediair  $I_1$  die *A. vinelandii* apoflavodoxine populeert tijdens evenwichtsontvouwning niet op de directe vouwingsroute ligt. Deze vraag kan misschien beantwoord worden door de structuur van deze intermediair te bestuderen met o.a. NMR spectroscopie. Daarnaast kan de bestudering van eventuele residuele structuur in de ontvouwen toestand van apoflavodoxine informatie geven over de manier waarop vouwende apoflavodoxine moleculen verdeeld worden over de twee voor hen toegankelijke vouwingsroutes: één naar natief apoflavodoxine en één naar de vouwingsintermediair  $I_1$ .

Het kinetische model voor vouwing van apoflavodoxine dat in dit proefschrift wordt geïntroduceerd houdt in dat apoflavodoxine moleculen die de intermediair  $I_1$  hebben gevormd eerst moeten ontvouwen voordat ze natief apoflavodoxine kunnen vormen. Door de bestudering van het gedrag van individuele *A. vinelandii* apoflavodoxine moleculen met behulp van gevoelige fluorescentie technieken kan onthuld worden in welke mate een apoflavodoxine molecuul moet ontvouwen in dit proces. Tot op heden zijn dergelijke technieken alleen gebruikt om de vouwingskinetiek van kleine eiwitten die in een enkele stap vouwen te bestuderen. Door de vouwing van individuele apoflavodoxine moleculen te bestuderen kan informatie verkregen worden van meer algemene aard, namelijk over de dynamica die de scheiding van individuele eiwit moleculen in twee aparte vouwingsroutes bepaalt.

Tenslotte zal het zeer interessant zijn om de vouwing van eiwitten in hun natuurlijke omgeving te bestuderen. Zoals in het eerste hoofdstuk van dit proefschrift is vermeld, beïnvloedt de hoge concentratie aan biomacromoleculen in een cel waarschijnlijk het vouwingsgedrag van eiwitten. Daarom zal bestudering van de invloed van macromoleculen op eiwitvouwing belangrijke informatie opleveren. In dit proefschrift is een solide basis gelegd voor toekomstig onderzoek naar vouwing van apoflavodoxine in levende cellen.

## Nawoord

Mijn promotietijd is een aangename en uitermate leerzame periode geweest. Allereerst wil ik Carlo daarvoor bedanken. Als copromotor heb je me alle ruimte gegeven om het project naar eigen inzicht in te vullen. Je deur stond altijd open, en ik kijk met plezier terug op de vele, vaak lange discussies over flavodoxine, eiwitvouwing, en wetenschap in het algemeen. In het laatste stadium heb je met een niet aflatend enthousiasme mijn manuscripten rood gekleurd, waarvan ik veel heb geleerd. De andere leden van de werkgroep, Maarten en Hans, wil ik bedanken voor de prettige jarenlange samenwerking. De vele discussies en werkbesprekingen hebben regelmatig tot nieuwe inzichten geleid.

Willem, door je grote kennis van de biochemie en met name van flavine-bevattende eiwitten kon je altijd nuttig commentaar leveren. Ton, jouw kennis van de fluorescentie spectroscopie is me een aantal maal zeer goed van pas gekomen. Bovendien wil ik je bedanken voor het feit dat je me in contact hebt gebracht met de Structuurbiologie groep van de VU. Walter, bedankt voor je hulp op moleculair-biologisch vlak. Mijn promotor Sacco wil ik bedanken voor de getoonde interesse in mijn werk, en voor zijn functie als stok-achter-de-deur in de laatste fase.

Ik heb tijdens mijn onderzoek veel gebruikt gemaakt van geavanceerde apparatuur, en ik heb dan ook veel te danken aan de uitstekende technici binnen de leerstoelgroep en daarbuiten. Willy, Jan-Willem, Sjef, Adrie Westphal, Adrie de Jager, Arie en Frank, zonder jullie hulp was mijn promotietijd lang niet zo productief geweest. Een speciaal woord van dank voor Laura, niet alleen voor het draaiend houden van de leerstoelgroep, maar ook voor de gezellige thee-pauzes en andere gesprekken.

Alle (oud)medewerkers van de leerstoelgroep wil ik bedanken voor de goede sfeer, zowel tijdens het werk als tijdens pauzes, borrels, kerstdiners, darts- en bowlingavonden etcetera. Mariëlle wil ik met name bedanken voor het organiseren van vele sociale evenementen en het optrommelen van iedereen voor de gezamenlijke lunchpauze.

Tijdens mijn onderzoek heb ik het genoeg gehad om twee uitstekende studenten te begeleiden bij hun afstudeeronderzoeken. Monique en Stan, ik vond het leuk om met jullie samen te werken en ik heb er veel van geleerd.

A small but crucial part of my research project was done in the Biozentrum in Basel. I am grateful to Thomas Kiefhaber for his advise, and for the opportunity to work in his lab during several occasions. I am especially grateful to Ignacio Sánchez for the pleasant collaboration we had. Nacho, you made my trips to Basel very inspiring. I would like to thank Andy for his assistance with the analysis of the data. To all the other members of the group: thanks for the hospitality and the nice atmosphere.

Naast het werk was er ook tijd voor ontspanning. Meestal gebeurde dat met mijn muziekvrienden bij Nozel, de Ontzetting en de Rijnpijpers. Ook met The Crew heb ik veel plezier gehad achter de schermen van de AID en diverse bevrijdingsfestivals. Plezier was er ook altijd thuis op Droefkonijn. Lianne, Jos, Mireille, Els, Vincent en Christy, bedankt voor de leuke tijd. Ook Bas, Clasien, Peter-Leon en alle andere studiegenoten wil ik bedanken voor veel gezelligheid.

De basis voor mijn wetenschappelijke opleiding is lang geleden gelegd. Ik wil mijn ouders en Raoul bedanken voor hun jarenlange steun en vertrouwen, en ook voor de kritische vragen. Het is altijd prettig thuiskomen in het "kuuroord".

Iris, bedankt voor je hulp en steun, zowel in praktisch als in moreel opzicht. Ik hoop voor jou te kunnen betekenen wat jij de afgelopen tijd voor mij hebt betekend, zowel in de nabije toekomst wanneer jij je proefschrift gaat schrijven, als nog lang daarna.

*Yves*



## **Curriculum vitae**

Yves Jean Marie Bollen is op 14 december 1974 geboren in Cadier en Keer. Na een gymnasium-bèta opleiding gevolgd te hebben aan het Jeanne d'Arc College in Maastricht begon hij in 1993 aan de opleiding Moleculaire Wetenschappen aan de Wageningen Universiteit. Zijn afstudeervakken deed hij bij de leerstoelgroep Biochemie, onder leiding van Elles Steensma en Dr. Carlo van Mierlo, en bij de leerstoelgroep Moleculaire Fysica, onder leiding van Chantal Houbiers en Dr. Marcus Hemminga. Vervolgens liep hij stage bij het laboratorium voor biochemie van Sydney University, Australië, onder toezicht van Dr. Joel Mackay. Na zijn afstuderen in 1998 werd hij onderzoeker in opleiding bij Dr. Carlo van Mierlo in Wageningen. Het onderzoek dat hij deed gedurende zijn promotietijd staat beschreven in dit proefschrift. Sinds februari 2004 werkt Yves als onderzoeker in de Structuurbiologie groep van de Vrije Universiteit in Amsterdam, onder leiding van Prof. Holger Lill.



## **Curriculum vitae**

Yves Jean Marie Bollen was born on December 14<sup>th</sup> 1974 in Cadier en Keer, The Netherlands. He graduated from grammar school at the Jeanne d'Arc College in Maastricht in 1993. He studied Molecular Sciences at Wageningen University. His undergraduate theses were done at the Laboratory of Biochemistry under the supervision of Elles Steensma and Dr. Carlo van Mierlo, and at the Laboratory of Molecular Physics under the supervision of Chantal Houbiers and Dr. Marcus Hemminga. A practical training period was done at the laboratory of Biochemistry of Sydney University, Australia, with Dr. Joel Mackay. He graduated in 1998 and started to work as a PhD student under the supervision of Dr. Carlo van Mierlo at the Laboratory of Biochemistry in Wageningen. The research carried out during his PhD period is described in this thesis. Since February 2004, Yves works as a postdoctoral research fellow in the group of Prof. Holger Lill at the Department of Structural Biology of the Free University in Amsterdam, The Netherlands.



## List of publications

Steensma, E., M. J. Nijman, Y. J. M. Bollen, P. A. de Jager, W. A. van den Berg, W. M. van Dongen and C. P. M. van Mierlo (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**(2): 306-17.

Houbiers, M. C., C. J. Wolfs, R. B. Spruijt, Y. J. M. Bollen, M. A. Hemminga and E. Goormaghtigh (2001). "Conformation and orientation of the gene 9 minor coat protein of bacteriophage M13 in phospholipid bilayers." Biochim Biophys Acta **1511**(2): 224-35.

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

Bollen, Y. J. M. and C. P. M. van Mierlo (2004). "Is the presence of intermediates observed in the folding kinetics of *Azotobacter vinelandii* apoflavodoxin governed by protein topology?" in preparation.

Bollen, Y. J. M., M. B. Kamphuis and C. P. M. van Mierlo (2004). "Partially unfolded forms of *Azotobacter vinelandii* apoflavodoxin identified by native state hydrogen exchange and NMR spectroscopy are off the productive folding route." in preparation.

Bollen, Y. J. M., W. J. H. van Berkel and C. P. M. van Mierlo (2004). "Last in, first out: the role of cofactor binding in *Azotobacter vinelandii* flavodoxin folding." in preparation.

Bollen, Y. J. M. and C. P. M. van Mierlo (2004). "The influence of cofactor binding on local unfolding of *Azotobacter vinelandii* flavodoxin as determined by H/D exchange and NMR spectroscopy." in preparation.



The research described in this thesis was part of the NWO-Unilever programme “Computational Chemistry of Biosystems”.

Cover picture: “Wave” © Bruce Teleky. Used with permission.

Printed by Pons en Looijen, Wageningen

