PROTEIN FOLDING, METAL IONS AND CONFORMATIONAL STATES

THE CASE OF A DI-CLUSTER FERREDOXIN

Dissertation presented to obtain the PhD degree in Biochemistry
at the Instituto de Tecnologia Química e Biológica,
Universidade Nova de Lisboa

Sónia Cristina Alves Dickson Leal Solano

Supervisor
Cláudio Emanuel Moreira Gomes

Opponents
Pernilla-Wittung Stafshede and Eduardo P. Melo

Instituto de Tecnologia Química e Biológica,
Universidade Nova de Lisboa

Oeiras, September 2008

Protein Folding, Metal Ions and Conformational States:
The case of a di-cluster Ferredoxin by Sónia S. Leal

Second Edition, December 2008

ITQB - Protein Biochemistry Folding and Stability Laboratory
Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa
Av. da República (EAN), 2781-901 Oeiras, PORTUGAL

Foreword

This dissertation describes the work performed under the supervision of Cláudio M. Gomes, in the Protein Biochemistry Folding and Stability Laboratory, at the Instituto de Tecnologia Química e Biológica, from October 2004 to May 2008.

The studies presented here aim at understanding the role of metal ions in protein stability and on the folding-unfolding pathways of a zinc-containing ferredoxin holding a [3Fe-4S] and a [4Fe-4S] cluster.

The thesis is organized in three parts. The introduction comprises two chapters; the first addresses the current state of knowledge on fundamental and general aspects of protein folding; the second provides an overview on iron-sulfur and metalloproteins proteins in general, as well as a more specific review on the characteristics of the target ferredoxin. The second part of the thesis is organized in six chapters and reports the main experimental results obtained. The last part and chapter corresponds to a general discussion, integrating the described results on the perspective of the role of iron-sulfur clusters in protein folding.

Acknowledgements

I would like to express my sincere appreciation to whom professionally but in many cases, also personally helped and supported me during my Ph.D contributing to make my aspiration a pleasant stage of my life.

Cláudio M. Gomes, my supervisor for his knowledge, encouragement, and patience that made this work possible. I also want to thank him for his enthusiasm for science and continuous new challenges that always kept me motivated. He has been an inspiring example of commitment and pragmatism that I would much ambition for myself.

My wonderful and dear colleagues at the Protein Biochemistry Folding and Stability group, for their friendship, constant support and help in the laboratory and for providing a most pleasant working environment: Barbara Henriques, Raquel Correia, Hugo Botelho, Vesna Prosinecki and João Rodrigues.

To Rita Rocha from PBFS group at ITQB for the studies on the stability of the di-cluster ferredoxin isoforms from *Sulfolobus metallicus*

To Vitor H. Teixeira, António Baptista and Cláudio Soares, from ITQB for the molecular dynamics simulations on the di-cluster ferredoxin isoforms from *Sulfolobus metallicus*

To Smilja Todorovic from ITQB, Ingo Zebger and Peter Hildebrandt from Technische Universität, Berlin, and Daniel Murgida from the Universidad de Buenos Aires for the FT-IR and RR studies.

To Carlos Salgueiro from the Departamento de Quimíca, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, for the NMR study and useful advices.

To Carlos Frazão, David Aragão, Ricardo Coelho and Maria Arménia Carrondo from ITQB for the crystallographic structure of the di-cluster ferredoxin from *Acidianus ambivalens*

To Manuela Regalla from ITQB for sequencing the di-cluster ferredoxin isoforms from *Sulfolobus metallicus*

To Arnulf Kletzin for receiving me in his laboratory at Darmstadt University and for his collaboration on molecular biology experiments.

To Miguel Teixeira for kindly providing the soluble extract of *Acidianus ambivalens* for ferredoxin purification

To Harald Huber, for having provided the *Sulfolobus metallicus* cells.

Fundação para a Ciência e Tecnologia is acknowledged for financial support, by awarding a PhD Grant SFRH/BD/18653/2004

This work has been funded by the projects POCTI/QUI/37521; POCTI/QUI/45758 and PTDC/QUI/70101 all to Cláudio M. Gomes.

Thesis Publications

Leal S.S., Teixeira M., Gomes C.M. (2004)
Studies on the degradation pathway of iron-sulfur centres during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability
J. Biol. Inorg. Chem. *9*: 987-996.

Leal S.S. and Gomes C.M. (2005)
Linear three-iron centres are unlikely clusters degradation intermediates during unfolding of iron-sulfur
Biol. Chem. *386*: 1295-1300.

Rocha R.M., **Leal S.S.**, Teixeira V.H., Regalla M., Huber H., Batista A.M., Soares C.M., Gomes C.M. (2006)
Natural domain design: enhanced thermal stability of zinc lacking ferredoxin isoforms shows that a hydrophobic core efficiently replaces the structural metal site
Biochemistry *45*: 10376-10384.

Leal S.S. and Gomes C.M. (2007)
Towards the understanding of the folding process and cofactor assembly in iron-sulfur proteins: structural characterisation of a ferredoxin molten globule state
Proteins *68: 606-616.*

Todorovic S., **Leal S.S.**, Salgueiro C.A., Zebger I., Hildebrandt P., Murgida D.H., Gomes C.M. (2007)
A spectroscopic study of the temperature induced modifications on ferredoxin folding and iron-sulfur moieties
Biochemistry *46:* 10733-10738

Frazão, C., Aragão, D., Coelho, R., **Leal S.S.,** Gomes, C.M., Teixeira, M., Carrondo, M.A. (2008)
Crystallographic analysis of the intact metal centres $[3Fe\text{-}4S]^{1+/0}$ and $[4Fe\text{-}4S]^{2+/1+}$ in a Zn^{2+} - containing ferredoxin"
FEBS Lett. *582*: 763-767

Leal S.S. and Gomes C.M. (2008)
On the relative contribution of ionic interactions over iron-sulfur clusters to ferredoxin stability
BBA – Proteins and Proteomics *in Press* doi:10.1016/j.bbapap.2008.05.001

Other publications not included in this thesis
Boscolo B., **Leal S.S.**, Ghibaud E.M., Gomes C.M. (2007)
Lactoperoxidase folding and catalysis relies on the stabilization of the α-helix rich core domain
Bioch. Biophys. Act. 1774 (9): 1164-72

Dissertation Abstract

Metal ions are present in over thirty percent of known proteins. Apart from a well established function in catalysis and electron transfer, metals and metal centres are also important structural elements which may as well play a key role in modulating protein folding and stability. In this respect, cofactors can act not only as local structural stabilizing elements in the native state, contributing to the maintenance of a given specific structural fold, but may also function as potential nucleation points during the protein folding process.

This dissertation addresses the role of iron-sulfur clusters and a zinc centre in the folding and overall stability of an Archaeal family of di-cluster ferredoxins commonly found among hyperthermophiles. These ferredoxins are small proteins of about 100 amino acids that present a [3Fe-4S] and a [4Fe-4S] centre, as well as an N-terminal extension of ~30 residues that holds a zinc centre and wraps the Fe-S containing core. Interestingly, some highly identical ferredoxin isoforms are simultaneously expressed in some organisms that either contain or lack the zinc site. In this respect, the conformational properties of the zinc-containing (FdA) and zinc-lacking (FdB) isoforms present in *Sulfolobus metallicus* were used as models, to investigate the effect of the absence of this site on ferredoxin folding and stability. The studies have indicated that FdB assumes a fold identical to that of the Zn^{2+} containing isoform. The thermal stability of the isoforms was investigated in a broad pH range (2<pH<10), and surprisingly the Zn^{2+} lacking isoform was always found to be more stable than its Zn^{2+} containing counterpart: a $\Delta T_m \approx 9°C$ is determined at pH 7 a difference which becomes even more significant at extreme pH values, reaching a $\Delta T_m \approx 24°C$ at pH 2 and 10. The contribution of the Zn^{2+} site to ferredoxin stability was further resolved using selective metal chelators: during thermal unfolding the zinc scavenger TPEN significantly lowers the T_m in FdA (≈ 10°C) whereas it has no effect in FdB. This shows that the Zn^{2+} site contributes to ferredoxin stability but that FdB has devised a structural strategy that accounts for an enhanced stability without using a metal cross linker. Analysis of the FdB sequence and structural model leads us to propose that the higher stability of this ferredoxin that lacks the zinc centre results from van der Waals contacts formed between the residues that occupy the

same spatial region where the zinc ligands are found in FdA. These favour the formation of a novel local stabilizing hydrophobic core and illustrate a strategy of natural fold design.

One of the aspects focused in this work was the behavior of the Fe-S clusters during ferredoxin unfolding. This fallowed previous studies to this dissertation that suggested that upon protein unfolding the iron-sulfur clusters degraded via linear three-iron sulfur centre species with a transient absorption spectra resembling the one observed in purple aconitase. In order to appraise this hypothesis we proceeded with a detailed kinetic and spectroscopic investigation on the alkaline chemical denaturation of the protein, in an attempt to elucidate the degradation pathway of the iron-sulfur centres in respect to protein unfolding events. For this purpose we investigated cluster dissociation, iron release and protein unfolding by complementary biophysical techniques. We found that shortly after initial protein unfolding, iron release proceeds monophasically at a rate comparable to that of cluster degradation, and that during the process no typical EPR features of linear three-iron sulfur centres are observed. Further, it was observed that EDTA prevents formation of the intermediate and that sulfide significantly enhances its intensity and lifetime, even after protein unfolding. Altogether, our data suggests that iron sulfides, which are formed from the release of iron and sulfide resulting from cluster degradation during protein unfolding in alkaline conditions, are in fact responsible for the observed intermediate spectral species, thus disproving the hypothesis suggesting the presence of a linear three-iron centre intermediate.

In sequence of the results obtained for the di-cluster ferredoxin, we further extended these studies to other proteins to further clarify that linear three-iron centres are most likely not degradation intermediates of iron-sulfur clusters proteins. In agreement we performed studies on proteins containing iron-sulfur clusters, iron-sulfur centres and di-iron centres in respect to their chemical degradation kinetics at high pH, in the presence and absence of exogenous sulfide, to investigate the possible formation of linear three-iron centres during protein unfolding. Our spectroscopic and kinetic data shows that in these different proteins an identical visible absorption spectra to the one that have been suggested to arise from linear three-iron centres, are formed. Iron release and

protein unfolding kinetics show that these bands result from formation of iron sulfides at pH 10, produced from the degradation of the iron centres, and not from rearrangements leading to linear three-iron centres. Thus, at this point, any relevant functional role of linear three-iron centres as cluster degradation intermediates in iron-sulfur proteins remains elusive.

In order to evaluate the relative contribution of iron–sulfur clusters in respect to ionic interactions to the overall protein stability, we have investigated the interplay between structural features and metal cofactors integrity as a function of pH in the ferredoxin model. Changes in protonation affect both the stability and the conformational dynamics of the protein fold. In the pH 5.5–8 interval, the protein has a high melting temperature (T_m ~ 120°C), which decreases towards pH extremes. Acidification triggers events in two steps: down to the isoelectric point (pH 3.5) the Fe-S clusters remain unchanged, the secondary structure content increases and the single Trp becomes more solvent shielded, denoting a more compact fold. Further acidification down to pH 2 sets off exposure of the hydrophobic core and Fe-S cluster disintegration, yielding a molten globule state. The relative stabilising contribution of the clusters becomes evident when stabilising ionic interactions are switched off as a result of poising the protein at pH 3.5, at an overall null charge: under these conditions, the Fe-S clusters disassemble at T_m= 72 °C, whereas the protein unfolds at T_m= 52°C. Overall, this ferredoxin denotes a considerable structural plasticity around its native conformation, a property which appears to depend more on the integrity of its metal clusters rather than on the status of its stabilising electrostatic interactions. The latter however play a relevant role in determining the protein thermal stability.

The biological insertion of iron-sulfur clusters involves the interaction of (metallo) chaperons with a target polypeptide that is not yet folded. In this respect, the study of nonnative protein conformations in iron-sulfur proteins is relevant for the understanding of the folding process and cofactor assembly. We have further investigated the formation of an apo molten globule state in the di-cluster ferredoxin. Biophysical studies have shown that, at pH 2.5, ferredoxin retains structural folding and metal centers. However, upon increasing the temperature, a series of successive modifications occur within the protein structure: Fe-S

disassembly, loss of tertiary contacts and dissociation of the Zn^{2+} site, which is simultaneous to alterations on the secondary structure. Upon cooling, an apo-ferredoxin state is obtained, with characteristics of a molten globule: compactness identical to the native form; similar secondary structure evidenced by far-UV CD; no near-UV CD detected tertiary contacts; and an exposure of the hydrophobic surface evidenced by 1-anilino naphthalene-8-sulfonic acid (ANS) binding. In contrast to the native form, this apo ferredoxin state undergoes reversible thermal and chemical unfolding. Its conformational stability was investigated by guanidinium chloride denaturation and this state is ~1.5 kcal mol^{-1} destabilised in respect to the holo ferredoxin. The single tryptophan located nearby the Fe-S pocket probed the conformational dynamics of the molten globule state: fluorescence quenching, red edge emission shift analysis and resonance energy transfer to bound ANS evidenced a restricted mobility and confinement within a hydrophobic environment. This rather structured, stable and flexible apo molten globule state lead us to speculate on the possible physiological relevance of this conformational state in the folding of Fe-S cluster proteins and speculate on the possible relevance of this apo state as a on-pathway apo conformation for metal center insertion.

Subsequently, we have assigned in detail the specific alterations occurring on the secondary structure during the formation of the molten globule state, and we have individually screened the transformations occurring on the [3Fe-4S] and [4Fe-4S] clusters along the process. Our main goal with this work was to infer on the possible relevance of this molten globule state in the folding pathway and discriminate the contribution of each cluster for the overall stability of the ferredoxin. In agreement, thermal perturbation of the ferredoxin was investigated employing a toolbox of spectroscopic methods. FTIR and visible CD were used for assessing changes of the secondary structure and coarse alterations of the [3Fe-4S] and [4Fe-4S] cluster moieties, respectively. Fine details of the disassembly of the metal centers were revealed by paramagnetic NMR and resonance Raman spectroscopy. Overall, thermally-induced unfolding of *Aa*Fd is initiated with the loss of α-helical content at relatively low temperatures (T_m^{app} ~ 44 °C) followed by the disruption of both iron-sulfur clusters (T_m^{app} ~ 53-60 °C). The degradation of the metal centers triggers major structural

changes on the protein matrix, including the loss of tertiary contacts (T_m^{app} ~ 58 °C) and a change, rather than a significant net loss, of secondary structure (T_m^{app} ~ 60 °C). In fact, this latter process triggers a secondary structure reorganization that leads to an increase in the β-structure content. The combined spectroscopic approach here reported illustrates how changes in the metalloprotein organization are intertwined with disassembly of the iron-sulfur centers, denoting the conformational interplay of the protein backbone with cofactors. In addition it seems to evidence that both clusters disintegrate in a single event and therefore most likely contribute equally to overall stability. Moreover, the pathway for the formation of this molten globule state clearly evidenced the formation of non-native β-structure along the process which strongly suggests that this conformation is most likely an off-pathway specie.

Altogether, the evidences suggest that the integrity of Fe-S clusters do not only stabilizes the native state but is likely to participate and be determinant in the folding process of this family of ferredoxins. In addition, this determinant role of the clusters in the fold of the protein appears to be directly correlated with the ferredoxin hydrophobic core, thus suggesting a simultaneous assembly process.

Resumo da Dissertação

Os iões metálicos são parte integrante de mais de trinta por cento de todas as proteínas identificadas. Para além do seu proeminente envolvimento em funções de catálise e transferência de electrões, os metais e os centros metálicos são também importantes elementos estruturais, podendo simultaneamente desempenhar um papel chave no enrolamento e estabilidade das proteínas. De facto, os cofactores podem não só actuar como importantes elementos locais estabilizadores do estado nativo contribuindo para a manutenção específica da estrutura, mas também funcionar como determinantes pontos de nucleação durante o processo de enrolamento das proteínas.

Esta dissertação visa essencialmente o estudo do papel dos centros de ferro e enxofre mas tambem do centro de zinco no enrolamento e estabilidade de ferredoxinas de sete ferros. Estas proteínas pertencem á família de ferredoxinas, vulgarmente encontradas em organismos hipertermofilicos da ordem Archaea e caracterizadas por conterem um centro [3Fe-4S] e um centro [4Fe-4S], bem como um centro de zinco numa extensão de ~30 a.a no N-terminal. Curiosamente, isoformas sequencialmente muito idênticas destas ferredoxinas mas que não contêm o centro de zinco são simultaneamente expressas. No sentido de investigar o efeito da ausência deste metal na estrutura e estabilidade da ferredoxina, as propriedades conformacionais destas isoformas expressas em *Sulfolobus metallicus* com o centro de zinco (FdA) e que não apresentam o centro (FdB) foram utilizadas como modelo. Os estudos indicam que a FdB assume um enrolamento estrutural muito semelhante á isoforma que contem o centro de zinco FdA. A estabilidade térmica das isoformas foi investigada a diferentes pHs no intervalo 2 <pH <10 e surpreendentemente a FdB foi para todas as condições testadas sempre mais estável que a sua isoforma, a FdA que contem o centro de zinco: a pH 7 foi observado um $\Delta T_m \approx 9°C$, uma diferença que se revelou ainda mais prenunciada a pHs extremos atingindo um $\Delta T_m \approx 24°C$ a pH 2 e 10. A contribuição do centro de zinco para a estabilidade da ferredoxina foi ainda aprofundada utilizando quelantes com selectividade para diferentes metais: a presença do quelante – TPEN caracterizado por uma acentuada especificidade para o zinco, durante a desnaturação térmica das ferredoxinas

teve como efeito na estabilidade térmica uma diminuição acentuada do valor de T_m na FdA (≈ 10°C), não provocando qualquer alteração na FdB. Isto sugere que o centro de zinco contribui para a estabilidade da FdA mas também que a estratégia estrutural alternativa da FdB sem o centro de zinco é mais estável. Analise á sequência e modelo estrutural da FdB sugere que o aumento de estabilidade desta isoforma pode resultar de contactos to tipo van der Waals entre os resíduos que ocupam a mesma região espacial onde os ligandos do zinco se encontram na FdA. Estes contactos favorecem a formação de um centro hidrofóbico estabilizador no local e ilustra uma estratégia natural de enrolamento.

Um dos vários aspectos focados neste trabalho incidiu no comportamento dos centros de Fe-S durante a desnaturação da ferredoxina. Esta análise surgiu em seguimento de estudos anteriormente publicados a esta dissertação que sugeriam que a desnaturação destas ferredoxinas passava pela formação transiente de um centro linear de três ferros como o que foi identificado para a aconitase. Esta hipótese foi essencialmente baseada na elevada semelhança encontrada entre a formação transiente de espécies espectofotometricas durante a desnaturação das ferredoxinas e o espectro de absorção característico do centro linear identificado na aconitase. Esta especulação foi aqui avaliada através de uma investigação espectroscópica e cinética detalhada da desnaturação química alcalina da ferredoxina na tentativa de esclarecer o processo de degradação dos centros de Fe-S em relação á desnaturação da proteína. Neste sentido, investigamos a relação da dissociação dos centros com a libertação de ferro para a solução e a desnaturação da ferredoxina através de técnicas biofísicas complementares. Os resultados claramente evidenciaram que com a desnaturação da ferredoxina o ferro é monofasicamente libertado a uma taxa idêntica á da degradação dos centros de Fe-S e que durante este processo não é observado nenhum sinal de EPR típico de um centro de Fe-S linear. Subsequentemente foi observado que a desnaturação da ferredoxina na presença do quelante EDTA inibe a formação da espécie espectral transiente e que a presença de sulfuretos aumenta significativamente a intensidade bem como tempo de vida destas formas espectrais, mesmo depois de a proteína já ter desnaturado. No seu conjunto estes resultados sugerem que a degradação dos centros de Fe-S durante a desnaturação da ferrdoxina resulta na libertação do

ferro e sulfuretos e sequente formação de sulfetos de ferro em solução que muito provavelmente estão na origem da formação das espécies espectrais intermediarias observadas. Neste sentido os dados apresentados claramente estão em desacordo com a hipótese que sugeria a formação de um centro linear intermediário de três ferros. No seguimento destas evidências com a ferredoxina alargamos este estudo a outras proteínas no sentido de esclarecer duvidas relativamente á possível formação de um centro linear de três ferros durante a desnaturação de proteínas de Fe-S. Nesse sentido, analisamos a cinética de desnaturação química de proteínas com centros de Fe-S contendo e não contendo sulfureto inorgânico bem como proteínas com apenas centros de ferro, na presença e ausência de sulfuretos exógenos. De facto, os resultados claramente evidenciam que é possível reproduzir em todas as proteínas estudadas as espécies espectrais intermediarias associadas á formação de um centro linear de três ferros. A libertação de ferro e a cinética de desnaturação proteica claramente indicam que a formação destas espécies espectrais resultam da formação de sulfuretos de ferro que são produzidos em solução alcalina e não devido ao rearranjo dos centros de Fe-S com formação de um centro linear de três ferros.

No sentido de avaliar o envolvimento das ligações iónicas relativamente á contribuição dos centros de Fe-S para a estabilidade geral da ferredoxina, investigamos a interacção existente entre as características estruturais da proteína e a integridade dos centros de Fe-S em função do pH. Os resultados sugerem que alterações no estado de protonação da proteína afectam a estabilidade e a dinâmica conformacional da estrutura da ferredoxina. No intervalo de pH entre 5.5 e 8 a ferredoxina evidenciou uma média elevada na temperatura de transição (T_m ~120 °C) que contudo diminui significativamente nos extremos de pH, particularmente com a acidificação. A acidificação até ao ponto isoeléctrico da ferredoxina (pH 3.5) não tem qualquer efeito na integridade dos centros de Fe-S e evidencia um ligeiro aumento no conteúdo da estrutura secundaria da proteína e o único Trp da ferrdoxina encontra-se menos exposto ao solvente. Sequente acidificação até pH 2 provoca a exposição do centro hidrofóbico e a desintegração dos centros de Fe-S cluster e dá origem á formação de um estado molten globule. A proeminente contribuição dos centros

de Fe-S para a estabilização relativa da ferredoxina torna-se evidente quando a proteína é incubada ao pH correspondente ao seu ponto isoeléctrico (pH 3.5) e as interacções iónicas nativas são "desligadas": nestas condições os centros de Fe-S dissociam-se a uma temperatura de transição T_m= 72 °C, enquanto que outros elementos estruturais apresentam uma temperatura de transição significativamente inferior com um T_m= 52°C. No geral a ferredoxina evidencia uma elevada plasticidade estrutural em relação á estrutura nativa, uma propriedade que aparenta estar mais dependente da integridade dos centros de Fe-S do que no estado das suas interacções iónicas. No entanto as interacções iónicas nativas evidenciaram um papel determinante na estabilidade térmica da proteína.

A inserção biológica dos centros de Fe-S envolve a interacção de *chaperons* moleculares com uma dada forma proteica parcialmente enrolada. Neste sentido o estudo de formas não nativas em proteínas de Fe-S pode contribuir para uma melhor compreensão do processo de enrolamento destas proteínas bem como do processo de inserção dos centros. Assim procedemos á investigação da formação e caracterização de um estado *apo molten globule* na ferredoxina de sete ferros. Estudos biofísicos revelaram que a pH 2.5 a proteína mantém a integridade dos centros metálicos e a sua estrutura nativa. No entanto ocorrem na estrutura da proteína uma serie de modificações com o aumento da temperatura: nomeadamente os centros de Fe-S bem como o de zinco dissociam-se alem de ocorrer uma perda acentuada dos contactos terciários e alterações ao nível da estrutura secundária. Após o arrefecimento obtém-se uma forma apo da ferredoxina com todas as características de um estado *molten globule*: compacticidade idêntica ao estado nativo, elevado conteúdo de estrutura secundária e redução acentuada de contactos terciários bem como uma exposição significativa do centro hidrofóbico. Em contraste com o estado nativo esta apo forma da ferredoxina evidencia uma desnaturação térmica e química reversível. A análise da sua estabilidade conformacional através da desnaturação química evidenciou que esta apo forma encontra-se ~1.5 kcal mol^{-1} destabilizada em relação á forma holo da ferredoxina. O estudo das características do único Trp que constitui a ferredoxina e que se encontra localizado muito próximo dos centros de Fe-S foi utilizado na avaliação da dinâmica configuracional desta

região da proteína na ausência dos centros metálicos. Resultados na análise de diferentes propriedades de fluorescência deste resíduo evidenciam que o Trp apresenta uma mobilidade muito restrita num ambiente altamente hidrofóbico. Este estado molten globule bastante estruturado e estável levou-nos a especular na possível relevância fisiológica deste estado configuracional no enrolamento e inserção dos centros de Fe-S nestas proteínas. Neste sentido, avaliamos posteriormente em detalhe as alterações que ocorrem ao nível da estrutura secundária da ferrdoxina durante a formação do estado *molten globule* e monitorizamos individualmente as transformações que ocorrem nos respectivos centros [3Fe-4S] e [4Fe-4S] durante o processo. O principal objectivo com este trabalho foi simultaneamente inferir na possível relevância deste estado molten globule durante o enrolamento da proteína e descriminar individualmente a contribuição relativa de cada centro de Fe-S para a estabilidade global da proteína. Neste sentido, a estrutura secundária da ferredoxina foi detalhadamente monitorizada por FT-IR durante a desnaturação térmica da proteína. A dissociação do centro [3Fe-4S] e do [4Fe-4S] foi individualmente analisada através de espectroscopia de RMN e RR. Os resultados evidenciaram que a desnaturação térmica da ferredoxina tem início com a perda de α-hélices a relativamente baixa temperatura (T_m^{app} ~ 44 °C) e que é seguida da desintegração simultânea de ambos os centros (T_m^{app} ~ 53-60 °C). Esta dissociação despoleta alterações estruturais mais significativas na matriz da ferredoxina incluindo a perda de contactos terciários (T_m^{app} ~ 58 °C) e uma alteração no conteúdo da estrutura secundária (T_m^{app} ~ 60 °C). De facto este último evento é caracterizado por uma reorganização acentuada da estrutura secundária da ferredoxina e que resulta num aumento bastante significativo no seu conteúdo de folhas-β. A combinação desta abordagem espectroscópica ilustra claramente como a dissociação dos centros de Fe-S se encontra estreitamente ligada a alterações na organização estrutural da ferredoxina, salientando a elevada interacção da estrutura da proteína com os cofactores. Este estudo evidencia também que ambos os centros de Fe-S se dissociam simultaneamente da proteína sugerindo que contribuem igualmente para a estabilidade geral da ferredoxina. Foi também demonstrado que a origem deste estado molten globule passa pela formação de estrutura β não nativa o que

sugere que esta espécie pode ser o resultado de um enrolamento incorrecto da ferredoxina e portanto é pouco provável que possa representar a forma apo da ferredoxina onde os centros são inseridos.

No seu conjunto os resultados sugerem que a integridade dos centros de Fe-S não é apenas determinante para a estabilidade conformacional da ferredoxina mas também presumivelmente essencial no processo de enrolamento da proteína. Além disso o papel determinante dos centros na estrutura da proteína aparenta estar directamente relacionado com a integridade do centro hidrofóbico o que pode sugerir um mecanismo de formação concomitante.

TABLE OF CONTENTS

1 PROTEIN FOLDING, STABILITY AND METAL IONS: AN INTRODUCTION

CONTENTS

1.1 OVERVIEW

Protein folding is a puzzling process. Proteins appear to fold into a unique native conformation, in spite of an astronomical number of alternative configurations. In regard, the introduction of a theoretical framework based on the global properties of the energy landscape helps to explain the simplest models of protein folding and provides an estimable approach for the understanding of more complex protein molecules. This approach to the folding process has in fact already contributed to a substantial improvement in protein structure prediction and design [1-4], as well as in the construction of metal sites [5], in addition to a emergent ability in engineering efficient foldable artificial polymers for practical applications [6-9]. Even so, a full comprehension of much of the priorities of folding forces and a quantitative understanding of the factors that define single low-energy fold remains yet rather limited. In fact, it appears that not all amino acids and therefore guiding folding forces, are equally important in specifying which fold is adopted; as proteins with high sequence identity can result in very distinct folds [10].

One of the main goals of protein folding studies relies not only in the ability to predict the protein structure from its amino acid sequence but also, and in particular, to quite understand the pathways and mechanisms of the folding process. In fact, a detailed knowledge on protein folding can greatly contribute to a better comprehension of several human pathologies associated with misfolded and aggregated proteins [11], as well as improve our understanding on the role of co-factors in protein folding, stability as well as metals assembly. In this respect, the elucidation of the kinetic and thermodynamic properties of a protein *in vitro* can be a step forward on the way to characterize its complete folding pathway. Subsequent steps can comprise the characterization of transiently formed intermediates and of the transition states between the various states of the protein and examining the factors that may influence *in vivo* folding, like crowding effects of the cellular environment and the presence of folding assistants [12].

This chapter presents the state of knowledge on several aspects of the challenging process of protein folding, raising simultaneously relevant issues

regarding the instituted approaches to the phenomena. The contextualization of protein folding events within the biological and evolutionary scenario is also discussed.

1.2 THE PROTEIN FOLDING EVENT

Protein folding is the remarkable process by which a polypeptide chain acquires a specific and organised three-dimensional structure. Although the cellular environment can enclose various factors involved in the development of an efficient folding process [13], the code-key for folding is strictly contained in the amino-acid sequence itself. In fact, many proteins are capable to properly self-assemble *in vitro* within a biologically relevant time scale [14-17]. The emerging approach to explain this phenomena is largely established on the energy landscape theory and the funnel perspective of the folding process [18-25]. This concept, based on theoretical and experimental data, proposes that the polypeptide chain folds up under the driving force of a free energy gradient throughout a progressive assembly of partially folded structures, until the protein reaches its thermodynamically most stable conformation. The required landscape to enable a protein to fold efficiently has been likened to a funnel because the conformational space accessible to the polypeptide chain is reduced as the native state is approached. In essence, the high degree of disorder of the polypeptide chain is reduced as folding progresses, as the enthalpy associated with stable native-like interactions can offset the decreasing entropy as the structure becomes more ordered. The conceptual basis of such mechanism is shown in figure 1.1. Nevertheless, if a sequence is chosen at random the specificity of the resulting structure can be very low and result in an ensemble of distinct ground states [26, 27]. This behaviour within the biologic scenario, would bring up serious restrains on the viability of proteins and raise disastrous consequences in the evolution of organism [23]. Thence, naturally occurring proteins must have evolved to have sequences able to achieve efficient folding under biological conditions. In fact, the robustness and predictability of biological self-assembly in comparison to related processes in non-biological systems is arguably the most remarkable feature of living systems. As a result of evolutionary pressure, the

interactions present in the functional native state are not in conflict and do not give rise to the concurrent presence of competing interactions that may result in distinct ground states, like it was observed for a random polymer. Instead, folding interactions must be mutually supportive and cooperatively lead to a single low-energy structure. By avoiding the frustrating conflicts between different energetic biases, proteins have most likely evolved a funnelling energy landscape that optimizes native structure-seeking interactions, while selecting against interactions leading to non-productive traps.

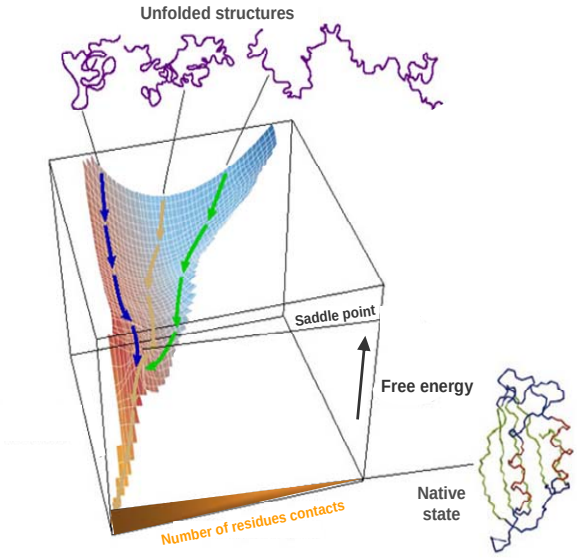


Figure 1.1 | Schematic energy landscape for protein folding.
The surface is derived from a computer simulation of the folding of a highly simplified model of a small protein. The surface 'funnels' the multitude of denatured conformations to the unique native structure, where highly simplified trajectories are indicated. The critical region on a simple surface such as this one is the saddle point corresponding to the transition state, the barrier that all molecules must cross if they are to fold to the native state. Adapted from [28]

Accordingly, the folding process does not necessarily involve a series of mandatory steps between specific partially folded states, but rather consist of a stochastic search of the many conformations accessible to a polypeptide chain [29, 30]. This model reduces significantly the difficult problem of configurational search through traps to a much smaller parallel search of the configuration space.

Once the energetic frustrations associated with conflicting interactions have been minimized, the topology of the protein becomes the key determinant of the folding mechanism, encoding the interplay between stabilizing interactions and chain entropy. The funneled organization of the energy landscape therefore dominates the kinetics of folding, and is responsible for the robust ability of proteins to fold. In fact, it seems rather evident that this process is extremely efficient for those special sequences that have been selected during evolution to fold to globular structures, where only a very small number of all possible conformations need to be sampled during the search process. However, the funnel theory does not necessarily implicate a single pathway; rather, a multiplicity of downhill folding routes is likely to co-exist in parallel, via distinct mechanisms [31]. In agreement, the heterogeneity in the folding kinetics which is in some cases observed when the process is monitored using distinct probes, or a non-exponential kinetics, has been interpreted as an evidence for a downhill folding mechanism [32]. In addition, it is established that in some cases, the same protein may fold via different pathways for the reason that the folding conditions changed, and can affect differently the stabilities of the diverse substructures along the folding pathway [33-35]. Also, the binding of cofactors can be accountable to induce a change in the folding pathway of the protein [36]. Moreover, the occurrence of mutations may be overcome by a simple shift in the folding route [37, 38], as long as native interactions prevail over the non-native ones [39]. Evolution is likely to have selected a funneling landscape with multiple folding pathways as a folding model, as a way to ensure successful *in vivo* protein folding under the varying conditions observed in the cellular context.

Structure-based Gō-type minimalist models represent the perfectly funneled energy landscape that is established on a folding process with minimal energetic frustration. This model implies that native interactions are attractive and non-native interactions are essentially repulsive which precede the absence of energetic conflicts during the folding pathway. Such models have impressively reproduced the experimentally observed folding mechanism of many single-domain proteins [40, 41] and folding/binding of dimers [42, 43]. This quantitative form of minimal frustration of proteins has also been successfully used to improve the energy functions used in protein structure prediction, and in protein design.

Although detailed mechanisms of folding for complex proteins is still to be achieved, in general the funnel landscape is likely to be enforced in all proteins. Such common patterns can be inferred from the perfect funnel models — as most naturally occurring proteins has to have sufficiently reduced energetic frustration, or otherwise would had been deleted by evolution. The funnel landscape idea implies the notion, as a general guideline, that topology determines the mechanism of folding, and that the topography of the global landscape has been determined by evolutionary pressure [24]

Protein folding is addressed by multiple approaches, both experimental and theoretical or computational. Among these are the experimental monitoring of protein unfolding and refolding using biophysical techniques, frequently involving the use of site-directed mutagenesis to probe the roles of individual residues during the folding process [44]. Theoretical approaches are also important in the understanding of the protein folding event, particularly those based on computer simulations of the events occurring during folding, or more often unfolding, as this process is easier to simulate and can be related to the complementary folding reaction [21, 45]. The ultimate objective of such studies is to define the complete energy landscape for the folding reaction, and to understand in detail how this is defined by the primary sequence.

Setting off protein folding: mechanisms to fold

Given the diversity of protein structures and the evolutionary pressure on function, a unique mechanism for folding, although rather appealing, seems nevertheless very unlikely. In fact, a variety of distinct folding mechanisms has been described even for small single-domain proteins, where secondary structures can form before or after collapse and side-chains can order before or after the main chain topology. In this regard, several models have emerged in order to aid in categorizing the different mechanisms of proteins folding (figure 1.2).

The *diffusion-collision model*, proposes that precedence is given to the formation of a diffuse transition state with some secondary structure that further nucleates the tertiary contacts [46, 47]. *Nucleation models*, suggest that a folding nucleus forms initially, probably in the early unfolded protein, and that this nucleus can act

as a scaffold for the build up of the rest of the structure [48], whereas other views suggests that proteins may fold via the stepwise assembly of structural subunits called *foldons* [49]. In the *hydrophobic collapse models*, a non-specific hydrophobic collapse is expected to occur as the initial step [50]. Clearly, a unified mechanism able to describe the protein folding reaction cannot be envisaged. Even so, an investigation on possible common features in folding reaction has suggested the possibility that some distinct pathways can be in fact the manifestation on an underlying common feature [51].

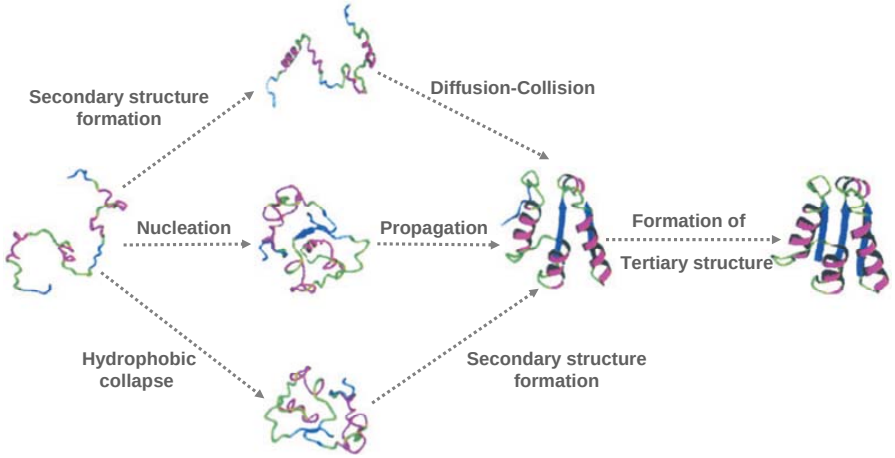


Figure 1.2 | Pathways for protein folding.
In the framework model, precedence is given to the formation of secondary structural units. In the hydrophobic collapse model, precedence is given to an initial chain collapse. In the nucleation-condensation model, an extended nucleus is formed early during folding. Molten globule-like intermediates accumulate during the folding of many proteins. For some proteins, particularly those following nucleation mechanisms, a molten globule intermediate does not usually accumulate. Although different folding pathways are usually discussed in the context of different proteins, can a single protein utilize fundamentally different folding pathways in different folding conditions? For example at very low temperatures, at which hydrophobic interactions are weakened, a protein could conceivably switch from a hydrophobic collapse to a framework mechanism. Adapted from [52]

In general, different folding pathways are usually discussed in the context of different proteins and as no particular evolutionary advantage is apparent for any of these mechanisms, it is reasonable to expect to see examples of them all. The big remaining question is whether a single protein exploits fundamentally different folding mechanisms in different folding conditions, or just uses a different pathway of the same mechanism. Also it would be rather interesting to determine if a

particularly folding mechanism is directly associated with the amino-acid sequence or with any specific structural features in the protein.

|Two state versus multi-state protein folding

The most elementary models of protein folding are the apparent two-state systems. In this minimalist framework, protein unfolding and refolding are considered monophasic processes, where only the unfolded state (U) and the folded native state (F) are populated (figure 1.3 a|). Simultaneously, a single energetic barrier separates the folded from the unfolded state (figure 1.3 b|).

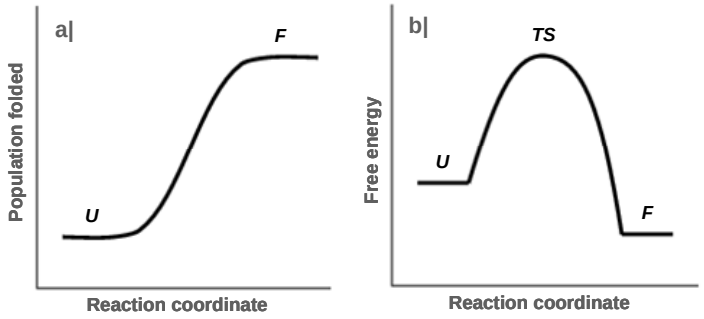


Figure 1.3 | Outline of a hypothetical two-state protein folding reaction
a| Side view of a cooperative folding/unfolding process where only the unfolded (*U*) and the folded (*F*) state are populated species **b|** Thermodynamic profile where a single energy barrier separates the unfolded (*U*) from the folded (*F*) state with the transition state (*TS*) as point along the reaction coordinate with the highest free energy.

This classical model ought to present an agreement between the free energies for folding inferred from equilibrium and kinetic measurements, and reveal a linear relationship of the free energy with protein unfolding. When these principles do not apply, this occurrence is often associated with an intermediate (on or off the folding pathway) and therefore, to a multi-state folding process. However, in some cases this border does not appear to be well delimited. In fact, it has been observed a non-linear free energy relationship for proteins that fold without apparent populating intermediate states [53, 54] and proteins which have been referred to fold via two-state processes that present kinetic bursts, which are normally attributed to the formation of intermediate states [55, 56]. On the other

hand, fast kinetics can reveal intermediates in apparent two-state unfolding reactions, by showing unfolding intermediates in the msec time range even though both the major folding and unfolding reactions follow a single exponential time course in the second's time range. In agreement, NMR methods were able to identify intermediates in otherwise thought to be two-state folding reactions [57]. These evidences appear to suggest that the two-state folding model, at least in some cases maybe under estimated. In fact, a nonlinear energy relationship in protein unfolding from an apparent two-state folding protein has been associated with the existence of unstable intermediates [58-60]. Moreover, these high-energy intermediates can be stabilized relatively to the native state by a change in the folding conditions [61, 62] or by mutation [63], and therefore an apparent two-state folding protein can be turned into a multi-state folder. However, is still very difficult in these conditions to evaluate whether the intermediate or transition state that is being stabilized is on the same pathway as it was before the folding perturbation, or if in fact it is just the exploitation of an alternative folding route [35].

Conversely, several proteins have been shown to fold typically to their native states via a populated intermediate [64-67]. Interestingly, in RNase H from *Escherichia coli*, it was also possible to alter the folding kinetics, converting a multi-state folder into an apparent two-state kinetic process [68], simply by destabilizing the intermediate state relatively to the unfolded one. The key question is if there are any general features that determine the fact that a protein folds through a stable intermediate state, or not. Actually, only the chain length and the protein stability appear to differentiating features in proteins that fold by an apparent two state or a multi-state pathway [69]. This is consistent with several studies where proteins that fold via multi-state kinetics are generally more stable and feature a longer amino-acid chain (> ~100 a.a) than proteins known to fold with apparent two-state kinetics. Moreover, a detailed comparison on the folding rates of the multi-state and apparent two-state folders, determined that they are both similarly dependent on the parameters that reflect the native backbone topology, like the absolute contact order and sequence-distant native pairs [70]. Thus, this may suggest that the mechanisms behind multi-state and apparent two-state folding are essentially identical and just reflect the folding hierarchy of

the native three-dimensional structure. Accordingly, the two-state folding process probably represents a merely simplified version of hierarchical folding. In overview, it appears that the folding pathway of a given protein is strongly dependent on the folding conditions and not exclusively on the characteristics of the protein itself: given that it is possible to convert an apparent two-state folder to a multi-sate folding reaction, and *vice-versa*. In agreement, it comes out that intermediates can act as valuable signposts for identifying possible changes in the apparent multiplicity of folding pathways. Moreover, it seems very likely that an apparent two-state folding pathway does not imply for certain that no intermediate exists, but simply that no intermediate could be detected. Therefore, it is most likely that the folding mechanism is in fact a unified one that depends on the native backbone topology and proceeds through a range of multi-states, with no single sequential route [52, 71].

Role of intermediates

The role of intermediate states on folding pathways has been at the least controversial. It has been suggested that the presence of intermediate states on folding pathways may slow the folding process, as these constitute a energetic trap on the pathway that otherwise needs to be reversed in order to get the proper fold [72, 73]. As well, and in complete disagreement, it was proposed that the partially folded states can in fact enhance the rate of protein folding, by guiding the folding polypeptide chains to low-energy transition states that otherwise may not become accessible directly from the unfolded state [74]. Moreover, it has been noted that intermediates must play a productive role to a correct folding [75-77], but also pointed out that intermediates can be off-pathway species leading to misfolded species [78, 79].

Contradictory points of view and experimental results also arise from the way that intermediates populate the folding pathway. A classical view suggests that all of the protein population folds through a sequence of intermediates predetermined by the foldon substructure of the target protein, and a sequential stabilization principle according to which prior native-like structure templates the formation of subsequent complementary structure(s). Thus, the folding pathway must be

determined by the same cooperative interactions that determine the target native structure. Several experimental data was shown to support this view of protein folding in which all of the molecules in a refolding population fold essentially through the same intermediate structures [80, 81]. On the other hand, intermediates are often interpreted not has discrete pre-determined conformations in the folding pathway, but in terms of an ensemble of conformations, where the transition from one ensemble of structures to the next one on the folding pathway can happen on independent parallel routes [82, 83]. This model of multiple pathways further suggests that specific populated intermediates do not need necessarily to exist and that partially folded conformations can in fact represent the slower multi-state fractions [84-86]. This heterogeneity has been evidenced within several proteins as distinct coexisting subpopulations [87-90] and also within families of proteins with similar folds [91]. In order to address these conflicting evidences on folding intermediates, a convergent hypothesis was recently proposed that merges the opposite explanations [92]. This unifying approach suggests that these discrepant interpretations can be resolved by modifying the predetermined pathway model to include probabilistic misfolding errors that can block the forward progress of normally occurring intermediates. Chance misfolding errors can corrupt different intermediates and insert optional error-repair barriers at different points in a pathway. When the error probability is zero at all steps of the pathway, folding appears to be a two-state process. When it is unity at one particular step, three-state folding occurs. Any other values or combinations will produce mixed behavior in which different population fractions display different naturally occurring but partially corrupted intermediates, or none at all, and fold at different rates. This heterogeneous behavior, when detected by the usual spectroscopic observations of kinetic phases, will appear to represent multiple alternative pathways. This hypothesis seems able to explain a varied folding behavior of proteins quite generally grounded in the observation that folding errors are ubiquitous. Well known misfolding errors include prolyl and non-prolyl peptide bond mis-isomerization, transient aggregation, formation of non-native hydrophobic clusters, disulfide shuffling, cofactor misligation, and perhaps nonnative domain docking modes. These errors are optional, not intrinsic to the folding process, and they can often be inserted or removed by the manipulation of

folding conditions. This goes in agreement with the evidence that the folding conditions can preferentially influence the sub-populations that become most populated during the folding pathway [33], thus implying that the folding pathway and observed intermediates for a given protein can be different under different conditions. Altogether, these findings seem to substantiate that the task of monitoring and characterizing intermediates can be a rather slippery job. As a result, it is not yet possible to ascertain if folding intermediates are a mere result of protein folding that characterizes a given folding pathway or if in fact they play a determinant role in guiding and defining the pathway.

Off-pathway phenomena: misfolding and aggregation

A misfolded protein is a protein that failed to fold properly. The misfolding process results when the protein acquires a number of persistent non-native interactions that affect its overall architecture and/or its properties in a biologically significant manner, like the loss of function. In addition, misfolded conformations often expose hydrophobic amino acid residues and segments of unstructured polypeptide backbone to the solvent thus promoting inter-chain hydrogen bonding and hydrophobic interactions that can lead to aggregation.

Oddly, folding errors in vivo appear to be a common event, since thirty percent or more of all synthesized polypeptides do not reach the final native folded conformation [93]. In order to correct these mistakes and prevent them to cause damages, biology finds it cost-effective to elaborate multiple helper proteins and error repair systems. Only when these misfolded proteins escape the cellular quality control mechanisms that the folding error can become a serious problem to the cells and lead to highly debilitating and prevalent pathologies. These pathologies can result from mutations that lead to misfolded forms (like mutations in CFTR in cystic fibrosis), or simply result from normally soluble proteins that are suddenly converted to insoluble aggregates (often presenting a well-defined fibrillar nature known as amyloid), which occur in more than twenty diseases, like Alzheimer's and Parkinson's disease. Remarkably, despite the range of distinct proteins involved in these amyloid diseases, the fibrils in which they are found in the disease states are extremely similar [94]. Moreover, amyloid fibrils can also

be observed in soluble proteins with no recognized connection to any known disease, frequently just by lowering the pH or increasing the temperature above that required for unfolding [95, 96]. In agreement, it is also possible in many cases to reproduce under laboratory conditions the structural transitions of the disease-associated molecules, by exposing the folded proteins to mildly denaturing conditions [97]. Altogether, these evidences suggest that the ability to form amyloid fibrils can in fact be an intrinsic property of polypeptide chains. In agreement with this suggestion is also the fact that the intermolecular bonds that stabilize amyloid fibrils are known to involve the peptide backbone, which is common to all proteins. Therefore, amyloid conformation can be a possible common state for all proteins regarding that favourable conditions are provided [98]. Thus, except when the protein is exposed to specific conditions, the peptide backbone is not accessible to form the inter-chain hydrogen bonds associated with amyloid fibrils.

The reason why only a few proteins actually form amyloid aggregates *in vivo* is likely to be influenced by the protein amino-acid sequence that defines the degree of exposed surfaces prone to aggregation in the native state [99], or become exposed by specific conditions in the cell. It is interesting to speculate that avoidance of aggregation, particularly to highly insoluble amyloid fibrils, might be an equally important driving force in the evolutionary design of natural proteins. Biology must has found a way to avoid the formation of this unwanted material under normal physiological conditions where proteins that persist to form amyloid aggregates must likely represent the exception that need further evolutionary improvements. In fact, many factors must be involved in this protective mechanism, but the selection of sequences during evolution that can fold efficiently to a globular form in which the polypeptide chain and the hydrophobic residues are hidden in the interior is likely to be particularly important [100, 101].

Assistants in folding

Within the cells of living organisms there is a large numbers of auxiliary factors that assist in the folding process of many protein [102]. Although the native structure of a protein is encoded in its amino-acid sequence, the process of *in*

vivo folding often requires assistance to either promote proper folding, improve the yield in many folding reactions or to play a role in a post-translational quality control system and maintain the proper conformation of proteins under changing environmental conditions. Such helping factors can include a vast network of *i) molecular chaperones* [103], as well as the effect of *ii) chemical chaperoning* [104]–in addition to the ambiguous influence of the *iii) molecular crowding* environment in the cell over proteins folding and assembly [105].

i) Molecular chaperones are a group of structurally diverse and mechanistically distinct proteins that share the ability to interact with the non-native conformation of other proteins (hydrophobic residues and/or unstructured backbone regions, i.e., structural features typically exposed by non-native proteins normally buried upon completion of folding). The chaperone binding to the yet non-native proteins does not input any conformational information to the folding process, but mainly serves to shield the interactive surfaces of non-native polypeptides from intermolecular aggregation and to prevent, or reverse, intramolecular misfolding [106]. This chaperone assistance is particularly active during protein co- and post-translation events but also under conditions of stress during which proteins may unfold and aggregate. Interestingly, a large family of molecular chaperones, the heat-shock proteins (Hsps) has an increased level of synthesis under stress conditions, such as high temperature [102]. Hsps, are often classified based on their molecular weight (hsp10, hsp40, hsp60, hsp70, hsp90, etc) and are highly conserved in all domains of life. In fact, a chaperone-assisted folding mechanism is present in the three domains of life, and it is mainly differentiated by the individual characteristics of the intervening chaperone proteins in each domain (figure 1.4). The conserved working strategy of chaperones within all domains for de novo protein folding relies on the cooperation of two distinct classes of chaperones in a topologically and timely ordered manner. The first class of chaperones, usually addressed as the nascent chain-binding chaperones includes chaperones that bind to the ribosome, such as trigger factor (TF) in eubacteria, and the nascent chain associated complex (NAC) in eucarya; as well as chaperones that do not associate directly with the ribosome, like the bacterial Hsp70 system (DnaK with its co-chaperone Hsp40_DnaJ) and the chaperone prefoldin (PFD) present in archaeal and eucaryotic cytosol. The role of this class

of chaperones is associated with the ability to hold nascent and newly synthesized chains in a flexible state competent for subsequent folding.

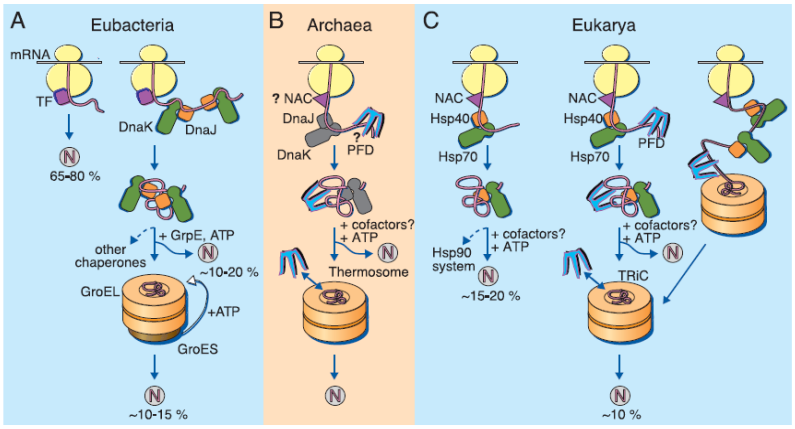


Figure 1.4 | Models for the chaperone-assisted folding of newly synthesized polypeptides in the cytosol.
(**A**) Eubacteria. TF, trigger factor; N, native protein. Nascent chains probably interact generally with TF, and most small proteins (~65 to 80% of total) fold rapidly upon synthesis without further assistance. Longer chains (~10 to 20% of total) interact subsequently with DnaK and DnaJ and fold upon one or several cycles of ATP-dependent binding and release. About 10 to 15% of chains transit the chaperonin system GroEL and GroES for folding. GroEL does not bind to nascent chains and is thus likely to receive an appreciable fraction of its substrates after their interaction with DnaK. (**B**) Archaea. PFD, prefoldin; NAC, nascent chain associated complex. Only some archaeal species contain DnaK/DnaJ. The existence of a ribosome-bound NAC homolog, as well as the interaction of PFD with nascent chains, has not yet been confirmed experimentally. (**C**) Eukarya - the example of the mammalian cytosol. Like TF, NAC probably interacts generally with nascent chains. The majority of small chains may fold upon ribosome release without further assistance. About 15 to 20% of chains reach their native states in a reaction assisted by Hsp70 and Hsp40, and a fraction of these must be transferred to Hsp90 for folding. About 10% of chains are co- or posttranslationally passed on to the chaperonin TRiC in a reaction mediated by PFD. Adapted from [103]

Additionally the non-binding chaperones also assist in co- or post-translation folding, or facilitate chain transfer to downstream chaperones, the chaperonins. This second class of chaperones, the chaperonins, usually addressed as folding machines, sequesters the proteins from the cytosol to their cage and provides a defined physical compartment suitable for the protein or a protein domain to properly fold. Chaperonins occur in two subgroups: the group that is generally present in eubacteria and in organelles of endosymbiotic origin (mitochondria-chloroplasts), known as Hsp60 or complexe GroEl with co-factor GroES; and the group that is found in the archaeal and the eukaryotic cytosol like the thermosome

and the TCP-1 ring complex (TriC) respectively. The assistance from the chaperonins is achieved via an active ATP-dependent mechanism that undergoes large scale ATP-driven conformational changes crucial for their protein folding function in contrast to the ribosome binding chaperones, usually addressed as holders and folding catalysts that act via a passive, ATP-independent mechanism. Approximately ten percent of all newly synthesized polypeptides transit to chaperonins in the cell [107, 108] where only a very few were identified with an absolute chaperonin dependence for correct folding [109]. The majority of these proteins were identified with complex topologies that often fold slowly and are strongly aggregation prone, owing to the exposure of extensive hydrophobic surfaces in their non-native states. These evidences seem to substantiate the role of chaperonins in preventing aggregation, but in addition it was also evidenced that chaperonins can also speed up the folding reaction substantially [110]. This acceleration or increased efficiency in protein folding has been widely explained by the affect of confinement of proteins in the cage of the chaperonins. This effect is proposed to smooth the energy landscape of folding for larger proteins, either by preventing the formation of certain kinetically trapped intermediates or by facilitating their progression toward the compact native state. In addition, an alternative model has also been proposed, where the chaperonins are proposed to speed up folding by a mechanism of “iterative annealing” [111]. In this mechanism the chaperonin is suggested to facility folding by cycles of unfolding kinetically trapped states, fallowed by repartitioning of the unfolded protein between productive and non-productive folding pathways.

To ensure an efficient use of the cytosolic folding machinery in thousands of different proteins in the cell, protein synthesis must be coordinated with the activities of the various chaperone systems in stabilizing nascent chains and promoting proper folding. In this respect, the mechanistic principles underlying this functional cooperation are not yet well understood and careful must be taken when performing in vitro studies to approach in vivo activity [112].

ii) Chemical chaperones, are a group of small molecular weight compounds that share the ability to promote proper folding or rescue folding defects in proteins by increasing the stability of the native state or assist refolding of polypeptides. This chemically induced stabilisation can result from the indirect activity of osmolytes

over proteins, or by ligands, inhibitors and cofactors that afford a more specific action over a particular target protein. Osmolytes includes polyols, sugars, methylamines, free aminoacids or its derivatives. These molecules can promote protein folding and increase thermodynamic stability from the resulting hydration effect on proteins. This effect relies on the concept that the interaction of the polypeptide with osmolytes is mostly unfavourable, which indirectly promotes preferential protein hydration by water molecules. As a result the free energy of both the native and the unfolded state increase. Nonetheless the raised energy is much more significant in the unfolded state on the account of the increased protein surface area exposed and therefore favours folding-promoting contacts and the native state configuration [113]. Osmolytes can therefore provide some protection to proteins against denaturation, and in fact some of the osmolytes are significantly present or untaken by certain organisms in order to adapt to harsh environments or stressful conditions [114, 115]. Moreover it was also describe the ability of these molecules to even compensate molecular chaperoning deficiency [116] or regulate the folding activities of molecular chaperones under stressful conditions [117]. On the other hand, chemical chaperoning by ligands, inhibitors and cofactors that bind reversibly to a specific protein are also suggested to be able to restore protein function and folding. This heterogeneous group of molecules bind weakly to a specific target protein (due to the specificity of action are often classified as pharmacological chaperones) that either can induce protein refolding, stabilisation, or contribute to structuring a particular region or domain within the protein. This effect is attained when the binding of the chemical agent to the protein has an energetic contribution favouring the native state. In this situation the absence of the binding factor would imply the loss of stabilizing interactions which increases the number of possible conformations accessible to that conformation and therefore result in protein destabilization and misfolding.

iii) Macromolecular crowding is the term attributed to the influence that excluded volume interactions that are present in all solutions of high macromolecular content, play in the assembly of proteins. Theory predicts that crowded solutions (like the intracellular environment) can influence the folding/unfolding equilibrium and kinetic rates by preferentially destabilizing the less compact states, and providing a non-specific force for macromolecular compactness and association.

This crowding effect suggests that a sufficient increase in the concentration of inert globular macrosolutes can decrease the mean size of unfolded polypeptides, and ultimately shift the equilibrium between native state and all unfolded polypeptides substantially towards the native state [118]. In this respect, it has been reported that molecular crowding can contribute considerably to the stability of proteins native state [119, 120], in addition to promote proper folding by enhancing the collapse of polypeptide chains into functional proteins and favouring the interaction of non-native conformations with molecular chaperones [121]. In contrast, molecular crowding is also predicted to enhance the rate and equilibrium constants for the formation of aggregates of non-native protein chains [122]. This phenomenon is often associated with the increased concentration and therefore enlarged interface within non-native conformations, that can lead to aberrant interactions like misfolding and aggregation [123]. In this context it was even speculated that the increase incidence of amyloid disease in the elder may be a direct consequence of increased total intracellular protein concentration (caused by the evidence of a significant decrease in water content in the brain of cells with advanced aged [124]) in the cells of aging tissues [105]. Molecular crowding undoubtedly affects protein stability and plays a controversial role in protein folding by influencing in distinct directions the free energy landscape. Crowding contributes positively to a correct and downstream protein folding by enhancing the equilibrium stability of the native state relative to less compact non-native conformations of a polypeptide chain. On the other hand, crowding can also contribute to slow the rate of conversion of compact intermediate states to the native state, leading to a transient increase in the abundance of these non-native conformations and contribute for the formation of off-pathway species like aggregates. Overall consequences of crowding in protein folding and assembly seem to depend on the sharp balance between these competitive effects. From an evolutionary point of view, the crowding environment in the cell may as well highly contribute to dictate the requirement for molecular chaperones to avoid the aggregation effect and maintain proteins in a folding competent conformation.

1.3 PROTEINS CONFORMATIONAL STATES DYNAMICS AND STABILITY

Proteins are known to undergo conformational transitions, and to be able to withstand some structural variability, which frequently depends on the protein environment. This amazing property of proteins is proven vital to regulate biological activity and targeting of proteins to different cellular locations *in vivo*. In fact, biological systems have probably become more robust with the ability to control and regulate the various states accessible to a given polypeptide chain, at given times and under given conditions.

Accordingly with their structural features, proteins can generally be classified within four major conformational states [125]: *i.* the *native state*, in which the side chains are tightly interdigitated, leading to a densely packed overall three-dimensional structure; *ii*) the *molten globule state*, corresponding to a structured, compact intermediate state with relatively high content of native-like secondary structure, native-like topology and some tertiary contacts, but lacking the tight packing and rigid tertiary structure found in the native state; *iii*) *pre-molten globule states*, that relies on ensembles of relatively unstructured intermediates that have regions of at least transient native-like secondary structure and are condensed relative to the unfolded state; *iv*) the *unfolded state*, which typically encompasses a very large ensemble of disordered conformations.

Although, the most stable and biologically active form of the protein under relevant conditions is generally considered the native state; It follows that protein function in some cases can arise from any of the four distinct conformational states or with transitions between them [126, 127]. In agreement, a growing number of intrinsically disorder proteins with biological function has been identified, experimentally characterised and classified in a database [128]. Presently, these proteins are often considered common in nature [125, 129, 130] and typically involved in regulation, signalling and control pathways [131, 132] where conformational changes associated with function may also be brought about by alterations in environmental or cellular conditions.

From a genomic perspective, this may not sound surprising given that a large portion of gene sequences appear to code not for folded proteins but for long stretches of amino-acids that are likely to be either unfolded in solution or adopt a

disorder conformational structure. In fact, “the high proportion of gene sequences in the genomes of all organisms argues for important, as yet unknown functions, since there could be no other reason for their persistence throughout evolution” [133]. Within this scenario, it appears that proteomics studies will have to be re-evaluating since whether the lack of specific three-dimensional structure occurs wholly or in part, such proteins do not fit the standard paradigm that an organised structure is a prerequisite to function [134]. Instead, the function of a protein seems to depend on its ability to adopt a specific conformation, structured or not.
On the other hand, proteins can shift through these various conformational states as a result of changes in pH [135, 136], temperature or pressure [137], solvent alterations [138, 139] as well as ligand binding [140] and protein-protein interactions [141].Despite the fact that environmental conditions strongly influence the structural features of proteins, almost all studies on protein folding do not mimic the cellular environment. In fact, as previously mentioned, the inside of cells is extremely crowed achieving concentrations of 300-400gL^{-1} [142], whereas most studies on proteins are performed in dilute solutions < 1 gL^{-1}. Given that macromolecular crowding has been suggested to influence the structure and compactness of proteins [143, 144] , it may be quite possible that some considered disordered proteins, are just artefacts resulting from the way those proteins are studied. Eventually, some putatively disordered proteins will be structured inside the cell, while others will not.

The native state

The forces accounting for the structural features and stability of a native protein are in fact the same that determine protein folding. These central forces arise mainly from the intrinsic propensities of the polypeptide chain to form hydrogen-bonding, van der Waals and hydrophobic interactions, as well as electrostatic charge repulsion and ion pairing interactions. All these interactions seem to act in concert to define the most stable conformation or native state, but nevertheless native proteins are known to be only 5-10 Kcal/mol more stable than the unfolded state [145]. This general feature in proteins strongly suggests that no type of intermolecular force can probably be neglected when it comes to define the

overall stability of the protein, even if it appears that some of these forces must play a more dominant role in driving protein folding then others. In fact, in order to rationalize that different sequences give rise to distinct native structures, the leading key to protein folding has to be in the side chains characteristics and not in the backbone hydrogen bonding, because it is through the amino acid side chains that one protein differs from another. In agreement, there is considerable evidence that hydrophobic interactions must play a major role during protein folding. For example, the natively disorder proteins mentioned in the previous section, which fail to form persistent, well organized three-dimensional structures, are significantly depleted of hydrophobic residues [146], especially the aromatic residues that are usually found in the interior of well structured proteins. In fact,, proteins sequences that are altered and retain only their correct hydrophobic and polar patterning, usually fold to their expected native states [147], even in the absence of efforts to re-design packing, charges, or hydrogen bonding. In addition, several studies have suggested [148, 149] that secondary structures are substantially stabilised by the chain compactness, which is an indirect consequence of the hydrophobic force to collapse. Although hydrogen bonds are key components of all secondary structures, among backbone amide and carbonyl groups, the leading driving force to achieve protein folding and stability are hydrophobic interactions. Furthermore, it is possibly not a coincidence that hydrophobic and polar patterning also appears to be a key to encoding of amyloid-like fibril structures [150]. Electrostatic interactions among charged side chains also play a role in the stability of the native state. Nevertheless, most structured proteins have relatively few charged residues, and usually concentrated in high-dielectric regions on the protein surface. Besides, protein stability tend to be independent of pH (near neutral) and salt concentration, and mutations in these charged residues typically lead to only small effects on structure and stability. Hence, native folding is not likely to be dominated by electrostatic interactions.

The molten globule state

Molten globules are conformations with physical properties between those of the native and the unfolded states, and it is frequently assumed that most proteins can form such species if the appropriate experimental conditions can be found. In fact, changes in the protein environment can reduce, abolish or alter part of the native conformational interactions of the protein. These conformational alterations that concurrently can define the conversion to a molten globule, can be usually obtained under mildly denaturing conditions, such as pH extremes [151, 152], low concentrations of denaturants (chemicals, salts, alcohols, sugars) [153], thermal perturbation [154], or by the removal of prosthetic groups (like metal ions) [155, 156].

The definition of molten globule state covers an ensemble of sub-conformations ranging from those that are quite compact and afford substantial native-like contacts, to those that are less-structured and retain relatively few native contacts. Therefore it does not look surprising that experimental evidence regarding thermodynamic analyses and detailed structural characterization on molten globules forms can greatly diverge. In fact, for some molten globules it was possible to detect a significant calorimetric transition relatively to the unfolded state, while for others no difference in the heat capacity and enthalpy change could be observed, as well only for the most structured cases it was possible to detect a cooperative transition for the unfolding of a molten globule. These thermodynamic differences are inherent to the diverse residual structure present in molten globules forms, particularly on the degree of collapse of the hydrophobic core. In agreement, theoretical calculations evidenced that the differences in the heat capacity between two conformational states (like the molten globule to the unfolded sate) is proportional to the change in solvent-accessible apolar area with the transition [157]. This suggests that only molten globule forms that retain significant hydrophobic core will be able to show a calorimetric transition.

Regardless of the heterogeneity within the residual structure of molten globule forms, some common hallmarks have been established to identify these non-native states [157]. Thence, molten globules must present little or non detectable native-like tertiary structure but feature substantial native-like secondary structure (although it was demonstrated by H/D exchange NMR studies that the secondary

structure present in molten globules is significantly less stable than the corresponding regions in the native state). Besides, molten globule states ought to present a near-native level of compactness typically expanded no more than twenty percent. In addition, these partially folded conformations must reveal an increased solvent-exposed hydrophobic surface area relative to the native state, where hydrophobic residues which are buried in the native state have become exposed to solvent. This last characteristic of molten globules most likely affords an explanation for the fact that the transition of molten globule forms to the unfolded state usually represents a minor enthalpic transition when compared with the unfolding of the native state to the molten globule state.

Interestingly, folding intermediates of proteins often have several of the described characteristics in common [158] and molten globules have been detected as stable equilibrium intermediates as well as transient kinetic intermediates during the refolding process of proteins [159-161]. Indeed, a close structural similarity has been established between the equilibrium molten globules and the burst-phase kinetic intermediates [67, 162, 163], thus suggesting that molten globules may represent a genuine *in vivo* folding intermediate, that can provide insights into the nature of transient folding intermediates. In this respect, molten globule conformations have been proposed to play distinct roles during protein folding. It has been suggested that molten globules may be a general intermediate for proteins to fold [164] and therefore a precious tool to understand the mechanism of protein folding (see figure 1.5). In agreement, several studies have demonstrated that molten globule states afford native-like fold, thus suggesting that the role of molten globule conformations in folding can be to maintain an approximate native backbone topology in late folding, while still allowing structural rearrangements to occur [165]. On the contrary, it has also been suggested that molten globules states can be in fact a distinct state, an off-pathway folding conformation that leads to a kinetic trap [78, 166, 167]. Based on the gathering of distinct results for different proteins, and the strong dependency on the surrounding conditions to obtain a molten globule like conformation, the mien of molten globule states in protein folding appears to be an upshot of the folding route that was adopted by the particular sequence in those specific conditions. In fact, distinct molten globule conformations have been evidenced in distinct

unfolding conditions, within the same protein [168]. This points out for the plasticity of molten globule forms even within the same protein and suggests that ultimately the folding/unfolding route can lead to an off or on-pathway molten globule conformation. These species can therefore represent a folding intermediate for some proteins under particular conditions, but are not necessarily a mandatory intermediate for all proteins to fold. Nevertheless it is important to always keep in mind that most of the studies in protein folding do not mimic the cellular environment. The biological stage includes a multitude of chaperones that can lead some proteins to different folding routes; besides the heavy molecular crowding that is present in the cell is known to also be able to influence proteins conformational states by stabilizing the native state, or in contrast, induce the molten globule formation [169].

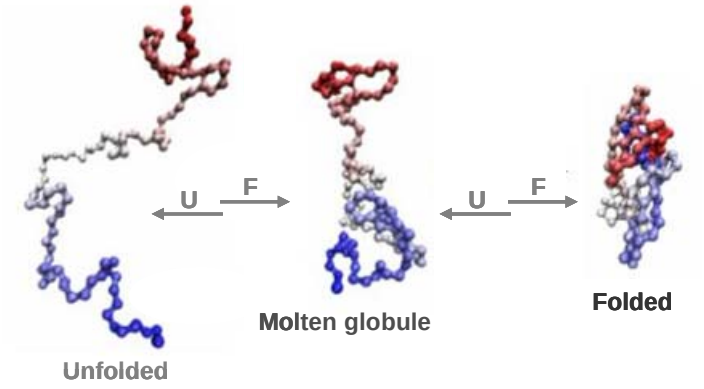


Figure 1.5 | Scheme on protein folding via a molten globule state as a general strategy for protein assembly.

U – unfolding; F - folding

In addition to a possible role of molten globule states in protein folding, this partially folded conformation has also been suggested to participate in important physiological processes like protein membrane translocation [170] and regulation of protein activity [171-173]. It was also pointed out that these partly unfolded conformations usually present a strong propensity to aggregate, due to their exposed hydrophobic surfaces [174] and have been suggested to act as precursors to amyloid fibril formation [175, 176] and on the onset of protein folding

diseases [28, 177]. The experimental and computational study of molten globules continues to be a stimulating area of research within the field of protein science. Elucidating the structural features of these partially folded conformers can provide important insights into the determinants of protein topology, the mechanism of protein folding, and biological function.

Bulk of unfolded states: outcome on the refolding pathway

The unfolded state of a protein is not actually a single type of structure but corresponds to an immense ensemble of distinct unfolded conformations. This heterogeneous and flexible population yields residual dynamic structures that are very sensitive to the unfolding environment conditions [178]. In fact, the distribution of the population of conformations can substantially depend on the type and concentration of the denaturant [179]. Moreover, a thermally denaturated protein is usually recognised to be able to retain significant residual structure, whereas chemical unfolding is generally related to a higher degree of structural elimination [180]. In this respect, this heterogeneity may sound unsurprising given that it has been suggested that a protein can unfold in multiple ways starting at different points in the structure [137, 181] and therefore it is most likely that may end up reaching distinct unfolded conformations. Even within the same unfolding conditions the multitude of unfolded configurations can be kinetically different and give rise to subpopulations that interconvert at rates slower than those of most folding reactions, and therefore each population can fold independently of the other [182]. For some proteins, this multi-kinetics pathways are usually ascribed to slow isomerization events in the denatured protein, such as *cis*/*trans* isomerization of peptidyl prolyl bonds [87]. But in other cases, it has been suggested that this may not be the sole reason for the observation of heterogeneous refolding. In fact, if most of the different subpopulations of the unfolded state are kinetically efficient to fold, then multiple parallel folding pathways may arise [183, 184]. In these cases, it would be expected to notice fast and slow refolding proceeding in a concurrent manner, where the presence of non-native residual structures could probably lead up to hold up folding, and subpopulations with native residual structures could possibly

contribute to a faster folding.

The characteristics of the residual structure present in the unfolded state of a given protein are most likely not the product of a random event that only depends on the unfolding conditions. In fact, it has been reported that the unfolded form has a direct bias in respect to the conformation of the native state [185]. This suggests a link between these two states on the unfolding route, as well as in the possible refolding pathway. Altogether, it seems that the protein dynamics and residual structural characteristics of the unfolded state can play a critical role in determining the refolding pathway of the protein. Therefore, this must be seriously taken in account when approaching a protein folding pathway through unfolding studies.

Protein unfolding occurs when the stabilizing non-covalent interactions or conformational forces that hold up the unique three-dimensional structure of the native state are broken and frequently the protein is no longer able to perform its known biologic function. Nevertheless is important to keep in mind that the loss of biological function can not be necessarily associated to the complete unfolding of the protein, but can result in the appearance of altered or new conformations with intermediate properties between those of the native and the completely unfolded states that impair function.

Hyperthermostability in proteins

Extremophilic organisms have developed some specific characteristics and components within the intracellular environment and extra-cellular membrane to cope with diverse harsh conditions like extremes of pH and salinity [186]. For temperature adaptations, however, the stress induced by elevated temperatures can not be entirely avoided by compensatory mechanisms within the cell environment. Proteins from thermophiles (growth temperatures above 45°C) and hyperthermophiles (growth temperatures above 80°C) are obliged to withstand elevated temperatures to function. In fact purified proteins from such organisms often evidence an extreme thermal stability but also chemical stability *in vitro* [187]. Undoubtedly, these hyperthermostable proteins, which are coded by the same universal twenty aminoacids, have evolved strategies to achieve an

enhanced structural stability.

By comparing the structural characteristics of proteins from hyperthermophiles and mesophiles, it became clear that protein thermostability does not arise from any peculiar or unique structural features. Moreover, it was not possible to identify any single dominating factor present among hyperthermostable proteins that would account for their thermostability [188]. In fact, it appears that hyperthermostability can arise with several distinct and combined strategies, in different hyperstable proteins. The stabilizing mechanisms can be associated with a high packing density of the molecule and low flexibility, found mainly by a prominent and solvent shielded hydrophobic core and tight surface loops. As well, it was also evidenced for a number of thermophilic proteins that an abundant presence of proline residues in the sequence, can contribute significantly to reduce the flexibility of the polypeptide chain [189]. This strategy appears to rely on the decrease of the entropy of the unfolded state, therefore stabilizing the protein. Moreover, an elevated number of ion pairs and hydrogen bonds, in particular buried polar contacts, are proposed to afford in several cases a more substantial contribution to thermostability, in addition to an improved secondary structure stabilization such as N-terminal and C-terminal helix capping [190, 191]. The availability of genomic information has allowed a comparison of the properties of the proteomes of (hyper)thermophiles and mesophile, on a global basis [192, 193]. It has been reported that the proteomes of a large number of hyperthermophiles, are biased in favour of charged rather than polar residues, a prominent difference between the proportions of charged versus polar amino acids. This observation is consistent with the thermodynamic advantage resulting from the increased significance of coulomb interactions with the increasing temperature (as the dielectric constant of water decreases). The simultaneous increase of oppositely charged residues further allows for more ion pairs to be formed at the surface of hyperthermostable proteins. So, it appears that increased ion-pair formation is one evolutionary strategy which has been adopted to increase the thermostability of soluble proteins.

In relation to thermodynamic parameters, it appears that there is no general unique formula able to explain the extreme additional stability of all hyperthermostable proteins relatively to their mesophile counterparts. In fact, it

appears that hyperthermostability in different thermostable proteins can be achieved with distinct or combined thermodynamic strategies [194]. Some proteins are thermodynamically more stable throughout the temperature range by presenting a higher ΔG_U at all temperatures as the principal way to increase their stability [195]. Others achieve higher thermal stability by presenting its maximum free-energy only at the higher temperatures. In this strategy, the maximum value for ΔG_U is rather similar to that of the mesophilic proteins but occurs only at higher temperatures, thus implying that although at higher temperatures the thermostable protein is more stable, at lower temperatures the protein from the mesophile is the most stable one. Moreover it is also possible to obtain additional thermostability in proteins having equal maximal ΔG_U values at similar temperatures as the mesophiles, by just presenting a more shallow dependence of ΔG_U on temperature [196]. Nevertheless when it comes to kinetic data of protein unfolding it appears that hyperthermostability is always associated to a very slow kinetics of protein unfolding [197]. This may suggest that this common property to all hyperstable proteins may have played a relevant role in the evolution of hyperthermophiles by selection of proteins mutants that unfold slowly. It is however important to always keep in mind that most of these studies were carried out *in vitro* and that in the intracellular environment, proteins can be further or differently stabilised by molecular crowding [120, 198] or even specific stabilizing organic solutes, the so called compatible solutes which are very common in organism thriving at high temperatures [199]. Therefore these determinations may not mimic completely the truly behaviour of the protein *in vivo*. Besides, structural information regarding hyperthermostable proteins is mainly obtain through crystallography studies and fixed crystals of structures do not allow to take in consideration the potential role of dynamics on protein stability. In fact it has been suggested that the dynamic nature of hyperthermostable proteins may be the key to unravelling the mechanism responsible for the delicate balance between rigidity, which is associated to heat resistance and molecular fluctuations at high temperatures that can account for biological function [200]. Protein thermostabilty arises from a concerted action of structural, dynamic and other physicochemical attributes that ensures the delicate balance between stability and functionality at elevated temperatures.

1.4 PROTEINS ALLIANCE WITH METAL IONS

Metal ions are an integral and indispensable part of more than thirty percent of all known proteins [201]. This is most likely due to a unique combination of physicochemical properties, namely: charge, rigid/adaptable coordination sphere, varying valence state, electron spin configuration combined with a small volume, simple structure, high mobility and some specific ligand affinity. Thus, nature has employed these building blocks, in addition to the twenty natural amino-acid residues, to finely tune, enhance, and/or diversify the properties of proteins. Metal cofactors are engaged in performing a number of tasks spanning from protein structure stabilization to enzyme catalysis, signal transduction, nitrogen fixation, photosynthesis, and respiration. In this respect, metal ions such as Zn^{2+}, Mg^{2+} and Ca^{2+} are in many cases usual stabilizers of the structure of folded proteins, while in other cases they help to maintain a particular physiologically active conformation of the protein. Transition metals, such as Fe, Cu, and Mn, are usually involved in many redox processes requiring electron transfer, while alkali metal ions, especially Na^{+}, and K^{+}, generally play a vital role in triggering cellular responses.

In most metalloproteins, one metal ion bound to a given (mononuclear) binding site is capable of fulfilling the “catalytic” and/or “structural” roles. In some cases, however, two or even more metal ions, which are embedded in binuclear or polynuclear binding sites, respectively, are needed to carry out the catalytic and/or structural function, and act synchronously during the catalytic process. Binuclear metal centres can contain metal ions of the same chemical type (e.g., Zn-Zn, like in some β-lactamases) or contain chemically different metal ions (e.g., Cu-Zn, like in superoxide dismutase-1). In many bimetallic proteins the two metal centres can be functionally and structurally unequivalent, having different ligand surroundings, solvent exposure, and binding affinity.

Regardless of the metal and its precise pattern of ligation to the protein, there seems to be a frequent qualitative feature to the binding site: the metal is ligated by a shell of hydrophilic groups (containing oxygen, nitrogen, or sulfur atoms) and this hydrophilic shell is embedded within a larger shell with predominating hydrophobic groups. That is, metals tend to bind at sites with a high contrast in

hydrophobicity [202]. Moreover, metals seem to preferentially bind directly to protein ligands. In agreement, metal binding sites are usually located inside cavities and pockets in the protein structure where usually solvent is inaccessible, and that have a low dielectric constant that enhances the electrostatic metal-protein ligand interactions, thus favouring inner- sphere metal binding [202] . A particularly intriguing question in metalloprotein chemistry is how a protein selects a specific metal from the bulk of metal ions that are present in the cellular environment. This selectivity is known to be in part associated with the properties of the metal (e.g., its stereochemical and charge to size requirements) and to properties of the protein (e.g., its unique set of amino acid residues forming the metal-binding pocket and the stereochemistry of this pocket). Although it is not completely clear if metal-binding sites in proteins are generally rigid or flexible, and the extent to which the protein can adjust to the stereochemical requirements of the incoming metal ion, or, conversely, the extent to which the metal ion can comply with the constraints of the protein matrix. Interestingly, some proteins such as those with the EF-hand motif binds only specifically to calcium ions [203], while others such as CheY [204] and rubredoxin [205, 206] can bind several ions with similar affinities. Apparently this discrepancy in the metal affinity of proteins is highly related to protein function and priorities. In fact, proteins involved in signal transduction events, tend to have a more specific binding site towards a given metal ion, thus enabling the protein to selectively bind its natural cofactor against a background of other metals. In this case, the type of metal ligands, their coordination mode and side chain interactions, as well as the overall charge and shape of the cavity govern the binding-site specificity. On the other hand, in the case of proteins with low affinity to a specific metal ion, the cell machinery also plays a determinant role in the process of metal binding preference, by regulating the concentration of metal ions in various biological compartments and making specific chaperones to mediate metal insertion into proteins, the so called metallochaperone proteins. Despite the tremendous body of information regarding the structure and function of metal sites in metalloproteins, general physical bases governing metal binding and selectivity in proteins have been derived. On the other hand, in most cases the interplay of the metal cofactors with protein folding, and the contribution to structural stability still remains poorly understood.

Residues involved in cross-linking with metal ions

The most commonly occurring metals in proteins, are the cations Mg, Zn, Ca, Fe, Cd, Mn, Cu, Co and Ni [207]. Knowledge of the type and number of amino-acid residues that coordinate a metal ion is determinant to clarify the protein-metal recognition and affinity, as well as to obtain insight into their possible connection to the function of the protein. The amino acid residues that most often coordinate metal ions in proteins are His, Asp, Cys and Glu followed by Asn, Gly, Thr, Ser, Tyr, Met and Gln [208]. With the exception of Gly, they usually participate in coordination using the electron donor group present in the side chain. It seems logic that divalent metal cations have high affinity towards the anionic protein ligands such as deprotonated Asp or Glu side chains due to the strong charge-charge interactions, and thus high free energy gain upon metal binding in a solvent-inaccessible, low dielectric medium. There is, however, an upper limit on the number of carboxylates that can coordinate to the metal. This number is predicted to be three for the hardest metal coordination. This hard amino-acid coordination usually only occurs when the protein matrix does not require significant relaxation or reorganization upon metal binding. Other rather neutral protein ligands, such as the Asn, Gln, His, Ser, or Thr side chain and the backbone carbonyl oxygen, can also add to the stability of the metal complex, although the free energy gain is less than that upon binding a Asp/Glu side chain. Nevertheless these ligands play an important role in shaping and fine-tuning the geometry of the binding site so that the protein can selectively sequester a particular metal cofactor from the cell moiety. The affinity of individual amino acid side chains to certain metals is often characteristic, resulting in a typical selectivity pattern with preferred amino acid/metal ion combinations (Table 1.1). Although there may be a tendency for some amino-acid towards binding some metals, there is not specificity in any residue that can solely dictate the selectivity for a particular metal. In this respect, the metal coordination number also plays an important role in metal selectivity, and can be a key determinant of the structure and properties of metals in proteins. Factors like the dielectric medium or solvent accessibility, properties of the metal (mainly by its ability to accept charge from its ligands), and the chemical characteristics of the ligands are known to influence the resulting coordination number of a metal in a protein [201].

Despite the preference of some metal ions towards a specific coordination number, some metal ions can adopt different coordination geometries at a relatively low free energy cost [209], which may suggest the assignment of a distinct function accordingly with the number of electron-pair donating ligands. In the case of Zn, the metal ion is usually tetrahedrally coordinated in Zn-finger proteins, but in some catalytic sites, it is found penta-coordinated and, rarely, hexa-coordinated [210]. For example, a survey of Zn proteins in the PDB shows that for structural binding sites the ratio between tetra/penta/hexa-coordinated Zn is 79:6:12%, respectively, whereas for catalytic sites the ratio is 48:44:6%, respectively [211].

Metal ions

(grouped according to the similarity of their coordination spheres)

Preferential electron-donor ligand	Cu^{II}	Fe^{II}	Ni^{II}	Zn^{II}	Cd^{II}	Ca^{II}	Mn^{II}	Mg^{II}	Co^{II}
	His	His	His	Cys	Cys	Asp	Asp	Asp	His
	Cys	Glu,Cys	Cys,Asp	His	His,Glu	Glu	His	Glu	Asp
	Met	Met,Asp	Glu	Asp,Glu	Asp	Asn	Glu	His	Glu
	Asp,Glu	Tyr	Ser	Asn,Gln	Ser,Asn	Thr,Ser	Asn	Thr,Ser Asn	Cys
	Gln,Asn	Asn	Gln ,Asn	Ser,Met Tyr,Thr	Gln,Thr Met,Tyr	Tyr,His Gln,Met	Ser,Thr Gln,Cys	Gln,Tyr Met	Ser,Gln Thr,Tyr

Table 1.1 | Relative occurrence of amino-acid residues in the coordination of the most abundant metals in proteins.
Amino-acid residues are displayed in a descendent order of frequency to each metal. Based on [208].

New metal sites can be introduced into designed proteins to add structural specificity in the binding of metallic cofactors, or to modulate the oligomeric state of the protein. Metal sites can also be engineered into naturally occurring folds in proteins normally devoid of metal centres, by rational design [5]. New sites forming a given ligation geometry are found by searching existing structures for backbone geometries consistent with the conformational needs of the ligating groups. Recent studies suggest that functional metal sites are accessible through design strategies, moving the field closer to the goal of novel metal-enzyme design.

Metal ions in protein folding

Metal ions play an important role in protein folding and stability [212]. In this respect, cofactors can act not only as local structural stabilizing elements in the native state [213-215], contributing to the maintenance of a given specific structural fold, but also can function as potential key nucleation points during *in vivo* folding, as its binding to the unfolded polypeptide is likely to impose a conformational restriction which lowers the entropy of the unfolded state thus favouring a specific folding pathway and speeding up radically the folding reaction [216, 217].

Metal-dependent protein folding involves the coordination of one or more metal ions to drive the folding of the polypeptide chain into the fully functional conformation. One of the most typical examples are the independent folding motifs of the Zn-finger domains. These coordinate Zn^{II} ions tetrahedrally via a combination of cysteinyl thiolate and histidinyl imidazole nitrogen ligands to form an extension from the globular protein conformation. In these cases, the metal-free 'finger' is initially largely unstructured emphasizing the requirement for Zn^{II} to provide the structural denominator that triggers the folding of the peptide into the functional form of the protein [218, 219]. Other examples can be found within *c*-type heme proteins that afford an iron-porphyrin metal group covalently bound at two cys residues. In some cases the removal of this heme group causes disruption of the native fold and loss of most of the secondary structure [220], while in other heme proteins, the apo state stands a significant loss of structure preferentially around the prosthetic binding site [221]. The heme group is most likely critical for maintaining the native structure of these proteins because it usually forms an extensive net of interactions with the surrounding protein. In addition to its axial ligands, interactions between the two heme propionate groups and polar residue side chains nearby, and hydrophobic interactions between the porphyrin ring and buried nonpolar side chains are observed in the native structures. These interactions reflect common contributions of the heme group to structural maintenance in the hosted proteins.

On the other hand metal ions are also evidenced to bind to proteins in a later stage of the folding reaction, specific binding sites are already defined, like for example in ribonuclease and lysosome [222, 223]. In such cases the

conformational change that is induced by the ligand usually affords an increase in the protein stability but is not determinative to achieve the general native fold and is quite often a general strategy for controlling protein function. A rather representative case is found in the copper chaperone proteins CopZ and AtoX1 where a switch between the apo and holo forms takes place as a part of their functional cycle and biophysical properties of both states are of biological relevance. These proteins, both apo and holo forms, were shown to unfold reversibly, in addition to clearly evidence a substantial increase in stability in the metal containing form [214]. An interesting case is found within hyperthermophile rubredoxins, which are proteins typically involved in electron transfer processes and no known biological function related to it apo form. In these proteins the holo and apo forms (obtained by holo demetallation) present the same overall structure and significant distinct thermal stabilities [224], thus suggesting that metal binding probably occurs in the final stage of the protein folding. Nevertheless, both apo and holo forms of the protein are unable to refold spontaneously due to the lack or disintegration of the metal centre, respectively. An exception to this irreversibility though is observed in the engineered rubredoxin mutated for the four Cys ligands, also lacking the iron centre [225]. This variant was shown to fold reversibly with a decreased melting temperature (T_m = 82ºC) comparatively with the wild type (wt ~ 200ºC [226]), and retain identical structure of the holo rubredoxin [225]. With these findings, one would be tempted to say that iron binding ligands are the cause for the irreversibility, possibly due to aberrant interactions that in the cell are avoid with the proper orientation of metallochaperone activity. An ambiguous example relatively to the modulation of co-factors in protein folding comes from a copper protein azurin [227]. Crystal structures of apo- and holo-azurin have shown that the overall three-dimensional structure is identical with and without the metal cofactor. In addition it was also shown that the affinity of the cofactor was significantly higher for the folded than for the unfolded azurin polypeptide. Thus, altogether the data suggests a minor contribution of the metal in the folding of this protein. On the other hand it was simultaneously noted that copper bind much more rapidly (i.e., 4 orders of magnitude) to the unfolded protein than to folded apo-protein, thus evidencing that the fastest route to holo and functional azurin is through copper

binding before polypeptide folding. This may suggests that cofactor binding in different stages of the protein folding can also occur in the cell and is likely to be determined by different biological requirements.

1.5 REFERENCES

1. Baker, D., *Prediction and design of macromolecular structures and interactions.* Philos Trans R Soc Lond B Biol Sci, 2006. **361**(1467): p. 459-63.
2. Bradley, P., K.M. Misura, and D. Baker, *Toward high-resolution de novo structure prediction for small proteins.* Science, 2005. **309**(5742): p. 1868-71.
3. Dahiyat, B.I. and S.L. Mayo, *De novo protein design: fully automated sequence selection.* Science, 1997. **278**(5335): p. 82-7.
4. Looger, L.L., et al., *Computational design of receptor and sensor proteins with novel functions.* Nature, 2003. **423**(6936): p. 185-90.
5. Coldren, C.D., H.W. Hellinga, and J.P. Caradonna, *The rational design and construction of a cuboidal iron-sulfur protein.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6635-40.
6. Goodman, C.M., et al., *Foldamers as versatile frameworks for the design and evolution of function.* Nat Chem Biol, 2007. **3**(5): p. 252-62.
7. English, E.P., et al., *Rational development of beta-peptide inhibitors of human cytomegalovirus entry.* J Biol Chem, 2006. **281**(5): p. 2661-7.
8. Wu, C.W., et al., *Helical peptoid mimics of lung surfactant protein C.* Chem Biol, 2003. **10**(11): p. 1057-63.
9. Razeghifard, R., et al., *Creating functional artificial proteins.* Curr Protein Pept Sci, 2007. **8**(1): p. 3-18.
10. Scott, K.A. and V. Daggett, *Folding mechanisms of proteins with high sequence identity but different folds.* Biochemistry, 2007. **46**(6): p. 1545-56.
11. Chiti, F. and C.M. Dobson, *Protein misfolding, functional amyloid, and human disease.* Annu Rev Biochem, 2006. **75**: p. 333-66.
12. Ignatova, Z., et al., *From the test tube to the cell: exploring the folding and aggregation of a beta-clam protein.* Biopolymers, 2007. **88**(2): p. 157-63.
13. Ellis, R.J. and F.U. Hartl, *Principles of protein folding in the cellular environment.* Curr Opin Struct Biol, 1999. **9**(1): p. 102-10.
14. Schuler, B. and W.A. Eaton, *Protein folding studied by single-molecule FRET.* Curr Opin Struct Biol, 2008. **18**(1): p. 16-26.
15. Eaton, W.A., et al., *Submillisecond kinetics of protein folding.* Curr Opin Struct Biol, 1997. **7**(1): p. 10-4.
16. Villegas, V., et al., *Evidence for a two-state transition in the folding process of the activation domain of human procarboxypeptidase A2.* Biochemistry, 1995. **34**(46): p. 15105-10.
17. Jackson, S.E. and A.R. Fersht, *Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition.* Biochemistry, 1991. **30**(43): p. 10428-35.
18. Bryngelson, J.D., et al., *Funnels, pathways, and the energy landscape of protein folding: a synthesis.* Proteins, 1995. **21**(3): p. 167-95.
19. Dill, K.A., Ozkan, S. B., Shell, M. S., Weikl, T.R., *The Protein Folding Problem.* Annu. Rev. Biophys., 2008. **37**: p. 289–316.
20. Dinner, A.R., et al., *Understanding protein folding via free-energy surfaces from theory and experiment.* Trends Biochem Sci, 2000. **25**(7): p. 331-9.

21. Dobson, C.M. and M. Karplus, *The fundamentals of protein folding: bringing together theory and experiment.* Curr Opin Struct Biol, 1999. **9**(1): p. 92-101.
22. Jin, W., et al., *De novo design of foldable proteins with smooth folding funnel: automated negative design and experimental verification.* Structure, 2003. **11**(5): p. 581-90.
23. Onuchic, J.N., Z. Luthey-Schulten, and P.G. Wolynes, *Theory of protein folding: the energy landscape perspective.* Annu Rev Phys Chem, 1997. **48**: p. 545-600.
24. Onuchic, J.N. and P.G. Wolynes, *Theory of protein folding.* Curr Opin Struct Biol, 2004. **14**(1): p. 70-5.
25. Dill, K.A. and H.S. Chan, *From Levinthal to pathways to funnels.* Nat Struct Biol, 1997. **4**(1): p. 10-9.
26. Hardin, C., et al., *Folding funnels: the key to robust protein structure prediction.* J Comput Chem, 2002. **23**(1): p. 138-46.
27. Bryngelson, J.D., Wolynes, P.G., *Intermediate and barrier crossing in a random energy model (with applications to protein folding).* J. Phys. Chem., 1989. **93**: p. 6902-6915.
28. Dobson, C.M., *Protein folding and misfolding.* Nature, 2003. **426**(6968): p. 884-90.
29. Karplus, M., *The Levinthal paradox: yesterday and today.* Fold Des, 1997. **2**(4): p. S69-75.
30. Wolynes, P.G., J.N. Onuchic, and D. Thirumalai, *Navigating the folding routes.* Science, 1995. **267**(5204): p. 1619-20.
31. Wright, C.F., et al., *Parallel protein-unfolding pathways revealed and mapped.* Nat Struct Biol, 2003. **10**(8): p. 658-62.
32. Gruebele, M., *Downhill protein folding: evolution meets physics.* C R Biol, 2005. **328**(8): p. 701-12.
33. Sridevi, K., et al., *Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble.* J Mol Biol, 2004. **337**(3): p. 699-711.
34. Pradeep, L. and J.B. Udgaonkar, *Differential salt-induced stabilization of structure in the initial folding intermediate ensemble of barstar.* J Mol Biol, 2002. **324**(2): p. 331-47.
35. Rami, B.R. and J.B. Udgaonkar, *pH-jump-induced folding and unfolding studies of barstar: evidence for multiple folding and unfolding pathways.* Biochemistry, 2001. **40**(50): p. 15267-79.
36. Krantz, B.A. and T.R. Sosnick, *Engineered metal binding sites map the heterogeneous folding landscape of a coiled coil.* Nat Struct Biol, 2001. **8**(12): p. 1042-7.
37. Lowe, A.R. and L.S. Itzhaki, *Rational redesign of the folding pathway of a modular protein.* Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2679-84.
38. Garcia, C., et al., *Changes in the apomyoglobin folding pathway caused by mutation of the distal histidine residue.* Biochemistry, 2000. **39**(37): p. 11227-37.
39. Linse, S. and B. Linse, *Protein folding through kinetic discrimination.* J Am Chem Soc, 2007. **129**(27): p. 8481-6.
40. Koga, N. and S. Takada, *Roles of native topology and chain-length scaling in protein folding: a simulation study with a Go-like model.* J Mol Biol, 2001. **313**(1): p. 171-80.
41. Clementi, C., H. Nymeyer, and J.N. Onuchic, *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, 2000. **298**(5): p. 937-53.
42. Levy, Y., P.G. Wolynes, and J.N. Onuchic, *Protein topology determines binding mechanism.* Proc Natl Acad Sci U S A, 2004. **101**(2): p. 511-6.
43. Papoian, G.A. and P.G. Wolynes, *The physics and bioinformatics of binding and folding-an energy landscape perspective.* Biopolymers, 2003. **68**(3): p. 333-49.
44. Matouschek, A., et al., *Mapping the transition state and pathway of protein folding by protein engineering.* Nature, 1989. **340**(6229): p. 122-6.

45. Fersht, A.R. and V. Daggett, *Protein folding and unfolding at atomic resolution.* Cell, 2002. **108**(4): p. 573-82.
46. Myers, J.K. and T.G. Oas, *Preorganized secondary structure as an important determinant of fast protein folding.* Nat Struct Biol, 2001. **8**(6): p. 552-8.
47. Karplus, M. and D.L. Weaver, *Protein folding dynamics: the diffusion-collision model and experimental data.* Protein Sci, 1994. **3**(4): p. 650-68.
48. Fersht, A.R., *Nucleation mechanisms in protein folding.* Curr Opin Struct Biol, 1997. **7**(1): p. 3-9.
49. Krishna, M.M., et al., *Order of steps in the cytochrome C folding pathway: evidence for a sequential stabilization mechanism.* J Mol Biol, 2006. **359**(5): p. 1410-9.
50. Agashe, V.R., M.C. Shastry, and J.B. Udgaonkar, *Initial hydrophobic collapse in the folding of barstar.* Nature, 1995. **377**(6551): p. 754-7.
51. Gianni, S., et al., *Unifying features in protein-folding mechanisms.* Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13286-91.
52. Udgaonkar, J.B., *Multiple Routes and Structural Heterogenity in Protein Folding.* Annu. Rev. Biophys., 2008. **37**: p. 489-510.
53. Dalby, P.A., M. Oliveberg, and A.R. Fersht, *Movement of the intermediate and rate determining transition state of barnase on the energy landscape with changing temperature.* Biochemistry, 1998. **37**(13): p. 4674-9.
54. Matouschek, A., et al., *Movement of the position of the transition state in protein folding.* Biochemistry, 1995. **34**(41): p. 13656-62.
55. Chan, C.K., et al., *Submillisecond protein folding kinetics studied by ultrarapid mixing.* Proc Natl Acad Sci U S A, 1997. **94**(5): p. 1779-84.
56. Scalley, M.L., et al., *Kinetics of folding of the IgG binding domain of peptostreptococcal protein L.* Biochemistry, 1997. **36**(11): p. 3373-82.
57. Korzhnev, D.M., et al., *Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR.* Nature, 2004. **430**(6999): p. 586-90.
58. Jonsson, T., C.D. Waldburger, and R.T. Sauer, *Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates.* Biochemistry, 1996. **35**(15): p. 4795-802.
59. Sanchez, I.E. and T. Kiefhaber, *Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding.* J Mol Biol, 2003. **325**(2): p. 367-76.
60. Zaidi, F.N., U. Nath, and J.B. Udgaonkar, *Multiple intermediates and transition states during protein unfolding.* Nat Struct Biol, 1997. **4**(12): p. 1016-24.
61. Hamid Wani, A. and J.B. Udgaonkar, *HX-ESI-MS and optical studies of the unfolding of thioredoxin indicate stabilization of a partially unfolded, aggregation-competent intermediate at low pH.* Biochemistry, 2006. **45**(37): p. 11226-38.
62. Gorski, S.A., et al., *Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9.* J Mol Biol, 2001. **312**(4): p. 849-63.
63. Nath, U. and J.B. Udgaonkar, *Perturbation of a tertiary hydrogen bond in barstar by mutagenesis of the sole His residue to Gln leads to accumulation of at least one equilibrium folding intermediate.* Biochemistry, 1995. **34**(5): p. 1702-13.
64. Bycroft, M., et al., *Detection and characterization of a folding intermediate in barnase by NMR.* Nature, 1990. **346**(6283): p. 488-90.
65. Matouschek, A., et al., *Transient folding intermediates characterized by protein engineering.* Nature, 1990. **346**(6283): p. 440-5.
66. Parker, M.J., J. Spencer, and A.R. Clarke, *An integrated kinetic analysis of intermediates and transition states in protein folding reactions.* J Mol Biol, 1995. **253**(5): p. 771-86.

67. Raschke, T.M. and S. Marqusee, *The kinetic folding intermediate of ribonuclease H resembles the acid molten globule and partially unfolded molecules detected under native conditions.* Nat Struct Biol, 1997. **4**(4): p. 298-304.

68. Spudich, G.M., E.J. Miller, and S. Marqusee, *Destabilization of the Escherichia coli RNase H kinetic intermediate: switching between a two-state and three-state folding mechanism.* J Mol Biol, 2004. **335**(2): p. 609-18.

69. Jackson, S.E., *How do small single-domain proteins fold?* Fold Des, 1998. **3**(4): p. R81-91.

70. Kamagata, K., M. Arai, and K. Kuwajima, *Unification of the folding mechanisms of non-two-state and two-state proteins.* J Mol Biol, 2004. **339**(4): p. 951-65.

71. Zwanzig, R., *Two-state models of protein folding kinetics.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 148-50.

72. Wildegger, G. and T. Kiefhaber, *Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate.* J Mol Biol, 1997. **270**(2): p. 294-304.

73. Fersht, A.R., *Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 10869-73.

74. Ellison, P.A. and S. Cavagnero, *Role of unfolded state heterogeneity and en-route ruggedness in protein folding kinetics.* Protein Sci, 2006. **15**(3): p. 564-82.

75. Udgaonkar, J.B. and R.L. Baldwin, *Early folding intermediate of ribonuclease A.* Proc Natl Acad Sci U S A, 1990. **87**(21): p. 8197-201.

76. Capaldi, A.P., et al., *Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate.* Nat Struct Biol, 2001. **8**(1): p. 68-72.

77. Raschke, T.M., J. Kho, and S. Marqusee, *Confirmation of the hierarchical folding of RNase H: a protein engineering study.* Nat Struct Biol, 1999. **6**(9): p. 825-31.

78. Bollen, Y.J., I.E. Sanchez, and C.P. van Mierlo, *Formation of on- and off-pathway intermediates in the folding kinetics of Azotobacter vinelandii apoflavodoxin.* Biochemistry, 2004. **43**(32): p. 10475-89.

79. Wu, Y., et al., *Specific structure appears at the N terminus in the sub-millisecond folding intermediate of the alpha subunit of tryptophan synthase, a TIM barrel protein.* J Mol Biol, 2005. **351**(3): p. 445-52.

80. Chamberlain, A.K. and S. Marqusee, *Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms.* Adv Protein Chem, 2000. **53**: p. 283-328.

81. Maity, H., et al., *Protein folding: the stepwise assembly of foldon units.* Proc Natl Acad Sci U S A, 2005. **102**(13): p. 4741-6.

82. Pande, V.S., et al., *Pathways for protein folding: is a new view needed?* Curr Opin Struct Biol, 1998. **8**(1): p. 68-79.

83. Wallace, L.A. and C.R. Matthews, *Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions.* Biophys Chem, 2002. **101-102**: p. 113-31.

84. Baldwin, R.L., *The nature of protein folding pathways: the classical versus the new view.* J Biomol NMR, 1995. **5**(2): p. 103-9.

85. Plotkin, S.S. and J.N. Onuchic, *Understanding protein folding with energy landscape theory. Part II: Quantitative aspects.* Q Rev Biophys, 2002. **35**(3): p. 205-86.

86. Plotkin, S.S. and J.N. Onuchic, *Understanding protein folding with energy landscape theory. Part I: Basic concepts.* Q Rev Biophys, 2002. **35**(2): p. 111-67.

87. Georgescu, R.E., et al., *Proline isomerization-independent accumulation of an early intermediate and heterogeneity of the folding pathways of a mixed alpha/beta protein, Escherichia coli thioredoxin.* Biochemistry, 1998. **37**(28): p. 10286-97.

88. Goldbeck, R.A., et al., *Multiple pathways on a protein-folding energy landscape: kinetic evidence.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2782-7.

89. Pletneva, E.V., H.B. Gray, and J.R. Winkler, *Snapshots of cytochrome c folding.* Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18397-402.

90. Wintrode, P.L., et al., *An obligatory intermediate controls the folding of the alpha-subunit of tryptophan synthase, a TIM barrel protein.* J Mol Biol, 2005. **347**(5): p. 911-9.

91. Chavez, L.L., et al., *Multiple routes lead to the native state in the energy landscape of the beta-trefoil family.* Proc Natl Acad Sci U S A, 2006. **103**(27): p. 10254-8.

92. Krishna, M.M. and S.W. Englander, *A unified mechanism for protein folding: predetermined pathways with optional errors.* Protein Sci, 2007. **16**(3): p. 449-64.

93. Yewdell, J.W., *Serendipity strikes twice: the discovery and rediscovery of defective ribosomal products (DRiPS).* Cell Mol Biol (Noisy-le-grand), 2005. **51**(7): p. 635-41.

94. Sunde, M. and C. Blake, *The structure of amyloid fibrils by electron microscopy and X-ray diffraction.* Adv Protein Chem, 1997. **50**: p. 123-59.

95. Guijarro, J.I., et al., *Amyloid fibril formation by an SH3 domain.* Proc Natl Acad Sci U S A, 1998. **95**(8): p. 4224-8.

96. Litvinovich, S.V., et al., *Formation of amyloid-like fibrils by self-association of a partially unfolded fibronectin type III module.* J Mol Biol, 1998. **280**(2): p. 245-58.

97. Kelly, J.W., *The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways.* Curr Opin Struct Biol, 1998. **8**(1): p. 101-6.

98. Chiti, F., et al., *Designing conditions for in vitro formation of amyloid protofilaments and fibrils.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3590-4.

99. Tartaglia, G.G., et al., *Life on the edge: a link between gene expression levels and aggregation rates of human proteins.* Trends Biochem Sci, 2007. **32**(5): p. 204-6.

100. Pawar, A.P., et al., *Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases.* J Mol Biol, 2005. **350**(2): p. 379-92.

101. Tartaglia, G.G., et al., *Prediction of aggregation-prone regions in structured proteins.* J Mol Biol, 2008. **380**(2): p. 425-36.

102. Gething, M.J. and J. Sambrook, *Protein folding in the cell.* Nature, 1992. **355**(6355): p. 33-45.

103. Hartl, F.U. and M. Hayer-Hartl, *Molecular chaperones in the cytosol: from nascent chain to folded protein.* Science, 2002. **295**(5561): p. 1852-8.

104. Leandro, P. and C.M. Gomes, *Protein misfolding in conformational disorders: rescue of folding defects and chemical chaperoning.* Mini-Reviews in Medicinal Chemistry, 2008. **in press**.

105. Minton, A.P., *Implications of macromolecular crowding for protein assembly.* Curr Opin Struct Biol, 2000. **10**(1): p. 34-9.

106. Ben-Zvi, A.P. and P. Goloubinoff, *Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones.* J Struct Biol, 2001. **135**(2): p. 84-93.

107. Houry, W.A., et al., *Identification of in vivo substrates of the chaperonin GroEL.* Nature, 1999. **402**(6758): p. 147-54.

108. Thulasiraman, V., C.F. Yang, and J. Frydman, *In vivo newly translated polypeptides are sequestered in a protected folding environment.* Embo J, 1999. **18**(1): p. 85-95.

109. Fayet, O., T. Ziegelhoffer, and C. Georgopoulos, *The groES and groEL heat shock gene products of Escherichia coli are essential for bacterial growth at all temperatures.* J Bacteriol, 1989. **171**(3): p. 1379-85.

110. Brinker, A., et al., *Dual function of protein confinement in chaperonin-assisted protein folding.* Cell, 2001. **107**(2): p. 223-33.

111. Shtilerman, M., G.H. Lorimer, and S.W. Englander, *Chaperonin function: folding by forced unfolding.* Science, 1999. **284**(5415): p. 822-5.

112. Giese, K.C., et al., *Evidence for an essential function of the N terminus of a small heat shock protein in vivo, independent of in vitro chaperone activity.* Proc Natl Acad Sci U S A, 2005. **102**(52): p. 18896-901.

113. Bolen, D.W. and I.V. Baskakov, *The osmophobic effect: natural selection of a thermodynamic force in protein folding.* J Mol Biol, 2001. **310**(5): p. 955-63.

114. Santos, H. and M.S. da Costa, *Compatible solutes of organisms that live in hot saline environments.* Environ Microbiol, 2002. **4**(9): p. 501-9.

115. Yancey, P.H., et al., *Living with water stress: evolution of osmolyte systems.* Science, 1982. **217**(4566): p. 1214-22.

116. Chattopadhyay, M.K., et al., *The chemical chaperone proline relieves the thermosensitivity of a dnaK deletion mutant at 42 degrees C.* J Bacteriol, 2004. **186**(23): p. 8149-52.

117. Diamant, S., et al., *Chemical chaperones regulate molecular chaperones in vitro and in cells under combined salt and heat stresses.* J Biol Chem, 2001. **276**(43): p. 39586-91.

118. Minton, A.P., *Effect of a concentrated "inert" macromolecular cosolute on the stability of a globular protein with respect to denaturation by heat and by chaotropes: a statistical-thermodynamic model.* Biophys J, 2000. **78**(1): p. 101-9.

119. Charlton, L.M., et al., *Residue-level interrogation of macromolecular crowding effects on protein stability.* J Am Chem Soc, 2008. **130**(21): p. 6826-30.

120. Stagg, L., et al., *Molecular crowding enhances native structure and stability of alpha/beta protein flavodoxin.* Proc Natl Acad Sci U S A, 2007. **104**(48): p. 18976-81.

121. Martin, J. and F.U. Hartl, *The effect of macromolecular crowding on chaperonin-mediated protein folding.* Proc Natl Acad Sci U S A, 1997. **94**(4): p. 1107-12.

122. Ellis, R.J. and A.P. Minton, *Protein aggregation in crowded environments.* Biol Chem, 2006. **387**(5): p. 485-97.

123. van den Berg, B., R.J. Ellis, and C.M. Dobson, *Effects of macromolecular crowding on protein folding and aggregation.* Embo J, 1999. **18**(24): p. 6927-33.

124. Nagy, I.Z., K. Nagy, and G. Lustyik, *Protein and water contents of aging brain.* Exp Brain Res, 1982. **Suppl 5**: p. 118-22.

125. Uversky, V.N., *Natively unfolded proteins: a point where biology waits for physics.* Protein Sci, 2002. **11**(4): p. 739-56.

126. Dunker, A.K., et al., *Intrinsic disorder and protein function.* Biochemistry, 2002. **41**(21): p. 6573-82.

127. Dunker, A.K., et al., *Intrinsically disordered protein.* J Mol Graph Model, 2001. **19**(1): p. 26-59.

128. Sickmeier, M., et al., *DisProt: the Database of Disordered Proteins.* Nucleic Acids Res, 2007. **35**(Database issue): p. D786-93.

129. Ilbert, M., et al., *The redox-switch domain of Hsp33 functions as dual stress sensor.* Nat Struct Mol Biol, 2007. **14**(6): p. 556-63.

130. Sohl, J.L., S.S. Jaswal, and D.A. Agard, *Unfolded conformations of alpha-lytic protease are more stable than its native state.* Nature, 1998. **395**(6704): p. 817-9.

131. Dunker, A.K., et al., *Flexible nets. The roles of intrinsic disorder in protein interaction networks.* Febs J, 2005. **272**(20): p. 5129-48.

132. Iakoucheva, L.M., et al., *Intrinsic disorder in cell-signaling and cancer-associated proteins.* J Mol Biol, 2002. **323**(3): p. 573-84.

133. Wright, P.E. and H.J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm.* J Mol Biol, 1999. **293**(2): p. 321-31.

134. Radivojac, P., et al., *Intrinsic disorder and functional proteomics.* Biophys J, 2007. **92**(5): p. 1439-56.

135. Gatewood, J.M., et al., *Zinc-induced secondary structure transitions in human sperm protamines.* J Biol Chem, 1990. **265**(33): p. 20667-72.

136. Pesenti, M.E., et al., *Structural basis of the honey bee PBP pheromone and pH-induced conformational change.* J Mol Biol, 2008. **380**(1): p. 158-69.

137. Li, L., et al., *Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type III module.* J Mol Biol, 2005. **345**(4): p. 817-26.

138. Johansson, J., et al., *Conformation-dependent antibacterial activity of the naturally occurring human peptide LL-37.* J Biol Chem, 1998. **273**(6): p. 3718-24.

139. Henkels, C.H., et al., *Linked folding and anion binding of the Bacillus subtilis ribonuclease P protein.* Biochemistry, 2001. **40**(9): p. 2777-89.

140. Strunk, J.J., et al., *Ligand binding induces a conformational change in ifnar1 that is propagated to its membrane-proximal domain.* J Mol Biol, 2008. **377**(3): p. 725-39.

141. Zhou, P., et al., *Solution structure of DFF40 and DFF45 N-terminal domain complex and mutual chaperone activity of DFF40 and DFF45.* Proc Natl Acad Sci U S A, 2001. **98**(11): p. 6051-5.

142. Ellis, R.J. and A.P. Minton, *Cell biology: join the crowd.* Nature, 2003. **425**(6953): p. 27-8.

143. Perham, M., L. Stagg, and P. Wittung-Stafshede, *Macromolecular crowding increases structural content of folded proteins.* FEBS Lett, 2007. **581**(26): p. 5065-9.

144. Tokuriki, N., et al., *Protein folding by the effects of macromolecular crowding.* Protein Sci, 2004. **13**(1): p. 125-33.

145. Waters, M.L., *Aromatic interactions in peptides: impact on structure and function.* Biopolymers, 2004. **76**(5): p. 435-45.

146. Dunker, A.K., C.J. Brown, and Z. Obradovic, *Identification and functions of usefully disordered proteins.* Adv Protein Chem, 2002. **62**: p. 25-49.

147. Kamtekar, S., et al., *Protein design by binary patterning of polar and nonpolar amino acids.* Science, 1993. **262**(5140): p. 1680-5.

148. Banavar, J.R. and A. Maritan, *Physics of proteins.* Annu Rev Biophys Biomol Struct, 2007. **36**: p. 261-80.

149. Chikenji, G., Y. Fujitsuka, and S. Takada, *Shaping up the protein folding funnel by local interaction: lesson from a structure prediction study.* Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3141-6.

150. Wurth, C., W. Kim, and M.H. Hecht, *Combinatorial approaches to probe the sequence determinants of protein aggregation and amyloidogenicity.* Protein Pept Lett, 2006. **13**(3): p. 279-86.

151. Rao, D.K., et al., *The alkali molten globule state of ferrocytochrome c: extraordinary stability, persistent structure, and constrained overall dynamics.* Biochemistry, 2006. **45**(10): p. 3412-20.

152. Rooki, H., et al., *Partially folded conformations of bovine liver glutamate dehydrogenase induced by mild acidic conditions.* J Biochem, 2007. **142**(2): p. 193-200.

153. Qureshi, S.H., et al., *Conformational and thermodynamic characterization of the molten globule state occurring during unfolding of cytochromes-c by weak salt denaturants.* Biochemistry, 2003. **42**(6): p. 1684-95.

154. Nakamura, S., T. Baba, and S. Kidokoro, *A molten globule-like intermediate state detected in the thermal transition of cytochrome c under low salt concentration.* Biophys Chem, 2007. **127**(1-2): p. 103-12.

155. Aitio, H., et al., *Characterization of apo and partially saturated states of calerythrin, an EF-hand protein from S. erythraea: a molten globule when deprived of Ca(2+).* Protein Sci, 2001. **10**(1): p. 74-82.

156. Christova, P., J.A. Cox, and C.T. Craescu, *Ion-induced conformational and stability changes in Nereis sarcoplasmic calcium binding protein: evidence that the APO state is a molten globule.* Proteins, 2000. **40**(2): p. 177-84.

157. Fink, A.L., *Molten Globule.* Encyclopedia of Life Sciences, 2001.

158. Sandberg, A., J. Leckner, and B.G. Karlsson, *Apo-azurin folds via an intermediate that resembles the molten-globule.* Protein Sci, 2004. **13**(10): p. 2628-38.

159. Arai, M. and K. Kuwajima, *Role of the molten globule state in protein folding.* Adv Protein Chem, 2000. **53**: p. 209-82.

160. Matthews, C.R., *Pathways of protein folding.* Annu Rev Biochem, 1993. **62**: p. 653-83.

161. Ptitsyn, O.B., *Kinetic and equilibrium intermediates in protein folding.* Protein Eng, 1994. **7**(5): p. 593-6.

162. Fujiwara, K., et al., *Folding-unfolding equilibrium and kinetics of equine beta-lactoglobulin: equivalence between the equilibrium molten globule state and a burst-phase folding intermediate.* Biochemistry, 1999. **38**(14): p. 4455-63.

163. Jennings, P.A. and P.E. Wright, *Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin.* Science, 1993. **262**(5135): p. 892-6.

164. Ptitsyn, O.B., et al., *Evidence for a molten globule state as a general intermediate in protein folding.* FEBS Lett, 1990. **262**(1): p. 20-4.

165. Peng, Z.Y., et al., *Does the molten globule have a native-like tertiary fold?* Philos Trans R Soc Lond B Biol Sci, 1995. **348**(1323): p. 43-7.

166. Bhakuni, V., *Alcohol-induced molten globule intermediates of proteins: are they real folding intermediates or off pathway products?* Arch Biochem Biophys, 1998. **357**(2): p. 274-84.

167. Osvath, S., et al., *Hierarchic finite level energy landscape model: to describe the refolding kinetics of phosphoglycerate kinase.* J Biol Chem, 2006. **281**(34): p. 24375-80.

168. Nishimura, C., H.J. Dyson, and P.E. Wright, *The kinetic and equilibrium molten globule intermediates of apoleghemoglobin differ in structure.* J Mol Biol, 2008. **378**(3): p. 715-25.

169. Roque, A., I. Ponte, and P. Suau, *Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1.* Biophys J, 2007. **93**(6): p. 2170-7.

170. van der Goot, F.G., J.H. Lakey, and F. Pattus, *The molten globule intermediate for protein insertion or translocation through membranes.* Trends Cell Biol, 1992. **2**(11): p. 343-8.

171. Baker, B.Y., D.C. Yaworsky, and W.L. Miller, *A pH-dependent molten globule transition is required for activity of the steroidogenic acute regulatory protein, StAR.* J Biol Chem, 2005. **280**(50): p. 41753-60.

172. Bose, H.S., et al., *The active form of the steroidogenic acute regulatory protein, StAR, appears to be a molten globule.* Proc Natl Acad Sci U S A, 1999. **96**(13): p. 7250-5.

173. Twigg, P.D., et al., *Disordered to ordered folding in the regulation of diphtheria toxin repressor activity.* Proc Natl Acad Sci U S A, 2001. **98**(20): p. 11259-64.

174. Kjellsson, A., I. Sethson, and B.H. Jonsson, *Hydrogen exchange in a large 29 kD protein and characterization of molten globule aggregation by NMR.* Biochemistry, 2003. **42**(2): p. 363-74.

175. Dumoulin, M., et al., *Reduced global cooperativity is a common feature underlying the amyloidogenicity of pathogenic lysozyme mutations.* J Mol Biol, 2005. **346**(3): p. 773-88.

176. Safar, J., et al., *Scrapie amyloid (prion) protein has the conformational characteristics of an aggregated molten globule folding intermediate.* Biochemistry, 1994. **33**(27): p. 8375-83.

177. Cremades, N., J. Sancho, and E. Freire, *The native-state ensemble of proteins provides clues for folding, misfolding and function.* Trends Biochem Sci, 2006. **31**(9): p. 494-6.

178. Gebel, E.B. and D. Shortle, *Characterization of denatured proteins using residual dipolar couplings.* Methods Mol Biol, 2007. **350**: p. 39-48.

179. Saxena, A.M., J.B. Udgaonkar, and G. Krishnamoorthy, *Characterization of intra-molecular distances and site-specific dynamics in chemically unfolded barstar: evidence for denaturant-dependent non-random structure.* J Mol Biol, 2006. **359**(1): p. 174-89.

180. Tanford, C., *Protein denaturation.* Adv Protein Chem, 1968. **23**: p. 121-282.

181. Perez-Jimenez, R., et al., *Mechanical unfolding pathways of the enhanced yellow fluorescent protein revealed by single molecule force spectroscopy.* J Biol Chem, 2006. **281**(52): p. 40010-4.

182. Wedemeyer, W.J., E. Welker, and H.A. Scheraga, *Proline cis-trans isomerization and protein folding.* Biochemistry, 2002. **41**(50): p. 14637-44.

183. Gianni, S., et al., *Parallel pathways in cytochrome c(551) folding.* J Mol Biol, 2003. **330**(5): p. 1145-52.

184. Kamagata, K., et al., *Multiple parallel-pathway folding of proline-free Staphylococcal nuclease.* J Mol Biol, 2003. **332**(5): p. 1143-53.

185. Ding, F., R.K. Jha, and N.V. Dokholyan, *Scaling behavior and structure of denatured proteins.* Structure, 2005. **13**(7): p. 1047-54.

186. Jaenicke, R., *Protein stability and molecular adaptation to extreme conditions.* Eur J Biochem, 1991. **202**(3): p. 715-28.

187. Szilagyi, A. and P. Zavodszky, *Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey.* Structure, 2000. **8**(5): p. 493-504.

188. Unsworth, L.D., J. van der Oost, and S. Koutsopoulos, *Hyperthermophilic enzymes--stability, activity and implementation strategies for high temperature applications.* Febs J, 2007. **274**(16): p. 4044-56.

189. Vieille, C. and G.J. Zeikus, *Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability.* Microbiol Mol Biol Rev, 2001. **65**(1): p. 1-43.

190. Matsui, I. and K. Harata, *Implication for buried polar contacts and ion pairs in hyperthermostable enzymes.* Febs J, 2007. **274**(16): p. 4012-22.

191. Goldman, A., *How to make my blood boil.* Structure, 1995. **3**(12): p. 1277-9.

192. Cambillau, C. and J.M. Claverie, *Structural and genomic correlates of hyperthermostability.* J Biol Chem, 2000. **275**(42): p. 32383-6.

193. Suhre, K. and J.M. Claverie, *Genomic correlates of hyperthermostability, an update.* J Biol Chem, 2003. **278**(19): p. 17198-202.

194. Luke, K.A., C.L. Higgins, and P. Wittung-Stafshede, *Thermodynamic stability and folding of proteins from hyperthermophilic organisms.* Febs J, 2007. **274**(16): p. 4023-33.

195. Razvi, A. and J.M. Scholtz, *Lessons in stability from thermophilic proteins.* Protein Sci, 2006. **15**(7): p. 1569-78.

196. Kumar, S., C.J. Tsai, and R. Nussinov, *Thermodynamic differences among homologous thermophilic and mesophilic proteins.* Biochemistry, 2001. **40**(47): p. 14152-65.

197. Wittung-Stafshede, P., *Slow unfolding explains high stability of thermostable ferredoxins: common mechanism governing thermostability?* Biochim Biophys Acta, 2004. **1700**(1): p. 1-4.

198. Ai, X., et al., *15N NMR spin relaxation dispersion study of the molecular crowding effects on protein folding under native conditions.* J Am Chem Soc, 2006. **128**(12): p. 3916-7.

199. Santos, H. and M.S. da Costa, *Organic solutes from thermophiles and hyperthermophiles.* Methods Enzymol, 2001. **334**: p. 302-15.

200. Tehei, M. and G. Zaccai, *Adaptation to high temperatures through macromolecular dynamics by neutron scattering.* Febs J, 2007. **274**(16): p. 4034-43.

201. Fraústo da Silva, J.J.R.a.W., R.J.P, *The Biological Chemistry of the Elements*. 1997, New York: Oxford University Press.

202. Yamashita, M.M., et al., *Where metal ions bind in proteins.* Proc Natl Acad Sci U S A, 1990. **87**(15): p. 5648-52.

203. Falke, J.J., et al., *Quantitating and engineering the ion specificity of an EF-hand-like Ca2+ binding.* Biochemistry, 1991. **30**(35): p. 8690-7.

204. Needham, J.V., T.Y. Chen, and J.J. Falke, *Novel ion specificity of a carboxylate cluster Mg(II) binding site: strong charge selectivity and weak size selectivity.* Biochemistry, 1993. **32**(13): p. 3363-7.

205. Dauter, Z., et al., *Zinc- and iron-rubredoxins from Clostridium pasteurianum at atomic resolution: a high-precision model of a ZnS4 coordination unit in a protein.* Proc Natl Acad Sci U S A, 1996. **93**(17): p. 8836-40.

206. Maher, M., et al., *Metal-substituted derivatives of the rubredoxin from Clostridium pasteurianum.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 2): p. 298-303.

207. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.

208. Dokmanic, I., M. Sikic, and S. Tomic, *Metals in proteins: correlation between the metal-ion type, coordination number and the amino-acid residues involved in the coordination.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 3): p. 257-63.

209. Dudev, M., et al., *Factors governing the metal coordination number in metal complexes from Cambridge Structural Database analyses.* J Phys Chem B, 2006. **110**(4): p. 1889-95.

210. Dudev, T. and C. Lim, *Principles governing Mg, Ca, and Zn binding and selectivity in proteins.* Chem Rev, 2003. **103**(3): p. 773-88.

211. Alberts, I.L., K. Nadassy, and S.J. Wodak, *Analysis of zinc binding sites in protein crystal structures.* Protein Sci, 1998. **7**(8): p. 1700-16.

212. Wittung-Stafshede, P., *Role of cofactors in protein folding.* Acc Chem Res, 2002. **35**(4): p. 201-8.

213. Coyne, H.J., 3rd, et al., *The characterization and role of zinc binding in yeast Cox4.* J Biol Chem, 2007. **282**(12): p. 8926-34.

214. Hussain, F. and P. Wittung-Stafshede, *Impact of cofactor on stability of bacterial (CopZ) and human (Atox1) copper chaperones.* Biochim Biophys Acta, 2007. **1774**(10): p. 1316-22.

215. Ikeguchi, M., K. Kuwajima, and S. Sugai, *Ca2+-induced alteration in the unfolding behavior of alpha-lactalbumin.* J Biochem, 1986. **99**(4): p. 1191-201.

216. Bushmarina, N.A., et al., *Cofactor effects on the protein folding reaction: acceleration of alpha-lactalbumin refolding by metal ions.* Protein Sci, 2006. **15**(4): p. 659-71.

217. Apiyo, D. and P. Wittung-Stafshede, *Presence of the cofactor speeds up folding of Desulfovibrio desulfuricans flavodoxin.* Protein Sci, 2002. **11**(5): p. 1129-35.

218. Berg, J.M. and Y. Shi, *The galvanization of biology: a growing appreciation for the roles of zinc.* Science, 1996. **271**(5252): p. 1081-5.

219. Cox, E.H. and G.L. McLendon, *Zinc-dependent protein folding.* Curr Opin Chem Biol, 2000. **4**(2): p. 162-5.

220. Fisher, W.R., H. Taniuchi, and C.B. Anfinsen, *On the role of heme in the formation of the structure of cytochrome c.* J Biol Chem, 1973. **248**(9): p. 3188-95.

221. Eliezer, D., et al., *Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding.* Nat Struct Biol, 1998. **5**(2): p. 148-55.

222. Goedken, E.R., et al., *Divalent metal cofactor binding in the kinetic folding trajectory of Escherichia coli ribonuclease HI.* Protein Sci, 2000. **9**(10): p. 1914-21.

223. Van Dael, H., P. Haezebrouck, and M. Joniau, *Equilibrium and kinetic studies on folding of canine milk lysozyme.* Protein Sci, 2003. **12**(3): p. 609-19.

224. Zartler, E.R., et al., *Structural basis for thermostability in aporubredoxins from Pyrococcus furiosus and Clostridium pasteurianum.* Biochemistry, 2001. **40**(24): p. 7279-90.

225. Strop, P. and S.L. Mayo, *Contribution of surface salt bridges to protein stability.* Biochemistry, 2000. **39**(6): p. 1251-5.

226. Hiller, R., et al., *Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200 degrees C.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11329-32.

227. Wittung-Stafshede, P., *Role of cofactors in folding of the blue-copper protein azurin.* Inorg Chem, 2004. **43**(25): p. 7926-33.

2 IRON-SULFUR PROTEINS AND THE DI-CLUSTER FERREDOXINS

CONTENTS

This chapter was partially published in

Carlos Frazão, David Aragão, Ricardo Coelho, **Sónia S. Leal**, Cláudio M. Gomes, Miguel Teixeira and Maria Arménia Carrondo
"Crystallographic analysis of the intact metal centres $[3Fe\text{-}4S]^{1+/0}$ and $[4Fe\text{-}4S]^{2+/1+}$ in a Zn^{2+} - containing ferredoxin"
FEBS Lett. 582: 763-767 (2008)

2.1 IRON-SULFUR CLUSTERS AND PROTEINS

Iron-sulfur proteins are a multifaceted class of proteins containing iron-sulfur clusters (Fe-S) as a prosthetic group. In contrast to most other bio-organic cofactors, these compounds are of inorganic nature and consist simply of iron cations and sulfide anions. These clusters are modular structures, meaning that given the proper conditions they can assemble spontaneously without the intervention of proteins. Withal, the Fe-S clusters are inherently labile due to an extreme sensitivity to oxidants. Therefore each organism has to assemble them inside the cell, de novo. The Fe-S proteins enclose a vast diversity of molecular structures within a large range of molecular weight sizes, ranging from 6 to over 500 kDa, and accomplish a wide array of functions in many metabolic reactions. The most prominent Fe-S proteins include complexes I, II, and III of the bacterial and mitochondrial respiratory chains; photosystem I and ferredoxin of photosynthesis; nitrogenase in azototrophic bacteria; aconitase of the Krebs cycle; and the iron regulatory protein 1 (IRP1) involved in iron-uptake regulation in mammals [1].

The classic Fe-S clusters in proteins are those in which multiple Fe atoms are coordinated by cysteinates from the protein and linked to each other through sulfide bridges. The most common types of clusters in proteins are the [4Fe-4S], [2Fe-2S] and [3Fe-4S] clusters (see figure 2.1; scheme A, B and C respectively). The two former clusters constitute the basic structures from where more complex structures with two or more of these simple Fe-S clusters can be formed, like the [8Fe-7S] P-cluster (see figure 2.1; scheme D). Another example is given by the [3Fe-4S] clusters, a cubane structure which is derived from a [4Fe-4S] cluster, after loss of a coordinating cysteine and of one iron atom. Nevertheless, Fe-S proteins family also encloses proteins that are bound to a FeS4 site, that is unique among Fe-S active sites in being devoid of inorganic sulfur (see figure 2.1; scheme E). As well, includes proteins that are bound to clusters, in which other metals, in addition to Fe, (such as molybdenum, vanadium, or nickel) and/or other ligands, in addition to cysteine, (such as nitrogen of histidine residues) are present.

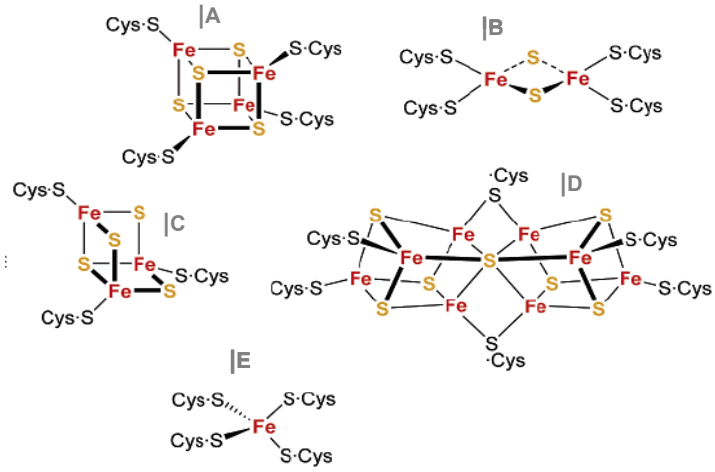


Figure 2.1 |Schematic representations of Fe-S cluster structures identified in biologic proteins.
A| Cubic tetranuclear cluster; **B|** Rhombic binuclear cluster; **C|** Cuboidal trinuclear cluster; **D|** P-cluster; **E|** Mononuclear Fe center.
Fe – iron; **S** – inorganic sulfur; **Cys.S** – cysteine thiolate ligand
Based on [2].

|Functional and structural diversity

Iron-sulfur (Fe-S) cluster proteins are found in all life domains, within all kind of cells and cellular compartments. In agreement, these ubiquitous proteins are involved in a plethora of biological functions (Table 2.1). This versatility of functions is directly correlated to the structural diversity of its clusters, as well as with their chemical plasticity and reactivity properties [3]. Moreover, it is not uncommon to encounter multiple clusters present in the same protein [4], as well as clusters containing metals other than iron (e.g., nickel or molybdenum) [1] or bridged to additional complex sites like a heme group [5], thus further contributing to increase the complexity and the multiplicity of structures and functions of Fe-S cluster proteins. Besides this natural diversity of the clusters structures, also the intrinsic ability of these molecules to interconvert contributes for the multiplicity of attributed functions in Fe-S proteins. In fact, several biological processes depend on interconversions between different types of Fe-S clusters. For example, that is the case of the FNR (fumarate nitrate reduction) protein from *Escherichia coli*, a transcription activator of genes encoding for components of anaerobic pathways,

in which exposure to oxygen inactivates the protein as a result of the reversible conversion of its [4Fe-4S] clusters to a [2Fe-2S] one, leading to dissociation of the catalytically proficient dimeric form [6, 7]. Also, in aconitase, cluster rearrangements between the active [4Fe-4S] and the inactive [3Fe-4S] containing form, modulate the protein function [8, 9].

Table 2.1 |Biological function of different Fe-S cluster proteins.

Functions	Fe-S proteins	Cluster type(s)
Electron transfer	Ferredoxins; redox enzymes	[2Fe-2S]; [3Fe-4S]; [4Fe-4S]
Coupled electron/proton transfer	Rieske protein Nitrogenase	[2Fe-2S] [8Fe-7S]
Substrate binding and activation	Acetyl-CoA synthase Radical SAM enzymes (de)Hydratases	Ni-Ni-[4Fe-4S]; [Ni-4Fe-5S] [4Fe-4S] [4Fe-4S]
Fe or cluster storage	Ferredoxins	[4Fe-4S]
Structural	Endonuclease III MutY	[4Fe-4S] [4Fe-4S]
Regulation of gene expression	SoxR FNR IRP IscR	[2Fe-2S] [4Fe-4S] / [2Fe-2S] [4Fe-4S] [2Fe-2S]
Regulation of enzyme activity	Ferrochelatase Glutamine PRPP amidotransferase	[2Fe-2S] [4Fe-4S]
Disulfide reduction	Ferredoxin:thioredoxin reductase Heterodisulfide reductase	[4Fe-4S] [4Fe-4S]
Sulfur donor	Biotin synthase	[2Fe-2S]

Abbreviations: **SAM**, S-adenosylmethionine; acetyl-CoA, acetyl coenzymeA; **FNR**, fumarate nitrate reduction; **IRP**, iron-regulatory protein; **IscR**, iron-sulfur cluster assembly regulatory protein; **PRPP**,phosphoribosylpyrophosphate. Adapted from [10]

The extreme ability to delocalize electron density over both Fe and S atoms makes Fe-S clusters a prosthetic group ideally suited for their most likely primary role, which is mediating biological electron transport by accepting, donating, shifting and storing electrons. For those Fe-S proteins preferentially engaged in electron transport, the redox potentials of their clusters are thus of central importance to their biological function. In this respect, the redox potential of a

given Fe-S cluster in a protein is strongly influenced by the structure and composition of the surrounding protein matrix where parameters such as cluster-protein hydrogen bonding, cluster solvent accessibility, and the number of charged residues play a determinant role. Therefore, it is not surprising that redox potentials of Fe-S clusters in Fe-S proteins vary widely, ranging from > +400mV to < -600 mV) [11], and is rather common to find similar Fe-S clusters with distinct redox potentials within different proteins. One clear example of how protein scaffolding can influence the redox potential of the cluster is found in HiPIPs (high potential iron proteins). These proteins contain a [4Fe-4S] bound to four conserved cysteines within the protein interior [12]. As a result of its hydrophobic environment and hydrogen-bonding network, the cluster implements the $[4Fe\text{-}4S]^{2+/3+}$ transition; hence a substantial increase in the redox potential (+100 to 400mV) is observed, in contrast to the low-potential (-150 to -700 mV) electron carriers implementing the $[4Fe\text{-}4S]^{1+/2+}$ redox transition in ferredoxins [13]. Other important factor likely to play a role in the modulation of the redox potential of the clusters, gating electron transport, or coupling proton and electron transport, are the coordinating ligands [10]. Although Cys residues dominate the preferences of clusters binding to proteins, other residues such as Asp, His, Ser, or backbone amide ligation at a unique Fe site are occasionally encountered in clusters that function in electron transport. Such an example is found in Rieske proteins, which are coordinated to the protein by two Cys and two His residues. These proteins contain the same cluster and share the same redox transition as the $[2Fe\text{-}2S]^{1+/2+}$ plant type ferredoxins. Nevertheless, the His ligation of the cluster in Rieske proteins is likely to be the cause of the upshift of the redox potential to (+100 to +400 mV) [14] in contrast to the low potential (-150 to -450 mV) of the [2Fe-2S] ferredoxins. On the other hand by straddling protein structural elements, these metal clusters are known to contribute to maintain the conformational stability and control the protein structure in the vicinity of the centre. Hence, it is not surprising that Fe-S clusters have also important structural and regulatory roles. In the case of the DNA repair enzymes, endonuclease III [15] and glicosylase MutY [16], the available data indicates that redox-inactive [4Fe-4S] clusters play purely structural roles, similar to that of Zn in Zn-finger proteins, in which the cluster controls the structure of a protein loop essential for recognition and repair of damaged DNA.

In other cases the integrity and disassembly of the cluster clearly provides a regulatory signal. An example is given by amidotransferase, an enzyme which requires the presence of one [4Fe-4S] cluster to be stable. When the protein is exposed to dioxygen *in vivo*, the rates of inactivation and protein degradation are the same, and the inactivation step is the destructive reaction of the cluster by molecular oxygen [17].

Assembly of Fe-S clusters: an evolutionary perspective

Iron-sulfur clusters are proposed to have played an important role in the evolution of protocellular systems and in the emergence of life on earth [18]. This speculation relies essentially on the intrinsic chemical properties of the clusters, in association with the dominant geochemical conditions of the primordial pre-biotic era [19]. At that time, an anoxic world with high concentrations of available iron and sulfur was the most likely scenario, thus favouring the spontaneous assembly of Fe-S active sites within primitive macromolecules (e.g. early proteins) possessing adequately positioned thiolate ligands. In fact, the simpler Fe-S clusters are known to form spontaneously from ferric iron, thiol, and sulfide under exclusion of oxygen. Such Fe-S structures share basic features observed with those of the protein-bound clusters [2], and are therefore believed to constitute one of the most ancient types of prosthetic groups. Another supporting fact for this speculation relies on the demonstration that the simplest Fe-S clusters can be assembled efficiently *in vitro* from an apoproteins, starting from ferrous iron (Fe^{II}) and sulfide, using chemical reconstitution methods and strictly anaerobic conditions [16]. Moreover, genome and metagenome of all known sequences including the most primitive ones confirms the pervasiveness of Fe-S clusters in the living world. Meanwhile, when dioxygen appeared in the biosphere (~ 2 billion years ago) the environmental conditions on earth have profoundly changed. The appearance of oxidizing conditions decreased the availability of iron (through precipitation as ferric oxides and hydroxides) and frustrated immensely the spontaneous assembly of Fe-S clusters. In response, life forms have adapted and evolved to these new circumstances. In fact, cluster biosynthesis in any present living cell does not occur spontaneously and requires surprisingly complex

biochemical assembly systems. The need for assisted, rather than spontaneous, Fe-S cluster assembly and insertion into apoproteins may mainly be the result of exigent anaerobic conditions, and the toxicity resulting from the high concentrations of iron and sulfide ions required for efficient chemical reconstitution in side the cell. Furthermore, the protein-mediated assembly of Fe-S clusters is likely to be more tightly regulated and efficient under biological control; in addition, protein mediation is likely to assure an increased specificity in the transfer of the formed Fe-S centre to the target peptide. Presently, among eukaryotes, three distinct assembly systems involved in the maturation of cellular Fe-S proteins are known: i) the ISC (iron-sulfur-cluster) assembly machinery, working in the mitochondria and suggested to be inherited from an evolutionary ancestor of these organelles, a alfa-proteobacteria [20]. In fact, this system is also present in the biogenesis of Fe-S proteins in eubacteria and homologs of all the isc operon–encoded components from eukaryotes are present in bacterial genomes [21]; ii) the SUF (sulfur mobilization) machinery present in the cytosol of eukariotes, but also found in some Eubacteria, Archaea and in the plastids of plant cells and algae [22], where they perform crucial roles in the biogenesis of Fe-S proteins of the photosynthetic apparatus [23]. It has been suggested that this machinery was probably transferred to the eukaryotic cell by endosymbiosis of a photosynthetic bacterium, inherited from the cyanobacterial ancestor of plastids [24]. In agreement with this hypothesis, homologues of this system are present in plastids of plants and algae; iii) the CIA (cytosolic iron-sulfur protein assembly) machinery, for maturation of cytosolic and nuclear Fe-S proteins that is unique among eukaryotes. Nevertheless, biogenesis of cytosolic and nuclear Fe-S proteins still depends on the mitochondrial ISC assembly system. The organelles appear to export a component to the cytosol (the so-called ISC export machinery) that is essential for extra-mitochondrial Fe-S protein maturation [25]. To date, the eukaryotic ISC assembly, ISC export, and CIA machineries encompass 21 components, impressively underlining the increased complexity of this biosynthetic process. The systems ISC, SUF and CIA are highly conserved in eukaryotes from yeast, to humans, and plants [26].

Bacteria share the ISC assembly machinery and the SUF-sulfur mobilization machinery with eukaryotes but have also developed one additional distinct

biosynthetic system for Fe-S protein formation. The nitrogen fixation system, NIF, that is dedicated to the assembly of the Fe-S clusters of nitrogenase, the enzyme responsible for the conversion of N_2 to NH_3 in nitrogen-fixing bacteria. Key components of the NIF system are the cysteine desulfurase NifS [27] and the scaffold protein NifU [28] on which the Fe-S clusters are preassembled before insertion to the nitrogenase.

The ISC machinery is required for the generation of the majority of cellular Fe-S proteins and thus performs a general housekeeping biosynthetic function in bacteria [29]. The two central proteins of the ISC machinery are IscS and IscU, and they are homologous in function to NifS and NifU, respectively. These two ISC proteins are encoded by the isc operon, which contains additional genes involved in biogenesis. The SUF or sulfur mobilization assembly plays a similar general role as the ISC system but operates mostly under stress conditions such as iron-limitation or oxidative stress [30]. The SUF system also contains a cysteine desulfurase (SufS) as the sulfur donor and an alternative scaffold protein termed SufA.

These different machineries, NIF, ISC and SUF, have in common the involvement of a cysteine desulfurase (NifS, IscS and SufS), which allows the utilization of *L*-cysteine as a source of sulfur atoms. In addition, they all contain scaffold proteins like NifU, IscU, SufU and many others, which provide an intermediate assembly site for Fe-S clusters or precursors, from which these inorganic species can be donated to the target apoproteins. The basic concept of Fe-S protein biogenesis relies therefore on the transient assembly of Fe-S clusters on a scaffold protein, with sulfur derived from a cysteine desulfurase.

Iron-sulfur clusters and hosting folds

Iron-sulfur clusters are modular structures that can increase the stability and shape the protein structure by preferential side chain ligation. On the other hand, the polypeptide chain accommodates the cluster and generates a protective ligand framework against the oxidative degradation of the centre. In this respect, Fe-S cluster proteins usually hold iron-sulfur clusters in a rather hydrophobic moiety and Fe-S clusters are generally hosted within a single domain which

provides a relatively rigid and protective ligand support; Fe-S sites located at exposed domains or subunit interfaces appear rather rare. The known exceptions are generally associated to scaffold or regulatory proteins like the IscA-type protein from *Synechocystis* [31] and poplar glutaredoxin C1, which has a bridging iron-sulfur cluster at the active site [32], respectively.

From this successful interplay between proteins and Fe-S clusters, evolution has created different suitable folds around the clusters, in order to diversify and optimize the influence of the centres on the protein structure and function. This has resulted in various Fe-S proteins structures that accomplish many, and very distinct functions.

The possibility for Fe-S clusters to be hosted in various ways in a given protein fold can in principle offer nearly unlimited opportunities for novel Fe-S folds, nevertheless the achievement of the already functional folds appears in fact “precious” to life. In this respect, the structure of a protein is generally much more conserved throughout evolution than the protein sequence [33]. Among protein families, the amino acid sequence identity can be as low as 5% whereas structural identity is at least of 50%, mainly at the protein core [34]. A good example can be found within the class of copper chaperone proteins, where CopZ from *Bacillus subtilis* and Atox1 from *Homo sapiens* share only 15% sequence identity, but retain the same overall fold [35]. This strongly establishes the importance of the protein domain as a fundamental unit in evolution, and reveals the astonishing diversity of proteins that can be assembled by duplicating domains combining them in different ways, or as a result of few variations. This speculation appears to suit the example of the Clostridial 2x[4Fe-4S] ferredoxin fold. This small Fd (only 55 residues) is suggested to hold a very ancient fold, as evidenced from analysis of amino acid sequences, from which an ancestral gene duplication event could be traced [36]. Concurring, this ferredoxin displays a conspicuous two-fold symmetry, around the two [4Fe-4S] clusters, and where each half of the sequence contributes with Cys ligands to both clusters. Moreover, the $(\beta\alpha\beta)_2$ topology is unique among Fe-S proteins, and one of the most common topology among ferredoxins. In fact, and in agreement with its suspected ‘antiquity’, the 2x[4Fe-4S] Fd fold is one of the most widespread folds in Fe-S proteins, and the one that has undergone the most extensive branching [37].

Some examples of the considerable diversification of this fold are outlined in Figure 2.2, which includes insertions of polypeptide segments, loss of one of the clusters, and occasional incorporation of a disulfide bond, as well as the inclusion of a [3Fe-4S] cluster replacing a [4Fe-4S] cluster one.

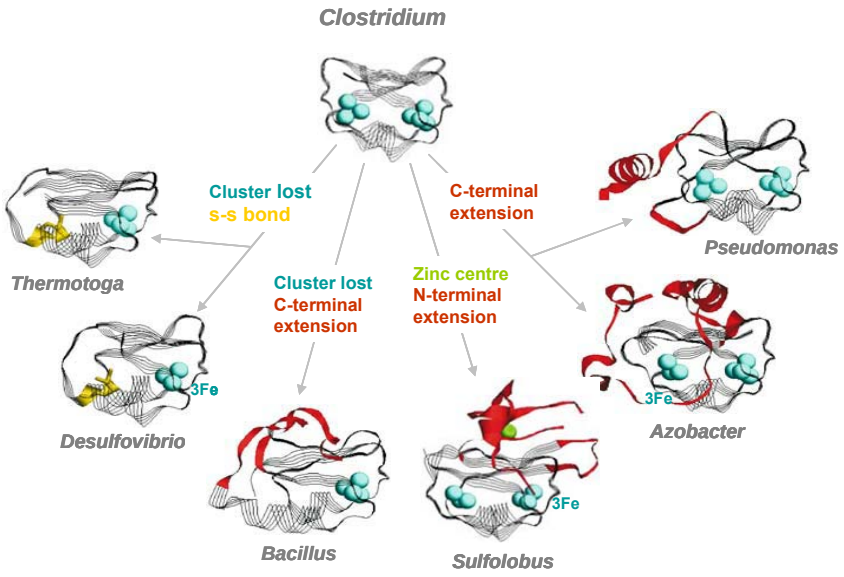


Figure 2.2 | Evolutionary scheme of the 2x[4Fe–4S] ferredoxin fold.
Only Ferredoxins (Fd) are shown. Even greater variations are displayed by homologous domains in redox enzymes. The generic names of the microorganisms are indicated. The ancestral Clostridial Fd is shown at the top, and the most significant variations are indicated on the arrows. Iron atoms are shown as cyan spheres, while sulfide atoms were removed for clarity sake. Clusters labelled 3Fe are of the [3Fe–4S] type. The core polypeptide fold common to all proteins is shown as grey strands, insertions are shown as red ribbons, disulfide bridges are shown in yellow, and the zinc atom is shown in green. Data Bank (PDB) entries: *Thermotoga* (1vjw); *Desulfovibrio* (1fxd); *Bacillus* (1iqz); *Sulfolobus* (1xer); *Azobacter* (7fd1); *Pseudomonas* (2fgo). Adapted from [38]

Proteins holding [2Fe-2S] clusters present a distinct fold from the one previously described. Within the family of the [2Fe-2S] clusters proteins, one of the most common is the one of plant- and vertebrate-type ferredoxins. The basic structural scaffold of these ubiquitous proteins, the β-grasp [39], is not unique to Fe-S proteins, but its Fe-S version has been suggested to feature among the most ancient of its varieties [40]. These proteins usually present the [2Fe-2S] cluster relatively near the surface and protected by a long loop including three of the four Cys ligands [41, 42]. The opposite side of the molecule consists of a four

stranded β-sheet strand covered by an α-helix, which form the β-grasp. Additional lateral α-helices connect the cluster binding region and the β-grasp. These proteins are structurally much less diverse than the [4Fe-4S] ferredoxins. Variations include N-terminal extensions [41] and differences in the redox partner interaction surfaces between the plant type and the vertebrate type ferredoxin [43]. These proteins usually function as low-potential electron carriers (e.g. in photosystem I, adrenoxin) but are also involved in the biosynthesis of Fe-S clusters (e.g. Isc Fd) or committed to the activation of some oxygenases (Xy1T).

The rubredoxin type Fe-Cys_4 site stands out as a basic building block of Fe-S clusters, although interestingly the same coordination unit can be found naturally holding a zinc metal ion in stead of iron, like in alcohol dehydrogenase and aspartate transcarbamylase [44]. Nevertheless, this type of centre is generally associated with the structural motifs of Rubredoxins fold that consist of a single iron atom coordinated to four cysteinyl sulfurs, occurring on two CysXXCys segments, and belonging to two symmetry-related loops. A three-stranded β-sheet involves both the N and C-termini, and a mostly aromatic hidrophobic core [45]. These are small proteins (~6 kDa), usually involved in electron transfer processes, and acting as electron carriers. The rubredoxin fold is found as a structural domain in diverse complex multi-domain metalloproteins, this includes proteins that contain Rd-like domains in combination with other active sites like in rubrerythin and flavorubredoxin [46], as well as duplication of the Rd like fold in homodimeric proteins [47].

Other Fe-S folds are most likely more severe adaptations of these pre-existing folds, to diversify or/and improve proteins functions. Examples supporting this speculation are probably the HiPIP and Rieske proteins. These Fe-S proteins hold a [4Fe-4S] and [2Fe-2S] cluster respectively, and present folds distinct from those previously associated for each of these clusters. There are also classes of proteins where the presence of a Fe-S cluster is an exceptional occurrence of unknown functional relevance. A significant case is the tryptophanyl transfer RNA synthetase from *Thermotoga maritima*, which is at least rare, if not unique, as a Fe-S protein among aminoacyl transfer RNA synthetases. The [4Fe-4S] cluster has no known function in this particular case, and its presence may simply result from the happenstance of at least three appropriately positioned cysteine

residues. This occurrence points to a possible way for Fe-S clusters to appear in novel sites where they may subsequently be stabilized, acquire new functions, gain a selective advantage, and eventually lead to the emergence of a new family of Fe-S proteins [38].

Besides the overall features associated to each specific Fe-S cluster protein, there are many folds for a protein to accommodate a given Fe-S clusters when considering the close vicinity of the metal centre [38]. This highlights the fact that proteins have evolved to become adjusted to the geometry of the clusters, by allowing a more plastic folding to occur, to suit better their functional demands. In contrast, a given fold was shown to host almost in exclusive a specific Fe-S cluster. The main reason for this exclusiveness most likely relies on the geometry of the clusters, which are very distinct (Figure 2.3).

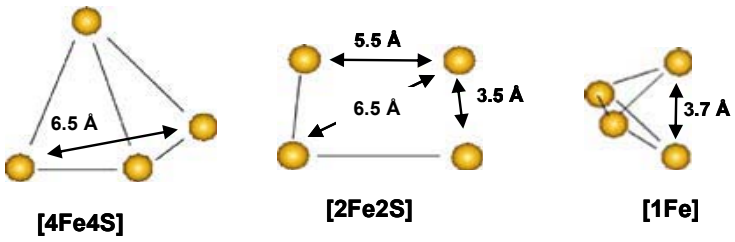


Figure 2.3 |Displacement of the Cys ligands accordingly with the nuclearity of Fe-S clusters. Positions of the Sy atoms of the Cys ligands. The cluster [3Fe–4S] (not shown here) is derived from the [4Fe–4S] framework by removal of one iron and its Cys ligand and involves only minor rearrangements not relevant to be displayed [2].

This prominent characteristic of the clusters makes it almost impossible to host clusters of different nuclearities in the same fold. One of the few documented exceptions is the Rubredoxin fold (holding a FeS4 site), that is superimposable to that of a Rieske protein in the region holding the [2Fe-2S], [48] even though different ligand sets are implemented in the two cases. However, promiscuity between the two clusters only actually occur when there is a replacement (a histidine for cysteine) and reorientation of two of the active-site ligands. Indeed, only mutating a single Cys ligand of the iron can result in the installation of a [2Fe-2S] cluster in the Rubredoxin [49], and conversely the Rieske site can only be converted into a Fe-Cys_4 site by a triple mutation [50]. Examples of cluster inter-

conversions usually happens when the protein allows flexibility and is able to undergo structural changes: this can be evidenced on the nitrogenase iron protein, which normally contains a [4Fe-4S] cluster [1], but can under some circumstances host a [2Fe-2S] cluster. This conversion however, happens only in conditions where the protein is known to undergo structural changes: either on the way to irreversible denaturation [51], or as a result of ATP binding in the presence of glycerol [52]. Other proteins having the potential to bind either binuclear or tetranuclear clusters include the fumarate nitrate regulator [53] and IscU [10], both of which are predicted from their function to be flexible. A more undemanding conversion can be observed between a [4Fe4S] and a [3Fe4S] cluster since these clusters share a close geometry. This is the case of aconitase where the inter-conversion of these two centres is one of the steps in a regulatory mechanism of activity in the citric-acid-cycle enzyme, as well as the iron-responsive element mRNA binding protein involved in iron homoeostasis [54].
There are a few protein folds that are specifically Fe-S folds (e.g. 2x[4Fe-4S] Fd, or HiPIP) while many others also occur in other contexts with other metals. The most evident example is the fold of rubredoxins that presents high similarity with the fold of proteins belonging to zinc-finger families. In fact, Rubredoxin exhibits affinities for iron and zinc ions, and the specificity appear to depend essentially on the concentration of the available metal [44].

2.2 THE ARCHEAL DI-CLUSTER FERREDOXINS

As discussed, ferredoxins are among the simplest iron-sulfur proteins and one of the most prevailing classes of mobile electron carriers in biology. These proteins are well-known represented in all life forms, where interestingly di-cluster ferredoxins holding a [3Fe-4S] and [4Fe-4S] centres appear particularly abundant within the Archaea domain [55-57]. Archaea are considered one of the older lineages of life and present valuable examples of primitive times [58, 59]. In addition, the extremely thermophilic branch of this ancient domain is recognized to have intrinsically highly stable ferredoxins. Taken together, the di-cluster ferredoxins from hyperthermophilic Archaea constitute excellent working models for studies addressing the role of iron–sulfur centres in proteins folding and

stability as well as to clarify ancestral strategies for protein stabilization.

The di-cluster archaeal seven-iron ferredoxins are a family of small (~11 kDa) monomeric and acidic (*p*I ~3.5) proteins, essentially characterized by holding a $[3Fe4S]^{1+/0}$ and a $[4Fe4S]^{2+/1+}$ cluster, in addition to a Zn^{2+} site and a N-terminal extension of ~30 residues. Interestingly, some archaeal organisms (e.g. *Acidianus ambivalens* and *Sulfolobus metallicus)* are known to express isoforms lacking the Zn ligands [56, 60]. The zinc centre in these ferredoxins is suggested to play a structural role, essentially by holding the protein core to the N-terminal extension. In fact, no catalytic activity was described for this zinc centre, and the mutagenesis of any of its coordinating His ligands is known to result in a destabilized protein [61, 62]. Studies on the redox properties of these ferredoxins showed that, whereas the trinuclear centre has a -250 to -310 mV reduction potential, the tetranuclear one is usually more negative, -550 mV [56, 63, 64], which led to the suggestion that this centre would have a structural, rather than a catalytic and functional role [65].

This family of proteins is in general very identical in sequence and structure and usually very abundant in the cytosol of the host organism, thus suggesting a rather important physiological role. These proteins are intrinsically very stable, both very resistant to chemical and thermal denaturation [66, 67]. In fact guanidine hydrochloride only induces an unfolding transitition at concentrations above 6.5 M (pH 7, 20ºC). Also, the midpoint of thermal denaturation, at neutral pH, is above the boiling point of water, at ~ 120 ºC [67, 68]. Withal it was described that this extreme stability is significantly decreased at low pH [66], thus evidencing a rather relevant role of ionic interactions in the overall stability of the protein. Although these are small proteins, di-cluster ferredoxins are described to unfold *in vitro* in an irreversible manner. This behaviour is suggested to occur as a result of the degradation of the Fe-S clusters and the subsequent lack of the appropriate conditions to refold properly; whether these conditions are an oxidative environment or the absence of the proper metallochaperones machinery, remains to be fully understood.

Crystal structure of a di-cluster ferredoxin holding a zinc centre

Detailed structural models of di-cluster seven-iron ferredoxins constitute a valuable resource for folding and stability studies relating the role of the metal cofactors with protein stability. Two crystal structures are currently available, one from the protein from *Sulfolobus tokodaii* (*St*Fd) [65] and the other from *Acidianus ambivalens* (*Aa*Fd) [69].

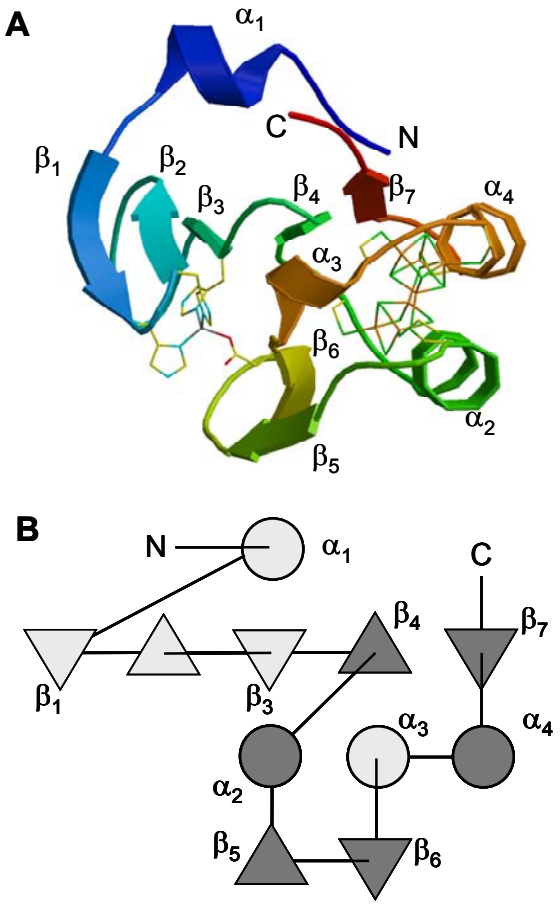


Figure 2.4 | Structure of the *Acidianus ambivalens* ferredoxin.
(**A**) Cartoon of the AaFd structure colored from blue (N-term) to red (C-term). (**B**) Topology diagram with β-chains as triangles and helices as circles, highlighting in white the N-terminal extension (residues 1-38) and in grey the canonical 2x(βαβ) ferredoxin fold (residues 38-103), whose second helix shows a kink which includes a 3_{10}-helix insert, in white.

The former allowed the finding of the zinc site, which was until then, unknown to exist in the protein. The latter, provided an accurate structural description of the Fe-S clusters and surrounding residues, as the *St*Fd structure reported a protein with the Fe-S damaged. Apart from the important detail that the AaFd structure refers to a protein with intact Fe-S clusters, the two structures are very similar, which is not surprising considering the fact that they share a very high amino acid identity (95%). For this reason, subsequent descriptions and amino acid numbering refer to the AaFd. The AaFd structure (figure 2.4) contains an extended antiparallel β-sheet, spanning over the entire protein and including three β-strands (β_{1-3}) of the N-terminal extension and two further β-strands (β_4 and β_7) of the core region. The core region also contains a second, smaller, anti-parallel β-sheet (β_5 and β_6), approximately perpendicular to the larger extended sheet, separating the Zn^{II} binding region from the Fe-S clusters. These are nested between the helices that flank the core region and the two sheets.

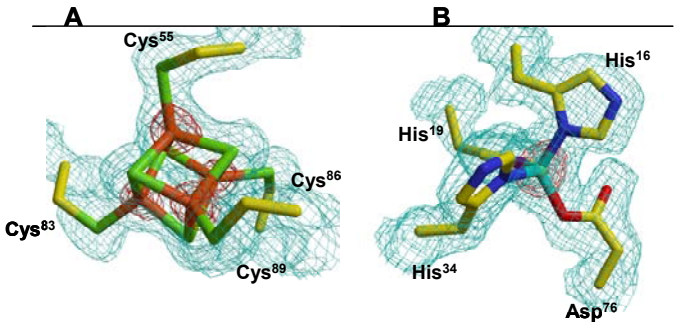


Figure 2.5 |2.0 Å resolution 2m|Fo|-D|Fc| sigma A Fourrier maps of electron density at 1.5σ (cyan mesh) and 8σ (red mesh), at metals centres and coordinating ligands.
(**A**) the [4Fe-4S] cluster and (**B**) the Zn centre with respective ligands are plotted as sticks, with carbon in yellow, nitrogen in blue, oxygen in red, sulfur in green, iron in orange and Zn in cyan.

The *Aa*Fd N-terminal extension provides the ligands for a zinc centre, which is exceptional among archaeal ferredoxins. The metal has a tetrahedral coordination, with three His and one Asp as ligands. The electron density clearly defines this fully occupied metal centre (figure 2.5, B), in accordance with the chemical analysis (0.8 ± 0.2 Zn^{II}/mol). The arrangement of this terminal extension

suggests that it plays an important role in ferredoxin conformational stability. Site-directed mutagenesis studies have shown that the zinc centre stabilizes the protein at high temperatures [61, 62]. This ferredoxin contained 6.6 ± 0.5 Fe/mol and showed EPR spectra characteristic of each type of centres (not shown), strongly suggesting that the iron–sulfur centres were intact, as corroborated by the X-rays diffraction analysis. The irons of each cluster are bound to the polypeptide chain via Fe–S bonds to Cys residues 45, 51 and 93 for cluster I, and to Cys residues 55, 83, 86 and 89 for cluster II (see figure 2.5, A). The obtained maps clearly show well defined densities and respective ligands.

|Stability and Unfolding of di-cluster ferredoxins: state of the art

Prior to this thesis, the *Aa*Fd had already been the focus of stability studies. The initial studies on the protein has shown that this Fd is equally expressed in aerobic as well as anaerobic conditions and that the protein is intrinsically very stable with respect to high temperatures [56]. Further studies revealed that the melting temperature of this protein at physiological pH was ~122ºC and simultaneously evidenced that *Aa*Fd was also highly resistant to chemical denaturation, as no unfolding occurred in the presence of 8 M GuHCl, pH 7 [67]. In sequence of this extreme stability it was suggested that a packed core fold and the presence of the zinc centre in the ferredoxin could significantly contribute to the observed stability. However, both thermal and chemical stability was shown to be severely affected at low pH, and thus suggested that interactions involving deprotonated Asp or Glu side chains were also significantly

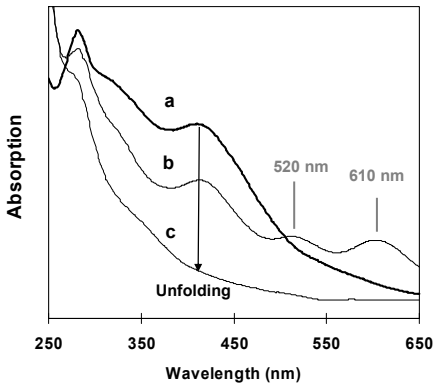


Figure 2.6 |Visible absorption spectra on a di-cluster ferredoxin during chemical alkaline unfolding.
a - Holo native protein; **b** -transient species observed during the chemical alkaline unfolding of ferredoxin; **c** - denaturated protein in a apo-state.

contributing to the high stability at neutral pH [66]. Subsequently, a more detailed analyse on the chemical denaturation of the *Aa*Fd at pH 10, evidenced that the characteristic peak of the Fe-S clusters at 410 nm decreased and transient absorption bands at 610 e 520 nm appeared after mixing with the denaturant (see figure 2.6). These observed transient spectroscopic features on ferredoxin, highly resembles those observed during the formation of a linear three-iron sulfur centre in the Fe-S cluster protein – aconitase. In aconitase the formation of the linear cluster results from the rearrangement of the cubic [3Fe-4S] centre, observed by absorption, EPR, and Mossbauer spectroscopy, under specific mild unfolding conditions; namely alkaline incubation (pH>9.5) or when the protein is treated with 4-8 M urea [70]. Although, no direct physiologic role was attributed to this linear form, structure-induced cluster rearrangements are known to be able to play an important role in many regulatory processes. Based on all previous evidences it was suggested in the literature that within this family of di-cluster ferredoxins, protein unfolding could result in a transitory rearrangement of the iron sulfur clusters into a presumably linear three iron sulfur centre intermediate followed by their dissociation upon complete unfolding (figure 2.7) [66, 67, 71].

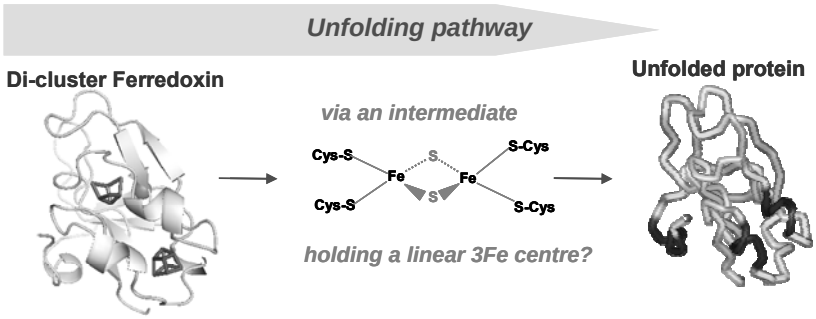


Figure 2.7 |Scheme on the hypothesis of the formation of a linear 3Fe centre as a general degradation intermediate of di-cluster ferredoxins unfolding.

Fe – iron; **S** – inorganic sulfur; **Cys.S** – cysteine thiolate ligand

In agreement it was evidenced that the transient species of ferredoxin characterized by the 520 and 610 nm bands formed in parallel with secondary-structure disappearance and shown by gel filtration to remain bound to the

unfolded polypeptide [71]. These observations were subsequently further extended to several others iron-sulfur proteins, suggesting a common strategy for the iron-sulfur proteins to unfold under alkaline conditions [72].

2.3 REFERENCES

1. Rees, D.C., *Great metalloclusters in enzymology.* Annu Rev Biochem, 2002. **71**: p. 221-46.
2. Venkateswara Rao, P. and R.H. Holm, *Synthetic analogues of the active sites of iron-sulfur proteins.* Chem Rev, 2004. **104**(2): p. 527-59.
3. Beinert, H., R.H. Holm, and E. Munck, *Iron-sulfur clusters: nature's modular, multipurpose structures.* Science, 1997. **277**(5326): p. 653-9.
4. Sazanov, L.A. and P. Hinchliffe, *Structure of the hydrophilic domain of respiratory complex I from Thermus thermophilus.* Science, 2006. **311**(5766): p. 1430-6.
5. Crane, B.R., L.M. Siegel, and E.D. Getzoff, *Sulfite reductase structure at 1.6 A: evolution and catalysis for reduction of inorganic anions.* Science, 1995. **270**(5233): p. 59-67.
6. Khoroshilova, N., et al., *Iron-sulfur cluster disassembly in the FNR protein of Escherichia coli by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity.* Proc Natl Acad Sci U S A, 1997. **94**(12): p. 6087-92.
7. Lazazzera, B.A., et al., *DNA binding and dimerization of the Fe-S-containing FNR protein from Escherichia coli are regulated by oxygen.* J Biol Chem, 1996. **271**(5): p. 2762-8.
8. Kent, T.A., et al., *Mossbauer studies of aconitase. Substrate and inhibitor binding, reaction intermediates, and hyperfine interactions of reduced 3Fe and 4Fe clusters.* J Biol Chem, 1985. **260**(11): p. 6871-81.
9. Cammack, R., *Iron-sulphur cluster interconversions and the activation of aconitase.* Nature, 1982. **298**(5877): p. 792-3.
10. Johnson, D.C., et al., *Structure, function, and formation of biological iron-sulfur clusters.* Annu Rev Biochem, 2005. **74**: p. 247-81.
11. Stephens, P.J., D.R. Jollie, and A. Warshel, *Protein Control of Redox Potentials of Ironminus signSulfur Proteins.* Chem Rev, 1996. **96**(7): p. 2491-2514.
12. Liu, L., et al., *Ultrahigh-resolution structure of high-potential iron-sulfur protein from Thermochromatium tepidum.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 7): p. 1085-91.
13. Dey, A., et al., *Solvent tuning of electrochemical potentials in the active sites of HiPIP versus ferredoxin.* Science, 2007. **318**(5855): p. 1464-8.
14. Link, T.A., et al., *Isolation, characterisation and crystallisation of a water-soluble fragment of the Rieske iron-sulfur protein of bovine heart mitochondrial bc1 complex.* Eur J Biochem, 1996. **237**(1): p. 71-5.
15. Kuo, C.F., et al., *Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III.* Science, 1992. **258**(5081): p. 434-40.
16. Porello, S.L., M.J. Cannon, and S.S. David, *A substrate recognition role for the [4Fe-4S]2+ cluster of the DNA repair glycosylase MutY.* Biochemistry, 1998. **37**(18): p. 6465-75.
17. Grandoni, J.A., et al., *Evidence that the iron-sulfur cluster of Bacillus subtilis glutamine phosphoribosylpyrophosphate amidotransferase determines stability of the enzyme to degradation in vivo.* J Biol Chem, 1989. **264**(11): p. 6058-64.
18. Russell, M.J. and W. Martin, *The rocky roots of the acetyl-CoA pathway.* Trends Biochem Sci, 2004. **29**(7): p. 358-63.
19. Huber, C. and G. Wachtershauser, *Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions.* Science, 1997. **276**(5310): p. 245-7.

20. Dyall, S.D., M.T. Brown, and P.J. Johnson, *Ancient invasions: from endosymbionts to organelles.* Science, 2004. **304**(5668): p. 253-7.

21. Schilke, B., et al., *Evidence for a conserved system for iron metabolism in the mitochondria of Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1999. **96**(18): p. 10206-11.

22. Takahashi, Y. and U. Tokumoto, *A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids.* J Biol Chem, 2002. **277**(32): p. 28380-3.

23. Balk, J. and S. Lobreaux, *Biogenesis of iron-sulfur proteins in plants.* Trends Plant Sci, 2005. **10**(7): p. 324-31.

24. McFadden, G.I., *Endosymbiosis and evolution of the plant cell.* Curr Opin Plant Biol, 1999. **2**(6): p. 513-9.

25. Kispal, G., et al., *The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins.* Embo J, 1999. **18**(14): p. 3981-9.

26. Lill, R. and U. Muhlenhoff, *Iron-sulfur protein biogenesis in eukaryotes: components and mechanisms.* Annu Rev Cell Dev Biol, 2006. **22**: p. 457-86.

27. Zheng, L., et al., *Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis.* Proc Natl Acad Sci U S A, 1993. **90**(7): p. 2754-8.

28. Yuvaniyama, P., et al., *NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein.* Proc Natl Acad Sci U S A, 2000. **97**(2): p. 599-604.

29. Takahashi, Y. and M. Nakamura, *Functional assignment of the ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster involved in the assembly of Fe-S clusters in Escherichia coli.* J Biochem, 1999. **126**(5): p. 917-26.

30. Fontecave, M. and S. Ollagnier-de-Choudens, *Iron-sulfur cluster biosynthesis in bacteria: Mechanisms of cluster assembly and transfer.* Arch Biochem Biophys, 2008. **474**(2): p. 226-37.

31. Wollenberg, M., et al., *A dimer of the FeS cluster biosynthesis protein IscA from cyanobacteria binds a [2Fe2S] cluster between two protomers and transfers it to [2Fe2S] and [4Fe4S] apo proteins.* Eur J Biochem, 2003. **270**(8): p. 1662-71.

32. Feng, Y., et al., *Structural insight into poplar glutaredoxin C1 with a bridging iron-sulfur cluster at the active site.* Biochemistry, 2006. **45**(26): p. 7998-8008.

33. Chothia, C. and A.M. Lesk, *The relation between the divergence of sequence and structure in proteins.* Embo J, 1986. **5**(4): p. 823-6.

34. Orengo, C.A., et al., *Review: what can structural classifications reveal about protein evolution?* J Struct Biol, 2001. **134**(2-3): p. 145-65.

35. Hussain, F. and P. Wittung-Stafshede, *Impact of cofactor on stability of bacterial (CopZ) and human (Atox1) copper chaperones.* Biochim Biophys Acta, 2007. **1774**(10): p. 1316-22.

36. Eck, R.V. and M.O. Dayhoff, *Evolution of the Structure of Ferredoxin Based on Living Relics of Primitive Amino Acid Sequences.* Science, 1966. **152**(3720): p. 363-366.

37. Moulis, J.M., et al., *Crystal structure of the 2[4Fe-4S] ferredoxin from Chromatium vinosum: evolutionary and mechanistic inferences for [3/4Fe-4S] ferredoxins.* Protein Sci, 1996. **5**(9): p. 1765-75.

38. Meyer, J., *Iron-sulfur protein folds, iron-sulfur chemistry, and evolution.* J Biol Inorg Chem, 2008. **13**(2): p. 157-70.

39. Orengo, C.A. and J.M. Thornton, *Protein families and their evolution-a structural perspective.* Annu Rev Biochem, 2005. **74**: p. 867-900.

40. Burroughs, A.M., et al., *Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold.* Biol Direct, 2007. **2**: p. 18.

41. Frolow, F., et al., *Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe-2S ferredoxin.* Nat Struct Biol, 1996. **3**(5): p. 452-8.

42. Kakuta, Y., et al., *Crystal structure of Escherichia coli Fdx, an adrenodoxin-type ferredoxin involved in the assembly of iron-sulfur clusters.* Biochemistry, 2001. **40**(37): p. 11007-12.

43. Grinberg, A.V., et al., *Adrenodoxin: structure, stability, and electron transfer properties.* Proteins, 2000. **40**(4): p. 590-612.

44. Dauter, Z., et al., *Zinc- and iron-rubredoxins from Clostridium pasteurianum at atomic resolution: a high-precision model of a ZnS4 coordination unit in a protein.* Proc Natl Acad Sci U S A, 1996. **93**(17): p. 8836-40.

45. Day, M.W., et al., *X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium Pyrococcus furiosus.* Protein Sci, 1992. **1**(11): p. 1494-507.

46. Gomes, C.M., et al., *A novel type of nitric-oxide reductase. Escherichia coli flavorubredoxin.* J Biol Chem, 2002. **277**(28): p. 25273-6.

47. Archer, M., et al., *Crystal structure of desulforedoxin from Desulfovibrio gigas determined at 1.8 A resolution: a novel non-heme iron protein structure.* J Mol Biol, 1995. **251**(5): p. 690-702.

48. Iwata, S., et al., *Structure of a water soluble fragment of the 'Rieske' iron-sulfur protein of the bovine heart mitochondrial cytochrome bc1 complex determined by MAD phasing at 1.5 A resolution.* Structure, 1996. **4**(5): p. 567-79.

49. Meyer, J., et al., *Assembly of a [2Fe-2S]2+ cluster in a molecular variant of Clostridium pasteurianum rubredoxin.* Biochemistry, 1997. **36**(43): p. 13374-80.

50. Iwasaki, T., et al., *Rational design of a mononuclear metal site into the archaeal Rieske-type protein scaffold.* J Biol Chem, 2005. **280**(10): p. 9129-34.

51. Anderson, G.L. and J.B. Howard, *Reactions with the oxidized iron protein of Azotobacter vinelandii nitrogenase: formation of a 2Fe center.* Biochemistry, 1984. **23**(10): p. 2118-22.

52. Sen, S., et al., *A conformational mimic of the MgATP-bound "on state" of the nitrogenase iron protein.* Biochemistry, 2004. **43**(7): p. 1787-97.

53. Kiley, P.J. and H. Beinert, *The role of Fe-S proteins in sensing and regulation in bacteria.* Curr Opin Microbiol, 2003. **6**(2): p. 181-5.

54. Klausner, R.D., T.A. Rouault, and J.B. Harford, *Regulating the fate of mRNA: the control of cellular iron metabolism.* Cell, 1993. **72**(1): p. 19-28.

55. Daas, P.J., et al., *Characterization and determination of the redox properties of the 2[4Fe-4S] ferredoxin from Methanosarcina barkeri strain MS.* FEBS Lett, 1994. **356**(2-3): p. 342-4.

56. Gomes, C.M., Faria, A., Carita, J.C., Mendes, J., Regalla ,M., Chicau ,P., Huber ,H.,Stetter ,K.O. and Teixeira ,M., *Di-cluster, seven iron ferredoxins from hyperthermophilic Sulfolobales.* JBIC, 1998. **3**: p. 499-507.

57. Hannan, J.P., et al., *Characterisation of oxidised 7Fe dicluster ferredoxins with NMR spectroscopy.* J Biol Inorg Chem, 2000. **5**(4): p. 432-47.

58. Pace, N.R., *A molecular view of microbial diversity and the biosphere.* Science, 1997. **276**(5313): p. 734-40.

59. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4576-9.

60. Janssen, S., et al., *Ferredoxins from the archaeon Acidianus ambivalens: overexpression and characterization of the non-zinc-containing ferredoxin FdB.* Biol Chem, 2001. **382**(10): p. 1501-7.

61. Kojoh, K., et al., *Zinc-coordination of aspartic acid-76 in Sulfolobus ferredoxin is not required for thermal stability of the molecule.* J Inorg Biochem, 2002. **89**(1-2): p. 69-73.

62. Kojoh, K., H. Matsuzawa, and T. Wakagi, *Zinc and an N-terminal extra stretch of the ferredoxin from a thermoacidophilic archaeon stabilize the molecule at high temperature.* Eur J Biochem, 1999. **264**(1): p. 85-91.

63. Breton, J.L., et al., *Identification of the iron-sulfur clusters in a ferredoxin from the archaeon Sulfolobus acidocaldarius. Evidence for a reduced [3Fe-4S] cluster with pH-dependent electronic properties.* Eur J Biochem, 1995. **233**(3): p. 937-46.

64. Teixeira, M., et al., *A seven-iron ferredoxin from the thermoacidophilic archaeon Desulfurolobus ambivalens.* Eur J Biochem, 1995. **227**(1-2): p. 322-7.

65. Fujii, T., et al., *The crystal structure of zinc-containing ferredoxin from the thermoacidophilic archaeon Sulfolobus sp. strain 7.* Biochemistry, 1997. **36**(6): p. 1505-13.

66. Moczygemba, C., et al., *High stability of a ferredoxin from the hyperthermophilic archaeon A. ambivalens: involvement of electrostatic interactions and cofactors.* Protein Sci, 2001. **10**(8): p. 1539-48.

67. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.

68. Griffin, S., et al., *High thermal and chemical stability of Thermus thermophilus seven-iron ferredoxin. Linear clusters form at high pH on polypeptide unfolding.* Eur J Biochem, 2003. **270**(23): p. 4736-43.

69. Frazao, C., et al., *Crystallographic analysis of the intact metal centres [3Fe-4S](1+/0) and [4Fe-4S](2+/1+) in a Zn(2+) -containing ferredoxin.* FEBS Lett, 2008. **582**(5): p. 763-7.

70. Kennedy, M.C., et al., *Evidence for the formation of a linear [3Fe-4S] cluster in partially unfolded aconitase.* J Biol Chem, 1984. **259**(23): p. 14463-71.

71. Jones, K., et al., *Formation of a linear [3Fe-4S] cluster in a seven-iron ferredoxin triggered by polypeptide unfolding.* J Biol Inorg Chem, 2002. **7**(4-5): p. 357-62.

72. Pereira, M.M., et al., *A ferredoxin from the thermohalophilic bacterium Rhodothermus marinus.* Biochim Biophys Acta, 2002. **1601**(1): p. 1-8.

2.4 ACKNOWLEDGMENTS

To Carlos Frazão, Maria Arménia Carrondo, David Aragão, and Ricardo Coelho from ITQB- Portugal for the crystallographic structure of the di-cluster ferredoxin from *Acidianus ambivalens.*

To Miguel Teixeira for the soluble extract of *Acidianus ambivalens*

To B.O. Kolbesen from Johann Wolfgang Goethe- Universität, Frankfurt, Germany for metal analysis.

3 STUDIES ON THE DEGRADATION PATHWAY OF IRON-SULFUR CENTRES DURING UNFOLDING OF A DI-CLUSTER FERREDOXIN: CLUSTER DISSOCIATION, IRON RELEASE AND PROTEIN STABILITY

CONTENTS

This chapter was published in

Sónia S. Leal, Miguel Teixeira, Cláudio M. Gomes
Studies on the degradation pathway of iron-sulfur centres during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability
J. Biol. Inorg. Chem 9: 987-996 (2004)

3.1 SUMMARY

Previous studies have shown that the archaeal family of di-cluster [3Fe-4S][4Fe-4S] ferredoxins are intrinsically very stable proteins and led to the suggestion that upon protein unfolding the iron-sulfur clusters degraded via linear three-iron sulfur centre species, with 610 and 520nm absorption bands, resembling those observed in purple aconitase. In this work we proceeded with a detailed kinetic and spectroscopic investigation on the alkaline chemical denaturation of the protein, in an attempt to elucidate the degradation pathway of the iron-sulfur centres in respect to protein unfolding events. For this purpose we investigated cluster dissociation, iron release and protein unfolding by complementary biophysical techniques. We found that shortly after initial protein unfolding, iron release proceeds monophasically at a rate comparable to that of cluster degradation, and that during the process no typical EPR features of linear three-iron sulfur centres are observed. Further, it was observed that EDTA prevents formation of the intermediate and that sulfide significantly enhances its intensity and lifetime, even after protein unfolding. Altogether, our data suggests that iron sulfides, which are formed from the release of iron and sulfide resulting from cluster degradation during protein unfolding in alkaline conditions, are in fact responsible for the observed intermediate spectral species, thus disproving the hypothesis suggesting the presence of a linear three-iron centre intermediate. Kinetic studies monitored by visible, fluorescence and UV second derivative spectroscopies have elicited that upon initial perturbation of the tertiary structure the iron-sulfur centres start decomposing and that the presence of EDTA accelerates the process. Also, the presence of EDTA lowers the observed melting temperature in thermal ramp experiments and the midpoint denaturant concentration in equilibrium chemical unfolding experiments, further suggesting that the clusters also play a structural role in the maintenance of the conformation of the folded state.

3.2 INTRODUCTION

Addressing the stability properties of proteins with bound prosthetic groups may pose particular difficulties. If the cofactor is covalently bound it is likely that the unfolding peptide is able to refold spontaneously to its native state; in fact, there are several available studies in which the working models are small monomeric proteins which contain a prosthetic group with a covalent linkage, such as cytochrome *c*, ribonuclease or lysozyme. On the other hand, conformational stability studies on proteins containing none covalently bound cofactors may be much more difficult to interpret as these proteins typically unfold irreversibly thus impairing a complete thermodynamic characterization of the unfolding-refolding equilibrium. The reasons for this irreversibility are likely to arise from the fact that upon lost of tertiary structure, cofactor lost in the form of metal release, cluster disintegration, or prosthetic group release such as a flavin, leads to a loss in a local structural element necessary for protein folding to proceed. Nevertheless, in recognition of the important role of metal containing proteins in living systems, which make up around 30% of the proteins in the cell [1], a considerable effort has been put in addressing the determinants of stability and folding properties of proteins containing metal cofactors. These studies include work on several different metalloproteins, such as azurin [2, 3], plastocyanin [4], ferredoxins [5], HiPIP [6, 7], cytochrome c_6 [8], desulfoferrodoxin [9] and adrenodoxin [10, 11].

In some cases of proteins containing non-covalent metal centres and in particular for proteins holding iron-sulfur clusters a detailed interpretation of the cluster assembly mechanism is not yet available. One approach towards the understanding of cluster assembly in iron sulfur proteins consists on the characterization of its disintegration during protein unfolding, which may elicit the involvement of intermediate cluster structures or highlight the role of particular structural elements in cluster stabilization. In this respect, the family of seven iron di-cluster ferredoxins from thermoacidophilic archaea constitute excellent working models, as in addition its study contributes towards the elucidation of stability determinants in thermophilic proteins, and also on the role of metal centres. Interestingly it was suggested in the literature that within this family, protein

unfolding results in a rearrangement of the iron sulfur clusters into a linear three iron sulfur centre intermediate [12], a feature which has been later generalized to other iron-sulfur proteins [13, 14], even those containing centres with lower nuclearities [15, 16]. In this chapter we characterize in detail the protein and metal centre unfolding kinetics on the seven iron ferredoxin from *Acidianus ambivalens*, using complementary spectroscopic methodologies, in order to establish the protein degradation pathway and in particular the fate of the metal clusters during the process. Our findings do not support the hypothesis of the formation of a linear three-iron centre, and rather suggest that the transient species observed are due to formation of iron sulfides resulting from cluster degradation. A combination of EPR and visible spectroscopies and metal quantification allows establishing that clusters degrade at identical rates and that cluster degradation and monophasic metal release are concomitant events, which occur shortly after protein unfolding.

3.3 MATERIALS AND METHODS

Chemicals

The chemical denaturant guanidine hydrochloride (GuHCl) was purchased from Promega. Bathophenanthroline-disulfonic acid (BPS), sodium sulfide (Na_2S) and iron(III) nitrate ($Fe(NO_3)_3$) were obtained from Sigma and ethylenedinitrilotetraacetic acid disodium salt (EDTA) from Merck. All reagents were of the highest purity grade commercially available.

Cell growth and protein purification

Acidianus ambivalens cells were grown in a 10L fermentor as previously described [17]. The ferredoxin (Fd) was purified according to published procedures [18]. Protein purity was confirmed by observation of a single band on 15% SDS/Page and a clean N-terminal sequence.

Spectroscopic methods

UV and visible absorption measurements were performed on a Shimadzu

MultiSpec-1501 diode array spectrophotometer equipped with cell stirring and temperature control by a computer-controlled water bath (Julabo). The integrity of iron-sulfur clusters was monitored following the decay at 410-610nm. The degree of exposure of tyrosyl residues was monitored by second-derivative UV spectroscopy monitoring the peak to peak distance (287-283nm) [19, 20]. Fluorescence spectroscopy was made on a Cary Varian Eclipse spectrofluorimeter, equipped with cell stirring and peltier temperature control. Tryptophan emission was monitored between 300-450 nm, with a maximum at 355nm, after excitation at 292 nm. EPR spectra were recorded as in [17] and for the time course experiments the EPR sample was thawed and incubated for a given period on a water bath prior to re-freezing in liquid N_2.

Iron quantification

The quantification of the released iron during unfolding of ferredoxin was achieved using BPS [21]. Briefly, the time course of iron release was obtained by collecting samples at various times during the ferredoxin alkaline unfolding reaction, and incubating them between 5-30 minutes with 1mM BPS and 20mM DTT in 200mM HEPES pH 7. After incubation, the absorbance which was independent of the incubation time, was measured at 534nm. Iron was quantified both from the extinction coefficient for the Fe^{2+}-BPS complex (ε^{534nm}= 22140 $M^{-1}.cm^{-1}$), and from calibration curves made with accurately titrated $Fe(NO_3)_3$ solutions (a kind gift from Rita Delgado, ITQB/UNL).

Kinetic experiments

Time resolved experiments were performed under vigorous stirring with a magnetic bar and specially designed cuvettes. The reaction mixture was allowed to stabilize at 25°C and the reaction was started with the addition of a concentrated ferredoxin solution. The reaction was allowed to proceed till completion and the obtained traces were fitted to a first order exponential decay (Origin 5.0).

Denaturant equilibrium unfolding

The chemical denaturant GuHCl was used to promote protein unfolding at 25ºC and pH 10. A stock solution of GuHCl at pH 10 was prepared in 200 mM glycine buffer and its accurate concentration was confirmed by refractive index measurements [22]. A two state model was used to analyse the unfolding transitions and data fitting (Origin 5.0) to a sigmoidal process allowed for the determination of midpoint denaturant concentrations ($[GuHCl]_{1/2}$). Unless otherwise stated, the ferredoxin concentration used in all assays was 10 μM.

Thermal denaturation

Experiments were carried out using a water bath (Julabo 4 MV) controlled by the freeware application for digital control allowing temperature ramping experiments. Heating proceeded at a scan rate of $1ºC.min^{-1}$ from 25 to 90º C. A two state model was used to analyse the unfolding transitions, and data fitting (Origin 5.0) to a sigmoidal process allowed the determination of the melting temperature (T_m).

3.4 RESULTS

Ferredoxin alkaline chemical unfolding

The ferredoxin from *A. ambivalens* is a remarkably stable protein which remains folded even in the presence of 8M GuHCl at pH 7. However, at extreme pH values, the protein is partly destabilized and can be chemically unfolded, an observation which is in line with the proposal that electrostatic interactions contribute significantly to its stability. The observed events monitored by visible spectroscopy for the *A. ambivalens* Fd upon alkaline chemical denaturation (7M GuHCl, pH 10) are identical to those reported for the

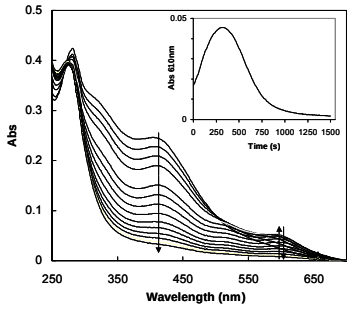


Figure 3.1 |Ferredoxin alkaline unfolding kinetics.
Spectral variations observed upon inducing ferredoxin chemical alkaline unfolding (7M GuHCl, 0.2M glycine pH 10, 10 μM protein). Inset: kinetics of the 610nm transient feature.

S. acidocaldarius Fd: while the 410nm band corresponding to the iron-sulfur clusters signature decreases in intensity throughout the process, two new species centred at 520 and 610nm are transiently formed (figure 3.1). It has been hypothesised that this feature could be attributable to a linear three-iron species identical to the one observed in purple aconitase, on the basis of the similarity between the visible spectra [12].

|Kinetics of cluster dissociation, iron release and protein unfolding

In order to fully elucidate the cluster degradation pathway in seven iron ferredoxins, and the relative contribution of the iron-sulfur centres to the formation of this putative intermediate, a detailed characterisation of the unfolding events in these conditions was undertaken, monitoring simultaneously iron release, protein unfolding and iron-sulfur centre degradation (figure 3.2, and also table 3.1). Different spectroscopic techniques were used: visible, fluorescence and UV second derivative spectroscopies. Whereas visible spectroscopy provides direct evidence for the degradation of the iron-sulfur clusters, the two latter methodologies are indicators of loss of protein structure: fluorescence spectroscopy reports modifications that are local to tryptophan residues [23] and in this system also indirectly iron-sulfur cluster integrity (see below); and UV second derivative spectroscopy allows to monitor subtle alterations in the microenvironments of tyrosine residues resulting in an increase of its exposure to a polar medium [20]. The disassembly of the metal clusters in 7M GuHCl and pH 10 was evaluated by monitoring the absorbance

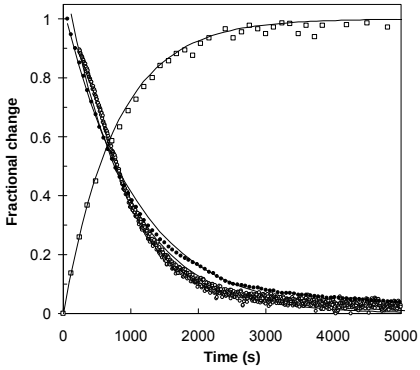


Figure 3.2 |Kinetics of iron release, protein unfolding and iron sulfur cluster degradation. Kinetics was determined upon inducing ferredoxin chemical alkaline unfolding (7M GuHCl, 0.2M glycine pH 10, 10 μM protein). Closed circles: iron-sulfur degradation determined from $Abs_{410\text{-}610nm}$ measurements. Open circles: Time course of fluorescence monitored protein unfolding. Open squares: Iron release determined by discontinuously sampling the reaction medium; the maximum of the released iron amplitude corresponds to 5.5±0.5 mol Fe/mol protein. The solid lines are mono-exponential fits (values reported in the text)

decay at 410 *minus* 610nm, which allows removing the contribution of the transient intermediate species. The kinetics of the iron-sulfur centres degradation was determined to be monoexponential having an observed rate (k_U^{FeS}= 16.9±0.3x10^{-4} s^{-1}) which is slightly higher than that determined by fluorescence spectroscopy (k_U^{fluo}=9.7±0.1x10^{-4} s^{-1}), but three-fold lower than the rate determined by second derivative UV analysis (k_U^{2UV}=50.4±3.9x10^{-4} s^{-1}). These differences in protein unfolding speed can be accounted by the fact that in this ferredoxin, fluorescence spectroscopy is considerably sensitive to the integrity of the iron-sulfur clusters, a situation that finds a parallel in other iron-sulfur proteins [24]. In fact, the close proximity between the single tryptophan and the clusters (about 4 Å as deduced from a molecular model of the protein [18]) and the superimposition of the tryptophan emission spectrum with the absorption spectrum of the iron sulfur centres greatly favours radiation less energy transfer and affects overall fluorescence [24]. Also, most of the three tyrosines that are probed by second derivative UV analysis are superficial residues and are thus more prone to quickly sense structural alterations.

Table 3.1 |Observed rate constants for the GuHCl promoted degradation of iron-sulfur clusters and protein unfolding in alkaline conditions.

	Observed rate (10^{-4} s^{-1})					
[GuHCl] (M)	**5**	**5.5**	**6**	**6.5**	**7**	**7.5**
k_U^{vis}	nd.	6.5	6.7	11.9	16.9	39.5
$k_U^{vis,EDTA}$	2.84	5.7	10.3	17.1	38.4	57.7
k_U^{UV2D}	3.66	18.0±4.2	15.7±1.4	nd.	50.4±3.9	nd.
$k_U^{UV2D,EDTA}$	8.11±1	15.0±2.7	16.2±1.3	30.1±2.3	56.1±7.3	nd.
k_U^{fluo}	nd.	nd.	3.64	7.86	9.68	13.2
$k_U^{fluo,EDTA}$	nd.	nd.	5.22	10.8	17.5±0.14	26.6

Experimental conditions: 0.2 M glycine pH 10, 10 µM Fd, 25°C.
nd. Not determined.
Legend: Vis, visible spectroscopy; UV2D, second derivative UV spectroscopy; Fluo, fluorescence spectroscopy; EDTA, assayed in the presence of 2.5 mM EDTA.
Only deviations above 0.5 10^{-4} s^{-1} are shown in the table.

The disintegration of the iron-sulfur centres was also evaluated by an independent method which consisted on quantifying discontinuously the amount of iron released into the solution, by sampling the reaction mixture at given time

points, and mixing with the iron chelator bathophenantroline (BPS). The time course of iron release was found to be monophasic, with an observed rate (k_U^{Fe}) of $12x10^{-4}$ s^{-1}, which is identical to that determined by visible and fluorescence spectroscopy. Thus, these results show that iron-sulfur cluster dissociation and iron release are concomitant events shortly preceded by protein unfolding. This data does not corroborate the hypothesis that an intermediate form of iron-sulfur centre is present. If this would be the case, a biphasic release of iron would be expected, with time courses agreeing with those observed for the appearance and disappearance of the 610nm transient band (see inset figure 3.1).

EPR analysis of iron-sulfur cluster degradation

EPR spectroscopy was utilized to obtain a more detailed insight into the transformations that occur on the iron-sulfur clusters during ferredoxin unfolding. Incubation of the oxidized protein in 6.7 M GuHCl at pH 10, results in modifications of the [3Fe-4S] environment which lead to the disappearance of more than 80% of its typical *g* 2.03 signal during the first 2 min after mixing. At an incubation time corresponding to the maximum of the 610nm absorbance (~10-15 min.), the EPR spectrum exhibits a modest band at *g* 4.3 and is featureless on the low field region (data not shown). Anaerobic incubation with sodium dithionite, does not lead to the appearance of any additional features. Further, in both conditions, the samples are silent in parallel mode EPR. In overall, these results show that in the conditions tested: i. early during unfolding the [3Fe-4S] centre is perturbed; ii. the [4Fe-4S] centre is already degraded when the 610nm band reaches it maximum; and iii. the characteristic features of a linear three-iron centre, namely a sharp resonance at *g* 4.3 and a broad low field signal as reported for purple aconitase [25], are not observed.

Cluster degradation and polypeptide unfolding

In order to evaluate the cluster degradation pathway in respect to protein unfolding, the unfolding kinetics in alkaline conditions was evaluated as a function of denaturant concentration, by visible, fluorescence and UV second derivative spectroscopy. The chevron plot of the obtained results (figure 3.3 and table 3.1)

allows determining the observed ferredoxin unfolding rate constants in water, at 25°C and pH 10, upon extrapolation to a zero concentration of GuHCl.

The value obtained for cluster degradation (k_U^{vis} (H_2O) = $6x10^{-7}$ s^{-1}) is lower in respect to that determined for protein unfolding by fluorescence spectroscopy (k_U^{fluo} (H_2O) = $33x10^{-7}s^{-1}$) and by UV second derivative analysis (k_U^{UV2D} (H_2O) = $52x10^{-7}$ s^{-1}). These differences can be rationalised considering a molecular model of the *A. ambivalens* ferredoxin: the single protein tryptophan is in a relatively superficial region near the iron-sulfur centres containing region, as well as the three tyrosine residues, two of which are located in the N-terminal extension region [18, 26]. Thus, the obtained results may reflect the different stability of distinct structural regions within the protein and may indicate that the N-terminal region is more sensitive to chemical denaturation and more readily perturbed than the protein core. As protein structure is lost, the cubic metal clusters degrade, indicating that they are particularly sensitive to subtle modifications of the supporting polypeptide chain agrees with their proposed role as important determinants on the structure of the folded state. This finding is in line with the reported findings for spinach ferredoxin [27] and other metalloproteins according

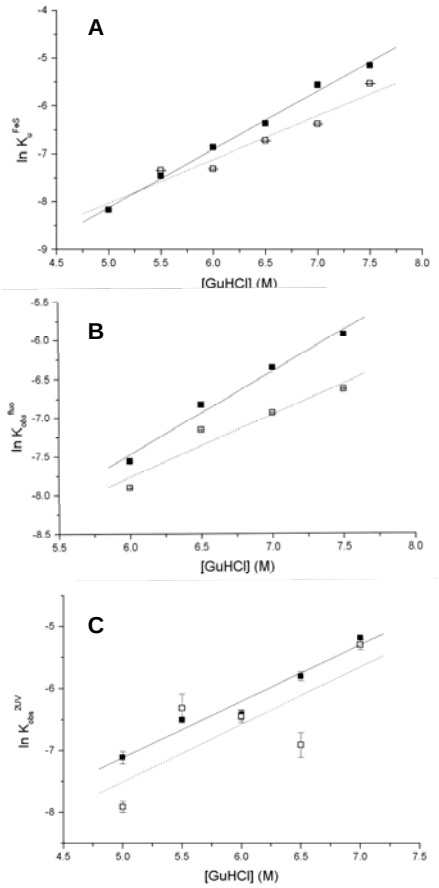


Figure 3.3 |Chevron plot for ferredoxin alkaline chemical denaturation. The natural logarithm of observed rates is represented as a function of denaturant concentration. Open squares, minus EDTA, closed squares, plus 2.5mM EDTA. Panel **A**, iron-sulfur degradation rates as monitored by $Abs_{410\text{-}610nm}$. Panel **B**, protein unfolding rates as monitored by fluorescence spectroscopy. Panel **C**, tyrosine exposure rates as monitored by second derivative UV spectroscopy. Experimental conditions: 0.2M glycine pH 10, 10 µM Fd, 7M GuHCl, 25°C.

to which cofactors stabilize the native form and are frequently required for the polypeptide to refold [28].

The effect of EDTA on metal centres and protein stability

Protein stability, metal centres integrity and iron release were investigated in comparative experiments performed in the presence of 7M GuHCl at pH 10 and 2.5 mM EDTA. Whereas protein unfolding remains essentially undisturbed by the presence of the chelator as inferred from UV second derivative analysis, cluster dissociation can go until two-fold faster in the presence of EDTA for the highest concentrations of denaturant (table 3.1, figure 3.3). Accordingly, iron release induced by chemical alkaline protein unfolding is also faster ($k_U^{Fe,\ plus\ EDTA}$ = 19x10^{-4} s^{-1} *versus* $k_U^{Fe,\ minus\ EDTA}$ = 12x10^{-4} s^{-1}). In the presence of EDTA, a lower amount of denaturant is also required to degrade the iron-sulfur clusters, as determined from visible spectroscopy monitored titration curves, performed at pH 10 and 20°C (table 3.2, figure 3.4). As ferredoxin unfolding is irreversible, most likely because the metal clusters are unable to be reformed from the unfolded polypeptide chain, the estimation of thermodynamic parameters is meaningless. However, the relative stability of the protein can be inferred from evaluating the denaturant midpoints ([GuHCl]$_{½}$) as a function of incubation time, which are expected to decrease as the incubation time is increased. From this data, it can be concluded that the apparent cluster stability is always higher in the absence of the chelator.

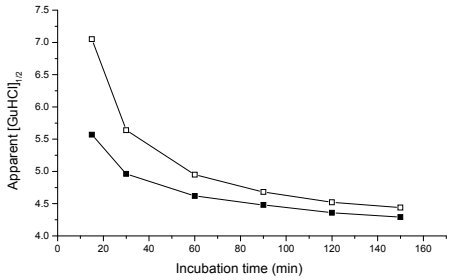


Figure 3.4 Time dependence of GuHCl induced midpoints.
Open squares, minus EDTA; Closed squares, plus 2.5mM EDTA. Midpoint denaturant concentrations were determined from titration experiments monitored at $Abs_{410\text{-}610nm}$, in which the obtained data was fitted according to a sigmoid curve, according to a two-state model. Experimental conditions: 0.2M glycine pH 10, 10 µM Fd, 25°C.

Table 3.2 | Observed midpoint denaturant concentrations as a function of incubation time.

Incubation time (min)		15	30	60	90	120	150
$[GuHCl]_{1/2}$	**0 EDTA**	7.05	5.64	4.95	4.68	4.52	4.44
	+2.5mM EDTA	5.57	4.96	4.62	4.48	4.36	4.29

Experimental conditions: 0.2 M glycine pH 10, 10 µM Fd, 25°C.

The effect of EDTA was also investigated by temperature ramp experiments which allow to gain insights on the relative thermal stability of the protein, from the measured melting temperatures (T_m). These experiments were performed in the presence of 3M GuHCl which is required to partly destabilize the protein thus allowing its thermal denaturation which otherwise would not occur [5]. The results obtained show that the difference between the observed melting temperatures in the absence (T_m^{-EDTA} =70.5°C) and in the presence of EDTA (T_m^{+EDTA}=66.1°C), at a heating rate of 1°C.min^{-1}, is 4.4°C (figure 3.5). In other experimental conditions such as different pH, denaturant concentrations and heating rates the effect of EDTA is even more pronounced (our own unpublished observations). Taken together, these results show that upon triggering denaturation and initial cluster perturbation, as the protein adopts a more open conformation and the iron-sulfur clusters and zinc site become more exposed, the presence of EDTA further promotes the degradation of the metal sites. However, with the present model system, it is not possible to estimate if the effect of EDTA is more significant over the iron-sulfur clusters or the zinc centre. This type of information can nevertheless be obtained by studying ferredoxin isoforms such as those characterized in *Sulfolobus metallicus* [18], one of which lack the ligands of the zinc site.

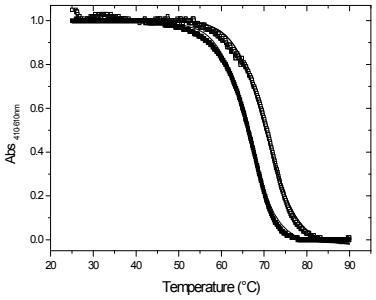


Figure 3.5 | Thermal degradation of iron-sulfur clusters.
Open squares, minus EDTA; Closed squares, plus 2.5mM EDTA. The melting temperatures were determined from temperature ramping experiments monitored at $Abs_{410\text{-}610nm}$. Experimental conditions: 0.2M glycine pH 10, 10 µM Fd, temperature ramping range: 30-90°C, heating rate: 1°C.min^{-1}

Chemical nature of the 610nm transient species

The chemical nature of the 610nm transient band was investigated by monitoring the unfolding kinetics in the presence of EDTA and sulfide (figure 3.6). The rational for these experiments was that by disturbing the degrading iron-sulfur centres with the addition of the iron chelator EDTA, or by adding excess sulfide which is also being released during cluster degradation, the formation of the species responsible for the 610nm absorption could be controlled and ideally its lifetime could be extended in order to make its study more amenable.

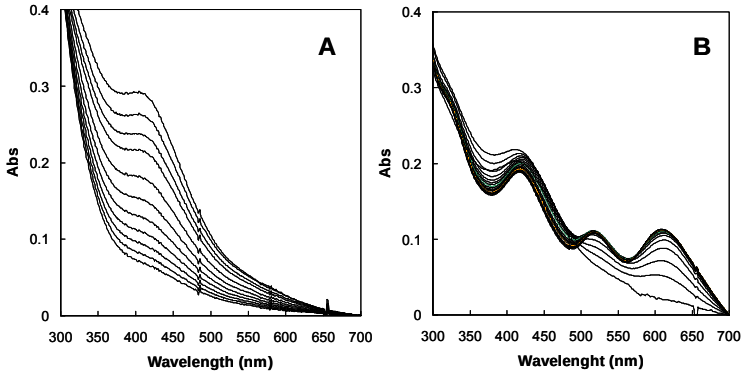


Figure 3.6 Effect of EDTA and sulfide on ferredoxin alkaline unfolding kinetics. Visible Spectrum variations observed upon inducing ferredoxin chemical alkaline unfolding (7M GuHCl, 0.2M glycine pH 10, 10 μM protein) in the presence of 2.5mM EDTA (panel **A**) and in the presence of 10mM sodium sulfide (panel **B**).

Surprisingly, there is no formation of the 520 or 610 nm species, nor of any other transient features. This indicates that the formation of the intermediate species involves utilization of the released iron, which under these experimental conditions is scavenged by EDTA. In what concerns the original hypothesis that this transient species is due to a linear three-iron centre, the present observation is in clear contrast to what is observed in aconitase. In this case, the presence of an iron chelating agent does not prevent the formation of the purple form at high pH nor affects the conversion reaction rate [25]. On the other hand, ferredoxin alkaline chemical unfolding in the presence of excess sulfide (7M GuHCl, pH 10, 10mM S^{2-}) results in the formation of the 610 and 520nm species, which however remain populated for a very long time scale (>60 min). In order to determine if an

intermediate form was stabilized under these conditions, the unfolding kinetics were investigated by fluorescence spectroscopy, thus monitoring directly the conformation of the polypeptide chain (figure 3.7). The polypeptide chain was found to be completely unfolded even before the maximal intensity of the 610nm band was achieved. However, the build up of the 610nm band was found to perfectly correspond to iron release measured independently and in a discontinuous fashion as described above. From this data it can be inferred that formation of the 610nm transient species is not related to any ferredoxin unfolding intermediate form, and that the 610nm feature can be directly correlated with iron release from the degrading clusters. This assumption is further corroborated by the observation that adding sulfide to unfolded chemically denatured ferredoxin results in the re-formation of the 520 and 610nm species, but does not lead to protein refolding, as inferred from fluorescence spectroscopy (data not shown). Altogether, these results show that the 520 and 610nm species result from the interaction between released iron and sulfide in alkaline conditions.

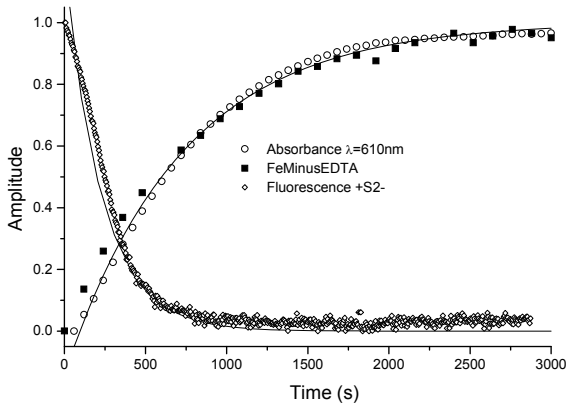


Figure 3.7 | Kinetics of iron release, 610nm band increase and protein unfolding. Kinetic traces were determined upon inducing ferredoxin chemical alkaline unfolding (7M GuHCl, 0.2M glycine pH 10, 10 µM protein). Open diamonds: Time course of fluorescence monitored protein unfolding in the presence of 10mM sodium sulfide. Closed squares: Iron release determined by discontinuously sampling the reaction medium; the maximum of the released iron amplitude corresponds to 5.5±0.5 mol Fe/mol protein. Open circles: amplitude variation of the 610nm band for an assay in the presence of 10mM sodium sulfide. The solid lines are mono-exponential fits (values reported in the text)

Formation of iron sulfides (Fe_xS_y) during ferredoxin unfolding

In overall, the reported data is compatible with formation of iron sulfides (Fe_xS_y) as a result of cluster degradation during protein unfolding in alkaline conditions, and do not necessarily involve a linear three-iron centre intermediate form. The transient species at 520 and 610 nm which are observed upon ferredoxin unfolding, and are associated to its gray color (figure 3.8, trace b), are identical to those obtained by mixing excess sulfide and ferrous iron at pH 10 (figure 3.8, trace d), and are enhanced when exogenous sulfide is added during protein unfolding (figure 3.8, trace c). On the other hand, the presence of a linear three-iron centre is further contradicted by comparing the ferredoxin spectra during unfolding with that of purple aconitase (figure 3.8, trace a), a protein form in which this type of cluster was demonstrated to be present [25]. In fact, during ferredoxin unfolding the iron-sulfur centres are degraded resulting in the release of iron ($Fe^{II/III}$) and sulfide (S^{2-}) to the solution. Metallic elements can readily react with sulfur, resulting in the formation of a great variety of binary compounds with a considerable complexity and variety; in particular, iron sulfides frequently have peculiar stoichiometries [29]. The nature of the formed compounds depends on the ratios of the reactants, temperature as well as other conditions [29]. During ferredoxin unfolding at high pH, both iron sulfides and iron hydroxides are expected to be formed. The fact that the presence of exogenous sulfides enhances the 520nm and 610nm species and the gray color, shows that formation of Fe_xS_y is being favored in relation to that of the hydroxides. When

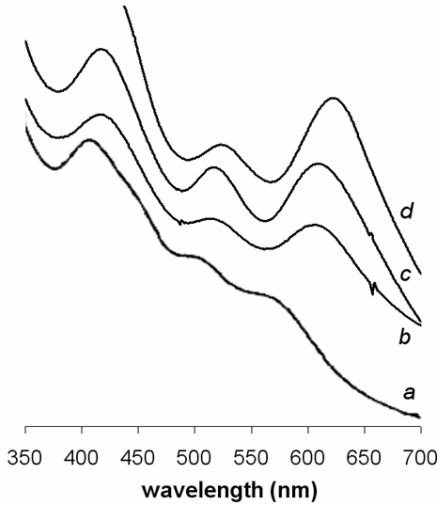


Figure 3.8 Comparison of the visible spectra of purple aconitase, degrading ferredoxin and iron sulfide in alkaline conditions.
Trace **a**, Purple aconitase (linear three-iron centre containing form, in 0.2M glycine pH 10, 6.3 μM protein, replotted from [25]); Trace **b**, Ferredoxin intermediate form (7M GuHCl, pH 10, 10 □M protein, taken at the maximum of the 610 nm band intensity); Trace **c**, same as trace b, but plus 10mM Na_2S; trace d, spectrum resulting from mixing excess sulfide and ferrous iron, at pH 10.

unfolding is performed in the presence of an excess of exogenous ferrous iron, then the intensity of the transient species does not increase but instead a yellowish coloration is observed, compatible with the formation of iron hidroxides which are yellow or light brown [30]. The fact that the transient species, and the gray color, decay after some time is compatible with the known instability of Fe_xS_y compounds. For example Fe_2S_3, which is formed from the direct combination of Fe^{3+} and S^{2-}, is an unstable black precipitate that oxidizes to Fe_2O_3 and S upon exposition to air. At low temperature (0°C or bellow) a black air-sensitive solid is formed. In complete agreement with these facts, when alkaline ferredoxin unfolding is interrupted by rapid freezing the sample at a time corresponding to the maximum intensity of the 520 and 610nm species, a black precipitate is observed which degrades upon thawing. Further evidence for the transient formation of Fe_xS_y compounds is given by the fact that re-addition of exogenous sulfide to an unfolded protein solution, in which there is free available iron, restores the gray color and the previously observed 520 and 610nm species. Also, the fact that in acidic conditions (pH 2.5) these transient spectral features are not observed (our own observations) supports the formation of Fe_xS_y compounds during ferredoxin unfolding: in these circumstances, the formation of hydroxides is greatly disfavored, and sulfides, due to its high solubility at low pH evolve into the formation of hydrogen sulfide (H_2S), which in fact can be scented quite intensely upon ferredoxin acid unfolding. Furthermore, incubation of sulfides and iron in alkaline conditions in the presence of EDTA, does not lead to formation of gray species nor the 610 and 520nm spectral species (our own observations). The fact that gel filtration experiments suggested that the gray transient forms remained bound to unfolded ferredoxin polypeptide [15, 31] can be accounted by unspecific binding of the formed iron sulfides to the polypeptide chain. In fact, in aconitase, attempts to prepare the apoprotein form by incubating the protein with excess of sulfide to remove iron from the cluster led to the observation that the resulting iron sulfides remained bound to the protein on desalting and that the colloidal state of these iron sulfides did prevent its separation from the protein by gel filtration or membrane centrifugation [32].

More generally, formation of iron sulfides is likely to occur during the unfolding of iron sulfur proteins in alkaline conditions. Originally, these transient spectral

features were observed in adrenodoxin, a [2Fe-2S] containing protein, upon its GuHCl mediated unfolding at pH 8.5 in the presence of exogenous sulfide [10]. However, the origins for the observed spectral complexity were not addressed. More recently, there are several reports describing transient 520 and 610nm species during the unfolding of iron-sulfur proteins in basic conditions [12], which were nevertheless assigned to the now disproved hypothesis of the presence of a linear three iron-sulfur centre. The fact that the *A. ambivalens* ferredoxin is a di-cluster seven-iron protein leads to the release of an abundant amount of iron and sulfide, which contributes to the production of a significant amount of iron sulfides leading to an intense grey coloration and spectral features. In other iron-sulfur proteins, given favourable reaction conditions and high enough concentrations of reactants, the same transients should be observed.

3.5 DISCUSSION

In this work we established that during the chemical alkaline degradation of a dicluster, seven iron ferredoxin, shortly after onset of protein unfolding iron is released monophasically at a rate which is comparable to that of the degradation of the iron-sulfur centres. The degradation pathway of the protein comprises an initial step in which the protein is destabilised to an open conformation, which results in an exposure of the iron-sulfur centres and zinc site, as inferred from the increase in the unfolding speed and general destabilising effect of EDTA in the unfolding process. Also, our kinetic experiments suggested that the protein N-terminal extension may be more sensitive to chemical denaturation than the protein core, a confirmation that this region of the protein may have a stabilising effect by covering the core region. During this process there is no detectable formation of any intermediate species involving the metal clusters. Instead we have demonstrated that the release of sulfide and iron from cluster degradation leads to the formation of iron sulfide species, which at basic pH, are responsible for the transient bands observed at 610 and 520nm. Thus, the previous hypothesis according to which decomposition of the iron sulfur clusters was proceeding via linear three iron sulfur centres is now disproved.

3.6 REFERENCES

1. Fraústo da Silva JJR, W.R., *The Biological Chemistry of Elements*. 1991, Oxford: Clarendon Press.
2. Sandberg, A., et al., *Effects of metal ligation and oxygen on the reversibility of the thermal denaturation of Pseudomonas aeruginosa azurin.* Biochemistry, 2002. **41**(3): p. 1060-9.
3. Leckner, J., et al., *The effect of the metal ion on the folding energetics of azurin: a comparison of the native, zinc and apoprotein.* Biochim Biophys Acta, 1997. **1342**(1): p. 19-27.
4. Gross, E.L., et al., *Thermal denaturation of plastocyanin: the effect of oxidation state, reductants, and anaerobicity.* Arch Biochem Biophys, 1992. **298**(2): p. 413-9.
5. Moczygemba, C., et al., *High stability of a ferredoxin from the hyperthermophilic archaeon A. ambivalens: involvement of electrostatic interactions and cofactors.* Protein Sci, 2001. **10**(8): p. 1539-48.
6. Foster, M.W., et al., *Elucidation of a [4Fe-4S] cluster degradation pathway: rapid kinetic studies of the degradation of Chromatium vinosum HiPIP.* J Biol Inorg Chem, 2001. **6**(3): p. 266-74.
7. Bentrop, D., et al., *Structural and dynamical properties of a partially unfolded Fe4S4 protein: role of the cofactor in protein folding.* Biochemistry, 1999. **38**(15): p. 4669-80.
8. Meyer, J., et al., *Assembly of a [2Fe-2S]2+ cluster in a molecular variant of Clostridium pasteurianum rubredoxin.* Biochemistry, 1997. **36**(43): p. 13374-80.
9. Apiyo, D., et al., *Equilibrium unfolding of dimeric desulfoferrodoxin involves a monomeric intermediate: iron cofactors dissociate after polypeptide unfolding.* Biochemistry, 2001. **40**(16): p. 4940-8.
10. Burova, T.V., et al., *Conformational stability of adrenodoxin mutant proteins.* Protein Sci, 1996. **5**(9): p. 1890-7.
11. Bera, A.K., A. Grinberg, and R. Bernhardt, *A step toward understanding the folding mechanism of bovine adrenodoxin.* Arch Biochem Biophys, 1999. **361**(2): p. 315-22.
12. Jones, K., et al., *Formation of a linear [3Fe-4S] cluster in a seven-iron ferredoxin triggered by polypeptide unfolding.* J Biol Inorg Chem, 2002. **7**(4-5): p. 357-62.
13. Griffin, S., et al., *High thermal and chemical stability of Thermus thermophilus seven-iron ferredoxin. Linear clusters form at high pH on polypeptide unfolding.* Eur J Biochem, 2003. **270**(23): p. 4736-43.
14. Pereira, M.M., et al., *A ferredoxin from the thermohalophilic bacterium Rhodothermus marinus.* Biochim Biophys Acta, 2002. **1601**(1): p. 1-8.
15. Higgins, C.L., J. Meyer, and P. Wittung-Stafshede, *Exceptional stability of a [2Fe-2S] ferredoxin from hyperthermophilic bacterium Aquifex aeolicus.* Biochim Biophys Acta, 2002. **1599**(1-2): p. 82-9.
16. Mitou, G., et al., *An Isc-type extremely thermostable [2Fe-2S] ferredoxin from Aquifex aeolicus. Biochemical, spectroscopic, and unfolding studies.* Biochemistry, 2003. **42**(5): p. 1354-64.
17. Teixeira, M., et al., *A seven-iron ferredoxin from the thermoacidophilic archaeon Desulfurolobus ambivalens.* Eur J Biochem, 1995. **227**(1-2): p. 322-7.
18. Gomes, C.M., Faria, A., Carita, J.C., Mendes, J., Regalla ,M., Chicau ,P., Huber ,H.,Stetter ,K.O. and Teixeira ,M., *Di-cluster, seven iron ferredoxins from hyperthermophilic Sulfolobales.* JBIC, 1998. **3**: p. 499-507.
19. Mach, H. and C.R. Middaugh, *Simultaneous monitoring of the environment of tryptophan, tyrosine, and phenylalanine residues in proteins by near-ultraviolet second-derivative spectroscopy.* Anal Biochem, 1994. **222**(2): p. 323-31.
20. Ragone, R., et al., *Determination of tyrosine exposure in proteins by second-derivative spectroscopy.* Biochemistry, 1984. **23**(8): p. 1871-5.

21. Cowart, R.E., F.L. Singleton, and J.S. Hind, *A comparison of bathophenanthrolinedisulfonic acid and ferrozine as chelators of iron(II) in reduction reactions.* Anal Biochem, 1993. **211**(1): p. 151-5.

22. Shirley, B.A., *Methods in Molecular Biology*. Urea and guanidine hydrochloride denaturation curves, ed. B.A. Shirley. Vol. 40. 1995, Totowa,NJ, USA: Humana Press Inc. p.337.

23. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. second ed. 1999, New York: Kluwer Academic / Plenum Publishers.

24. Dorovska-Taran, V., et al., *A comparative picosecond-resolved fluorescence study of tryptophan residues in iron-sulfur proteins.* FEBS Lett, 1994. **348**(3): p. 305-10.

25. Kennedy, M.C., et al., *Evidence for the formation of a linear [3Fe-4S] cluster in partially unfolded aconitase.* J Biol Chem, 1984. **259**(23): p. 14463-71.

26. Fujii, T., et al., *The crystal structure of zinc-containing ferredoxin from the thermoacidophilic archaeon Sulfolobus sp. strain 7.* Biochemistry, 1997. **36**(6): p. 1505-13.

27. Iametti, S., et al., *Reversible, non-denaturing metal substitution in bovine adrenodoxin and spinach ferredoxin and the different reactivities of [2Fe-2S]-cluster-containing proteins.* Eur J Biochem, 1996. **239**(3): p. 818-26.

28. Wittung-Stafshede, P., *Role of cofactors in protein folding.* Acc Chem Res, 2002. **35**(4): p. 201-8.

29. Cotton, F.A.a.W., G, *Advanced inorganic chemistry: a comprehensive text*. 3rd ed. 1972, NY: Interscience.

30. Russell, J., *General chemistry*. 1980, NY: McGraw-Hill.

31. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.

32. Kennedy, M.C. and H. Beinert, *The state of cluster SH and S2- of aconitase during cluster interconversions and removal. A convenient preparation of apoenzyme.* J Biol Chem, 1988. **263**(17): p. 8194-8.

3.7 ACKNOWLEDGMENTS

To Miguel Teixeira for the assistance with EPR studies and for providing the soluble extract of *Acidianus ambivalens*

To João Carita from ITQB analytical services is gratefully acknowledged for cell mass processing

To Ana Coelho from ITQB for preliminary mass spectrometry experiments.

4 LINEAR THREE-IRON CENTRES ARE UNLIKELY CLUSTER DEGRADATION INTERMEDIATES DURING UNFOLDING OF IRON-SULFUR PROTEINS

CONTENTS

This chapter was published in

Sónia S. Leal and Cláudio M. Gomes
Linear three-iron centres are unlikely clusters degradation intermediates during unfolding of iron-sulfur proteins
Biol. Chem. 386: 1295-1300 (2005)

4.1 SUMMARY

Previous studies on the chemical alkaline degradation of ferredoxins have contributed to establish in the literature the concept that linear three-iron centres were commonly observed as degradation intermediates of iron-sulfur clusters. In this work we assess the validity of this hypothesis: for that we have studied different proteins containing iron-sulfur clusters, iron-sulfur centres and di-iron centres in respect to their chemical degradation kinetics at high pH, in the presence and absence of exogenous sulfide, to investigate the possible formation of linear three-iron centres during protein unfolding. Our spectroscopic and kinetic data shows that in these different proteins, visible absorption bands at 530 and 620 nm identical to those that have been suggested to arise from linear three-iron centres, are formed. Iron release and protein unfolding kinetics show that these bands result from formation of iron sulfides at pH 10, produced from the degradation of the iron centres, and not from rearrangements leading to linear three-iron centres. Thus, at this point, any relevant functional role of linear three-iron centres as cluster degradation intermediates in iron-sulfur proteins remains elusive.

4.2 INTRODUCTION

Iron sulfur clusters are ubiquitously found in all life domains and are considered to be one of the most ancient forms of biological catalysts. The versatility of functions performed by iron sulfur proteins is directly correlated to the intrinsic structural diversity of its clusters, which ranges from the simplest rubredoxin-type iron sulfur centres, and linear and cubic clusters with higher nuclearities ([2Fe-2S],[3Fe-4S] and [4Fe-4S]), to the very complex assemblies in which a cuboid cluster is bridged to another component such as the cofactor cluster ([Mo7Fe-9S]) and P-cluster ([8Fe-S7]) of nitrogenase. In addition, the polypeptide chain plays itself an essential role in the fine-tuning of the cluster biological function, either by modulating the cluster reduction potential or by stabilising it. The latter has a particular importance in the non-redox functions of iron sulfur centres, in which assembly and disassembly of the cluster provides a regulatory signal. These are the cases of the FNR protein from *Escherichia coli*, a

transcription activator of genes encoding for components of anaerobic pathways, in which exposure to oxygen inactivates the protein as a result of the reversible conversion of its [4Fe-4S] to [2Fe-2S] clusters leading to dissociation of the catalytically proficient dimeric form. Also, in aconitase and in the iron regulatory protein, cluster rearrangements between the active [4Fe-4S] and the inactive [3Fe-4S] containing form modulate the protein function as transcription regulator of ferritin and transferrin receptor, two proteins involved in the cellular iron metabolism. Aconitase also constitutes the so far clearest example of a protein containing a linear three-iron sulfur centre, which is formed when the inactive [3Fe-4S] containing protein is exposed to pH > 9 or partly denatured with urea. The resulting cluster has a purple colour and characteristic spectroscopic signatures and can be re-converted to the active [4Fe-4S] containing form upon lowering the pH and incubation with an iron salt and a thiol [1]. Nevertheless, it has no known biological function and it remains as an interesting example of cluster interconversion within a protein environment. However, it has been suggested that a linear three-iron sulfur centre could be formed upon the chemical denaturation at pH 10 of a seven iron ferredoxin [2-4] . The original hypothesis was essentially based on spectroscopic evidence from visible spectroscopy, as a gray-purple compound was transiently formed during protein degradation and cluster dissociation. Subsequent studies in other ferredoxins, even those containing centres with lower nuclearities, have contributed to establish in the literature the concept that linear three-iron sulfur centres were commonly observed as degradation intermediates in iron sulfur proteins [5-7]. However, recent evidence has challenged this hypothesis, as studies on the degradation pathway of a seven iron ferredoxin indicate that the observed intermediate species are likely to result from the formation of iron sulfides [8]. In order to contribute to the clarification of this issue in the literature and to clearly establish if linear three-iron sulfur centres are formed during the degradation of iron sulfur proteins at alkaline pH or if the transient species observed are due to unspecific reactions of the cluster degradation products, we undertook a spectroscopic and kinetic study on the degradation of several iron proteins.

4.3 MATERIALS AND METHODS

Proteins

Proteins used in this study were purified as described elsewhere (see references below) and were kind gifts of M. Teixeira (*Escherichia coli* flavorubredoxin and the rubredoxin domain [9]) and I. Pereira (*Desulfovibrio gigas* hydrogenase [10, 11], both at Instituto Tecnologia Quıimica e Biologica.

Spectroscopic methods

Visible absorption measurements were performed on a Shimadzu MultiSpec-1501 diode-array spectrophotometer equipped with cell stirring and temperature control via a computer-controlled water bath (Julabo, Seelbach, Germany). The integrity of iron-sulfur clusters and iron centres was monitored by following the decay at 420–700 nm for hydrogenase, at 480–700 nm for the rubredoxin domain, and at 440–700 nm for flavorubredoxin. Fluorescence spectroscopy was carried out on a Cary Varian Eclipse spectrofluorimeter equipped with cell stirring and Peltier temperature control. Intrinsic tryptophan fluorescence emission was monitored between 300 and 450 nm (slit width 5 nm), after excitation at 280 nm (slit width 5 nm). All experiments were performed at room temperature.

Iron quantification

Quantification of the iron released during unfolding was performed using 1,2-dihydroxy-3,5-benzenedisulfonic acid (Tiron) [12] and bathophenanthroline disulfonic acid (BPS) [13]. The time course of iron release was obtained by continuously monitoring the formation of the Fe(III)-Tiron complex at 480–700 nm during protein unfolding. A control experiment, performed in the absence of Tiron, was used to subtract the contribution owing to protein variation at this wavelength. Iron was quantified from calibration curves made with accurately titrated $Fe(NO_3)^3$ solutions (a kind gift from Rita Delgado, ITQB/UNL).

Kinetic experiments

The chemical denaturant GuHCl was used to promote protein unfolding at 25ºC

and pH 10. A stock solution of GuHCl at pH 10 was prepared in 200 mM glycine buffer and its accurate concentration was confirmed by refractive index measurements [14]. The reaction mixture was allowed to stabilise at 25ºC and the reaction was started with the addition of a concentrated protein solution.

Data analysis

Optical decomposition was performed in Matlab, using the builtin singular value decomposition (SVD) algorithm. Briefly, in this procedure a data matrix of time-resolved spectra (*A*) is decomposed by SVD in the product *A*s*U*=*S*=*VT*, in which *U* is the basis spectra matrix, *S* the matrix of relative occupancies and *V* the time dependence of the basis spectra [15]. The kinetic experimental data were processed and fitted using the program Origin (OriginLab, Northampton, USA).

4.4 RESULTS

kinetics of iron centre degradation

The kinetics of chemical unfolding promoted by guanidinium hydrochloride (GuHCl) at pH 10 were monitored by UV-visible spectroscopy, which allows probing directly metal site integrity and inspection of the formation of bands in the 500-620nm region which have been attributed to the transient formation of linear three-iron centres [2-4]. Anaerobic conditions were used in order to protect possible reaction intermediates from degradative oxidation by molecular oxygen. The results obtained show that during unfolding of Hase and Fd, there is formation of new bands at 530 and 620nm (see figure. 4.1, panel A). On the other hand, during FlRd and Rd unfolding (figure 4.1, panels B and C), these new features are not observed; instead, there is a bleaching of the spectrum corresponding to the native protein form, which is compatible with the ongoing denaturation process. From these proteins hydrogenase is the only one in which there is release of sulfide (S^{2-}) as a result of iron-sulfur cluster decomposition during unfolding. A different scenario occurs during the degradation of the FlRd and Rd iron centres, which simply result in iron release. Similar results were obtained for all proteins when the experiment was made in aerobic conditions,

with the difference that the observed amplitude of the 620 and 530nm absorption bands in Hase was smaller.

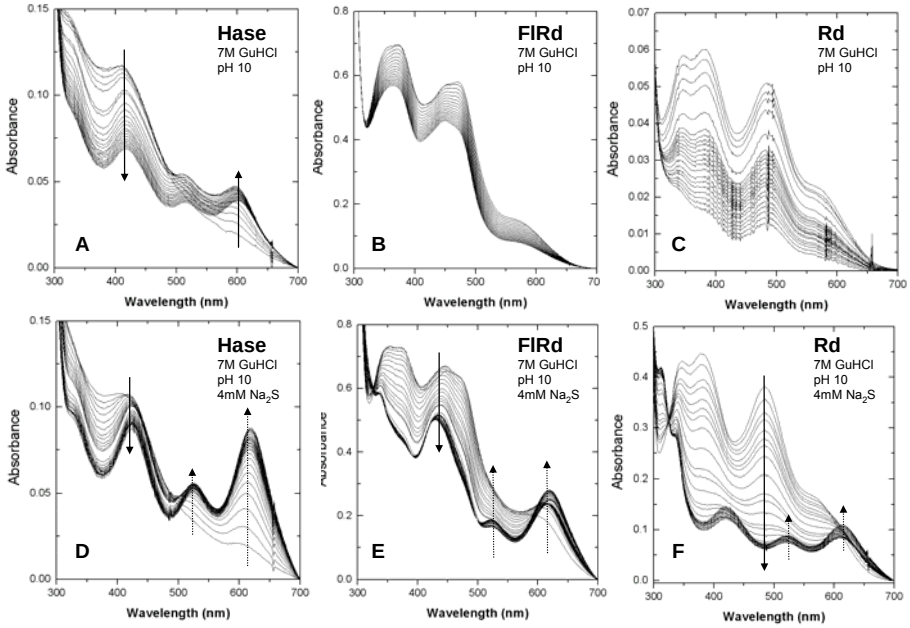


Figure 4.1 |Absorption spectra at different times upon chemical denaturation at pH 10. **A,D|** Hydrogenase (Hase); **B,E|** Flavorubredoxin (FlRd; **C,F|** Rubredoxin (Rd) were chemically denaturated upon mixing with 7 M GuHCl 0.2 M glycine, pH 10 at 25ºC in the absence (A-C) and presence of 4 mM sulphide (D-F). Protein concentrations used were 5.5 µM (Hase); 25 µM (FlRd) and 60 µM (Rb).

|Effect of exogenous sulfide

The same experiments were performed in the presence of exogenous sulfide (figure 4.1, panels D to F). In this case, not only the 620 and 530nm absorption bands are observed in all proteins, but the obtained spectra are identical in all proteins, suggesting that a common chemical species was formed upon the degradation of the iron sites (see below). Further, the formed species with absorption bands at 530 and 620nm is not transient but rather is stable in solution in a very long time scale (tens of minutes). Interestingly, adding sulfide to the fully degraded FlRd and Rd obtained in the unfolding experiments performed in the absence of exogenous sulfide (panels B and C), results in the immediate

formation of intense bands at 530 and 620nm, identical to those observed when sulfide is present during unfolding. Thus, it is clear that the bands at 530 and 620 nm are only formed when iron and sulfide are available in solution.

|Spectral decomposition of the transient species

The spectra of the species with bands at 530 and 620nm observed during unfolding of Hase, FlRd and Rd in the presence of sulfide, were obtained by two different analysis methodologies which gave identical results (figure 4.2, panel A).

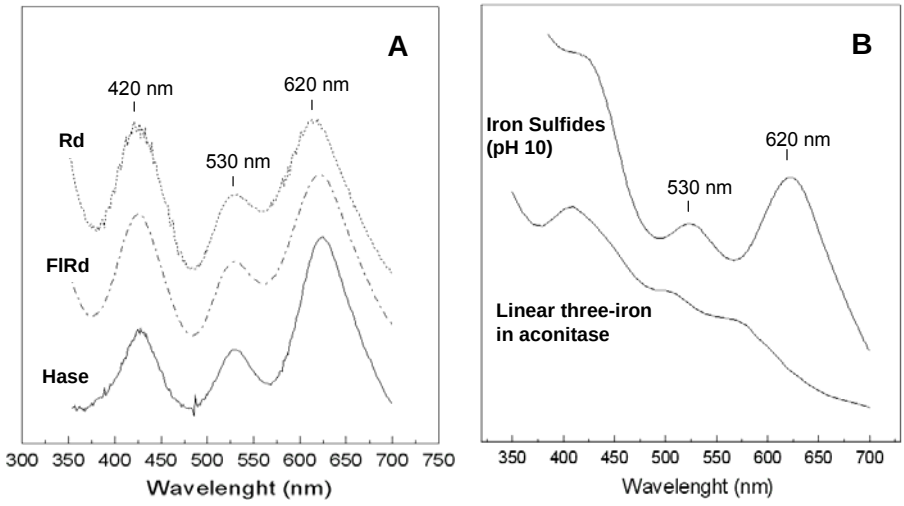


Figure 4.2 |Comparison of the decomposed spectra with those of iron sulfides at pH 10 and purple acotinase (linear three-iron cluster).
A| Decomposition spectra of the chemical species formed upon unfolding of Hydrogenase (Hase), Flavorubredoxin (FlRd) and Rubredoxin (Rd) in 7 M GuHCl, 4 mM Na_2S , pH 10 **B|** Absorption spectra of the linear three-iron centre in purple acotinase [in 0.1 M glycine pH 10.5, replotted from [1]] and of iron sulfides (Fe_xS_y) obtained by mixing excess sulfide and ferrous iron in 0.1 M glycine pH 10.

One was difference spectroscopy, which consisted in subtracting the final spectra of the proteins obtained upon denaturation in the absence of sulfide, to those obtained in its presence. The other was singular value decomposition analysis (SVD), which allowed to deconvolute the progression of the unfolding process on the basis of its basis spectral components and its relative occupancy during the reaction. This analysis clearly shows that identical bands are observed in all

proteins, with features at 420, 530 and 620nm. These are identical to those observed during the unfolding of several ferredoxins at high pH [3, 5, 6, 16], which have been suggested to arise from linear three-iron centres, on the basis of a spectral similarity (table 4.1). A comparison of these spectra with those of the linear three-iron centre from aconitase and iron sulfides (Fe_xS_y) produced at pH 10 is shown in figure 4.3, panel B. It is clear that the bands observed during unfolding of iron-containing proteins are a result of the formation of iron sulfides and not of linear three-iron centres, as previously postulated.

Table 4.1 |Iron-sulfur proteins exhibiting the appearance of absorption bands at 530 and 620 nm during chemical denaturation, assigned to linear three-iron centres

Protein (organism)	Denaturation conditions	Ref.
[2Fe-2S] (*Bos taurus*)	6 M GuHCl, pH 8.5, 10 mM Na_2S	[17]
[2Fe-2S] (*Aquifex aeolicus*)	6 a 8 M GuHCl, pH 10	[2]
[3Fe-4S] (*Rhodothermus marinus)*)	5.5 M GuHCl, pH 7	[7]
[3Fe-4S][4Fe-4S] (*Acidianus ambivalens*)	7 M GuHCl, pH 10	[4]
[3Fe-4S][4Fe-4S] (*Thermus thermophilus*)	6-7.9 M GuHCl, pH 10	[5]
[3Fe-4S][4Fe-4S] (*Sulfolobus acidocaldarius*)	6 - 7.5 M GuHCl pH 10 4 - 6 M GuSCN pH 11	[3]

|Protein unfolding and iron release

In order to correlate the formation of the bands at 530 and 620nm with protein unfolding and metal centre dissociation, the protein unfolding kinetics during GuHCl denaturation at pH 10 in the presence of sulfide, were also investigated by intrinsic tryptophan emission and iron release measurements (figure 4.3). The results show that protein unfolding, iron centre degradation, iron release and formation of the 530 and 620nm bands are simultaneous events (table 4.2).These events were also occurring at identical rates during the unfolding of a seven iron ferredoxin [8]. In *E. coli* flavorubredoxin, protein unfolding proceeds via a more complex mechanism, as a consequence of the fact that the protein is a fusion of three distinct structural domains: a β-lactamase like, a flaxodoxin-like and a rubredoxin-like fold [9, 18]. Overall these results show that protein unfolding leads to simultaneous metal centre degradation, without the formation of any detectable

protein unfolding intermediate; and that released iron readily reacts with available sulfide leading to the formation of the iron sulfides which are responsible for the bands at 530 and 620nm. Altogether, this agrees well with a re-interpretation of events occurring during the alkaline unfolding of iron-sulfur proteins excluding the formation of intermediate linear three-iron centres.

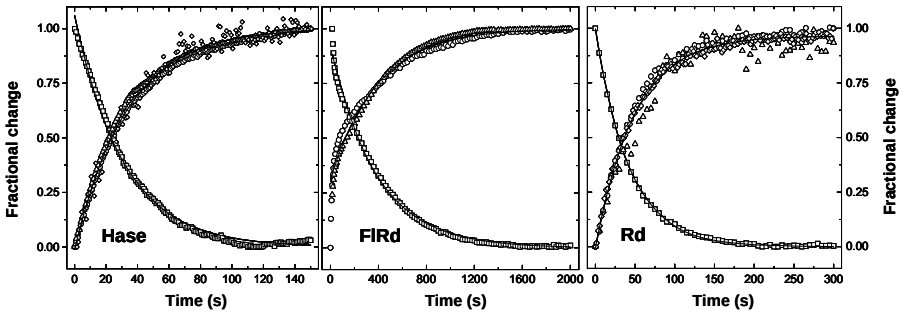


Figure 4.3 | Kinetic traces for protein unfolding, metal centre degradation, iron release and variation in the 620nm band.
Kinetics of protein unfolding was measured after mixing with 7 M GuHCl, 0.2 M glycine, 4 mM Na_2S at pH 10. Symbol key for all panels: (□) metal centre degradation; (◊) protein unfolding (tryptophan fluorescence emission); (○) iron release; (Δ) formation of iron sulfides (increase in 620 nm band). Protein concentrations used in this assay were 1.5 μM Hase, 5 μM FlRd and 0.9 Fe/mol for Rd. Solid lines correspond to best fits and the observed rate constants are given in table 4.2.

Table 4.2 | Observed kinetic rates for protein unfolding, metal centre dissociation, iron release and iron sulfide formation

	$K_{obs}(s^{-1})$			
	Hidrogenase	**Flavorubredoxin**	**Rubredoxin**	**Ferredoxin[a]**
Metal centre	$2.24x10^{-2}$	$1.85\ x10^{-3}$	$1.63\ x10^{-2}$	ND
Protein unfolding	$1.82\ x10^{-2}$	ND	$1.21x10^{-2}$	$39\ x10^{-4}$
Fe_xS_y formation	$2.09\ x10^{-2}$	$1.98\ x10^{-3}$	$1.11\ x10^{-2}$	$12\ x10^{-4}$
Iron Release	$1.95\ x10^{-2}$	$1.80\ x10^{-3}$	$1.37\ x10^{-2}$	$12\ x10^{-4}$

See legend figure 2.5 for conditions. ND, not determined.
[a]Data taken from [8]

4.5 | DISCUSSION

In this study we have investigated possible iron centre inter-conversions occurring upon metalloprotein unfolding in alkaline conditions. The pioneer work

by Kennedy and co-workers [1] has shown that the cubic iron-sulfur centre from beef heart aconitase, upon protein incubation at pH>9.5 or treatment with 4-8M urea, undergoes a reversible conversion into a linear three iron centre. The chemical nature of this new type of centre was then clearly established, both by comparative studies using model compounds and Mossbauer spectroscopy [1]. More recently, it has been postulated that identical linear iron-sulfur centres could be formed during ferredoxin chemical unfolding at high pH [3], although this hypothesis was subsequently revised [8]. In order to contribute to the understanding of metal centre transformations occurring during protein unfolding, we have studied several iron-proteins containing a variety of different iron centres. The issue of the biological synthesis of iron centres is of ultimate interest to the field of protein folding and catalysis, and in this respect, the investigation of cluster assembly and eventual inter-conversions is particularly challenging. In order to investigate these processes, it is however often necessary to induce non-native protein states to which correspond partly misfolded protein forms, which somehow may represent facilitated intermediate states in the cluster assembly or rearrangement mechanisms. This can be achieved by using non-physiological pH conditions, namely high pH. Therefore, we have studied the kinetics of the chemical unfolding of hydrogenase, flavorubredoxin and rubredoxin at pH 10, as linear three iron-sulfur centres are known to be preferentially formed in these conditions. Our results clearly show that in all these different proteins, visible absorption bands at 530 and 620nm are observed upon chemical unfolding, but that these result from iron sulfides, and not from the transient formation of linear three-iron centres. These spectroscopic bands are identical to those observed during the unfolding of several ferredoxins which have been suggested to arise from linear three-iron centres [3]. In fact, we observed that iron sulfides are formed from the reaction of the iron released from iron centres degradation, when available sulfide in solution. This is clearly shown by the experiments in which exogenous sulfide is present, in which the intensity of these spectral bands increases. In fact, as demonstrated, the observed transient spectral species are due to the formation of iron sulfides at pH 10. Providing that exogenous sulfides are available, these are even observed in iron proteins that contain no iron-sulfur centres. Thus, our work provides a new appraisal disproving the view suggesting

that these centres are general cluster degradation intermediates, transiently formed during degradation of iron-sulfur clusters, although the possibility of having residual cluster structures associated to the polypeptide chain during its unfolding cannot be completely excluded. At this point, any relevant biological function for linear three-iron centres remains unclear.

4.6 REFERENCES

1. Kennedy, M.C., et al., *Evidence for the formation of a linear [3Fe-4S] cluster in partially unfolded aconitase.* J Biol Chem, 1984. **259**(23): p. 14463-71.
2. Higgins, C.L. and P. Wittung-Stafshede, *Formation of linear three-iron clusters in Aquifex aeolicus two-iron ferredoxins: effect of protein-unfolding speed.* Arch Biochem Biophys, 2004. **427**(2): p. 154-63.
3. Jones, K., et al., *Formation of a linear [3Fe-4S] cluster in a seven-iron ferredoxin triggered by polypeptide unfolding.* J Biol Inorg Chem, 2002. **7**(4-5): p. 357-62.
4. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.
5. Griffin, S., et al., *High thermal and chemical stability of Thermus thermophilus seven-iron ferredoxin. Linear clusters form at high pH on polypeptide unfolding.* Eur J Biochem, 2003. **270**(23): p. 4736-43.
6. Higgins, C.L., J. Meyer, and P. Wittung-Stafshede, *Exceptional stability of a [2Fe-2S] ferredoxin from hyperthermophilic bacterium Aquifex aeolicus.* Biochim Biophys Acta, 2002. **1599**(1-2): p. 82-9.
7. Pereira, M.M., et al., *A ferredoxin from the thermohalophilic bacterium Rhodothermus marinus.* Biochim Biophys Acta, 2002. **1601**(1): p. 1-8.
8. Leal, S.S., M. Teixeira, and C.M. Gomes, *Studies on the degradation pathway of iron-sulfur centers during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability.* J Biol Inorg Chem, 2004. **9**(8): p. 987-96.
9. Gomes, C.M., et al., *Spectroscopic studies and characterization of a novel electron-transfer chain from Escherichia coli involving a flavorubredoxin and its flavoprotein reductase partner.* Biochemistry, 2000. **39**(51): p. 16230-7.
10. Teixeira, M., et al., *Electron paramagnetic resonance studies on the mechanism of activation and the catalytic cycle of the nickel-containing hydrogenase from Desulfovibrio gigas.* J Biol Chem, 1985. **260**(15): p. 8942-50.
11. Volbeda, A., et al., *Crystal structure of the nickel-iron hydrogenase from Desulfovibrio gigas.* Nature, 1995. **373**(6515): p. 580-7.
12. Berger, S.A. and J.B. McKay, *The photometric titration of Fe(III) with EDTA using Tiron indicator.* Mikrochim Acta, 1974(4): p. 665-70.
13. Cowart, R.E., F.L. Singleton, and J.S. Hind, *A comparison of bathophenanthrolinedisulfonic acid and ferrozine as chelators of iron(II) in reduction reactions.* Anal Biochem, 1993. **211**(1): p. 151-5.
14. Shirley, B.A., *Methods in Molecular Biology*. Urea and guanidine hydrochloride denaturation curves, ed. B.A. Shirley. Vol. 40. 1995, Totowa,NJ, USA: Humana Press Inc. p.337.
15. Henry, E.R.a.H., J., *Singular value decomposition: application to analyses of experimental data.* Methods Enzymol., 1992. **210**: p. 129-192.
16. Mitou, G., et al., *An Isc-type extremely thermostable [2Fe-2S] ferredoxin from Aquifex aeolicus. Biochemical, spectroscopic, and unfolding studies.* Biochemistry, 2003. **42**(5): p. 1354-64.

17. Burova, T.V., et al., *Conformational stability of adrenodoxin mutant proteins.* Protein Sci, 1996. **5**(9): p. 1890-7.

18. Frazão, C., et al., *Structure of a dioxygen reduction enzyme from Desulfovibrio gigas.* Nat Struct Biol, 2000. **7**(11): p. 1041-5.

4.7 ACKNOWLEDGMENTS

To Miguel Teixeira and Inês Pereira (both at ITQB/UNL) are gratefully acknowledged for providing the proteins used in this study, and for critically reading of the manuscript.

5 NATURAL DOMAIN DESIGN: ENHANCED THERMAL STABILITY OF A ZINC-LACKING FERREDOXIN ISOFORM SHOWS THAT A HYDROPHOBIC CORE EFFICIENTLY REPLACES THE STRUCTURAL METAL SITE

CONTENTS

This chapter was published in

Rita Rocha, **Sónia S. Leal**, Vitor H. Teixeira, Manuela Regalla, Harald Huber, António M. Batista, Cláudio M. Soares and Cláudio M. Gomes
"Natural domain design: enhanced thermal stability of zinc lacking ferredoxin isoforms shows that a hydrophobic core efficiently replaces the structural metal site"
Biochemistry 45: 10376-10384 (2006)

5.1 SUMMARY

Zinc centres play a key role as important structure determinants in a variety of proteins, including ferredoxins (Fd). We exploit the availability of two highly identical ferredoxin isoforms from the thermophile *Sulfolobus metallicus*, which differ in the residues involved in coordinating a His/Asp zinc site which ties together the protein core with its N-terminal extension, to investigate the effect of the absence of this site on ferredoxin folding. The conformational properties of the zinc-containing (FdA) and zinc-lacking (FdB) isoforms were investigated using visible absorption and tryptophan fluorescence emission. Comparative modelling and molecular dynamics simulations indicate that the FdB N-terminal extension assumes a fold identical to that of the Zn^{2+} containing isoform. The thermal stability of the isoforms was investigated in a broad pH range (2<pH<10), and at physiological pH conditions both proteins unfold above 100°C. Surprisingly, the Zn^{2+} lacking isoform was always found to be more stable than its Zn^{2+} containing counterpart: a $\Delta T_m \approx 9$°C is determined at pH 7, a difference which becomes even more significant at extreme pH values, reaching a $\Delta T_m \approx 24$°C at pH 2 and 10. The contribution of the Zn^{2+} site to ferredoxin stability was further resolved using selective metal chelators: during thermal unfolding the zinc scavenger TPEN significantly lowers the T_m in FdA ($\approx$ 10°C) whereas it has no effect in FdB. This shows that the Zn^{2+} site contributes to ferredoxin stability but that FdB has devised a structural strategy that accounts for an enhanced stability without using a metal cross linker. Analysis of the FdB sequence and structural model leads us to propose that the higher stability of the zinc containing ferredoxin results from novel non-polar contacts formed between the residues that occupy the same spatial region where the zinc ligands are found in FdA. These favour the formation of a novel local stabilising hydrophobic core and illustrate a strategy of natural fold design.

5.2 INTRODUCTION

The tertiary structure of a protein is mainly stabilized by weak non-covalent interactions which have an additive effect towards fold stabilization: salt bridges (12.5-17 $kJ.mol^{-1}$), hydrogen bonds (2-6 $kJ.mol^{-1}$) and van der Waals interactions (4-17 $kJ.mol^{-1}$) present within a protein contribute to a net stabilizing effect towards the folded structure, which is nevertheless only marginally stable in water at physiological temperature ($\Delta G_U \approx$ 20-40 $kJ.mol^{-1}$) [1]. However, the topological arrangement of secondary structure elements in a particular local motif can also involve additional covalent interactions that act as natural cross-linkers of a particular conformation. The most common protein covalent modifications are disulfide bridges (167 $kJ.mol^{-1}$) which involve the oxidation of two sulfhydryl groups from nearby cysteine residues, which can also have a functional role (reviewed in [2]).

Another common cross linking interaction in proteins are the coordinate covalent bonds formed between a metal ion and protein ligands. These structural metal sites which provide stabilization at a particular region of a protein are most frequently occupied by calcium (Ca^{2+}) and zinc (Zn^{2+}) [1]. The latter are typified by the so called zinc-fingers motifs [3, 4], in which various combinations of four cysteine or histidines residues in a short sequence stretch of around 30 residues binds a Zn^{2+} ion that stabilizes a local domain comprising a α-helix and three β-strands [3]. This is the most abundant structural domain in the human genome, and in the absence of the metal, it unfolds [5]. There is a plethora of examples in which engineering of a zinc site in a protein results in an important increase in its thermodynamic stability. For example, adenylate kinases (AK) from gram-positive bacteria hold a structural Zn^{2+} site, whereas those of gram-negatives lack it. Site direct mutagenesis in the AK from the gram-negative *Escherichia coli* introducing a zinc site results in a protein with an increased thermal stability (ΔT_m=9°C) [6, 7]. There are also several examples of Zn containing proteins in which metal removal results in decreased stability [8, 9].

Zinc structural sites encompass a versatile combination of soft (Cys, His) and harder (Asp, Glu) ligands, a feature also denoted in some catalytic zinc centers [3, 10]. In fact, protein coordination in structural zinc sites is diverse and ranges

from the preferred all cysteine coordination to sites in which Asp and Glu are found in combination with His residues. Such a His/Asp site has also been identified in ferredoxins [11, 12] isolated from organisms belonging to the Archaeal order of the Sulfolobales. These proteins are structurally characterized by having two Fe-S clusters (a [3Fe-4S] and a [4Fe-4S]) and a 36 amino acid long N-terminal extension comprising three β-strands: this part of the protein comprises the three His that bind Zn^{2+}, whereas the Asp comes from the protein core, completing a tetrahedral coordination. Thus, this zinc center has been suggested to play an important role in stabilizing the protein structure: in fact, mutagenesis of any of its His ligands results in a destabilized protein [13, 14]. Interestingly, some ferredoxins from this family were found to naturally lack the zinc site, and in some instances both the zinc containing and lacking isoforms are simultaneously expressed [12].

We have a long standing interest in the characterization of the biochemical, structural and stability properties of Fe-S proteins [12, 15-17], and at this point we turned our attention towards the two ferredoxin isoforms available from *Sulfolobus metallicus*, with and without the Zn^{2+} center. The rationale of our approach was to investigate the effect of the zinc site removal on the protein fold by determining the stability differences between the two ferredoxin isoforms, which have a very high percentage of amino acid identity (>80%) throughout most of their sequences, but differ precisely in the residues involved in coordinating the zinc site which ties together the protein core with its N-terminal extension. The results obtained in this work were a surprise with respect to initial expectations: the zinc lacking ferredoxin isoform, not only is not destabilized by the absence of the center, but it has in fact an enhanced stability with respect to the zinc-containing protein. These findings illustrate a natural fold design strategy, as during evolution nature has found multiple alternative strategies within the same protein family to achieve stabilization of a particular fold. This leads to the proposal that the higher stability of the zinc containing ferredoxin results from novel non-polar interactions present in the same spatial region where the zinc center is found in the other isoforms, which favour the formation of a local stabilizing hydrophobic core.

5.3 MATERIALS AND METHODS

Chemicals

All reagents were of the highest purity grade commercially available. The chemical denaturant Guanidinium hydrochloride (GuHCl) was obtained from Promega and the accurate concentration of the stock solutions in different buffers was confirmed by refractive index measurements. The chelator ethylenediaminetetraacetic acid (EDTA) was purchased from Merck. The chelators desferrioxamine (DFO) and N,N,N′,N′-tetrakis(2-pyridylmethyl) ethylenediamine (TPEN) were from Sigma.

Protein purification

The ferredoxins from *Sulfolobus metallicus* used in this work were purified as previously described [12]. Briefly, after cell disruption in 40mM potassium phosphate pH 6.5 (buffer A) the resulting cytosolic fraction was applied to a Q-sepharose fast flow column (V_c=20ml), and the Fd containing fractions eluted at ~200-350mM NaCl. After concentration in 5kDa amicons, this fraction was loaded into a Superdex G-50 column (V_c=150ml), eluted with buffer A + 150 mM NaCl and pooled according to the typical visible spectra of Fds. This fraction containing a mixture of FdA and FdB was dialyzed against buffer A, and further resolved in a High Performance Q-Sepharose column (V_c=20ml), after application of a flat salt gradient, from which two separated peaks were obtained. Protein purity was confirmed by SDS-PAGE and N-terminus sequencing, and the stocks were kept as concentrated aliquots at -20°C.

Protein sequencing

The ferredoxin sequencing procedures were identical to those previously described for the *Acidianus ambivalens* ferredoxin [12], with some minor modifications. For the complete sequence determination, ~100 µg of FdA and FdB were carboxymethylated following published procedures [18]. The resulting samples were divided into three aliquots which were processed in three distinct steps. Two of the steps were common for FdA and FdB and consisted in N-

terminal sequencing and chemical cleavage. The chemical cleavage with CNBr was carried by adding a CNBr crystal to the carboxymethylated polypeptides dissolved in 70% of formic acid. After an incubation period of 24 hours in the dark, the reagents were evaporated and the resulting peptides re-suspended. The third step consisted in an enzymatic digestion step using endoproteinases and was different for the two proteins: FdA was subjected to specific hydrolysis by Asp-N (Roche Diagnostics GmbH) whereas FdB was digested with Asp-N and Lys-C (Roche Diagnostics GmbH). Digestions were carried out following the manufacturer instructions, typically during 18 h and with a 1:100 enzyme to polypeptide mass ratio. The peptide fragments produced in each of the cleavage steps were separated by reverse-phase HPLC on a Beckmann Gold System using a 3.9x150 mm C_{18} column (Delta-Pak 300A, Waters), a gradient of 0–80% acetonitrile in 0.1% aqueous TFA as eluent and a flow rate of 0.5 ml min^{-1}. Peptide bands were detected in the effluent by UV absorption at 214 nm and sequenced. These sequences together with the N-terminal sequence (up to residue ~50) were used to build a peptide map from which the ferredoxin sequences were inferred. Sequencing was carried out on an Applied Biosystems Procise 491 HT sequencer.

Spectroscopic and Biochemical methods

UV/Vis spectra were recorded using a Shimadzu Multispec-1501 diode array spectrophotometer equipped with cell stirring and temperature control by a Julabo 4 MV water bath. Fluorescence spectroscopy was performed using a Cary Varian Eclipse instrument equipped with cell stirring and Peltier temperature control. Protein concentrations were determined either using the method of Bradford or the extinction coefficient (ε^{410nm}=30.4 $mM^{-1}cm^{-1}$). For pH variations, the buffers (20 mM) used were acetate (pH 4), HEPES or potassium phosphate (pH ~7) and glycine (pH 2 and 10); the effect of the variation of the pH with the temperature was verified or taken into consideration in buffer preparation. Metal analysis (iron and zinc) was carried out by total reflection X-ray Fluorescence (TXRF) at the laboratory of Prof. B.O. Kolbesen at the Inst. für Anorg. und Analyt. Chemie (Johann Wolfgang Goethe-Universität, Frankfurt, Germany)

Comparative modeling

Comparative modeling was carried out using two structures as structural templates: the ferredoxin from *Sulfolobus tokodaii* [11] (PDB: 1xer) and that from *Acidianus ambivalens* [19] (PDB: 2vkr), both determined at 2 Å resolution. While the former had the merit of eliciting the presence of a Zn site, it lacked density around the [4Fe-4S] site region, thus preventing a detailed description of an important part of the protein core structure. This limitation is circumvented with the use of the *A. ambivalens* ferredoxin structure, which provides a complete description of the Fe-S moieties. Some residues of the *Aa*FdA had alternative conformations and for those, the one with the largest occupation value was chosen. The FdA and FdB proteins from *S. metallicus* have very high amino acid identities with respect to both structural templates used (90% for FdA and 83% for FdB), and have also a high amino acid identity (82%) among themselves (see also results). Altogether, this makes us confident that a good structural model can be derived. The program MODELLER [20] (version 7v7), was used for deriving the structures, and the alignments were optimized until a good quality model for the unknown structure was achieved. A Ramachandran analysis of the final models was performed using PROCHECK [21]. In the final isoform models, the majority of the residues are in the most favored regions (89.3% for FdA and 87.8 %), whereas the remaining are found in additional allowed regions.

Molecular Dynamics simulations

Molecular dynamics simulations were performed using GROMACS [22] (version 3.2.1). The GROMOS96 43A1 force field [23, 24] was used together with the atomic partial charges for the iron sulfur centers, as calculated elsewhere [25]. The simulations started from the comparative model structures (FdA and FdB) using the protonation states at pH 7.0, as determined using previously described methods [26, 27]. The system was solvated in a dodecahedral box where a minimum distance of 8 Å between the protein and the box walls was imposed. The final FdA and FdB systems had 3151 and 2882 water molecules, respectively. The molecular dynamics simulations were performed using separate heat and pressure baths [28] at 300 K and 1 bar, respectively, for the solute and

solvent (water), and coupling constants of 0.1 and 0.5 ps, and an isothermal compressibility of $4.5x10^{-5}$ bar^{-1}. All bonds were constrained with the LINCS [29] algorithm. The equations of motion were integrated using a time step of 2 fs. Non-bonded interactions were treated with a twin-range method, using group-based cutoffs of 8 and 14 Å, updated every 5 steps. The electrostatic forces thus truncated were corrected with a reaction field using a dielectric constant of 54 [30], the dielectric constant of SPC water [31] under these conditions. The system was energy minimized with the steepest descent method, for optimization of the hydrogen atoms. A 10^5 kJ/(mol.nm^2) position-restraining force constant on heavy atoms was used in the minimization step. The initialization of each MD simulation was done in two steps. In the first step, a 100 ps simulation was run with all protein heavy atoms position-restrained with a 10^4 kJ/(mol.nm^2) force constant and with initial velocities taken from a Maxwellian distribution at 300 K and a temperature and pressure coupling constants of the baths of 0.01 and 0.05 ps, respectively**.** In the second step, a 100 ps simulation was run with all C^{α} atoms position-restrained with a 10^4 kJ/(mol nm^2) force constant and temperature and pressure coupling constants of 0.1 and 0.5 ps, respectively. After this initialization, the simulation continued with all atoms free. Conformations were saved every pico-second for subsequent analysis.

Thermal unfolding

Thermal denaturation was followed by monitoring the variation of the spectroscopic signals (Trp emission or Fe-S cluster integrity) as a function of temperature. Typically, 10µM ferredoxin was used and a heating rate of $1ºC.min^{-1}$ was used. Data were analyzed according to a two-state model described from which the transition midpoint temperature (T_m) was determined. As it happens with other Fe-S proteins, the cluster disintegration of the clusters is an irreversible process, limiting the thermodynamic analysis to a comparison of the T_m values. As a further control of the thermal transition we have verified that the Fe-S centers disintegrate at a T_m identical to that determined from Trp intrinsic fluorescence, thus leading to the conclusion that the thermal denaturation of the polypeptide is not influenced by the irreversible character of the denaturation of the Fe-S centers. For the experiments in the presence of different metal

chelators, the protein was allowed to incubate for 10 min with chelators (5 mM). This incubation prior to monitoring the thermal transition had no effect on the metal:protein stoichiometry. Curve fitting was carried out using Origin (Microcal).

5.4 RESULTS

Sequencing and homology modelling

Since its original definition, the number of members of the seven iron di-cluster ferredoxin family from archaeal crenarchaeotas has been increasing, mostly as a result of genome sequencing projects. The sequences of the two ferredoxin isoforms from *S. metallicus* used in this study were determined chemically after proteolytic digestion and sequencing of internal peptides (figure 5.1). Both proteins are ~100 amino acids long and have several of the characteristics of the di-cluster ferredoxin family, namely two iron-sulfur binding Cys motifs (figure 5.1, marked with #) and a N-terminal extension of ~30 residues. The two ferredoxin types present within the family differ precisely in this region, whereas the FdA variants contain a Zn^{2+} binding motif, the FdB isoforms do not (figure. 5.1, Zn centre ligands marked with *). In agreement, the purified ferredoxin isoforms differ in metal content: both contain a ~6.5 Fe/mol protein, corresponding to the presence of a [3Fe-4S] and a [4Fe-4S] centre, but whereas FdA contains 0.8 Zn/mol protein, the FdB isoform lacks it. Amino acid sequence alignment performed between homologous ferredoxins belonging to this family, confirms their high amino acid identities (79-98%) as noted in the original definition of the family [12]. In order to correlate ferredoxin conformation and dynamics studies with protein structure, molecular models of the two ferredoxin isoforms were produced by comparative modelling given the high sequence identity towards members with known structures. Most of the residues differing between the two isoforms under study (20 out of 103) occur at the N-terminus (figure 5.1).

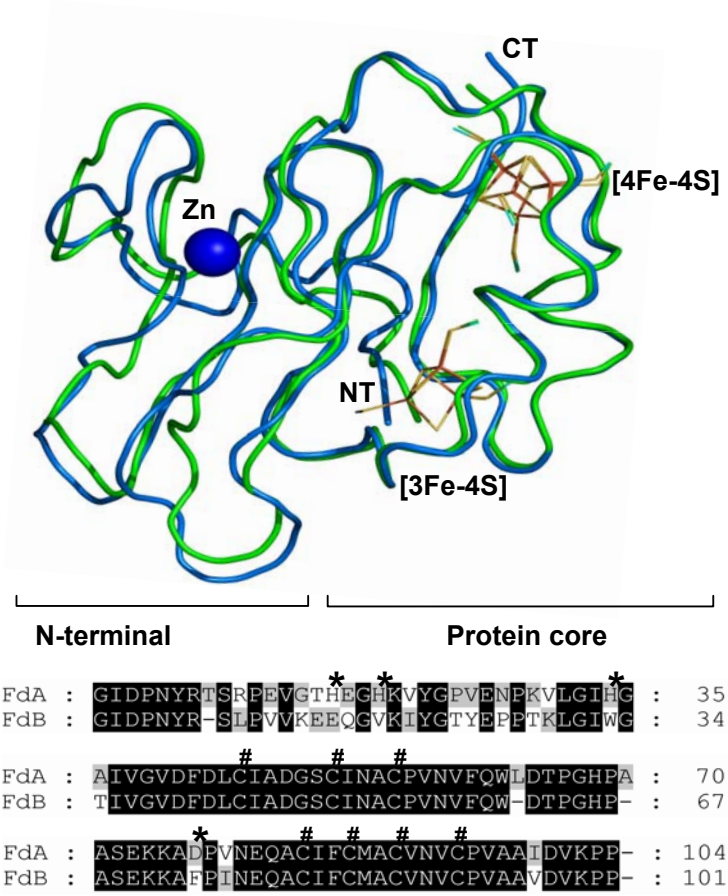


Figure 5.1 |Ferredoxin molecular models and amino-acid sequences.
Top: structural analyses of ferredoxins FdA (dark grey) and FdB (light grey) built by comparative modelling. The sphere represents the Zn^{II} center, which is only present in FdA. The proteins are depicted in an orientation that allows the visualization of the *N*-terminal extension (up to residue 35) and the protein Fe-S harbouring core. CT, *C*-terminus; NT, *N*-terminus. The figure was created with PyMol. Bottom: primary sequences of the ferredoxin isoforms. (*), zinc binding residues; (#), Fe-Scysteine ligands

|Stability of ferredoxin isoforms: a molecular dynamics simulation

Molecular dynamics (MD) simulations were used to investigate the stability of the ferredoxin isoforms structural models, in particular that of the zinc-lacking proteins. Since proteins are very complex systems with multiple minima where the system can easily become trapped during the simulation, the use of a high number of MD simulation replicates greatly improves the sampling of the system

[32]. Therefore, both FdA and FdB were independently simulated in five molecular dynamics replicates of 15 ns each. Inspection of the structural models of the two ferredoxins obtained by comparative modelling, which were the starting points for the simulations, evidences an overall identical fold (figure 5.1). RMSD values of 1.5 and 2.2 Å, respectively for FdA and FdB (results not shown).

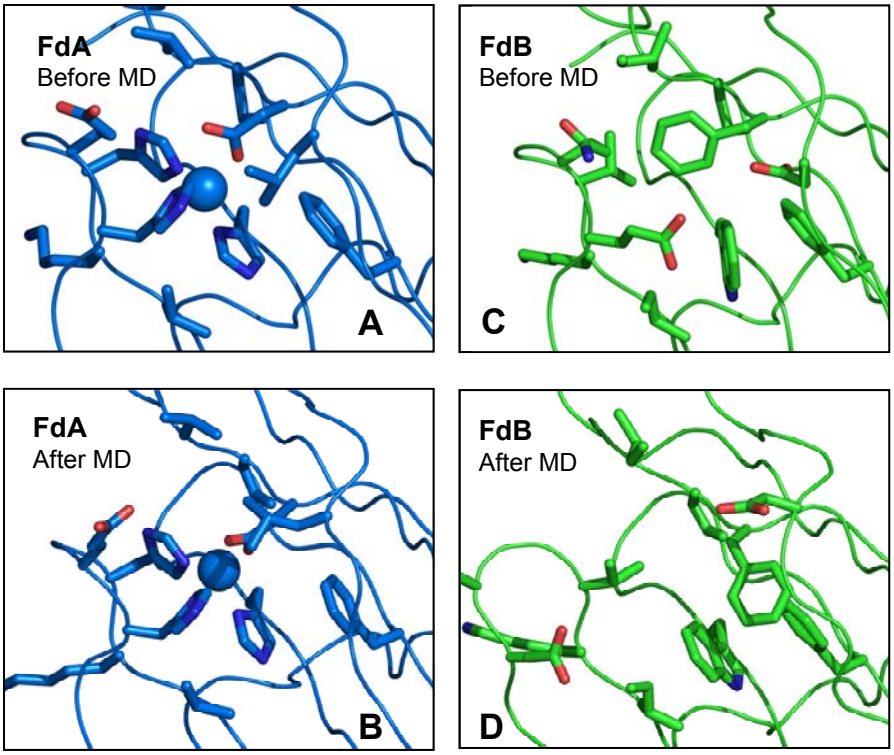


Figure 5.2 | Close-up of the FdA zinc region (A and B; dark gray), and the corresponding FdB region (C and D; light gray), showing the structures directly obtained by comparative modelling (A and C) and those obtained by minimization of the MD simulation average (last 12 ns of each MD replicate) (B and D).
The protein backbone is represented as a line, zinc as a sphere, and selected residues as sticks. The selected residues are within 6 Å from the zinc in FdA and the corresponding region in FdB. The figure was created with PyMOL [33]

During the simulation period, both proteins were relatively stable with maximum C^{α} ferredoxins showed structural changes in the molecular dynamics simulations, and despite the fact that their magnitudes are different, they both occur in the same spatial region. As noted from the analysis of the sequence alignments (figure 5.1), the N-terminal region of the isoforms is the region in which higher

amino acid variability is observed. In agreement, this is the part of the proteins in which the larger differences are observed before (figure 5.2, panel A and C) and after (figure 2.7, panel B and D) the simulations. Although the RMSD of the FdB ferredoxin is a little larger than that of FdA, the RMSD of the region shown in figure 5.2 panels B and D is stable (0.15 Å for FdB and 0.10 Å for FdA, within a sphere with a radius of 6Å from the metal). These computational results are compatible with a stable fold for both isoforms, especially at the N-terminal region of the zinc lacking isoform.

Thermal stability of ferredoxin isoforms

Thermal stability of the zinc-containing FdA and zinc-lacking FdB isoforms was investigated in a broad pH range, between pH 2 and 10. The two ferredoxins are hyperstable proteins: at physiological pH valu es (pH=7), these proteins remain folded even upon incubation with high concentrations of denaturants (e.g. 7M GuHCl or 5M GuSCN) or at extreme temperatures (up to 100°C). This observation is in agreement with previous studies on proteins belonging to this family [12, 16, 34], and it implies that for most thermal unfolding studies some additional destabilisation has to be introduced. Ferredoxin unfolding can be evaluated either by monitoring Fe-S cluster decomposition (bleaching of the 410nm band) or increase and red shift of the Trp fluorescence maximum (figure 5.3). Due to their high stability, only at very acidic pH values (typically below pH 4) are the ferredoxins sufficiently destabilised by extensive protonation to allow unfolding to be observed in the 30-90°C range. In the present study, and in order to study the thermal transition at pH 7 and 10, we had to increase the thermal susceptibility of the ferredoxins by incubating them with different concentrations of GuHCl during the thermal unfolding transitions (table 5.1). In these conditions the thermal

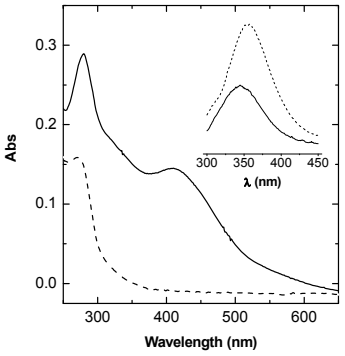


Figure 5.3 Spectroscopic fingerprint of Zn ferredoxin isoforms (FdA and FdB) in the native (–) and denaturated (- - -) states. In the UV-visible spectrum, the band at 410 nm is typical of Fe-S centres and bleaches upon unfolding. The inset shows the emission spectra in arbitary units upon excitation at 280 nm. Either probe yielded identical results in monitoring thermal transitions.

transitions were determined for each denaturant concentration used, and the T_m value in water was calculated by performing linear extrapolations (figure 5.4, table 5.1). For acid pHs the transition midpoints could be determined in the absence of denaturant.

Table 5.1 | Thermal midpoint transitions for ferredoxins determined in the presence of denaturant, at pH 7.

(M)	Fe-S, T_m (ºC)		Trp, T_m (ºC)	
	FdA	FdB	FdA	FdB
3.0	-	74.3	69.3	-
3.5	61.6	-	59.0	-
4.0	55.0	63.1	55.8	63.5
4.5	49.3	57.1	-	53.4
5	-	-	-	51.3
5.25	-	-	-	45.6
5.5	36.2	42.6	-	-
0*	105.9	113.1	108.6	114.4

* Calculated by extrapolation

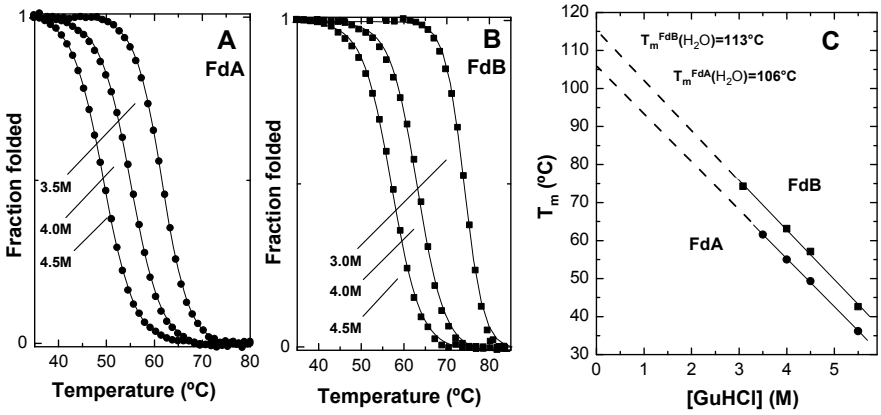
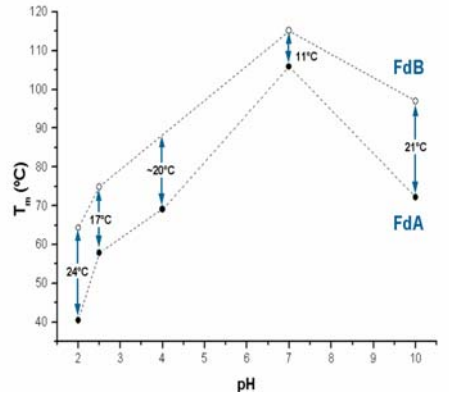


Figure 5.4 | Determination of the thermal midpoint transitions in water for ferredoxin isoforms at pH 7.
Thermal denaturation is monitored following Fe-S clusters collapse in the presence of different GuHCl concentrations as FdA (panel A) anf FdB (panel B) unfold at temperatures above 100ºC. The transition midpoints determined for each denaturant concentrations (Fe-S, T_m in table 5.2) are then used to determine the melting temperatures in water (T_m (H_2O)) which are obtained by linear extrapolation plots (panel C). Identical results are obtained by monitoring Trp emission variations.

The results obtained for the pH dependency study show that the ferredoxin variant lacking the zinc centre has consistently higher melting temperatures at all

tested conditions, and that the stability differences between the two homologous isoforms (ΔT_m) is rather impressive, ranging from 11 to 24°C (figure 5.5). Addition of excess zinc chloride (up to 10mM) during thermal unfolding has no effect on either isoform.



pH	$T_m(H_2O)$ (ºC)		$\Delta T_m(H_2O)$
	FdA	FdB	
2	40.5 ± 0.1	64.3 ± 0.1	23.8 ± 0.1
2.5	57.9 ± 0.1	74.9 ± 0.3	17.0 ± 0.3
4	69.2 ± 0.2	82.4 ± 0.4	13.2 ± 0.4
7*	105.9 ± 1.1	113.1 ± 2.8	9.2 ± 2.3
10*	72.2 ± 3.6	97.0 ± 1.6	24.8 ± 3.9

*Calculated by extrapolation as in figure 5.4 and table 5.1

Figure 5.5 | Variation of the thermal midpoint transitions in water for ferredoxins as a function of pH.
Melting temperature (T_m) dependency on pH for FdA and FdB was plotted in order to better visualize the differences between the two isoforms (ΔT_m) at each pH.

Effect of Iron and Zinc chelators on ferredoxins thermal stability

The role of the zinc centre in ferredoxin stability was investigated in a series of thermal unfolding experiments which were performed in the presence of different metal chelators. For this study, three different molecules with distinct metal specificities were used: DFO, EDTA, and TPEN. The rationale for this experiment was that by using either iron-specific (DFO, Fe^{3+}), zinc-specific (TPEN, Zn^{2+}) or ambivalent (EDTA) metal chelators, we could dissect the relative contribution of the Zn-centre in respect to a possible stabilizing role of the iron-sulfur clusters. Thus, if the zinc site has a stabilising role, the presence of zinc-specific chelators during the thermal unfolding reaction should decrease the T_m due to its scavenging effect. On the other hand, if the iron-sulfur clusters would be playing a stabilising role then, iron-chelating agents should have the same effect. These experiments were carried out at pH 7 and in the presence of 4M GuHCl, in order to achieve thermal transitions in the working temperature range without the need

for extrapolations. The results obtained upon unfolding the two ferredoxins in the presence of different metal chelators are show in table 5.2. It is clear that none of the chelators has a significant effect over FdB thermal stability, which has an overall constant melting temperature (T_m=65±2°C) in all tested conditions (table 5.2). On the contrary, FdA has a significantly lowered stability in the presence of the zinc scavengers TPEN and EDTA (ΔT_m = 12°) but is relatively unaffected by iron chelators. These data suggest that *i.* the iron-sulfur centres do not account for increased stability of FdB; and *ii.* the zinc centre plays an important role in stabilising the isoform that contains it. In summary, the Zn centre is contributing to the enhanced stability of the FdA ferredoxin but the FdB isoform, which lacks it, has devised a more efficient stabilising strategy, which does not seem to involve a direct contribution from the iron-sulfur centres.

Table 5.2 | Effect of zinc and iron chelators on the ferredoxin isoforms transition midpoint temperatures

	FdA		FdB	
	T_m (ºC)	ΔT_m(ºC)	T_m (ºC)	ΔT_m (ºC)
No Chelators	55.0	-	63.1	-
DFO	53.3	1.7	65.9	-2.8
TPEN	44.0	11.0	66.9	-3.8
EDTA	43.1	11.9	65.3	-2.2

Conditions: 10 mM Hepes pH 7, 4M GuHCl, 5mM chelator.
DFO - (desferrioxamine) chelator with specificity for Fe^{3+}; TPEN – N,N,N',N'-tetrakis(2-pyridylmethyl) ethylenediamine) chelator with specificity for Zn^{2+}; EDTA (ethylenediaminetetraacetic acid) ambivalent metal chelator.

5.5 | DISCUSSION

Zinc binding sites are very commonly found in proteins as structural elements, as typified by zinc-finger domains. Among the different types of structural zinc centres, the sites from the ferredoxins from thermophilic Sulfolobales constitute a somehow unique type. In these cases, the metal is tetrahedrally coordinated by His/Asp ligands and it seems to be involved in maintaining the three-dimensional integrity of a local structural domain comprising a α-helix and three β-strands that constitute the typical N-terminal extension found in these proteins. The three imidazole ligands (His 16,19,34 in FdA) are within the β-sheet whereas the fourth carboxylate ligand (Asp76 in FdA) spans

more than 40 residues away from this position and is already located at the surface of the ferredoxin core fold region, which is structurally conserved in respect to other di-cluster ferredoxins, and it is classified within the superfamily of [4Fe4S] ferredoxins (SCOP entry 54862). These zinc centres are unique in iron-sulfur proteins, attaching the N-terminal antiparallel β-sheet domain to the $(\beta\alpha\beta)_2$ core. In this respect this zinc centre, which is essentially solvent inaccessible and is involved in the interaction between these two domains of the ferredoxin, combines features of structural and interfacial zinc centres [3], as it stabilises the N-terminal domain promoting its stable interaction with the protein central region. Mutagenesis of the two first histidine ligands [13] result in loss of the zinc centre, whereas coordination by the aspartic acid was shown to be dispensable [14]. Removal of the zinc centre as well as truncation of the N-terminal extension results in a ferredoxin with decreased thermal stability [13, 14]. The availability of naturally occurring ferredoxin isoforms [12] in *S. metallicus* with (FdA) and without (FdB) the zinc centre constitutes a powerful tool for studying the contribution of this structural element on the ferredoxin fold stability. Here we have shown that at physiological pH, both FdA and FdB undergo thermal unfolding transitions above the boiling point of water. However, one of the most interesting results was the observation that at all pH values tested (2<pH<10) the zinc lacking isoform (FdB) was always more stable than its zinc containing counterpart (FdA). This unexpected observation prompted us to investigate the contribution of each metal cofactor on the protein stability. For that purpose, we added chelators specific either for zinc (TPEN) or iron (DFO) during the thermal unfolding transition. What we concluded was that these have no effect on the zinc lacking isoform (FdB), as the measured thermal transition midpoint was basically invariant in respect to the control (T_m = 65±2°C) showing that the increased stability of this isoform can not be accounted by a stabilising effect from the FeS centres. On the other hand, in the zinc containing isoform (FdA) zinc chelators (TPEN or EDTA) lead to a decrease in the thermal transition midpoint of ~ 11°C, whereas DFO has almost no effect. Thus, it is concluded that in ferredoxins containing His/Asp ligands such as FdA, the zinc centre contributes to the protein stability as its removal results in a less stable protein. On the other hand, in the zinc lacking isoform, not only the protein fold was kept but it was in fact more stable. The remaining key question

now is what could be the molecular determinants which account for the increased stability of the ferredoxin without the zinc centre. Analysis of the structural models at the region where the zinc centre is found (figure 5.6) shows that the spatial region in which the zinc ligands and the metal would be present is now mostly occupied by hydrophobic residues.

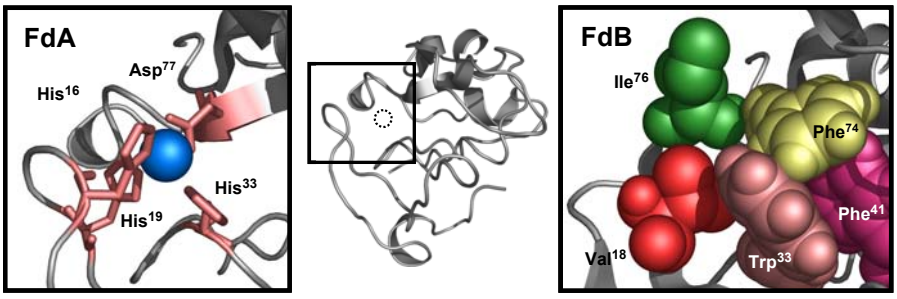


Figure 5.6 |Zinc domain coordination in ferredoxin isoforms.
The FdA isoform comprise a set of conserved His/Asp ligands that chelate a structural Zn^{2+} site (**A**), whereas in FdB, these residues have been replaced by others with higher hydrophobicity (**B**) The central figure shows the common fold of the proteins and the region that is blown up in the lateral panels. The dotted circle points to the location of the zinc ion in FdA.
The figure was prepared with PyMOL [33]

This leads to the suggestion that these novel non polar interactions arising from this unique spatial arrangement of replacing residues are, not only sufficient to stabilise the α-helix/β-sheet N-terminal domain, but in fact result in an enhanced and more efficient stabilisation of the "zinc" domain. It is very interesting to note that a similar strategy has been computationally outlined for stabilising a zinc-finger domain. Mayo and co-workers have designed, expressed and determined the structure of a variant peptide [35, 36] which should fold into a ββα motif based on the backbone structure of a zinc-finger domain. Indeed, the domain was found to be stable without the zinc centre, whose removal in zinc finger proteins results in unfolding [37]. The computed sequence which resulted in a protein with a ββα structure comprised two phenylalanines replacing previous zinc binding ligands which contributed to a well packed hydrophobic core [36]. In the case of the proteins used in our study, nature has made a similar design: in FdB a set of hydrophobic residues (Val18, Phe74, Phe41, Trp33, and Ile76) occupies the spatial region in which the zinc site in found is the FdA isoform (figure 5.6). This

core of hydrophobic residues results in stabilisation of the protein fold, which happens to be more efficient than that provided by the metal cross linking, illustrating the versatility of molecular strategies available to achieve a given three-dimensional structure.

5.6 REFERENCES

1. Petsko, G. and D. Ringe, *Protein Structure and Function*. Primers in Biology. 2004: New Science Press.
2. Hogg, P.J., *Disulfide bonds as switches for protein function.* Trends Biochem Sci, 2003. **28**(4): p. 210-4.
3. Auld, D.S., *Zinc coordination sphere in biochemical zinc sites.* Biometals, 2001. **14**(3-4): p. 271-313.
4. Lee, M.S., et al., *Three-dimensional solution structure of a single zinc finger DNA-binding domain.* Science, 1989. **245**(4918): p. 635-7.
5. Lee, M.S., J.M. Gottesfeld, and P.E. Wright, *Zinc is required for folding and binding of a single zinc finger to DNA.* FEBS Lett, 1991. **279**(2): p. 289-94.
6. Perrier, V., et al., *Zinc chelation and structural stability of adenylate kinase from Bacillus subtilis.* Biochemistry, 1994. **33**(33): p. 9960-7.
7. Perrier, V., et al., *Genetically engineered zinc-chelating adenylate kinase from Escherichia coli with enhanced thermal stability.* J Biol Chem, 1998. **273**(30): p. 19097-101.
8. Green, S.M., et al., *Roles of metal ions in the maintenance of the tertiary and quaternary structure of arginase from Saccharomyces cerevisiae.* J Biol Chem, 1991. **266**(32): p. 21474-81.
9. Olry, A., et al., *Insights into the role of the metal binding site in methionine-R-sulfoxide reductases B.* Protein Sci, 2005. **14**(11): p. 2828-37.
10. Gomes, C.M., et al., *Functional control of the binuclear metal site in the metallo-beta-lactamase-like fold by subtle amino acid replacements.* Protein Sci, 2002. **11**(3): p. 707-12.
11. Fujii, T., et al., *The crystal structure of zinc-containing ferredoxin from the thermoacidophilic archaeon Sulfolobus sp. strain 7.* Biochemistry, 1997. **36**(6): p. 1505-13.
12. Gomes, C.M., Faria, A., Carita, J.C., Mendes, J., Regalla ,M., Chicau ,P., Huber ,H.,Stetter ,K.O. and Teixeira ,M., *Di-cluster, seven iron ferredoxins from hyperthermophilic Sulfolobales.* JBIC, 1998. **3**: p. 499-507.
13. Kojoh, K., H. Matsuzawa, and T. Wakagi, *Zinc and an N-terminal extra stretch of the ferredoxin from a thermoacidophilic archaeon stabilize the molecule at high temperature.* Eur J Biochem, 1999. **264**(1): p. 85-91.
14. Kojoh, K., et al., *Zinc-coordination of aspartic acid-76 in Sulfolobus ferredoxin is not required for thermal stability of the molecule.* J Inorg Biochem, 2002. **89**(1-2): p. 69-73.
15. Leal, S.S. and C.M. Gomes, *Linear three-iron centres are unlikely cluster degradation intermediates during unfolding of iron-sulfur proteins.* Biol Chem, 2005. **386**(12): p. 1295-300.
16. Leal, S.S., M. Teixeira, and C.M. Gomes, *Studies on the degradation pathway of iron-sulfur centers during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability.* J Biol Inorg Chem, 2004. **9**(8): p. 987-96.
17. Moczygemba, C., et al., *High stability of a ferredoxin from the hyperthermophilic archaeon A. ambivalens: involvement of electrostatic interactions and cofactors.* Protein Sci, 2001. **10**(8): p. 1539-48.

18. Stone, K.L., Williams, K.R., *A Pratical Guide to Protein and Peptide Purification for Microsequencing*, ed. P. Matsudaira. 1993: Academic Press.

19. Frazao, C., et al., *Crystallographic analysis of the intact metal centres [3Fe-4S](1+/0) and [4Fe-4S](2+/1+) in a Zn(2+) -containing ferredoxin.* FEBS Lett, 2008. **582**(5): p. 763-7.

20. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.

21. Laskowski, A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* J. Appl. Cryst., 1993. **26**: p. 283-291.

22. Lindahl, E., B. Hess, and D. van der Spoel, *GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis.* J. Mol. Model., 2001. **7**: p. 306-317.

23. Scott, W.R.P., et al., *The GROMOS biomolecular simulation program package.* J. Phys. Chem. A, 1999. **103**: p. 3596-3607.

24. van Gunsteren, W.F., et al., *Biomolecular simulation: The GROMOS96 manual and user guide.* 1996, Zurich, Groninger: BIOMOS b.v.

25. Teixeira, V.H., Baptista,A.M, Soares, C.M., *Pathways of H2 Towards the Active Site of [NiFe]-Hydrogenase.* Biophys. J., 2006. **in press**.

26. Baptista, A.M. and C.M. Soares, *Some Theoretical and Computational Aspects of the Inclusion of Proton Isomerism in the Protonation Equilibrium of Proteins.* J. Phys. Chem. B, 2001. **105**: p. 293-309.

27. Teixeira, V.H., C.M. Soares, and A.M. Baptista, *Studies of the Reduction and Protonation Behavior of the Tetraheme Cytochromes Using Atomic Detail.* J. Biol. Inorg. Chem., 2002. **7**: p. 200-216.

28. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath.* J. Chem. Phys., 1984. **81**: p. 3684-3690.

29. Hess, B., et al., *LINCS: A Linear Constraint Solver for Molecular Dynamics.* J. Comp. Chem., 1997. **18**: p. 1463--1472.

30. Mitou, G., et al., *An Isc-type extremely thermostable [2Fe-2S] ferredoxin from Aquifex aeolicus. Biochemical, spectroscopic, and unfolding studies.* Biochemistry, 2003. **42**(5): p. 1354-64.

31. Hermans, J., et al., *A consistent empirical potential for water-protein interactions.* Biopolymers, 1984. **23**: p. 1513-1518.

32. Frazao, C., et al., *Structure of a dioxygen reduction enzyme from Desulfovibrio gigas.* Nat Struct Biol, 2000. **7**(11): p. 1041-5.

33. DeLano, W.L., *The PyMOL User's Manual*. 2002, San Carlos, CA: DeLano Scientific.

34. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.

35. Dahiyat, B.I., C.A. Sarisky, and S.L. Mayo, *De novo protein design: towards fully automated sequence selection.* J Mol Biol, 1997. **273**(4): p. 789-96.

36. Dahiyat, B.I. and S.L. Mayo, *De novo protein design: fully automated sequence selection.* Science, 1997. **278**(5335): p. 82-7.

37. Pfeil, W., et al., *Ferredoxin from the hyperthermophile Thermotoga maritima is stable beyond the boiling point of water.* J Mol Biol, 1997. **272**(4): p. 591-6.

5.7 ACKNOWLEDGMENTS

To Rita Rocha from PBFS group at ITQB- Portugal for the conjoined studies on the stability of the di-cluster ferredoxin isoforms from *Sulfolobus metallicus*

To Cláudio Soares, Vitor H. Teixeira and Antónia Baptista from ITQB-Portugal for the molecular dynamics simulations on the di-cluster ferredoxin isoforms from *Sulfolobus metallicus*

To Manuella Regalla from ITQB- Portugal for sequencing the di-cluster ferredoxin isoforms from *Sulfolobus metallicus*

To Harald Huber for providing the *Sulfolobus metallicus* cells.

6 On the Relative Contribution of Ionic Interactions Over Iron-Sulfur Clusters to Ferredoxin Stability

CONTENTS

This chapter was published in

Sónia S. Leal and Cláudio M. Gomes
"On the relative contribution of ionic interactions over iron-sulfur clusters to ferredoxin stability"
BBA – Proteins and Proteomics *in press* doi:10.1016/j.bbapap.2008.05.001 (2008)

6.1 SUMMARY

Metal centres play an important structural role in maintaining the native conformation of a protein. Here we use biophysical methods to investigate what is the relative contribution of iron–sulfur clusters in respect to ionic interactions in the thermophilic di-cluster ferredoxin model. Changes in protonation affect both the stability and the conformational dynamics of the protein fold. In the pH 5.5 to 8 interval, the protein has a high melting temperature (T_m~120°C), which decreases towards pH extremes. Acidification triggers events in two steps: down to the isoelectric point (pH 3.5) the Fe-S clusters remain unchanged, the secondary structure content increases and the single Trp becomes more solvent shielded, denoting a more compact fold. Further acidification down to pH 2 sets off exposure of the hydrophobic core and Fe-S cluster disintegration, yielding a molten globule state. The relative stabilising contribution of the clusters becomes evident when stabilising ionic interactions are switched off as a result of poising the protein at pH 3.5, at an overall null charge: under these conditions, the Fe-S clusters disassemble at T_m= 72°C, whereas the protein unfolds at T_m= 52°C. Overall, this ferredoxin denotes a considerable structural plasticity around its native conformation, a property which appears to depend more on the integrity of its metal clusters rather than on the status of its stabilising electrostatic interactions. The latter however play a relevant role in determining the protein thermal stability.

6.2 INTRODUCTION

Electrostatic interactions are essential to determine the conformational characteristics and stability of a protein. The extent of this contribution can be modulated by varying the pH, as a result of the changes in the protein net charge due to the alteration of the protonation state of key residues. Ultimately, at extreme pH values, stabilising ionic interactions are broken and destabilising repulsive forces may arise as a result of the spatial proximity of groups with identical charges. An alteration in the net charge of a protein can result in non-native conformational states such as the molten globules or result in protein

unfolding [1]. Hence, the determination of the effect of pH changes on the structure and conformation of a protein is a tool to investigate the contribution of electrostatic interactions on the overall stability of the protein. Additionally, in metalloproteins, metallic centres are also known to contribute significantly to protein stability, thus having an important role in the maintenance of a particular structural fold [2-4]. Understanding the intertwining of the stabilisation effect arising from a metal site, in the context of a network of electrostatic interactions in a protein, is essential towards a better understanding of metalloprotein folding and stability. Following this goal, the *Acidianus ambivalens* ferredoxin (*Aa*Fd) that contains a [3Fe-4S] and a [4Fe-4S] cluster in addition to a Zn^{II} His/Asp site was used as a model. This metalloprotein is composed of 103 amino acids and belongs to the archaeal family of di-cluster ferredoxins [5, 6]. Structurally, it comprises a core domain containing the two Fe-S clusters (residues 37–103) in a canonical $(\beta\alpha\beta)_2$ fold, and an N-terminal extension (residues 1–36) conjointly held by a Zinc centre (see figure 2.4 in chapter 2 ,PBD: 2vkr). The structural Zinc centre presents a tetrahedral coordination by three His residues in the N-terminal segment and one Asp located in the protein core [7], thus making the centre inaccessible from the protein surface. This ferredoxin has an extremely high thermal stability at neutral pH (T_m ~120°C), that nevertheless becomes significantly reduced in acidic conditions [8]. In this work we have designed a series of experiments aimed at contributing to understand to which magnitude electrostatic interactions and the Fe-S clusters modulate the conformational stability and folding of this ferredoxin.

6.3 MATERIALS AND METHODS

Protein Purification

The di-cluster ferredoxin was purified from *Acidianus ambivalens* grown at pH 2.5 and 80 °C [5]. A cytosolic fraction of the cells was processed and the protein purified in two consecutive chromatographic steps as previously described [6]. Protein purity was evaluated from the Abs410/Abs280 ratio (~0.6) and SDS-

PAGE. The integrity of both iron–sulfur clusters was evaluated by TXRF metal analysis (B.O. Kolbesen, Johann Wolfgang Goethe-Universität, Frankfurt), which yielded 6±0.5 Fe per mol of ferredoxin. Protein stocks were kept as concentrated aliquots at −20°C (40 mM potassium phosphate, 150 mM NaCl, pH 6.5).

Thermal denaturation

Thermal transitions were monitored by visible absorption, far and near-UV CD and fluorescence methods using between 5 and 20 μM of ferredoxin. The apparent midpoint transition (T_m^{app}) at each pH was determined by direct analysis of the spectroscopic signal vs. temperature and curve fitting to a sigmoid transition. Due to the high thermostability of the protein in the pH range 4 to 10, different GuHCl concentrations (at least 4 different concentrations) were used to determine the T_m^{app} in water for ferredoxin at each pH, from a linear extrapolation to 0M GuHCl. Typically, the temperaturewas increased from 25 to 95 °C at 1°C/min.

ANS binding

Ferredoxin samples (typically 5 μM) were incubated with 250 μM 1-anilinonaphthalene-8-sulfonic acid (ANS) during 15–60 min prior to measurements. ANS fluorescence emission enhancement of bound ANS at 480 nm was evaluated upon excitation at 370 nm, and corrected for the background emission of a protein-free ANS solution [9].

Equilibrium pH titration

Ferredoxin samples were incubated for 24 h at 4 °C at different pH values. For each pH, an adequate buffer system (40 mM) was chosen: glycine (pH 1.5 - 3.5), sodium acetate (pH 4 - 5.5), phosphate, (pH 6 - 8), and Tris (pH 8.5 - 9.5).

CD spectroscopy

CD spectra were recorded in a Jasco J-815 spectropolarimeter equipped with Peltier temperature control and cell stirring. Far UV-CD measurements (protein

concentration, 0.1 mg ml^{-1}) were obtained by monitoring the mean residue ellipticity at 222 nm and recorded using 1 mm path length polarimetrically certified cells (Hellma). Near-UV CD spectra (protein concentration 0.2 mg ml^{-1}) were measured using rectangular 10 mm path length cells.

Visible spectroscopy

Visible spectra were recorded in a Shimadzu Multispec-1501 diode array spectrophotometer equipped with temperature control by a Julabo 4 MV water bath. FeS cluster integrity was monitored following the decay of the absorption band at 410 nm (ε410=30,400 M^{-1} cm^{-1}, [8]). Protein concentration ranged from 5–10 μM.

Fluorescence spectroscopy

Fluorescence measurements were recorded on a Cary Varian Eclipse instrument ($slit_{ex}$: 5 nm, $slit_{em}$: 10 nm unless otherwise noted) equipped with Peltier temperature control. Protein concentration ranged from 5–10 μM for Trp emission studies.

6.4 RESULTS

Protein conformational changes occurring on *Aa*Fd upon altering the protein net charge through equilibrium pH titrations were monitored using different biophysical methods. These allowed analysis of events at different structural levels and changes at the metal clusters. Alterations on the protein secondary and aromatic moieties were investigated by far and near-UV CD, respectively, whereas fluorescence emission studies of the single Trp residue (Trp 62), located nearby the Fe-S clusters, allowed monitoring changes in the tertiary contacts within the metal cluster environment. Enhancement of 1,8- ANS fluorescence also provided a tool to investigate the exposure of hydrophobic regions. The status of the Fe-S clusters was directly monitored using visible absorption spectroscopy as well as the more detailed visible CD fingerprint.

Electrostatic effects on the thermal stability of Fe-S clusters

The effect of changes in the protein charge on the Fe-S clusters was investigated by monitoring directly the effect of pH variations on the thermal stability of the centres. For the purpose, the variation of the 410 nm absorption band as a function of temperature was used as a reporter for evaluating the integrity of the clusters in the pH 2–12 range. Since the protein is stable beyond the boiling point of water, approximately from pH 4 to 10, for these cases the melting transitions of the clusters (T_m Fe-S) were obtained by measuring the apparent T_m values at a given GuHCl concentration below the $[GuHCl]_{1/2}$. This destabilisation affords measuring the thermal transition within the working temperature range (≈25–95 °C) from which the T_m Fe-S in water are obtained by linear extrapolation (see figure 6.1). This is a commonly used method to experimentally determine the T_m in water for proteins unfolding above 100 °C (e.g. [10, 11]) although this could also be achieved using special high-pressure cells adapted to spectroscopic measurements, which allow thermal unfolding to be carried out at up to 180 °C, as in [12].

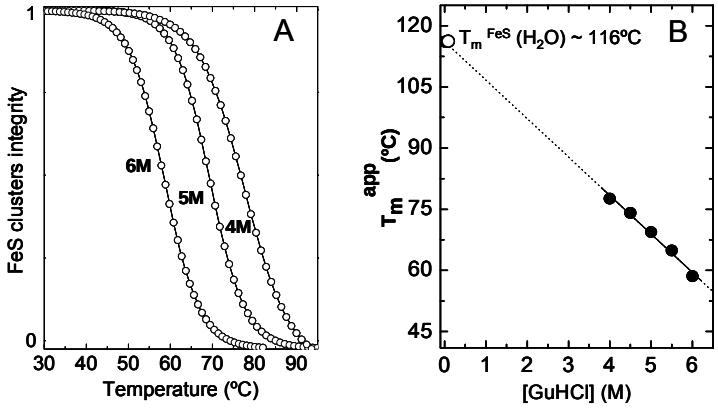


Figure 6.1 | Determination of the thermal midpoint transitions for ferredoxin between pH 5 to 10 where the protein is stable beyond the boiling point of water.
Panel **A** - Thermal denaturation at pH 7 monitoring FeS clusters collapse in the presence of different GuHCl concentrations in order to attain the apparent transition midpoints values. The solid lines represent a two state sigmoidal fit. Panel **B** – Apparent transition midpoints (T_m^{app}) as function of GuHCl concentration applied to determine the melting temperature in water ($T_m^{FeS}(H_2O)$) which is obtained by a linear extrapolation plot [T_m = 58.4ºC (6M GuHCl); T_m = 65ºC (5.5M GuHCl); T_m =69.5ºC (5M GuHCl); T_m = 74.3ºC (4.5M GuHCl); T_m = 78ºC (4M GuHCl);] Linear extrapolation yields a $T_m^{FeS}(H_2O)$ of 116.2ºC.

Thermal unfolding studies were carried out at different pH values. The plot of the

cluster's midpoint transition versus pH (figure 6.2), shows that the highest stability (T_m Fe-S ≈ 120 °C) is obtained between pH 5.5 and 8.

Considering that the centre of this interval is at ~pH 6.8, then acidification is more destabilising than alkalinisation: at pH 11.5, the melting point is 25 °C higher than the melting temperature measured at pH 2.5. Analysis of the *Aa*Fd structure [7] provides some rationale for this observation. First, the protein has no notorious superficial basic residues near the Fe-S clusters, whose deprotonation would be likely to influence the stability of the centres. On the other hand, we have previously hypothesised, based on a homology model, that some residue pairs could be involved in stabilising ionic interactions: Asp99–Arg7, Asp41–Arg10, Asp64–Lys73 and Asp43–His68 [13]. In light of the available crystal structure [7], the average distance between the donor and acceptor groups in the pair Asp64–Lys73 is relatively high (~6.8 Å) and these residues may not be involved in an electrostatic interaction. For the Asp41–Arg10 pair, inspection of the structure shows that the position of the side chains is very suggestive of an interaction, although the distance between groups (4.7 Å) is larger than the usually accepted value (<3.5 Å, [14]). Further, these residues are at the protein surface and peripheral salt bridges are known to be potentially stabilizing [15, 16]. These ionic interactions could be disrupted upon acidification, thus contributing more to protein destabilisation. An additional stabilising effect may arise also from the interaction of aliphatic parts of the side chains of charged residues at the protein surface. Finally, direct protonation of the His/Asp zinc site coordinating residues could also play a destabilising role. In fact, in the preceding thermal

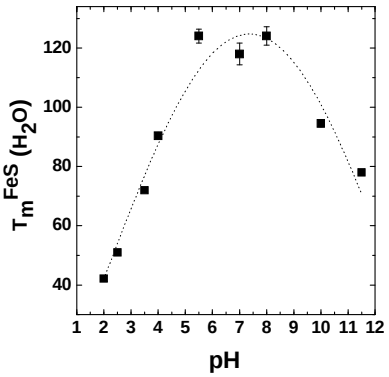


Figure 6.2 | pH dependence of [FeS] clusters melting temperature (T_m FeS).
Thermally induced equilibrium denaturation curves were obtained using the 410 nm absorption band as the probe for ferredoxin [FeS] cluster collapse. In the pH interval 5 to 10 the melting temperature in water was obtained through extrapolation (as in figure 2.11) [T_m=42.2 °C (pH2); T_m =52 °C (pH2.5); T_m =72 °C (pH 3.5); T_m =91±1 °C (pH 4); T_m =124±2 °C (pH 5); T_m =118±4 °C (pH 7); T_m =124±3 °C (pH 8); T_m =95±1 °C (pH 10); T_m =78 °C (pH 11.5)].
A dotted line is drawn to guide the eye.

stability study carried out on the related ferredoxin isoforms from *S. metallicus* (which differ almost exclusively on the presence or absence of the zinc centre ligands and zinc site) showed that the isoform carrying the Zinc centre (FdA) was considerably more destabilised by acidic conditions (T_m≈41 °C at pH 2) than the ferredoxin lacking the Zinc centre (FdB), which was more stable (T_m≈64 °C) [17]. In addition, it has also been shown that during thermal unfolding, the specific zinc scavenger TPEN significantly lowers the T_m in FdA by ~10 °C, clearly indicating that the zinc site contributes to ferredoxin stability. Thus, destabilisation of zinc binding as a result of protonation of its ligands is very likely to lead to centre disassembly.

Conformational Changes upon equilibrium pH titrations

In order to understand the nature of the conformational changes observed upon a modification of the electrostatic interactions, a direct investigation of the changes occurring at the protein structural level was carried out. For the purpose, and in contrast with other studies involving pH changes [13], a long term effect after overnight incubations at different pH values was investigated, rather than after a sudden pH change. This allows monitoring of effects after equilibrium is reached. Trp fluorescence emission and far-UV CD evidenced an identical behaviour as a function of the incubation pH (figure 6.3). On decreasing the pH from neutrality to approximately pH 3.5, variations on the AaFd native structure were observed. Monitoring of the protein secondary structure content showed a decreasing intensity of the CD signal ellipticity at 222 nm, which is suggestive of an

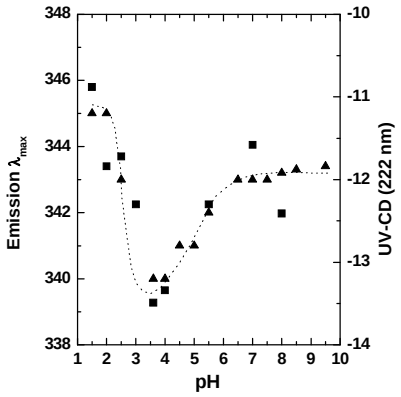


Figure 6.3 Ferredoxin conformational changes with pH.
Trp fluorescence maximum emission wavelength shift (▲) and far UV-CD ellipticity at 222 nm (■) are represented as a function of pH. The protein was previously incubated for 24 h in the respective buffering pH before each measurement. A dotted line is drawn to guide the eye.

increasing content in secondary structure. Also, a blue shift of the Trp fluorescence emission maximum, from 343 to 340 nm, was noted, indicating a less solvent exposed aromatic moiety of the amino acid. Further acidification until pH 2.5 appears to restore the ferredoxin native-like conformational characteristics. Acidification below pH 2–2.5 sets off substantial structural changes suggestive of further destabilisation events, as indicated by loss of secondary structure and a solvent exposed Trp moiety, as shown by the red shifted emission band with a maximum at 345 nm. Monitoring the near-UV CD spectrum showed minor conformational changes from pH 7.0 to 3.5 (figure 6.4) corroborating the evidence that tertiary contacts are kept in this pH range.

Additionally, the visible CD spectrum provided valuable information about changes in or near the [3Fe-4S] and [4Fe-4S] clusters. In comparison with the single broad band at 410 nm which is observed by visible absorption spectroscopy, a complex pattern of four bands with maxima at 365, 470 and 575 nm and minima at 422 nm is obtained in the visible CD region. These features are attributable to both Fe-S clusters, but a band cannot be specifically assigned to a particular cluster. At pH 3.5, most of the cluster's features are retained, with the exception of the band at 422 nm which is slightly less intense, suggesting some destabilisation of the clusters. A more detailed structural understanding of the conformational changes occurring near the isoelectric point can be provided by complementary biophysical studies such as NMR and FT-IR, which are currently underway in our laboratory; these studies will hopefully contribute to an atomistic understanding of the changes occurring in Fd conformation and dynamics at zero net charge. At pH 1.5 tertiary

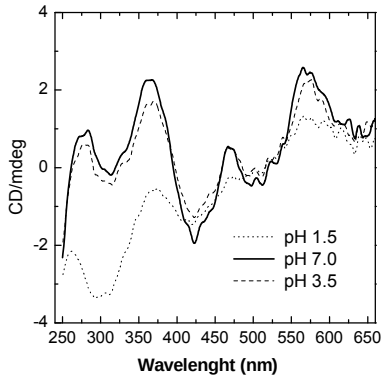


Figure 6.4 | Near-UV and visible CD spectra of ferredoxin at different pH.
The near-UV CD (250–300 nm) is due to the contribution of the aromatic residues probing variations on the protein tertiary structure contacts. In the visible CD region (300–700 nm) the observed four bands are ascribed to the [FeS] clusters; pH 7 (solid line); pH 3.5 (dashed line); pH 1.5 (dotted line).

contacts are absent and the Fe-S have disintegrated.

Influence of net charge on Fe-S cluster and hydrophobic core

A detailed analysis of the effect of acidification on the Fe-S clusters was obtained monitoring the 410 nm band in the visible absorption spectrum, which allowed to probe directly the integrity of Fe-S clusters. Previous studies have suggested that the release of the metal clusters results in the exposure of an hydrophobic surface, as evidenced by an enhancement of the 1-anilino naphthalene-8-sulfonic acid (ANS) fluorescence [18, 19]. Again, using samples after overnight equilibration at the desired pH, the stability of the clusters and ANS binding were investigated. The results obtained (figure 6.5) suggest that pH acidification from pH 7 down to 3.5 has no significant effect on the Fe-S cluster integrity, and that no hydrophobic patches are exposed. Nevertheless, further pH drop from 3.5 to 1.5 resulted in a marked increase in the ANS fluorescence emission that is concomitant with the disintegration of Fe-S clusters. In fact, at pH 1.5 the *Aa*Fd presents only ~10% of the intensity of the 410 nm band (data not shown), and a ~20 fold increase in the ANS emission, in comparison to the native state at pH 7. These results clearly indicate that the acidification below pH 3.5 is enough to disintegrate the Fe-S clusters, making hydrophobic regions simultaneously accessible.

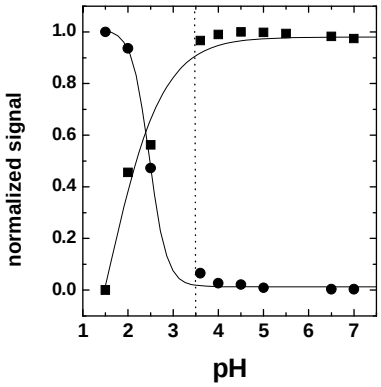


Figure 6.5 pH effect on the integrity of [FeS] clusters and hydrophobic core
Normalised variations of the ANS fluorescence emission enhancement (●) and absorption at 410 nm (■). The solid lines are drawn to guide the eye. The vertical dotted line is set at pH=3.5.

Contributions of protein folding and Fe-S clusters to thermal stability at zero net charge

At pH 3.5, and although the stabilising electrostatic interactions have been overturned due to protonation of protein side chain groups, the overall structure and Fe-S clusters of AaFd are retained in respect to the native state at physiological pH. This particular condition is therefore particularly useful to evaluate the relative contribution of the protein fold versus Fe-S clusters to the overall ferredoxin stability. Indeed, the apparent T_m determined for the Fe-S clusters (T_m Fe-S =72 °C) is substantially higher than that measured from the changes at the secondary structure level (T_m $CD_{216\ nm}$=52 °C) (figure 6.6) circles and squares, respectively). This difference suggests that abolishing electrostatic interactions by poising the protein at an overall null charge, lowers the stability of the protein fold more than the stability of the clusters. Under this particular condition, the clusters afford a stabilisation of ≈+20 °C, in addition to that of the protein structure. Interestingly, these two contributions are perceived when the thermal unfolding is monitored from the variation of the CD signal in the 360–365 nm region in which the contributions from aromatics and Fe-S clusters overlap (figure 6.6, triangles).

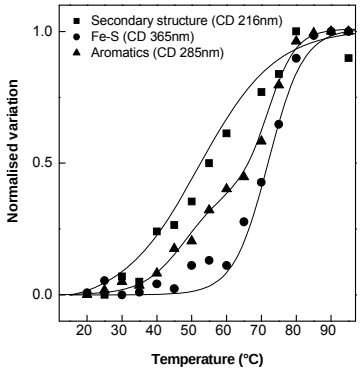


Figure 6.6 Ferredoxin thermal transitions at the isoelectric point (pH 3.5).
Thermal unfolding was followed monitoring the variation of the CD signal at 216 nm(■), the [FeS] CD band at 365 nm (●) and the aromatics CD signal at 285 nm (▲). (■, T_m =52 °C; ▲, T_m=52 °C (35%) and T_m =72° (65%); ●, T_m =72 °C).

6.5 DISCUSSION

Metal centres play an important structural role in maintaining the native conformation of a protein. This is well known for the classical structural metal ions such as for e.g. zinc centres, but also holds true for more complex cofactors such as hemes and iron–sulfur centres which are primarily involved in a catalytic or

electron transferring function. From a fundamental viewpoint, understanding the intertwining effect that arises from combining non-covalent forces such as electrostatic and hydrophobic interactions, with the structural stabilisation resulting from complexing a metal ion or a cluster to a polypeptide, is an essential issue towards the understanding of the folding and stability of metalloproteins. We have performed a study focusing on evaluating protein conformational changes and Fe-S cluster stability, upon poising the protein at different pH values. Changes in the protonation state of acid and basic residues will change the network of electrostatic interactions that are involved in the stabilisation of ferredoxin, which would allow discriminating the relative contribution of electrostatic interactions and metal centres to overall stability. Different structural and conformational changes are observed upon ferredoxin acidification. Acidification down to pH 3.5, which corresponds to the protein isoelectric point (pI 3.6, [5]), results in a conformational change as inferred from an increase of the 222 nm CD signal, a blue shift in Trp62 emission (figure 6.3), and an alteration on the pattern of near-UV CD spectrum (figure 6.4). These effects may result from structural packing due to protonation, presumably of Asp and Glu residues. In fact, Trp62 is located nearby Asp48 (≈4.7 Å) and Asp64 (≈5.4 Å). Protonation of the later could lead to a rearrangement of the β-harpin in which it is involved, and in which Trp62 is also located. Interestingly this pH change has a minor affect on the Fe-S clusters, as the intensity of the 410 nm band remains unchanged. The same was confirmed by preliminary ^{1}H NMR data, which showed that there were no changes in the chemical shifts of the β-CH_2 protons from the cluster ligand cysteines. The only exception was a slight shift in the position corresponding to Cys-93, which is involved in [3Fe-4S] coordination (C. Salgueiro et al., personal communication). This residue is very close to Asp48 (≈5.4 Å), so this observation is compatible with a conformational adjustment around the cluster moiety, as a result of protonation. At pH 3.5, which corresponds to zero net charge, the Fe-S clusters have clearly an important contribution to the overall stability. Under these conditions, the apparent Tm at which the metal centres disrupt (T_m≈ 72 °C), is substantially higher than that determined for the loss of overall secondary structure (T_m≈ 52 °C). This result shows that effectively, at a null protein charge, the Fe-S clusters afford an additional stabilising contribution to the protein.

Nevertheless, further acidification results in a severe decrease in the cofactor stability. In fact, dropping down to pH 1.5 results in the disassembly of the Fe-S clusters leading to the exposure of the protein hydrophobic core as evidenced by ANS enhancement and structure loss. Nevertheless, some secondary structure is retained in comparison to the guanidinium hydrochloride unfolded state (data not shown). The disintegration of the Fe-S clusters appears therefore to trigger a structural rearrangement of the apo-state, leading to a conformation with molten globule properties. In fact, *Aa*Fd incubation at pH 1.5 results in poor tertiary contacts, substantial secondary structure and strong ANS binding, which are structural characteristics of molten globule states [20]

6.6 REFERENCES

1. Fink, A.L., et al., *Classification of acid denaturation of proteins: intermediates and unfolded states.* Biochemistry, 1994. **33**(41): p. 12504-11.
2. Bushmarina, N.A., et al., *Cofactor effects on the protein folding reaction: acceleration of alpha-lactalbumin refolding by metal ions.* Protein Sci, 2006. **15**(4): p. 659-71.
3. Robinson, C.R., et al., *Energetics of heme binding to native and denatured states of cytochrome b562.* Biochemistry, 1997. **36**(51): p. 16141-6.
4. Wittung-Stafshede, P., *Role of cofactors in protein folding.* Acc Chem Res, 2002. **35**(4): p. 201-8.
5. Teixeira, M., et al., *A seven-iron ferredoxin from the thermoacidophilic archaeon Desulfurolobus ambivalens.* Eur J Biochem, 1995. **227**(1-2): p. 322-7.
6. Gomes, C.M., Faria, A., Carita, J.C., Mendes, J., Regalla ,M., Chicau ,P., Huber ,H.,Stetter ,K.O. and Teixeira ,M., *Di-cluster, seven iron ferredoxins from hyperthermophilic Sulfolobales.* JBIC, 1998. **3**: p. 499-507.
7. Frazao, C., et al., *Crystallographic analysis of the intact metal centres [3Fe-4S](1+/0) and [4Fe-4S](2+/1+) in a Zn(2+) -containing ferredoxin.* FEBS Lett, 2008. **582**(5): p. 763-7.
8. Moczygemba, C., et al., *High stability of a ferredoxin from the hyperthermophilic archaeon A. ambivalens: involvement of electrostatic interactions and cofactors.* Protein Sci, 2001. **10**(8): p. 1539-48.
9. Semisotnov, G.V., et al., *Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe.* Biopolymers, 1991. **31**(1): p. 119-28.
10. Burova, T.V., et al., *Conformational stability of adrenodoxin mutant proteins.* Protein Sci, 1996. **5**(9): p. 1890-7.
11. Zeeb, M., et al., *Folding and association of an extremely stable dimeric protein from Sulfolobus islandicus.* J Mol Biol, 2004. **336**(1): p. 227-40.
12. Uchiyama, S., et al., *Complete thermal-unfolding profiles of oxidized and reduced cytochromes C.* J Am Chem Soc, 2004. **126**(45): p. 14684-5.
13. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.
14. Petsko, G., Ringe, D., *Protein Structure and Function*. 2004: New Science Press.

15. Ragsdale, S.W. and M. Kumar, *Nickel-Containing Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase(,).* Chem Rev, 1996. **96**(7): p. 2515-2540.

16. Makhatadze, G.I., et al., *Contribution of surface salt bridges to protein stability: guidelines for protein engineering.* J Mol Biol, 2003. **327**(5): p. 1135-48.

17. Rocha, R., et al., *Natural domain design: enhanced thermal stability of a zinc-lacking ferredoxin isoform shows that a hydrophobic core efficiently replaces the structural metal site.* Biochemistry, 2006. **45**(34): p. 10376-84.

18. Carneiro, F.A., A.S. Ferradosa, and A.T. Da Poian, *Low pH-induced conformational changes in vesicular stomatitis virus glycoprotein involve dramatic structure reorganization.* J Biol Chem, 2001. **276**(1): p. 62-7.

19. Cavagnero, S., et al., *Response of rubredoxin from Pyrococcus furiosus to environmental changes: implications for the origin of hyperthermostability.* Biochemistry, 1995. **34**(31): p. 9865-73.

20. Fink, A.L., *Molten Globule*. 2001: Encyclopedia of Life Sciences.

6.7 ACKNOWLEDGMENTS

To Carlos Salgueiro from Universidade Nova de Lisboa-Faculdade de Ciências e Tecnologia-Departamanto de Química for preliminary NMR experiments.

7 FERREDOXIN MOLTEN GLOBULE STATE: STRUCTURAL CHARACTERISATION AND IMPLICATIONS ON PROTEIN FOLDING AND FE-S CENTRE ASSEMBLY

CONTENTS

This chapter was published in

Sónia S. Leal and Cláudio M. Gomes
Studies of the molten globule state of ferredoxin: Structural characterization and implications on protein folding and iron-sulfur center assembly
Proteins 68*: 606-616* (2007)

7.1 SUMMARY

The biological insertion of iron-sulfur clusters involves the interaction of (metallo) chaperons with a partly folded target polypeptide. In this respect, the study of nonnative protein conformations in iron-sulfur proteins is relevant for the understanding of the folding process and cofactor assembly. We have investigated the formation of a molten globule state in the [3Fe-4S][4Fe-4S] ferredoxin from the thermophilic archaeon *Acidianus ambivalens* (*Aa*Fd), which also contains a structural zinc site. Biophysical studies have shown that, at acidic pH, *Aa*Fd retains structural folding and metal centers. However, upon increasing the temperature, a series of successive modifications occur within the protein structure: [FeS] disassembly, loss of tertiary contacts and dissociation of the Zn^{2+} site, which is simultaneous to alterations on the secondary structure. Upon cooling, an apo-ferredoxin state is obtained, with characteristics of a molten globule: compactness identical to the native form; similar secondary structure evidenced by far-UV CD; no near-UV CD detected tertiary contacts; and an exposure of the hydrophobic surface evidenced by 1-anilino naphthalene-8-sulfonic acid (ANS) binding. In contrast to the native form, this apo ferredoxin state undergoes reversible thermal and chemical unfolding. Its conformational stability was investigated by guanidinium chloride denaturation and this state is ~ 1.5 kcal mol^{-1} destabilised in respect to the holo ferredoxin. The single tryptophan located nearby the Fe-S pocket probed the conformational dynamics of the molten globule state: fluorescence quenching, red edge emission shift analysis and resonance energy transfer to bound ANS evidenced a restricted mobility and confinement within a hydrophobic environment. The possible physiological relevance of molten globule states in Fe-S proteins and the hypothesis that their structural flexibility may be important to the understanding of metal center insertion are discussed.

7.2 INTRODUCTION

Protein conformations can range from the tightly packed, homogenous native state to the unstructured, unfolded form which comprises a very large ensemble of conformations. In between these extremes, more or less structured protein states can be found, which are presumably intermediates of the folding process and encompass the so called molten globules. Since their original descriptions [1, 2] a large number of protein molten globule states has been described and it is now established that these partly folded forms of proteins are characterised by a defined set of common properties. A typical molten globule is defined by: i. secondary structure content comparable to that of the native state; ii. absence of most of the tertiary structure; iii. large exposed hydrophobic surface area as a result of a loosely packed hydrophobic core; and iv. compactness comparable to that of the native state, with only slightly increased radius (<20%) [3, 4]. The role of molten globule states in the protein folding process is far from being consensual: whereas some models postulate that all proteins fold through a common type of intermediate related to molten globules [5], other views consider that folding intermediates depend exclusively on the protein primary sequence and are therefore dissimilar in respect to their structural properties [4, 6], or simply that there are no on-pathway intermediaries of folding, and that all apparent intermediates are either misfolded species or compact forms of the unfolded state (reviewed in [4]).

Folding and conformational stability studies of iron-sulfur proteins [7] have received an increasing attention in recent years [8-21], in recognition of the importance of addressing the conformational dynamics of these proteins and of the different types of inorganic Fe-S centres that they accommodate. Apart from their functional role, iron-sulfur centres (Fe-S) also are key structural and stabilising elements that act as relevant nucleation points in the folded protein state [17, 22]. In fact, many Fe-S proteins are unable to maintain a native-like conformation when converted to an apo form [23]. However, little is know about the conformational properties and structural characteristics of non-native states of Fe-S proteins, and in particular molten globule conformations are poorly characterised. A molten globule-like partly unfolded state of the [4Fe-4S]

containing high potential iron-sulfur protein from *Chromatium vinosum* obtained at ~4M GdmCl, has been characterised by NMR spectroscopy. In this state the protein has a collapsed secondary structure [10] but it remains in the holo form as shown by EXAFS fingerprinting of an intact cluster [22]. Other protein which is reported to exist in multiple discrete conformers which have some characteristics of molten globule is the IscU protein which is involved in the Fe-S biosynthetic pathway. This protein acts as a scaffold for the assembly of intermediate Fe-S clusters and it mediates its delivery to the final apo-targets. It was found to be a structured protein, hosting a [2Fe2S] centre [24] and it adopts a mobile molten-globule state [25] which is biologically relevant in the context of the protein-protein interactions in which IscU is involved.

We have directly tackled this aspect by investigating conditions that would result in molten globule formation in an iron sulfur protein. As a model we selected the thermophilic di-cluster ferredoxin (AaFd) from *Acidianus ambivalens* (Topt=80°C), which has been extensively characterised in biochemical, biophysical and conformational terms [12, 16, 17, 26-28]. Structurally, this protein comprises 105 amino acids arranged in two distinct regions: a core domain with a $(\beta\alpha\beta)_2$ fold comprising the [3Fe-4S] and [4Fe-4S] centres, which is wrapped by a 30 amino acid long N-terminal extension in which a His/Asp zinc site is found. This topological arrangement accounts for a very high stability of the AaFd fold: the protein has a very high thermal stability (Tm≈120°C at pH 7), a broad pH stability and a high resistance to chemical denaturants (it remains folded in 7M GdmCl at pH 7) [12, 17, 28]. Here we report our studies on the protein state which is obtained upon thermal unfolding at pH 2.5. The refolded apo-Fd obtained after this treatment was found to have all characteristics of a molten globule and its spectroscopic, dynamic and conformational properties were extensively characterised. The results obtained were framed in the context of a mechanistic model for molten globule formation and the implications of this conformational state in the folding of iron sulfur proteins is discussed, especially in what concerns insertion and assembly of the metal centres.

7.3 MATERIALS AND METHODS

Chemicals

All reagents were of reagent grade. Guanidinium chloride (GdmCl) was obtained from Promega and the accurate concentration of the stock solutions in different buffers was confirmed by refractive index measurements [29]. Fluozin and 1-anilinonaphthalene-8-sulfonic acid (ANS) were from Molecular Probes.

Protein purification

The ferredoxin from *Acidianus ambivalens* used in this work were purified as previously described [12]. Briefly, after cell disruption in 40mM potassium phosphate pH 6.5 (buffer A) the resulting cytosolic fraction was applied to a Q-sepharose fast flow column (V_c=20ml), and the Fd containing fractions eluted at ~200-350mM NaCl. After concentration in 5kDa amicons, this fraction was loaded into a Superdex G-50 column (V_c=150ml), eluted with buffer A + 150 mM NaCl and pooled according to the typical visible spectra of Fds. Protein purity was confirmed by SDS-PAGE and N-terminus sequencing.

Spectroscopic methods

UV/Vis spectra were recorded in a Shimadzu UVPC-1601 spectrometer equipped with cell stirring and an interfaced PC-controlled waterbath (JULABO). Fluorescence spectra were recorded Cary Varian Eclipse instrument ($slit_{ex}$: 5 nm, $slit_{em}$: 10 nm unless otherwise noted) equipped with cell stirring and Peltier temperature control. Far-UV CD spectra were recorded typically at 0.2 nm resolution on a Jasco J-815 spectropolarimeter fitted with a cell holder thermostated with a peltier device.

Thermal denaturation

Thermal transitions were monitored using different spectroscopic probes and release of metals. Typically, the temperature was changed from 25 to 95ºC and heating rates of 1-3ºC/min were used. A non-linear least-square analysis was used to fit the data to a two state model which allows the determination of the

midpoint (T_m) of the thermal transition curves. Data analysis was carried out in Origin (MicroCal).

Chemical denaturation

The denaturation curves were measured using the dilutions method, and two solutions with the same protein concentration were prepared, one with no denaturation agent and the other with a high concentration of denaturation agent. These were combined in different proportions yelling different denaturant concentrations. After mixing, the solutions were left equilibrating for two hours. A non-linear least-square analysis was used to fit the data to a two state model from which the denaturation agent concentration at the curve midpoint (C_m) and the free energy of unfolding could be determined. Data analysis was carried out in Origin (MicroCal).

Fluorescence quenching

Quenching experiments were preformed using acrylamide (Bio-Rad) and 2,2,2-tricloroethanol (TCE) (Fluka) as quenchers. The samples were excited at 295nm in order to ensure that the light was absorbed almost entirely by the tryptophanyl group and the fluorescence intensity decrease at the emission maximum was followed. Results were analysed according to the Stern-Volmer equation:

$$\frac{F_0}{F} = 1 + K_{SV}[Q]$$

where F_0 and F are the fluorescence intensities in the absence and presence of quencher, respectively, K_{SV} is the collisional quenching constant and [Q] is the quencher concentration [30].

ANS binding

Ferredoxin samples (typically 5 μM) were incubated with 250 μM ANS during 15-60 minutes prior to measurements. ANS fluorescence emission enhancement of bound ANS at 480nm was evaluated upon excitation at 370 nm, and always corrected for background emission at 520 nm of a control free ANS solution in buffer [30].

Differential scanning calorimetry

Differential scanning calorimetry (DSC) was performed on a MicroCal VP-DSC MicroCalorimeter controlled by the VP-viewer program and equipped with 0.51 ml cells. Studies were made using 50 µM *Aa*Fd in 0.2M glycine, pH 2.5. Heating rates from 0.5 - 1.5ºC.min^{-1} were used from 25 to 95°C. For each rate at least four blank measurements were performed with the respective buffer in both compartments. These runs were used as the baseline for the run with *Aa*Fd in the sample compartment and the corresponding buffer in the reference compartment. Calorimetric data were converted to heat capacity by subtracting the buffer baseline and dividing by the scan rate and protein concentration.

Biochemical methods

Ferredoxin concentrations were routinely determined using either the method of Bradford or spectroscopically (ε^{410nm} = 30400 $M^{-1}.cm^{-1}$). Size exclusion chromatography was used to determine variations in the elution volume of *Aa*Fd in different conformational states (native and molten globule state) were determined using a 70 ml Superdex S-75 (Amersham Biosciences) previously calibrated using a low molecular weight gel filtration calibration kit (Amersham Biosciences)

Metal analysis

Iron and zinc release during *Aa*Fd unfolding were determined using specific chelators. Briefly, samples were withdrawn during the thermal denaturation at different temperatures and analysed as follows. Iron was quantified as described [17] both from the extinction coefficient for the Fe^{2+}-BPS complex (ε^{534nm}= 22140 $M^{-1}.cm^{-1}$), and from calibration curves made with accurately titrated $Fe(NO_3)_3$ solutions. Zinc release was determined fluorimetrically: samples were incubated with Fluozin which emits at 523 nm upon excitation at 488 nm proportionally to the available zinc in solution. Typically experiments were carried out at a 2:1 Fluozin/Fd ratio. Total metal content (iron and zinc) in *Aa*Fd preparations was also determined by total reflection X-ray Fluorescence (TXRF) at the laboratory of Prof. B.O. Kolbesen at the Inst. für Anorg. und Analyt. Chemie.

7.4 RESULTS

Ferredoxin retains native like structure at pH 2.5

The *A. ambivalens* ferredoxin (AaFd) is a hyperstable protein, with a thermal transition midpoint of ~120°C at pH 7. In order to set the grounds for an investigation of the thermal transition under acidic conditions, a characterisation of the protein poised at pH 2.5 was undertaken. In overall, the UV-visible spectrum features are identical to those observed at pH 7: the typical Fe-S band at 410nm is unaltered and the Abs_{410}/Abs_{280} ratio is kept constant upon lowering the pH. Analysis of the *Aa*Fd metal content determined a stoichiometry of ~7 Fe/Fd and ~1 Zn/Fd at pH 2.5, which are identical to those determined at pH 7, thus showing that the metal centres remain intact. Both Trp fluorescence emission and near and far UV-CD show that the overall secondary and tertiary structure of the protein is kept under these conditions (see below). The fluorescent probe 1-anilinonaphthalene-8-sulfonic acid (ANS) was used to investigate the exposure of hydrophobic patches as a result of eventual structural modifications resulting from the acidic condition. Although a slight increase in the ANS emission intensity at pH 2.5 was noted, it is nevertheless insufficient to be representative of a molten globule state upon acidification. In overall, at pH 2.5 *Aa*Fd retains its native structural fold and metal centres: lowering the pH results in extensive side chain protonation of *Aa*Fd, which is nevertheless insufficient to cause a transition to a non-native conformational protein state or unfolding. Among Fe-S proteins, a similar effect has been noted in thermophilic rubredoxins which are destabilised at low pH values (pH 2), but retain both folding and the Cys_4Fe site [19, 20].

Ferredoxin thermal transition occurs in two steps

Differential scanning calorimetry (DSC) was used to monitor the *Aa*Fd thermal transition at pH 2.5. The thermogram resulting from increasing the temperature from 25 to 90ºC shows two consecutive transitions in the first scan: an exothermic one occurring at $T_m^{exo} \approx 60$ºC and an endothermic at $T_m^{endo} \approx 70$ºC (figure 7.1, trace a). Exergonic reactions observed by DSC are usually associated with

protein aggregates [31] , but in this case no precipitation is observed.

Exothermic transitions have been observed also during DSC studies on bovine adrenodoxin, a [2Fe2S] containing protein, and have been suggested to correspond to the decomposition of the FeS centre [21]. We have thoroughly investigated this possibility in *Aa*Fd using visible spectroscopy and quantitation of released iron using the chelator BPS. In fact, the thermal transition midpoints obtained from evaluating the Fe-S clusters integrity from the intensity of the 410 nm band (T_m^{FeS}=63ºC) and from iron discharge (T_m^{Fe}=62ºC) correlate very well with the first calorimetric transition (figure 7.1, trace c). On the other hand, the second transition at 70ºC is typical of the protein unfolding reaction, which involves heat energy uptake [31]. In order to investigate the structural variations occurring in the *Aa*Fd fold as a result of temperature increase, the thermal transition from 30 to 90°C was followed using other techniques. Monitoring the near UV-CD signal of the protein at 280 nm showed that loss of tertiary contacts starts occurring at a midpoint transition temperature ($T_m^{near\ UV\text{-}CD}$ = 61ºC), which is identical to that of Fe-S clusters disassembly (figure 7.1, trace b). This observation agrees well with the analysis of the *Aa*Fd crystal structure [39] which shows that the protein single Trp (W62) has contacts at Fe-S binding pocket, which are lost upon cluster disintegration. The shift of the Trp fluorescence emission maximum from 340 to 350 nm observed as the temperature increases (figure 7.1, trace e) allows monitoring

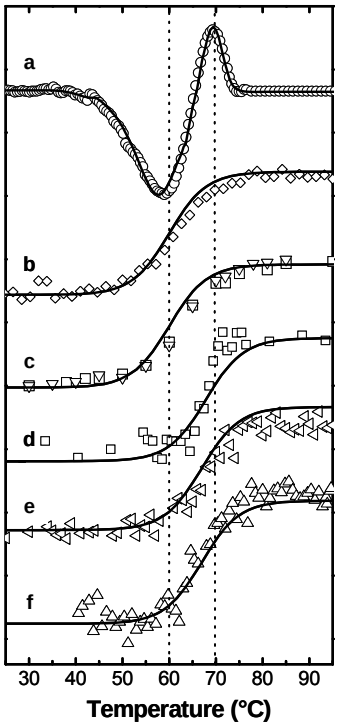


Figure 7.1 |Thermal transition of AaFd
a. DSC thermogram; **b**. near UV-CD signal at 280nm; **c**. Released Fe^{2+} (squares) and variation of the Abs_{410nm} (Fe-S centres, inverted triangles); **d**. Released Zn^{2+}; **e.** Ratio of Trp fluorescence emission at λ_{em}=340nm and λ_{em}=350nm; **f.** far UV-CD signal at 222 nm.

conformational changes which result in a higher exposure of the Trp to the solvent (T_m^{Trp} = 68ºC), and these occurs closer the value of the second calorimetric transition. The release of zinc from the structural site which tights together the N-terminal extension and the protein core was also monitored using fluozin, a fluorescent zinc probe. The determined midpoint for the zinc release during the thermal transition ($T_m^{Zn} \approx$ 69ºC) shows that disintegration of the His/Asp zinc site occurs at a temperature higher than that at which the Fe-S centres break up, and is associated to changes in the tertiary structure. Interestingly, monitoring secondary structure changes (figure 7.1, trace f) showed a transition ($T_m^{far\ UV\text{-}CD}$ = 67ºC) which corresponds to a secondary structure rearrangement, as evidenced by a slight decrease of the ellipticity at 222 nm. Altogether, these results show that exposing the AaFd at pH 2.5 at high temperature (90°C) does not result in a completely unfolded form of the protein but in an apo- form with secondary structure and no tertiary contacts.

Unfolding of metalloproteins is often an irreversible process, due to loss of metal cofactors or inability to reassemble metallic clusters. During the calorimetric studies we have however observed that, upon cooling down the thermally unfolded ferredoxin to 25ºC and performing a second temperature ramp, an endothermic transition persists at ~70ºC (figure 7.2, trace a). An identical behaviour was observed monitoring protein conformational changes in subsequent temperature ramps. The Trp fluorescence emission maxima shift, a cooperative transition is

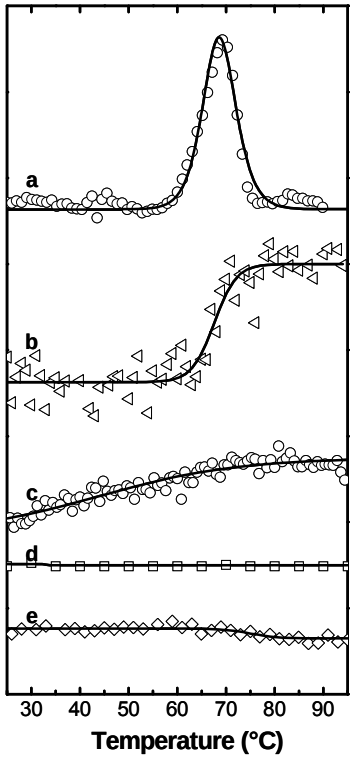


Figure 7.2 | Thermal transition of apo AaFd
a. DSC thermogram; **b**. Ratio of Trp fluorescence emission at λ_{em}=340nm and λ_{em}=350nm; **c**. far UV-CD signal at 222 nm; **d.** Abs_{410nm} (Fe-S centres); **e.** near UV-CD signal at 280nm.

observed (T_m^{fluo} ≈ 69ºC), although the amplitude of the signal is only approximately 50% of that observed in the first transition curve (figure 7.2, trace b). Far-UV CD evidenced a monotonic variation at 222nm and CD ellipticity at 280 nm had no variation during the second temperature ramp (figure. 6.2, trace e). The Fe-S centres were not reassembled in this refolded ferredoxin form, as the protein did not recover nor its brown colour, nor any of the bands of the visible spectrum typical of [FeS] clusters (figure 7.2, trace d).

Apo-ferredoxin is a molten globule

In order to completely characterise the apo-ferredoxin form, a detailed spectroscopic characterisation was undertaken at 25°C in order to compare its characteristics with those of native *Aa*Fd. Circular dichroism spectroscopy was used to study secondary and tertiary structure of the protein. The *Aa*Fd structure is dominated by β-strands (7 sheets *vs.* 3 helices) and in agreement, the far-UV CD spectrum of native ferredoxin is characteristic of α/βprotein (figure 7.3 A).

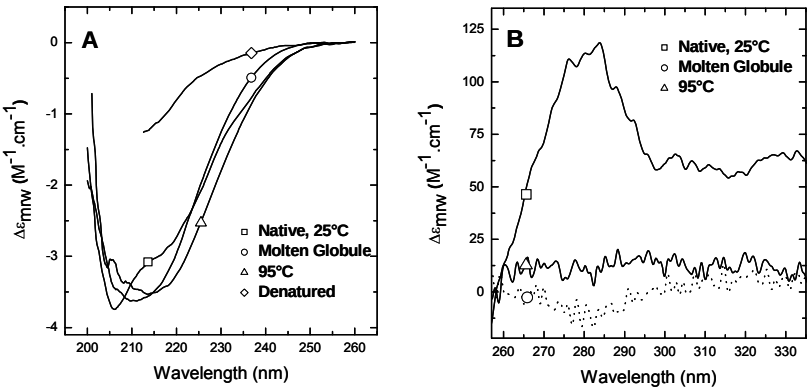


Figure 7.3 Circular dichroism spectra of AaFd in different conformational states: (***A***) far-UV CD and (***B***) near-UV CD spectra of native (N) *Aa*Fd at pH 2.5 (squares), at 95°C (triangles), after cooling down to 25°C (circles) and in the denatured state (90°C and 6M GdmCl). Conditions: *Aa*Fd 0.15 mg.ml^{-1} (far-UV CD) or 1 mg.ml^{-1} (near-UV CD) at pH 2.5.

In these cases, the α-helical component is predominant and negative bands at 222 and 205nm are obtained. Upon raising the temperature to 95°C, probably as a result of loss of the Fe-S centres, the protein undergoes moderate conformational changes which result in a slight gain of secondary structure as shown by a decrease of the ellipticity in the 205-225nm region (figure 7.3 panel

A). On the other hand, the near-UV CD bands of native ferredoxin which derive essentially from the hydrophobic amino acids Trp, Tyr and Phe, reflect the tertiary structure of the protein, as shown by features including fine structures in the 270-290nm range (figure 7.3B), which are not observed at 95°C. However, the near UV-CD spectrum of the cooled *Aa*Fd after obtained the thermal ramp is also silent, indicating that tertiary structure was not restored in the refolded *Aa*Fd state. Altogether, these results which show native-like secondary structure and little detectable tertiary structure are highly suggestive that the *Aa*Fd refolded state is a molten globule. In order to clarify if the refolded apo-ferredoxin form is in a molten globule state, a series of other properties typical of these partly folded protein forms were investigated. One of the characteristics of molten globules is the significantly increase in ANS binding, that is not observable in the native or denaturated states. This affinity of the fluorescent probe to molten globule states, is due to the fact that in this state the hydrophobic surface area is significantly exposed, as a result of the loss of tertiary contacts, but retaining compactness and secondary structure elements [5]. The results obtained showed that incubation of the refolded apo-ferredoxin with ANS results in a 30-fold increase of the chromophore emission in comparison to the native state at pH 7 (figure 7.4). As a control, binding of ANS was also investigated in other conformational states, namely native and unfolded ferredoxin: native AaFd at pH 7 or at pH 2.5 and unfolded AaFd were found to bind only residual amounts of ANS. Although a slight increase in ANS fluorescence emission is observed upon acidification, only in the apo form obtained after thermal denaturation a significant fluorescence enhancement is detected, reflecting a significant exposure of hydrophobic regions compatible with a molten globule

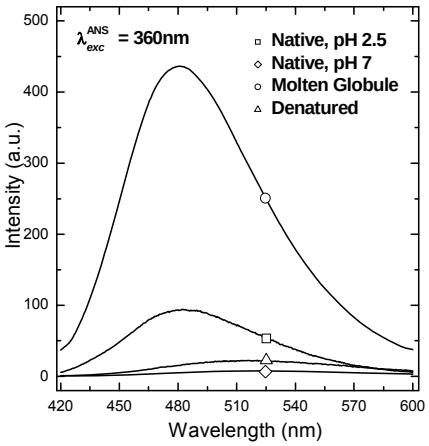


Figure 7.4 Emission spectra of ANS in the presence of different AaFd conformational states: native pH 2.5 (□), and pH 7(◊), molten globule (○) and denatured (△) forms.

state. Other important characteristics of molten globule states are compactness identical to that of the native state, resulting from a relatively stable hydrophobic core which retains a native-like structure. The compactness of the apo-Fd refolded state was investigated by size exclusion chromatography [5], and the results obtained showed that a fraction of the protein in this state has a K_{av} of 0.63 in comparison to 0.59 of the native AaFd. This slight increase corresponds well to the reported 10-20% expansion of the protein in the molten globule state [3, 4].

Conformational dynamics of the molten globule state

Intrinsic Trp fluorescence, red edge excitation shift (REES) and fluorescence quenching were used to evaluate the Trp conformational freedom and flexibility, and the chemical nature of its surroundings in the molten globule state. In particular, the ferredoxin has a single tryptophan (W62) in a relatively superficial region near the [3Fe4S] cluster. In the native state, Trp emission is partially quenched by the metal centre and an emission centred at 343 nm is observed (figure 7.5 panel A). In the molten globule state, the Trp emission maxima remains similar to the native state ($\lambda_{max} \approx 343$ nm) and only a slight increase on the emission intensity is observed (figure. 7.5 panel A). This indicates that in the molten globule the Trp remains shielded from the solvent:, if the Trp would become completely solvent-exposed, a significant red shifted emission maximum

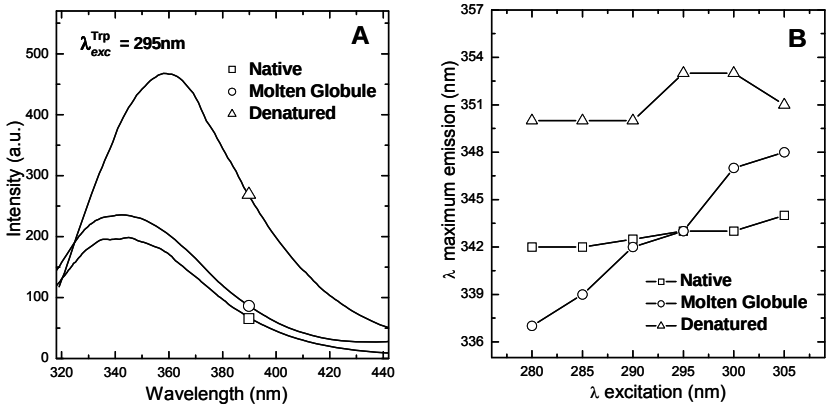


Figure 7.5 Conformational analysis of AaFd from Trp fluorescence data. Intrinsic Trp fluorescence spectra of AaFd in different conformation states
(**A**) and red edge excitation shift (**B**). AaFd was analysed in the native (□), molten globule (○) and denatured (△) states. Conditions: AaFd 5μM in 0.2M glycine pH 2.5.

to ≈ 360 nm would be expected and fallowed by a significative increase in emission intensity (as it is observed in other denaturing conditions, such chemical unfolding at pH 2.5 in the presence of 6M GdmCl, figure 7.5 panel A).

The conformational flexibility of the Trp in the molten globule state was further investigated by monitoring the red edge excitation shift (REES), which is a shift in the position of the emission maxima to higher wavelengths as a result of a change in the excitation wavelength towards the red edge of the absorption band. An effect is observed when the fluorophore is present in a conformational restricted moiety and arises from the slow rates of solvent relaxation around the excited state [32]. This effect was probed in several conformational states of ferredoxin: native, unfolded and molten globule form (figure 7.5 panel B). A minor shift was observed both on the native and denaturated states, indicating that the in these conditions the Trp is not in a motion-restricted environment, although in the denaturated state the chromophore undergone a reorientation towards a solvent exposed medium, as indicated by the shift of the emission maxima to ~350 nm. However, in the molten globule state, a significant red edge excitation shift of 11 nm is observed, indicating that in this case the mobility of the tryptophan moiety is restricted, an observation which is compatible with a structural confinement within a hydrophobic protein environment.

Fluorescence quenching studies were used to further discriminate the chemical nature of the microenvironment surrounding the Trp in the molten globule state. This was achieved comparing the quenching profiles obtained using molecules with different chemical characteristics, which interact differently with the protein. Acrylamide and trichloroethanol (TCE) were used as quenchers, and its combined use allowed evaluating the degree of exposure and the environment surrounding the Trp at different conformational states. Both molecules are neutral (uncharged) quenchers, with the difference that TCE which is less polar, interacts preferentially with hydrophobic regions leading to an exalted quenching of tryptophan when this residue is in the vicinity of hydrophobic domains [33]. The resulting Stern-Volmer plots show that acrylamide and TCE quench identically the Trp emission on the native and unfolded states, with a higher accessibility measured for the unfolded state, in agreement with an increased solvent exposure of the Trp in these conditions (figure 7.6; table 7.1)

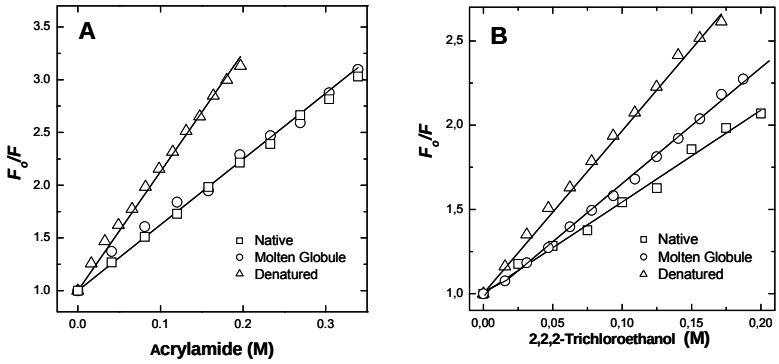


Figure 7.6 |Stern-Volmer plots of AaFd fluorescence quenching.
The quenching of acrylamide (**A**) and 2,2,2-trichloroethanol (**B**) was studied in different AaFd conformational states: native (□), molten globule (○), and unfolded (△). Conditions: AaFd 5μM in 0.2M glycine pH 2.5

Table 7.1 |Stern-Volmer Constants for AaFd in Different Conformational States Using Nonpolar Fluorescence Quenchers.

State	K_{sv} (M^{-1})	
	Acrylamide	2,2,2-Trichloroethanol
Native	6.3	5.4
Molten Globule	6.1	9.2
Denatured	9.8	10.8±

However, in the molten globule state, the quenching constant determined for TCE (9.2 M^{-1}) is higher that that determined for acrylamide (6.1 M^{-1}). This strongly suggests that in the molten globule state, TCE interacts with a hydrophobic pocket at the microenvironment surrounding the Trp, resulting in an enhanced quenching is comparison to acrylamide. It is also interesting to point out the fact that similar Stern-Volmer constants were determined when acrylamide was used to quench the native (6.3 M^{-1}) and molten globule state (6.1 M^{-1}) of ferredoxin (figure 7.6 panel A), which can be interpreted as further confirmation that the Trp did not became exposed to the solvent in the molten globule state, remaining in a relatively apolar environment. Additional evidence for the hydrophobic core surrounding the Trp in the molten globule state was obtained from investigating energy transfer processes from ANS to the single residue. Excitation of the molten globule at 295nm results in a Trp emission band at ~343 nm, whereas an ANS solution excited at the same wavelength has only background emission.

However, excitation of the Trp moiety in the molten globule with ANS bound abolishes Trp emission, while an ANS emission band with a maximum at 480 nm is obtained (figure 7.7). This emission band corresponds to protein-bound ANS, as in solution free ANS emits at 520 nm. In fact, fluorescence emission of the tryptophan can be transferred to ANS since the emission spectrum of the tryptophan (donor) overlaps with the absorption spectrum of the probe (acceptor). But in order to have this fluorescence resonance energy transfer (FRET) phenomenon occurring, it is also necessary that the donor is at the vicinity of the acceptor. Altogether, these data corroborate the view that the apo-ferredoxin molten globule state comprises an apolar core in which the protein Trp is located, nearby a hydrophobic pocket to which ANS can bind.

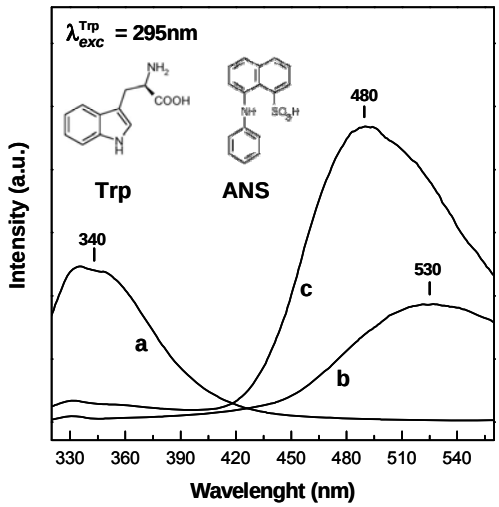


Figure 7.7 | Fluorescence resonance energy transfer between Trp and ANS.
Fluorescence emission spectra obtained upon excitation at 295nm of molten globule *Aa*Fd (*a*), ANS free in solution (*b*) and of the *Aa*Fd molten globule /ANS complex (*c*).

| Conformational stability of the molten globule

The molten globule state is a compact globular form of a protein, with a molten side chain. The conformational integrity of this state is kept by local interactions and non-specific hydrophobic contacts that stabilise this conformational state, identically to what is achieved by the tertiary packing interactions that stabilise the protein native structure. A comparison of the chemical unfolding of the native and molten globule states was carried out in order to characterise its relative stabilities. The Trp maximum emission wavelenght was monitored as a function of different concentrations of the chemical denaturant guanidinium chloride (GdmCl), and the obtained data was fit to a two-state model (figure 7.8, table 7.2). The

molten globule state is ~1.5 kcal.mol^{-1} destabilised in respect to the native form, an observation agrees well with the fact that its overall structure which has a decreased rigidity as a result of loss of tertiary interactions.

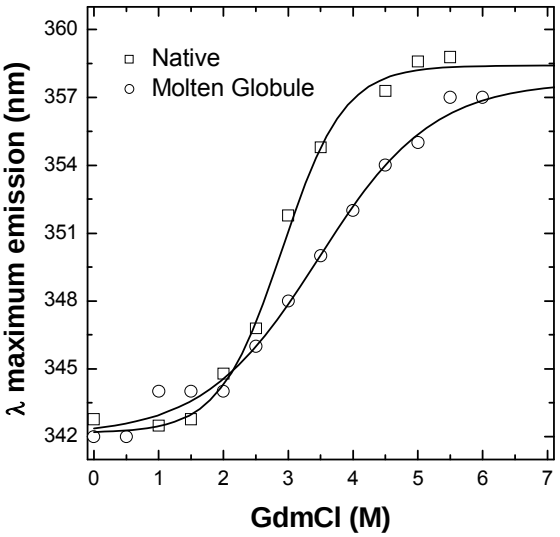


Figure 7.8|Chemical denaturation of AaFd in the native and molten globule state.
Trp fluorescence emission maxima wavelengths are plotted as a function of GdmCl for AaFd in the native (□) and molten globule (○) states.

Table 7.2 |AaFd GdmCl denaturation parameters

	$[D]_{1/2}$ (M)	*M* (cal.mol^{-1}.M.$^{-1}$)	ΔG_{H2O} (kcal.mol^{-1})
Native	2.8 ± 0.1	1241 ± 111	3.5
Molten Globule	3.2 ± 0.1	613 ± 44	1.9

7.5 |DISCUSSION

The usefulness of characterising protein molten globule states lies in the fact that their detailed structural and biophysical characterisation contributes to the understanding of the structural factors and forces involved in the formation of a particular protein fold. However, an additional interest arises from the study of molten globule states of proteins that contain prosthetic groups. In this case, the

structural characterisation of molten globules is useful to understand the mechanism of cofactor insertion and assembly into a given protein scaffold. Further, as cofactor insertion frequently involves interactions with molecular (metallo)chaperons, the molten globule states may seen as good working models for the nascent peptides with which they interact. In particular, examples of proteins containing iron-sulfur centres that can be converted to a molten globule state are scarce.

The [3Fe4S][4Fe4S] ferredoxin studied in the present work was found to undergo a conformational transition to a molten globule state in conditions which are somehow atypical. Molten globules are frequently observed under mild denaturing conditions such as low concentrations of chemical denaturants and at extremes of pH (pH<4 and pH>10), whereas thermal denaturation is less frequent as thermally unfolded proteins tend to aggregate. In this case, acidification of the medium down to pH 2.5 not only not resulted in molten globule formation but also had little impact on the protein conformation which retained the main characteristics of the native form, including its metal centres (the Fe-S clusters and the (His/Asp)Zn^{2+} site). However, the apo-ferredoxin obtained after increase of temperature rests in a molten globule state. The release of the metal centres, which occurs in a stepwise fashion during *Aa*Fd thermal unfolding at pH 2.5, may also be relevant for the formation of the molten globule state. There are examples of formation of molten globule states as a result of loss of metals [3, 4]. Disruption of the two Fe-S clusters occurs prior to release of the zinc, and this probably reflects the unfolding pathway of the protein: upon increase of the temperature, progressive loss of tertiary interactions, probably residues within sheets $\beta_{5\text{-}7}$ and helices $\alpha_{2\text{-}3}$ (figure 7.9) promotes disassembly of the Fe-S centres at the protein core (T_m ≈ 60°C) and it probably also affects the zinc coordinating Asp (D76) which lies at the protein core. Mutagenesis studies in a homologue ferredoxin have shown that this residue is dispensable for zinc coordination [15], thus supporting the hypothesis that the Zn centre remains intact at the expenses of the His ligands from the N-terminal extension (figure 7.9, dotted region). This part of the protein has a higher conformational flexibility and probably unfolds only at a higher temperature (T_m ≈ 70°C) with dissociation of the zinc ion. The data reported shows that the AaFd molten globule state has all properties which are

typical of a canonical molten globule [3, 4]. The loss of tertiary contacts evidenced by near UV-CD was somehow replaced by the novel hydrophobic interactions established within the exposed hydrophobic areas, which allowed maintenance of a compact native-like structure. The single Trp present in the protein revealed to be a very useful probe for characterizing the hydrophobic core formed in the molten globule state of ferredoxin. From the data available a working model for the formation of the molten globule can be put forward. Initially, disassembly of the Fe-S clusters creates a set of apolar contacts which are formed between hydrophobic residues located at the protein interior nearby the Fe-S centres: the Trp accommodates within this pocket which drives apo-ferredoxin refolding to a molten globule state. The combination of fluorescence quenching, red edge emission wavelength analysis and Forster energy exchange with bound ANS clearly indicated that in the molten globule state the Trp is in a motionally restricted apolar environment, i.e. within a hydrophobic core.

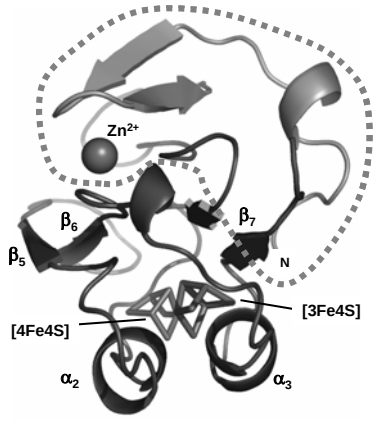


Figure 7.9 | Cartoon representation of AaFd structure
The dotted line encircles the protein N-terminal region which hosts a His/Asp zinc site. The protein core comprises the two iron sulfur clusters which are within a hydrophobic pocket. Metal centres and some secondary structure elements are labelled for clarity. Figure prepared using PyMOL and the *Aa*Fd crystal structure coordinates

The physiological role of molten globules and non native protein states is nowadays evident. Within a cell, non native protein states range from fully unfolded, partly (un)folded to nearly native conformations and respective alternative folds. Molten globules, due to their decreased thermodynamic stability in respect to the native state, can be seen as a form of the protein denatured state under physiological cellular conditions, and are therefore expected to be involved in relevant biological processes. In fact, partly denatured amyloidogenic intermediates with features reminiscent of those of molten globules have been reported in several proteins (see [3] and references within). However, maybe the

most relevant cellular role arises from interactions with molecular chaperons such as GroEL and hsp70 proteins, which recognize non native states and assist the folding process. For some metalloproteins which further require a concerted interaction with metallo-chaperons, molten globule-like states may be relevant to the understanding of the metal centre biosynthesis process. The folding process for small globular proteins is generally considered to proceed directly without molten globule intermediates, but for iron-sulfur proteins such as the ferredoxin here studied (105 amino acids) this is an unlikely scenario, as the intervention of proteins from the iron-sulfur cluster assembly machinery is certainly required to mediate assembly and insertion of cubic inorganic Fe-S structures into the protein polypeptide. For example, the scaffolding IscU protein from *Thermotoga maritima,* which serves as a template for FeS assembly and delivery to apo-proteins, has been thoroughly characterised from the structural and biophysical point of view [24, 25, 34, 35]. This protein which transiently hosts Fe-S clusters was suggested to exist in a dynamic equilibrium between multiple conformations which have some characteristics of molten globules [24]. Structural flexibility has been proposed to be necessary for IscU function *in vivo* [25], as it needs to recognize diverse apo-targets and partners (such as sulfur and iron donors and chaperones (see e.g. [36] for a review)). At this point it seems clear to us that some structural flexibility and a partly folded molten globule like state may also be a pre-requisite of the target apo iron-sulfur proteins which will receive the clusters. The available reports describing [FeS] cluster transfer to apo-ferredoxins, do not address if the conformational characteristics of the latter are those of molten globule states. In our opinion this could be a general feature which nevertheless remains to be experimentally verified. In this respect, the characterisation of the apo ferredoxin molten globule state here described provides a framework that allows tackling directly this aspect, bridging protein physics and cell biology. In fact, this conformational state may be seem as a an excellent working model for newly synthesised peptides of iron-sulfur proteins, as apo-forms undergo partial folding upon their biosynthesis while needing to interact with cellular metallochaperons.

7.6 REFERENCES

1. Ptitsyn, O.B., *How the molten globule became.* Trends Biochem Sci, 1995. **20**(9): p. 376-9.
2. Ohgushi, M. and A. Wada, *'Molten-globule state': a compact form of globular proteins with mobile side-chains.* FEBS Lett, 1983. **164**(1): p. 21-4.
3. Kunihiro, K., and Arai, M., *The molten globule state: the physical picture and biological significance*. Mechanisms of Protein Folding, ed. R.H. Pain. 2000, Oxford: Oxford University Press.
4. Fink, A.L. *Molten Globule*. Encyclopedia of Life Sciences 2001 [cited; Available from: www.els.net.
5. Ptitsyn, O.B., et al., *Evidence for a molten globule state as a general intermediate in protein folding.* FEBS Lett, 1990. **262**(1): p. 20-4.
6. Fink, A.L., et al., *Classification of acid denaturation of proteins: intermediates and unfolded states.* Biochemistry, 1994. **33**(41): p. 12504-11.
7. Beinert, H., R.H. Holm, and E. Munck, *Iron-sulfur clusters: nature's modular, multipurpose structures.* Science, 1997. **277**(5326): p. 653-9.
8. Agarwal, A., D. Li, and J.A. Cowan, *Role of aromatic residues in stabilization of the [Fe4S4] cluster in high-potential iron proteins (HiPIPs): physical characterization and stability studies of Tyr-19 mutants of Chromatium vinosum HiPIP.* Proc Natl Acad Sci U S A, 1995. **92**(21): p. 9440-4.
9. Bentrop, D., et al., *Structural and dynamical properties of a partially unfolded Fe4S4 protein: role of the cofactor in protein folding.* Biochemistry, 1999. **38**(15): p. 4669-80.
10. Bertini, I., et al., *Characterization of a partially unfolded high potential iron protein.* Biochemistry, 1997. **36**(31): p. 9332-9.
11. Foster, M.W., et al., *Elucidation of a [4Fe-4S] cluster degradation pathway: rapid kinetic studies of the degradation of Chromatium vinosum HiPIP.* J Biol Inorg Chem, 2001. **6**(3): p. 266-74.
12. Gomes, C.M., Faria, A., Carita, J.C., Mendes, J., Regalla ,M., Chicau ,P., Huber ,H.,Stetter ,K.O. and Teixeira ,M., *Di-cluster, seven iron ferredoxins from hyperthermophilic Sulfolobales.* JBIC, 1998. **3**: p. 499-507.
13. Griffin, S., et al., *High thermal and chemical stability of Thermus thermophilus seven-iron ferredoxin. Linear clusters form at high pH on polypeptide unfolding.* Eur J Biochem, 2003. **270**(23): p. 4736-43.
14. Jones, K., et al., *Formation of a linear [3Fe-4S] cluster in a seven-iron ferredoxin triggered by polypeptide unfolding.* J Biol Inorg Chem, 2002. **7**(4-5): p. 357-62.
15. Kojoh, K., et al., *Zinc-coordination of aspartic acid-76 in Sulfolobus ferredoxin is not required for thermal stability of the molecule.* J Inorg Biochem, 2002. **89**(1-2): p. 69-73.
16. Leal, S.S. and C.M. Gomes, *Linear three-iron centres are unlikely cluster degradation intermediates during unfolding of iron-sulfur proteins.* Biol Chem, 2005. **386**(12): p. 1295-300.
17. Leal, S.S., M. Teixeira, and C.M. Gomes, *Studies on the degradation pathway of iron-sulfur centers during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability.* J Biol Inorg Chem, 2004. **9**(8): p. 987-96.
18. Higgins, C.L., J. Meyer, and P. Wittung-Stafshede, *Exceptional stability of a [2Fe-2S] ferredoxin from hyperthermophilic bacterium Aquifex aeolicus.* Biochim Biophys Acta, 2002. **1599**(1-2): p. 82-9.
19. Henriques, B.J., L.M. Saraiva, and C.M. Gomes, *Combined spectroscopic and calorimetric characterisation of rubredoxin reversible thermal transition.* J Biol Inorg Chem, 2006. **11**(1): p. 73-81.

20. Henriques, B.J., L.M. Saraiva, and C.M. Gomes, *Probing the mechanism of rubredoxin thermal unfolding in the absence of salt bridges by temperature jump experiments.* Biochem Biophys Res Commun, 2005. **333**(3): p. 839-44.

21. Burova, T.V., et al., *Conformational stability of adrenodoxin mutant proteins.* Protein Sci, 1996. **5**(9): p. 1890-7.

22. Dilg, A.W., et al., *Dynamics of wild-type HiPIPs: a Cys77Ser mutant and a partially unfolded HiPIP.* J Biol Inorg Chem, 2002. **7**(7-8): p. 691-703.

23. Pilon, M., B. de Kruijff, and P.J. Weisbeek, *New insights into the import mechanism of the ferredoxin precursor into chloroplasts.* J Biol Chem, 1992. **267**(4): p. 2548-56.

24. Mansy, S.S., et al., *Iron-sulfur cluster biosynthesis. Thermatoga maritima IscU is a structured iron-sulfur cluster assembly protein.* J Biol Chem, 2002. **277**(24): p. 21397-404.

25. Bertini, I., et al., *Thermotoga maritima IscU. Structural characterization and dynamics of a new class of metallochaperone.* J Mol Biol, 2003. **331**(4): p. 907-24.

26. Teixeira, M., et al., *A seven-iron ferredoxin from the thermoacidophilic archaeon Desulfurolobus ambivalens.* Eur J Biochem, 1995. **227**(1-2): p. 322-7.

27. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.

28. Moczygemba, C., et al., *High stability of a ferredoxin from the hyperthermophilic archaeon A. ambivalens: involvement of electrostatic interactions and cofactors.* Protein Sci, 2001. **10**(8): p. 1539-48.

29. Pace, N.R., *A molecular view of microbial diversity and the biosphere.* Science, 1997. **276**(5313): p. 734-40.

30. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. second ed. 1999, New York: Kluwer Academic / Plenum Publishers.

31. Freire, E., et al., *Calorimetrically determined dynamics of complex unfolding transitions in proteins.* Annu Rev Biophys Biophys Chem, 1990. **19**: p. 159-88.

32. Raja, S.M., et al., *Localization and environment of tryptophans in soluble and membrane-bound states of a pore-forming toxin from Staphylococcus aureus.* Biophys J, 1999. **76**(3): p. 1469-79.

33. Eftink, M.R., J.L. Zajicek, and C.A. Ghiron, *A hydrophobic quencher of protein fluorescence: 2,2,2-trichloroethanol.* Biochim Biophys Acta, 1977. **491**(2): p. 473-81.

34. Adinolfi, S., et al., *Bacterial IscU is a well folded and functional single domain protein.* Eur J Biochem, 2004. **271**(11): p. 2093-100.

35. Mansy, S.S. and J.A. Cowan, *Iron-sulfur cluster biosynthesis: toward an understanding of cellular machinery and molecular mechanism.* Acc Chem Res, 2004. **37**(9): p. 719-25.

36. Frazzon, J. and D.R. Dean, *Formation of iron-sulfur clusters in bacteria: an emerging field in bioinorganic chemistry.* Curr Opin Chem Biol, 2003. **7**(2): p. 166-73.

8 A SPECTROSCOPIC STUDY OF THE TEMPERATURE INDUCED MODIFICATIONS ON FERREDOXIN FOLDING AND IRON-SULFUR MOIETIES

CONTENTS

This chapter was published in

Smilja Todorovic[§], **Sónia S. Leal[§],** Carlos A. Salgueiro, Ingo Zebger, Peter Hildebrandt, Daniel H. Murgida and Cláudio M. Gomes
A Spectroscopic Study of the Temperature Induced Modifications on Ferredoxin Folding and Iron-Sulfur Moieties
Biochemistry 46**,** 10733-10738 (2007)

[§] These authors have equally contributed to the work.

8.1 SUMMARY

Previously we have ascertained the formation of an apo molten globule state after acidic thermal denaturation of the di-cluster ferredoxin from *Acidianus ambivalens*. Here, we assign in detail the specific alterations occurring on the secondary structure during the evolution of the molten globule sate, as well as screen individually the [3Fe-4S] and [4Fe-4S] clusters transformations along the process. Our main goal with this work is to infer on the possible relevancy of this molten globule state in the folding pathway and discriminate the contribution of each cluster for overall stability of the ferredoxin. In agreement, thermal perturbation of the ferredoxin was investigated employing a toolbox of spectroscopic methods. FTIR and visible CD were used for assessing changes of the secondary structure and coarse alterations of the [3Fe-4S] and [4Fe-4S] cluster moieties, respectively. Fine details of the disassembly of the metal centers were revealed by paramagnetic NMR and resonance Raman spectroscopy. Overall, thermally-induced unfolding of *Aa*Fd is initiated with the loss of α-helical content at relatively low temperatures (T_m^{app} ~ 44°C), followed by the disruption of both iron-sulfur clusters (T_m^{app} ~ 53-60°C). The degradation of the metal centers triggers major structural changes on the protein matrix, including the loss of tertiary contacts (T_m^{app} ~ 58°C) and a change, rather than a significant net loss, of secondary structure (T_m^{app} ~ 60°C). In fact, this latter process triggers a secondary structure reorganization that leads to an increase in the β-structure content. The combined spectroscopic approach here reported illustrates how changes in the metalloprotein organization are intertwined with disassembly of the iron-sulfur centers, denoting the conformational interplay of the protein backbone with cofactors. In addition it seems to evidence that both clusters disintegrate in a single event and therefore most likely contribute equally to overall stability. Moreover, the pathway for the formation of this molten globule state clearly evidenced the formation of non-native β-structure along the process which strongly suggests that this conformation is most likely an off-pathway specie. Thus this molten globule state conformation apparently has no direct potential to represent a possible apo-intermediate state suitable for the assembly of the Fe-S clusters.

8.2 INTRODUCTION

The folding and stability of proteins relies on a series of weak interactions of variable magnitudes, whose additive effect results in a stabilized three dimensional fold. In the case of proteins that contain cofactors, especially metal centers, the impact and role of these cross-linkers in the protein overall folding, dynamics and interconversion between native and non-native states, turns out to be particularly relevant. A ferredoxin from the thermophile *Acidianus ambivalens* (*Aa*Fd) was shown to be an exceptionally good model for folding studies [1-3]. This protein, which contains one $[3Fe\text{-}4S]^{1+/0}$ and one $[4Fe\text{-}4S]^{2+/1+}$ center in a canonical $(\beta\alpha\beta)_2$ fold [4], belongs to the Archaeal family of di-cluster ferredoxins [5]. Typically, these ferredoxins harbor a structural Zn^{2+} site with His/Asp coordination that stabilizes the N-terminal domain extension [6], which is also the case in *Aa*Fd [5]. In some isoforms within the family, this centre is replaced by hydrophobic contacts, which results in a comparable stabilization of the domain [7]. *Aa*Fd has a remarkable thermal and pH stability: at neutral pH it remains folded in the presence of 5M GuHCl and it has a T_m ~120°C [1]. However, upon acidification down to pH 2.5 the protein folding and its iron-sulfur centers remain intact, although destabilized, making them amenable to thermal unfolding studies with T_m transitions well below the boiling point of water [8]. We have shown that a highly stable and structured molten globule (Fd^{MG}) state is formed by thermally denaturing the native protein (Fd^{N}), followed by its (Fd^{D}) subsequent cooling at low pH [8]:

$$Fd^{N} \xrightarrow{heating} Fd^{D} \xleftrightarrow{cooling} Fd^{MG}$$

In this work we address the thermal perturbation of *Aa*Fd at acidic pH using a set of complementary spectroscopic tools that provides a detailed description of the processes underlying the unfolding process and the simultaneous molten globule formation at different levels of the metalloprotein organization.

8.3 MATERIALS AND METHODS

Protein purification

The di-cluster ferredoxin was purified from *Acidianus ambivalens* grown at pH 2.5 and 80°C [9]. A cytosolic fraction of the cells was processed and the protein purified in two consecutive chromatographic steps as previously described [5]. Protein purity was evaluated from the Abs_{410}/ Abs_{280} ratio (~0.6) and SDS-PAGE. The integrity of both iron-sulfur clusters was evaluated by TXRF metal analysis, which yielded 6±0.5 Fe per mol of ferredoxin. Protein stocks were kept as concentrated aliquots at -20°C (40 mM potassium phosphate, 150 mM NaCl, pH 6.5).

Thermal perturbation

Prior to the spectroscopic studies the purification buffer was freshly exchanged for 200 mM glycine buffer, pH 2.5. For NMR experiments the protein was lyophilized and then re-suspended twice in 200mM glycine / D_2O (99.96% atom) buffer solution, at pH 2.5. Thermal perturbation was performed by successively incubating *Aa*Fd (pH 2.5) at increasing temperatures, typically in the 20-90°C range, during 5 min, with ca. 5-10°C interval. After each incubation step the protein was cooled down to the measurement temperature.

Dynamic light scattering

Measurements were carried out in a Malvern Zetasizer nano ZS instrument, equipped with a 4 MW He-Ne Laser (632 nm). The protein (0.2mM *Aa*Fd in 200mM glycine buffer, pH 2.5) was filtered though a 0.22 µm filter prior to measurements, and a 1x1 cm quartz cuvette (Hellma) was used for measurements, at different incubation stages between 25 and 90ºC. The operating procedure was set to 20 runs, each being averaged for 20s. Data was analysed by the DTS software (Malvern) in respect to the distribution of sizes by volume.

FT-IR spectroscopy

FTIR spectra were recorded on a Bruker Tensor 27 spectrometer equipped with a N_2(l)-cooled MCT detector at a spectral resolution of 4 cm^{-1}. About 10 µl of thermally treated sample (0.4 mM in 200 mM glycine buffer, pH 2.5) were injected into a gas-tight cell for liquids (pathlength 12 µm; CaF_2 windows). All measurements were done at 10 °C purging the sample compartment with dry air.

ANS binding

Temperature incubated ferredoxin samples (typically 5 µM) were incubated with 250 µM 1-anilinonaphthalene-8-sulfonic acid (ANS) during 15-60 minutes prior to measurements. ANS fluorescence emission enhancement of bound ANS at 480 nm was evaluated upon excitation at 370 nm, and corrected for the background emission of a protein-free ANS solution in buffer [10].

CD spectroscopy

CD spectra were recorded in a Jasco J-815 spectropolarimeter equipped with Peltier temperature control and cell stirring. Far UV-CD spectra were recorded using 1 mm pathlength polarimetrically certified cells (Hellma). Near-UV and visible CD measurements were carried out using rectangular 10 mm pathlength cells. Unless otherwise noted, experiments were carried out at 20°C.

Resonance Raman

About 2 µL of 1.2 mM oxidized *Aa*Fd (200 mM glycine buffer, pH 2.5) were introduced into a N_2(l)-cooled cryostat (Linkam) mounted on a microscope stage and cooled down to -190 °C. Spectra from the frozen sample were collected in backscattering geometry by using a confocal Raman microscope (Jobin Yvon, XY). The different excitation lines (413, 458 and 514 nm) were obtained either from a krypton ion laser (Coherent Innova 302) or from an argon ion laser (Coherent Innova 70). Typically, spectra were accumulated for 60 seconds with a laser power at sample of 9 mW. After polynomial background subtraction, spectra were treated with home made component analysis software. For quantification of

the thermally-induced perturbation the positions and widths of the four bands assigned to the iron-sulfur clusters [11-16] were maintained constant and only their intensities were adjusted for each incubation temperature. The band fitting parameters for [3Fe-4S], defined from spectra obtained with 514nm excitation, allowed unambiguous determination of the [4Fe-4S] band parameters in spectra obtained with the 413 nm laser.

^{1}H NMR

Spectra were collected at 500 MHz in a Bruker DRX500 spectrometer using a 5-mm inverse probe head with *z* gradients and a Eurotherm 818 temperature-control unit. The 1D ^{1}H-NMR spectra of the oxidized *Aa*Fd were recorded at 23 °C with 32 k data points and water pre-saturation. Spectral width of 38 kHz, pulse width of 16 μs, and 1.2 s^{-1} pulse repetition rate were used. A total of 1024 transients were acquired. Free induction decays were subject to a 40 Hz exponential line broadening factor before Fourier transform. Chemical shifts are reported in parts per million (ppm), relative to tetramethylsilane, and the proton spectra were calibrated using the water signal as an internal reference [17]. Control spectra recorded before and after lyophilization were identical (data not shown), indicating that the protein structure was not affected by the removal of water.

8.4 RESULTS

Thermally-induced structural perturbations of *Aa*Fd were investigated under acidic conditions by a variety of complementary spectroscopic methods which selectively monitor changes at different levels of the protein organization. A common perturbation protocol was adopted for the various spectroscopic studies. It consists of successive incubations of *Aa*Fd at increasing temperatures followed by cooling down to a reference temperature, prior to a spectroscopic measurement (see Materials and Methods). The apparent midpoint transitions (T_m^{app}) were determined by different spectroscopic techniques. The obtained T_m^{app} is the temperature at which the intensity of a given spectral feature characteristic of Fd^N falls to 50 % of its initial value or, alternatively, the temperature at which a

characteristic feature of the Fd^{MG} form has increased up to 50% of its maximum value.

Monitoring secondary structure alterations: FT-IR and far-CD

FTIR was used to probe perturbations at the level of the secondary structure. The spectra show distinct changes as a function of the incubation temperature in the amide I region which are better visualized in a second derivative representation (figure 8.1 panel A). Specifically, we note large amplitude changes at 1630 cm^{-1} and smaller ones at 1690 cm^{-1} and 1649 cm^{-1}. The rise of the first two negative bands is ascribed to a formation of extended β-structures and turns, respectively [18-21], and renders T_m^{app} = 60°C (figure 8.1 panel B) in both cases.

Note that the appearance of a signal at 1630 cm^{-1} may also be indicative of the formation of intermolecular β-sheet aggregates. This possibility was ruled out by dynamic light scattering experiments (DLS), which showed that no large size particles are formed during *Aa*Fd thermal perturbation. Rather, a 25% increase in the *Aa*Fd diameter is observed (from 6.9 ± 0.4Å to 9.2 ± 0.2Å) wich is within the size range typically

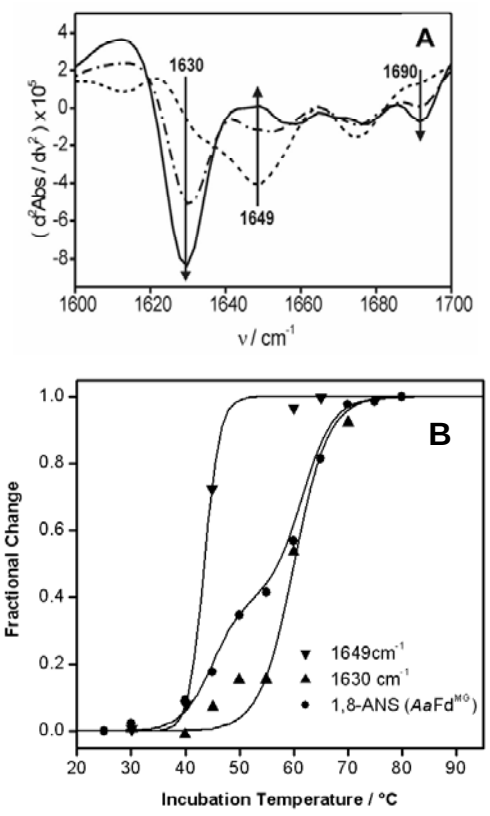


Figure 8.1 | FT-IR spectra and temperature dependence profile
(**A**) FT-IR spectra of *Aa*Fd recorded after thermal incubation at 30 °C (dash lines), 60 °C (dash-dot) and 70 °C (solid line); (**B**) Temperature dependence profile of the FTIR bands (Up-triangles, 1630 cm^{-1}; down-triangles, 1649 cm^{-1}) and normalized 1,8-ANS fluorescence enhancement at 480nm (circles).

observed in molten globule states, which have a compactness identical to the native state, but exhibit a ~10-30% increase in size. The growth of the positive band at 1649 cm^{-1} indicates a loss of α-helical content [18-21], which occurs with T_m^{app} = 44 °C (figure 8.1 panel B). This value is distinctly lower than all other transition temperatures determined by the various methods (table 8.1). Such a discrepancy cannot be attributed to the experimental error, although it is somewhat larger for the quantification of the 1649-cm^{-1} amplitude due to uncertainties in the background subtraction. Thus, regardless of the exact numerical value, it can be safely concluded that disorganization of the α-helices occurs at relatively low temperatures. Interestingly, no simultaneous changes are observed in the 1940-1945 cm^{-1} region, which would be indicative for the formation of an unordered (random coil) structure [18-21]. Instead, β-structures are present, and this fact agrees with previous work showing that cooled thermally unfolded *Aa*Fd in acidic conditions leads to a molten globule state, which retains partial secondary structure [8]. Changes in the secondary structure were also monitored by far-UV CD. Monitoring the signal at 222nm, which comprises contributions both from β and α structure, yielded a transition at T_m^{app} = 60°C (table 8.1). The very low α-helical content of the protein (~10%) impairs a detailed resolution of the two types of secondary structure by far-UV CD, which was achievable by FTIR.

Appraisal on the forming molten globule

The fraction of formed molten globule can be estimated from the increase of emission of 1,8-ANS at 480nm. This fluorescent probe exhibits an increased emission upon binding to hydrophobic patches exposed in the MG state [10]. A plot of the normalized 1,8-ANS emission enhancement as a function of incubation temperature denotes two transitions, T_m^{app} at 45°C and 60 °C (figure 8.1, panel B), which correlate with the FTIR transitions. Therefore, the onset of molten globule formation is associated with this initial disruption of the α-helices. From this incubation temperature onwards, *Aa*Fd undergoes further folding reorganization as the MG becomes significantly more populated in subsequent incubations, concomitantly with disruption of the Fe-S moieties (*below*). Additionally, loss of tertiary contacts as evidenced by monitoring the near-UV CD

band at 285 nm, occurs within the same temperature range T_m^{app} = 58 °C; data not shown).

Course alterations on the di-clusters integrity during the molten globule formation: visible-CD and RR

A first approach to evaluate course alterations of the Fe-S centers was based on visible CD spectroscopy. This technique provides better resolved features, eventually more informative than the visible absorption signature, which in *Aa*Fd comprises a main weak band at ca. 410 nm [9]. The visible CD spectrum of *Aa*Fd displays a complex pattern (figure 8.2, panel A), including four bands that are attributable to both iron-sulfur cluster domains [22]. The temperature-dependencies of the individual bands (figure 8.2, panel B) allow determination of different values, T_m^{app} as summarized in Table 8.1. These results indicate locally different thermal stability of the protein environment at the different metal centers. A detailed assignment, however, is precluded by the fact that none of the bands can be regarded as marker for an individual metal center.

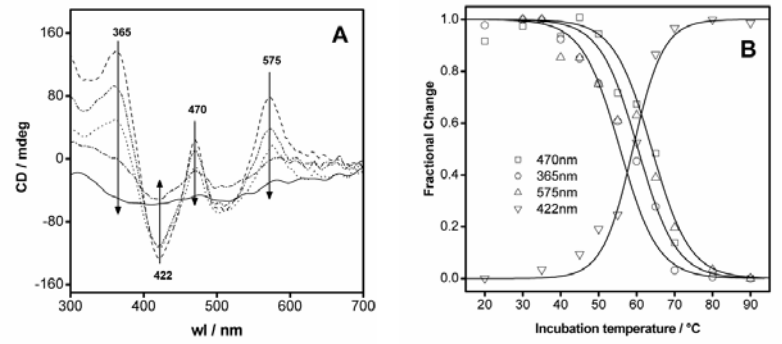
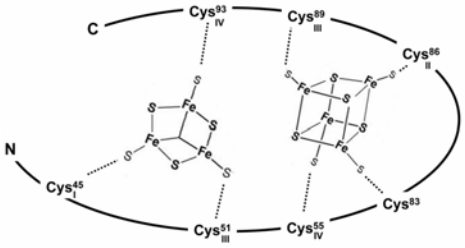


Figure 8.2 Visible CD spectra and temperature dependence profile
(**A**) *Aa*Fd after thermal incubation at 30 °C (dash), 50 °C (dot-dot), 55 °C (dot), 65 °C (dash-dot) and 70 °C (solid), and temperature dependence profile (**B**) visible CD bands at different wavelengths (squares, 470 nm; circles,365 nm; triangles, 575 nm and inverted triangles, 422 nm). The lines represent sigmoidal fits.
(results are summarized in table 8.1).

As a more specific probe, RR spectroscopy was employed to monitor the disassembly of the individual clusters. In order to separate the spectral contributions from the [3Fe-4S] and [4Fe-4S] centers of *Aa*Fd, three different

laser excitation lines (413, 458 and 514 nm) were used to achieve a differential enhancement of the RR bands associated with the two clusters. For all three excitation wavelengths, the most prominent RR band is observed at 346 cm^{-1} and is assigned to the stretching of the Fe-S-Fe bridges of the [3Fe-4S] center, [3Fe-$4S^B$] [12-14, 16, 23]. This band is accompanied only by a weaker band at 366 cm^{-1}, upon excitation at 514 nm (data not shown), which is thus assigned to the terminal Fe-S vibrations of the same cluster, [3Fe-$4S^T$] [11-16]. The RR spectra recorded with 413 nm and 458 nm excitation show two additional features at 336 cm^{-1} and 358 cm^{-1} that are assigned to the Fe-S vibrations involving the bridging and terminal ligands of the [4Fe-4S] cluster, respectively, [12-14, 16, 23] ([4Fe-$4S^B$] and [4Fe-$4S^T$]), Scheme 8.I. As shown in figure 8.3 panel A, the RR bands of the [3Fe-4S] center dominate the RR spectrum at 413 nm excitation [12].



Scheme 8.I |Cartoon representing the general ligation scheme for the di-cluster *Aa*Fd.

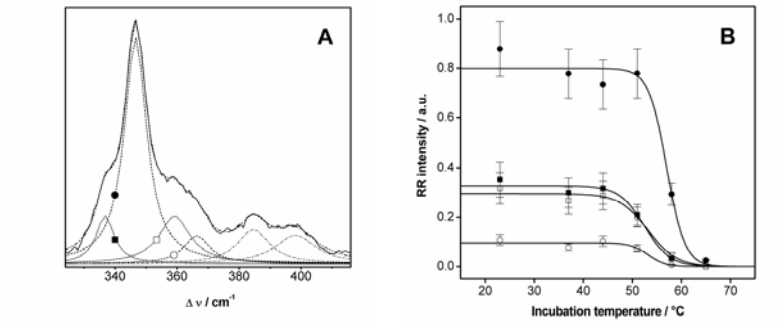


Figure 8.3 |Experimental and band fitted RR spectra of oxidized *Aa*Fd (A), and Relative intensities of [4Fe-4S] (squares) and [3Fe-4S] (circles) vibrational modes (B) as a function of the incubation temperature.
Experimental RR spectrum (solid line, A) was obtained from 1.2 mM *Aa*Fd (pH 2.5, no incubation), with 413nm excitation and laser power of 9 mW at -190 °C. Overall fitted spectrum: dash-dot-dot line. Component spectra, dash line / solid circles: [3Fe-$4S^B$], 346 cm^{-1}; dot line / solid square: [4Fe-$4S^B$], 336 cm^{-1}; dash line / open circles: [3Fe-$4S^T$], 366 cm^{-1}; dot line / open squares: [4Fe-$4S^T$], 358 cm^{-1}. The solid lines (B) represent sigmoidal fits (results are summarized in table 8.1).

Nevertheless, the spectral contributions of the [4Fe-4S] center can also be accurately determined by applying standard band fitting procedures. Besides the modes attributed to the bridging and terminal modes of [3Fe-4S] and [4Fe-4S] clusters, the spectrum also contains resolved bands at 386 cm^{-1} and 397 cm^{-1}. However, since these bands do not represent unambiguous pure vibrational modes they will not be further considered in the analysis. This excitation line was thus used to monitor the temperature-dependent RR spectral changes and thermal disassembly of the two clusters. Figure 8.3 panel B shows the intensity changes of the individual RR bands upon successive incubations of *Aa*Fd at increasing temperatures. An apparent transition at 53 °C is observed for all vibrational modes, except for [3Fe-4S^{B}] for which a T_m^{app} = 57 °C is determined (table 8.1). Although the difference of 4 °C is relatively small and close to the experimental error, it is reproducibly observed.

Probing Fe-S clusters bridges to the protein: ^{1}H-NMR evaluation

Complementary structural information concerning the interaction of the iron-sulfur clusters and the protein chain was obtained by ^{1}H-NMR (figure 8.4).

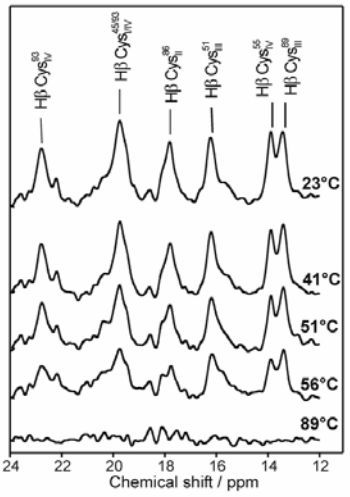


Figure 8.4 Low field region of a 1D ^{1}H-NMR proton spectrum of *Aa*Fd.
The assignments and incubation temperatures are indicated in the figure.

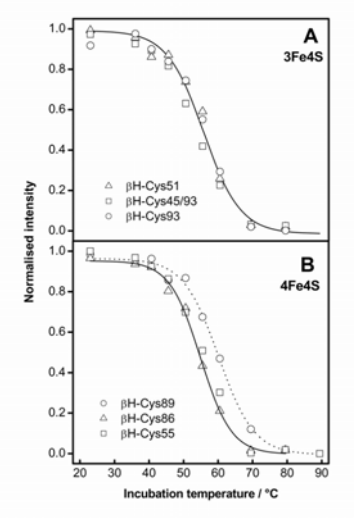


Figure 8.5 Thermal transitions of the cysteine ligands monitored from individual resonances for the [3Fe-4S] clusters ligands (A) and [4Fe-4S] cluster ligands (B) (see table 8.1).

Table 8.1 |Apparent midpoint transition temperatures obtained by the different spectroscopic methods.

Method	Assignment		T_m^{app}/ °C
FTIR	Extended β-structures (1630 cm^{-1})		60 ± 2
	β-turns (1690 cm^{-1})		60 ± 2
	α-helices (1649 cm^{-1})		44 ± 5
ANS	Molten globule formation (λ_{em}=480 nm)		45 ± 2 / 60 ± 2
CD	[FeS] clusters	(365 nm)	58 ± 2
		(422 nm)	60 ± 2
		(470 nm)	64 ± 2
		(547 nm)	62 ± 2
	Tertiary structure (285 nm)		58 ± 2
	Secondary structure (222 nm)		62± 2
RR	[3Fe4S^B] (346 cm^{-1})		57 ± 2
	[3Fe4S^T] (366 cm^{-1})		53 ± 2
	[4Fe4S^B] (336 cm^{-1})		53 ± 2
	[4Fe4S^T] (358 cm^{-1})		53 ± 2
NMR	[3Fe4S] βH-Cys^{93}		57 ± 2
	[3Fe4S] βH-Cys$^{45/93}$		56 ± 2
	[3Fe4S] βH-Cys^{51}		56 ± 2
	[4Fe4S] βH-Cys^{86}		55 ± 2
	[4Fe4S] βH-Cys^{55}		55 ± 2
	[4Fe4S] βH-Cys^{89}		60 ± 2

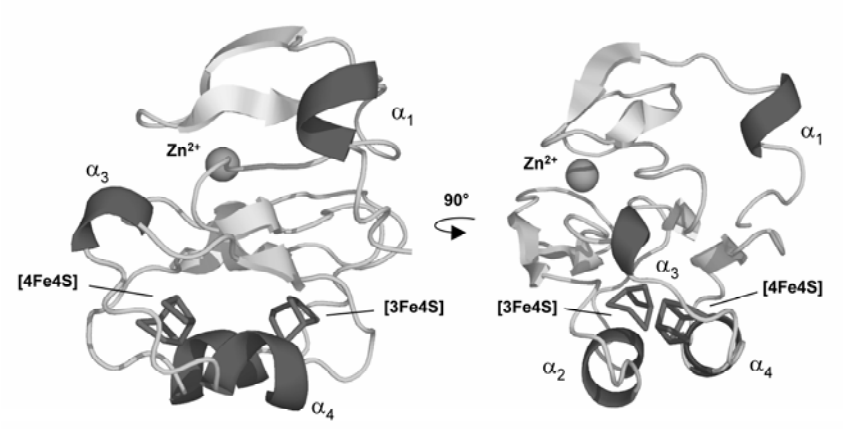
FTIR – Fourier Transform InfraRed; **ANS** - 1-anilinonaphthalene-8-sulfonic acid; **CD** – Circular Dichroism; **RR** – Resonance Raman; **NMR** - Nuclear Magnetic Resonance

Inspection of the 1D ^{1}H-NMR spectral window covering the downfield paramagnetic region of the *Aa*Fd at 23°C, shows broadened signals, which are clearly shifted outside the diamagnetic envelope, indicating that they correspond to the protons located in the proximity of the paramagnetic centers. These resonances have been assigned to the β-CH_2 protons of the seven cysteine residues involved in the coordination of the two cubic Fe-S clusters of *Aa*Fd, at pH 7 [11] (Scheme 8.I). The intensities and positions of the NMR signals are unaltered at pH 2.5 (not shown), corroborating that the protein and its metal centers remain structurally intact in acidic conditions. The thermal perturbation was followed by the decrease of intensity and areas of the assigned resonances as a function of the incubation temperature (Figure 8.5, table 8.1). The differencies in the T_m^{app} of the individual cysteine ligands are very small. Only

Cys^{89} from the [4Fe-4S] cluster has a slightly higher T_m^{app} then the average of the other ligands (~5°C). The enhanced stability of ligand may originate from stabilizing apolar interactions in its vicinity. Inspection of *Aa*Fd structure suggests that the adjacent Val^{90} and nearby Ile^{98} and Val^{100} may be involved in such stabilization.

8.5 DISCUSSION

The body of spectroscopic data accumulated in the present work provides a framework for understanding the structural and dynamical modifications of *Aa*Fd during the formation of an apo-molten globule state. Changes in metalloprotein organization are clearly intertwined with disassembly of the iron-sulfur centers, denoting the conformational interplay of the cofactors and the protein backbone (see scheme 8.II).



Scheme 8.II | Structural model of AaFd in two orthogonal views

Thermally-induced unfolding of *Aa*Fd is initiated with the loss of α-helical content at relatively low temperatures (T_m^{app} ~ 44 °C), followed by the disruption of both iron-sulfur clusters (T_m^{app} ~ 53-60 °C). The observed disruption of α-helices occurring at ~44°C may be hypothesized to involve mostly helices α1 and α3, which are located far from the protein $(\beta\alpha\beta)_2$ core and do not provide with any coordinating cysteines for the iron-sulfur clusters. This is clearly not the case of helices α2 and α4 that are directly involved in the clusters coordination and where

it is feasible to speculate that these helices undergo a conformational distortion concomitantly with the cluster disruption, at higher temperatures (see scheme 8.II). RR and NMR data indicate that the last steps in the disassembly of the clusters are the disruption of the terminal Cys^{89} ligand from the [4Fe-4S] center and of the bridging ligands of the [3Fe-4S] cluster. The degradation of the metal centers triggers major structural changes of the protein matrix, including the loss of tertiary contacts (T_m^{app} ~ 58 °C) and a increase, rather than a significant net loss, of β-structures (T_m^{app} ~ 60 °C).

Altogether, it seems evident that Fe-S clusters plays a crucial role in the overall native structure of this di-cluster ferredoxin but most likely also in the correct course of the native folding pathway. In fact, after removing thermal perturbation the protein appears to mis-refold without the clusters reintegration. A frustrated folding reaction is highly suggested by the evidence of the non-native β-structures that are formed during apo refolding. This misfolding force is most likely driven by non-native hydrophobic interactions. This speculation is also supported by the fact that these studies were always performed at pH 2.5 where all ionic residues are most likely protonated and ionic interactions can no longer contribute for re-folding. An appraisal on the effect of ionic interactions over the apo molten globule conformation was noticed when incubated the apo form at pH 7, where acidic residues are no longer protonated and therefore available for relating ionic interactions. In this circumstance the once structured acidic molten globule loses all the characteristics of a molten globule state and adopts the conformation of an unfolded protein (data not shown). This seems to further substantiate the importance of preferential side chain ligation by the Fe-S clusters assembly in the correct folding of the ferredoxin and suggest that metals insertion are needed in an early folding event given the impossibility of the protein to reach an structured intermediate apo state in physiologic conditions.

Although the resulting apo molten globule state presented promising features [8] for a potential on-pathway intermediate state for the assemble of Fe-S clusters, the confronted results shown here suggest that this is most likely a case of misfolded off-pathway intermediate in the absence of the coordinating metal cofactors.

8.6 REFERENCES

1. Wittung-Stafshede, P., C.M. Gomes, and M. Teixeira, *Stability and folding of the ferredoxin from the hyperthermophilic archaeon Acidianus ambivalens.* J Inorg Biochem, 2000. **78**(1): p. 35-41.
2. Leal, S.S., M. Teixeira, and C.M. Gomes, *Studies on the degradation pathway of iron-sulfur centers during unfolding of a hyperstable ferredoxin: cluster dissociation, iron release and protein stability.* J Biol Inorg Chem, 2004. **9**(8): p. 987-96.
3. Leal, S.S. and C.M. Gomes, *Linear three-iron centres are unlikely cluster degradation intermediates during unfolding of iron-sulfur proteins.* Biol Chem, 2005. **386**(12): p. 1295-300.
4. Fujii, T., et al., *The crystal structure of zinc-containing ferredoxin from the thermoacidophilic archaeon Sulfolobus sp. strain 7.* Biochemistry, 1997. **36**(6): p. 1505-13.
5. Gomes, C.M., Faria, A., Carita, J.C., Mendes, J., Regalla ,M., Chicau ,P., Huber ,H.,Stetter ,K.O. and Teixeira ,M., *Di-cluster, seven iron ferredoxins from hyperthermophilic Sulfolobales.* JBIC, 1998. **3**: p. 499-507.
6. Kojoh, K., H. Matsuzawa, and T. Wakagi, *Zinc and an N-terminal extra stretch of the ferredoxin from a thermoacidophilic archaeon stabilize the molecule at high temperature.* Eur J Biochem, 1999. **264**(1): p. 85-91.
7. Rocha, R., et al., *Natural domain design: enhanced thermal stability of a zinc-lacking ferredoxin isoform shows that a hydrophobic core efficiently replaces the structural metal site.* Biochemistry, 2006. **45**(34): p. 10376-84.
8. Leal, S.S. and C.M. Gomes, *Studies of the molten globule state of ferredoxin: structural characterization and implications on protein folding and iron-sulfur center assembly.* Proteins, 2007. **68**(3): p. 606-16.
9. Teixeira, M., et al., *A seven-iron ferredoxin from the thermoacidophilic archaeon Desulfurolobus ambivalens.* Eur J Biochem, 1995. **227**(1-2): p. 322-7.
10. Semisotnov, G.V., et al., *Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe.* Biopolymers, 1991. **31**(1): p. 119-28.
11. Bentrop, D., et al., *Paramagnetic NMR analysis of the seven-iron ferredoxin from the hyperthermoacidophilic archaeon Desulfurolobus ambivalens reveals structural similarity to other dicluster ferredoxins.* European Journal of Biochemistry, 1996. **236**(1): p. 92-99.
12. Czernuszewicz, R.S., et al., *Vibrational-Mode Structure and Symmetry in Proteins and Analogs Containing Fe4S4 Clusters - Resonance Raman Evidence for Different Degrees of Distortion in Hipip and Ferredoxin.* Journal of the American Chemical Society, 1987. **109**(23): p. 7178-7187.
13. Golinelli, M.P., et al., *Extensive ligand rearrangements around the [2Fe-2S] cluster of Clostridium pasteurianum ferredoxin.* Biochemistry, 1998. **37**(29): p. 10429-37.
14. Han, S., et al., *Fe2S2 Protein Resonance Raman-Spectra Revisited - Structural Variations among Adrenodoxin, Ferredoxin, and Red Paramagnetic Protein.* Journal of the American Chemical Society, 1989. **111**(10): p. 3505-3511.
15. Iwasaki, T., et al., *Spectroscopic investigation of selective cluster conversion of archaeal zinc-containing ferredoxin from Sulfolobus sp strain 7.* Journal of Biological Chemistry, 2000. **275**(33): p. 25391-25401.
16. Johnson, M.K., et al., *Resonance Raman-Spectroscopic Evidence for a Common [3Fe-4S] Structure among Proteins Containing 3-Iron Centers.* Journal of the American Chemical Society, 1983. **105**(22): p. 6671-6678.
17. Glasoe, P.K. and F.A. Long, *Use of Glass Electrodes to Measure Acidities in Deuterium Oxide.* Journal of Physical Chemistry, 1960. **64**(1): p. 188-190.
18. Arrondo, J.L., et al., *Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy.* Prog Biophys Mol Biol, 1993. **59**(1): p. 23-56.

19. Byler, D.M. and H. Susi, *Examination of the secondary structure of proteins by deconvolved FTIR spectra.* Biopolymers, 1986. **25**(3): p. 469-87.

20. Dong, A., et al., *Spectroscopic study of secondary structure and thermal denaturation of recombinant human factor XIII in aqueous solution.* Arch Biochem Biophys, 1997. **347**(2): p. 213-20.

21. Jackson, M. and H.H. Mantsch, *The use and misuse of FTIR spectroscopy in the determination of protein structure.* Crit Rev Biochem Mol Biol, 1995. **30**(2): p. 95-120.

22. Stephens, P.J., D.R. Jollie, and A. Warshel, *Protein Control of Redox Potentials of Ironminus signSulfur Proteins.* Chem Rev, 1996. **96**(7): p. 2491-2514.

23. Meyer, J., et al., *Assembly of a [2Fe-2S]2+ cluster in a molecular variant of Clostridium pasteurianum rubredoxin.* Biochemistry, 1997. **36**(43): p. 13374-80.

8.7 ACKNOWLEDGMENTS

Smilja Todorovic, Ingo Zebger, Peter Hildebrant and Daniel Murgida for the FT-IR and RR studies.

Carlos Salgueiro for the NMR studies.

9 Fe-S Clusters and Ferredoxin Folding

General Discussion and Conclusions

Our current understanding of the folding pathways of Fe-S proteins is still rather limited. Apart from the established fact that pre-assembled Fe-S clusters are inserted into a given apo protein state by scaffold proteins, little is known about the molecular mechanisms underlying cluster transfer. One particular aspect that remains unaddressed concerns the requirements of the receiving apo-state in respect to favourable conformations, and also on how this event couples to the overall protein folding process and to the overall stability of the fold.

In this thesis we have attempted to contribute to a better understanding of the role of Fe-S clusters in the folding and stability of Fe-S proteins and to elucidate the dynamics and features of apo-state conformations in these proteins. To address these studies we have used a set of small, monomeric and very stable ferredoxins from the family of the Archaeal di-cluster ferredoxins. This family of proteins contains a [3Fe-4S] as well as a [4Fe-4S] cluster and an N-terminal extension of ~30 residues that holds a zinc centre. The zinc centre is localised at the interface of the N-terminal domain with the Fe-S core region and suggested to play essentially a structural and rather stabilizing role in the ferredoxin structure. Interestingly, some highly identical isoforms of this class of proteins are conjointly expressed that either contain (FdA) or miss (FdB) the zinc site. Here, we have investigated the effect of the absence of this metal site on the conformation and folding of these ferrredoxins, using the isoforms as models before addressing the stabilising role of Fe-S clusters in overall fold. In this respect, we have observed that a set of hydrophobic residues from the ferredoxin isoform lacking the zinc centre was able to provide the same structural fold to the protein and additionally improve the overall stability of the protein. These residues occupy the same spatial region of the zinc site, and the established hydrophobic interactions are in fact leading to a very efficient stabilisation of the region. Although the zinc centre was in fact evidenced to play an important contribution for the stability of FdA, apparently no stabilizing advantage was gained with the insertion of the zinc centre when comparing with FdB. On the other hand is very unlikely that the protein has evolved in the direction to lose the metal centre. In fact, in an evolutionary context, zinc loss rarely occurs, and when it occurs it generally leads to the deletion of the part of the protein responsible for metal binding or results in major structural changes of the protein [1]. This outcome is most likely attributed

to the fact that structural zinc ions are frequently necessary for protein folding, so that the ability to bind a metal cannot be lost without significant change of the protein structure. This was clearly not the case within these ferredoxin isoforms where the absence of the zinc centre did not result in major structural alterations or the deletion of the part of the protein responsible for metal binding. In addition, it is largely accepted that this family of Archaeal di-cluster ferredoxins most likely descended from the *Clostridial* type 2x[4Fe-4S] ferredoxin fold that undergone posterior inclusion of the N-terminal extension containing the zinc ligands. Typical structural zinc sites are coordinated by four cysteines or a combination of cysteines and histidines in a relatively short linear sequence. In this case, the zinc centre is coordinated by three histidines (His^{16}, His^{19}, His^{34}) in the N-terminal region but also by a distant aspartic acid (Asp^{75}) in the core region which is usually characteristic of catalytic zinc binding sites. Evolution may have recruited this uncommon set of ligands in a structural site to alternatively achieve stabilization of the ferredoxin fold. Another remote possibility, for which there is no direct evidence, would be that this extension could have been recruited from a catalytic zinc-containing domain of another protein. In this respect, catalytic zinc centers are generally coordinated by three protein ligands (His-His-Asp) and a water molecule that is considered obligatory for the catalytic function. Nevertheless, exceptions have been found, like in the case of a β-carbonic anhydrase from an red alga, where the catalytic zinc site is coordinated by four ligands in the protein (Cys-Asp-His-Cys) and no water molecule [2]; as well it was also described for a metalloproteinase that a water molecule could be transitorily replaced by a protein residue in the inactive form of the protein that upon activation binds again to a water molecule [3, 4]. However, solely based on this data, suggesting the possibility that the zinc site in these ferredoxins could be a vestigial reminiscence of a catalytic site is solely speculation.

Within the di-cluster ferredoxins family, the [4Fe-4S] cluster, considering its very low redox potential (~ -550mV) has been suggested to represent an “evolutionary relic” [5] likely to play essentially a structural and stabilizing role in the protein. Unlike that of the [3Fe-4S] clusters (~-270 mV), such a low redox potential is unlikely to be functional after the evolutionary adaptation to an oxidative atmosphere. In order to analyse if the clusters would have a differential

contribution to protein stability, we have employed a combination of spectroscopic methods that monitored directly the integrity of the Fe-S clusters. In particular, the use of ^{1}H-NMR and Resonance Raman allowed discriminating the individual signatures of the [3Fe-4S] and of the [4Fe-4S] clusters. We have observed that during thermal denaturation the clusters disassemble under identical conditions, an observation suggesting that they exert a similar contribution to the protein overall stability. In addition we have also analysed the behaviour of the Fe-S clusters during ferredoxin unfolding. This followed an initial observation that in alkaline conditions the Fe-S clusters would degrade with the formation of an intermediate inorganic species, namely an aconitase-like linear 3Fe cluster. We have here ruled out this possibility, by evidencing that both clusters disintegration is concomitant with the released of iron to the solution and protein unfolding, as well as showing that the observed spectroscopic signature that has led to this assignment in fact results from the formation of iron sulfides in solution.

One aspect we particularly focused on, concerned the interplay between the Fe-S clusters and ferredoxin hydrophobic core. In respect, it has long been ascribed a close alliance between Fe-S centres and hydrophobic shields for clusters function and stability [6, 7]. In addition, from all the hydrophobic residues that constitute di-cluster ferredoxins more than 60% are concentrated in the vicinity of the coordinating ligands of both clusters, which could imply a possible role of these surrounding hydrophobic residues in the assembly mechanisms of the Fe-S clusters in the protein. Interestingly, we have observed that Fe-S clusters disintegration is concerted with the collapse of the hydrophobic core thus suggesting that packing of hydrophobic residues and the coordination of Fe-S clusters may be coupled. In agreement, this was also recently suggested within the folding mechanism of a zinc finger protein [8] where metal binding is known to be crucial for protein folding. The extent of interplay of hydrophobic with the overall stability of the Fe-S clusters was further evidenced when stabilizing electrostatic interactions were overturned at the isoelectric point and the di-cluster ferredoxin was shown to retain the integrity of the Fe-S clusters and the integrity of the hydrophobic core. Moreover we only noted major conformational alterations and protein denaturation with the disintegration of Fe-S clusters, and the collapse of the hydrophobic core. Put together these evidences seem to substantiate that

the integrity of the Fe-S clusters and that of the hydrophobic core are closely associated. In addition it also suggests that this association play a major role in sustaining backbone rigidity and likely to represent the major building framework for the overall structure fold. Furthermore, it was also shown that, unlike the apparent marginal effect on overall fold, ionic interactions must play a strong influence in ferredoxin thermal stability. In fact, the switching-off of ionic interactions results in a significant decrease in ferredoxin thermal stability thus evidencing that all forces are important and determinative for ferredoxin stability.

For most chemical and thermal unfolding conditions that were tested, ferredoxin unfolding and the disintegration of the Fe-S clusters lead to a non-reversible apo-unfolded state. Interestingly we evidenced an exception. Upon acidic thermal denaturation or a long period incubation at pH < 2, the di-cluster ferredoxin was found to undergo a conformational transition to a rather stable apo molten globule state. The usefulness of characterising protein molten globule states lies in the fact that their detailed structural and biophysical characterisation can contribute to the understanding of the structural factors and forces involved in the formation of a particular protein fold. In this particular case, the structural characterisation of apo molten globules can become rather useful to better understand the mechanisms underlying cofactor assembly and protein folding. In fact, it is recognised that Fe-S cluster proteins require a concerted interaction with metallo-chaperons in order to assemble the metal centres. In this respect, molten globule states may be involved in interactions with molecular chaperones which recognize these non native states and assist the folding process. Nevertheless, the available reports describing Fe-S cluster transfer to apo-ferredoxins, do not address if the conformational characteristics of the latter are those of molten globule states. Given that molten globule states were already described as a characteristic apo state for metal binding in other metalloprotein [9, 10], it seems reasonable to speculate that structural flexibility and partly folded molten globule like state could be a pre-requisite for scaffold proteins recognised and attach to the target apo iron-sulfur proteins. Yet little is known about the recognizing process between scaffold proteins and apo-states. However, it was demonstrated that the transfer reaction of a [4Fe-4S] clusters from a scaffold protein to the acceptor biotin synthase, proceeds in two observable steps: a first fast one

leading to a protein-protein complex between the cluster donor, and a slow one consisting of cluster transfer leading to the apoform of the scaffold protein and the holoform of the target protein. Mutation of cysteines in the acceptor protein specifically inhibits the second step of the reaction, showing that these cysteines are involved in the cluster transfer mechanism but not in complex formation [11]. Further, it was previously demonstrated a lack of specificity between the protein donor and acceptor [11]. In fact, there is no acceptor specificity for a given donor, since the latter is in general able to deliver its cluster to different targets with comparable rates, leading to the formation of both [2Fe-2S] and [4Fe-4S] complexes. For example, SufA can serve as a [2Fe-2S] cluster donor to ferredoxin and a [4Fe-4S] donor to biotin synthase [12]. As well, a given acceptor protein was proven to be able to accept clusters from different scaffold proteins; for example, ferredoxin can obtain its cluster in vitro from different donor proteins, [13, 14]. To our view this strongly suggests that the protein-protein interaction between the cluster donor and acceptor has to be set through non specific interactions. Eventually, it can be speculated that the abundant hydrophobic residues that surround the clusters and the scaffold protein may play a role in establishing hydrophobic interactions. Moreover, the transfer proceeds through a concerted and protective mechanism with no transient disassemble of the cluster from the scaffold proteins into the solution [15], thus further suggesting the need of a hydrophobic environment to protect the cluster during the transfer. In agreement with this speculation, we have observed that in the apo ferredoxin molten globule state, the cluster binding pocket was maintained by apolar contacts of the hydrophobic residues located at the protein interior, nearby the Fe-S centres. This would provide a protective hydrophobic environment for clusters to be transferred from a scaffold protein. Altogether these evidences lead us to speculate on the possibility that this molten globule conformation had the potential to represent the apo-state of di-cluster ferredoxins where metallochaperones can insert Fe-S clusters. Nevertheless, we have subsequently noted that non-native β-structures are formed during the formation of the molten globule, suggesting a frustrated folding reaction. The apo unfolded protein appears to mis-refold without the clusters reintegration. Eventually cys-cys crosslinking and/or non-native hydrophobic interactions between may play a role

in this process. This latter speculation is also supported by the fact that these studies were always performed at pH 2.5, where all ionic residues are protonated and ionic interactions can no longer contribute for re-folding. An appraisal on the effect of ionic interactions over the apo molten globule conformation was noticed when the apo form was incubated at pH 7, a condition that allows ionic interactions to be established. In these circumstances, the once structured molten globule formed under acidic conditions loses all the characteristics of a molten globule state and adopts the conformation of an unfolded protein. This supports the importance of preferential side chain ligation by the incorporation of the Fe-S clusters, and suggests that these may be an early stage of the folding process, given the impossibility of the protein to reach a partly structured intermediate apo state in physiologic conditions. Altogether, it seems evident that the Fe-S clusters play a crucial role in the overall native structure of this di-cluster ferredoxin but most likely also in the correct course of the folding pathway.

In sequence of our findings we propose that it is highly conceivable that the successful alliance between Fe-S centres and hydrophobic shields in proteins may also be extensive to play a rather relevant role in the assembly process of the clusters and folding of these proteins. This close association evidently can only be inferred within Fe-S proteins that hold the Fe-S clusters in the interior of the protein and can not be extrapolated for example to scaffold Fe-S proteins, known and expected to contain partially exposed Fe-S clusters given that their function is to assemble Fe-S clusters transiently and deliver them efficiently and rapidly to apoproteins [16-18]. In addition, the evidences suggest that the integrity of both Fe-S clusters do not only stabilizes the native state but is likely to participate and be determinant in an early folding process. Without the help of Fe-S clusters modelling, the ferredoxin cannot start to fold properly. Native folding is most likely driven by the combination of hydrophobic forces modelled by Fe-S clusters coordination which point out for the requirement of an early binding event of the clusters in this family of ferredoxins.

REFERENCES

1. Torrance, J.W., M.W. Macarthur, and J.M. Thornton, *Evolution of binding sites for zinc and calcium ions playing structural roles.* Proteins, 2008. **71**(2): p. 813-30.
2. Mitsuhashi, S., et al., *X-ray structure of beta-carbonic anhydrase from the red alga, Porphyridium purpureum, reveals a novel catalytic site for CO(2) hydration.* J Biol Chem, 2000. **275**(8): p. 5521-6.
3. Becker, J.W., et al., *Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme.* Protein Sci, 1995. **4**(10): p. 1966-76.
4. Morgunova, E., et al., *Structure of human pro-matrix metalloproteinase-2: activation mechanism revealed.* Science, 1999. **284**(5420): p. 1667-70.
5. Iwasaki, T., et al., *Functional and evolutionary implications of a [3Fe-4S] cluster of the dicluster-type ferredoxin from the thermoacidophilic archaeon, Sulfolobus sp. strain 7.* J Biol Chem, 1994. **269**(47): p. 29444-50.
6. Cheng, V.W., et al., *The iron-sulfur clusters in Escherichia coli succinate dehydrogenase direct electron flow.* J Biol Chem, 2006. **281**(37): p. 27662-8.
7. Dey, A., et al., *Solvent tuning of electrochemical potentials in the active sites of HiPIP versus ferredoxin.* Science, 2007. **318**(5855): p. 1464-8.
8. Li, W., et al., *Metal-coupled folding of Cys2His2 zinc-finger.* J Am Chem Soc, 2008. **130**(3): p. 892-900.
9. Aitio, H., et al., *Characterization of apo and partially saturated states of calerythrin, an EF-hand protein from S. erythraea: a molten globule when deprived of Ca(2+).* Protein Sci, 2001. **10**(1): p. 74-82.
10. Christova, P., J.A. Cox, and C.T. Craescu, *Ion-induced conformational and stability changes in Nereis sarcoplasmic calcium binding protein: evidence that the APO state is a molten globule.* Proteins, 2000. **40**(2): p. 177-84.
11. Ollagnier-de-Choudens, S., Y. Sanakis, and M. Fontecave, *SufA/IscA: reactivity studies of a class of scaffold proteins involved in [Fe-S] cluster assembly.* J Biol Inorg Chem, 2004. **9**(7): p. 828-38.
12. Wollenberg, M., et al., *A dimer of the FeS cluster biosynthesis protein IscA from cyanobacteria binds a [2Fe2S] cluster between two protomers and transfers it to [2Fe2S] and [4Fe4S] apo proteins.* Eur J Biochem, 2003. **270**(8): p. 1662-71.
13. Nishio, K. and M. Nakai, *Transfer of iron-sulfur cluster from NifU to apoferredoxin.* J Biol Chem, 2000. **275**(30): p. 22615-8.
14. Wu, G., et al., *Characterization of an iron-sulfur cluster assembly protein (ISU1) from Schizosaccharomyces pombe.* Biochemistry, 2002. **41**(15): p. 5024-32.
15. Ollagnier-de Choudens, S., et al., *SufA from Erwinia chrysanthemi. Characterization of a scaffold protein required for iron-sulfur cluster assembly.* J Biol Chem, 2003. **278**(20): p. 17993-8001.
16. Chandramouli, K., et al., *Formation and properties of [4Fe-4S] clusters on the IscU scaffold protein.* Biochemistry, 2007. **46**(23): p. 6804-11.
17. Cupp-Vickery, J.R., et al., *Crystal structure of IscA, an iron-sulfur cluster assembly protein from Escherichia coli.* J Mol Biol, 2004. **338**(1): p. 127-37.
18. Morimoto, K., et al., *The asymmetric IscA homodimer with an exposed [2Fe-2S] cluster suggests the structural basis of the Fe-S cluster biosynthetic scaffold.* J Mol Biol, 2006. **360**(1): p. 117-32.