PROTEIN DYNAMICS AND ITS CORRELATION TO ACTIVITY AND STABILITY

A THESIS

Submitted in partial fulfillment

of the requirements for the degree

of

Doctor of Philosophy

at

The University of Waikato

by

DAVID CLEMENT

The University of Waikato

2009

Abstract

The aim of this thesis is to investigate the role of fast protein dynamics (pico-second timescale) in enzyme activity and stability, and specifically to test the hypothesis that enzyme activity and stability are inversely related by their internal dynamics.

Activity ∝ Dynamics (flexibility) ∝ 1/Stability

In order to test this hypothesis, the well known anti-cancer drug: methotrexate was used as an informative ligand in the network established between these properties. A multidisciplinary approach combining neutron scattering, circular dichroism, UV absorption, isothermal titration calorimetry and X-ray crystallography was undertaken to examine the current paradigm using the enzyme: dihydrofolate reductase as a model.

As inferred by neutron spectroscopy, the binding of MTX influences the dynamical behavior of DHFR. Macromolecular dynamics such as the resilience: $\langle k \rangle$ (i.e. structural rigidity) was found to be increased and, inversely, the flexibility decreased upon MTX binding. In addition, as revealed by circular dichroism, this dynamical dependency upon MTX binding was correlated with an enhanced thermal stability. Compared to the free enzyme, the melting temperature was found to be increased by 13.8 °C in the presence of MTX. The inhibitory power of MTX was also examined by steady state kinetics and isothermal titration calorimetry. The K_i for MTX was found to be in the nanomolar range K_i= 10.9 nM. Using isothermal titration calorimetry, the binding thermodynamic signature between MTX and DHFR was characterized. The binding event was found to be largely favourable (ΔG_b=-12.1 Kcal mol^{-1}), enthalpy driven (ΔH_b= -16.8 Kcal mol^{-1}) with an unfavourable entropy ΔS_b=-15.6 cal $K^{-1}mol^{-1}$.

In conclusion, the modulation of the macromolecular dynamics may reflect how specific conformations are favoured for subsequent protein function in response of the binding of specific ligand and how conformational substates approach to protein function. In this context the unprecedented power of transition state analogs

such as MTX on protein function might therefore be dependent on fast protein dynamics.

Keywords: Incoherent neutron scattering, circular dichroism, isothermal titration calorimetry, steady state kinetics, dihydrofolate reductase, methotrexate.

Acknowledgements

J’aimerais tout d’abord remercier ma mère pour m’avoir toujours supporté dans mes choix et s’être sacrifiée pour l’éducation de mon frère et moi-même. J’aurais aimé qu’elle soit encore parmi nous, ainsi que mon oncle Jean Louis pour partager ce moment avec eux. Je tiens à remercier toute ma famille et en particulier mon frère : Fréderic pour leur soutien.

I would you to thank Roy Daniel for giving me the chance to do this PhD with so much freedom and independence in his lab, for allowing me to get involved with so much interesting people and projects. I’m also indebted to him for pointing out how politics and money can affect Science.

I want to greatly thank Colin Monk who he is the ‘MacGyver’ of Science and conceptually another kind of lab technician for his support and fruitful advices. Murielle Lopez for being a cool flatmate, and a sweet person to work with. Alena, Amy Craig, Charles, Michelle, Lee Anne and Rachel for their friendship.

I would to very sincerely thank the examiners for doing me the honour of judging my work, and for their helpful comments and corrections of the manuscript: Thank to Professor Joe Zaccai and Professor Juliet Gerrard.

This thesis was made possible with the excellent supervision of Moeava Tehei who never let me down over these five years. He has kept his enthusiasm intact on this project, his support. Honestly, if I think your deserve this PhD as much as me. I will never forget the famous Arrhenius and his legendary exponential, your rugby skill and love dance.

I’m also very much indebted to the people of the Deuteration Laboaratory in Grenoble, and LSS group at the ILL. Thanks to Martine Moulin for having me as a fermentor trainee during her pregnancy but the cost was relatively high in terms of croissant over the last year of my PhD. Thanks to Valerie Laux for her bike, needles and a tiny bit of her bunch. Marion Jasnin for her advices and joy of living. I have to say that more people with her philosophy and way of life are crucially needed in our

poor science social life. Thanks to Trevor Forsyth and Michael Hartlein for their kindness and welcoming in their lab at the ILL. Special thanks to Suzy for introducing me to the fascinating world of crystallography with her unique smiling touch. Phil Callow and Shirley for their English jokes or German jokes, I do not remember exactly. Estelle for her absolute exquisite Belgium chocolates, JB for keeping an eye on my cheese, Shona and Anna for her day to day happiness. Finally to Christian Vettier and Madame Le Pennec who have made possible this adventure at the Institute Laue Langevin.

Fortunately, during a PhD there is much than the scope of science.

Quel monde étrange la communauté scientifique. On en parle souvent à la télé ou dans les revues ou non spécialisées avec toujours cette idée que ces hommes sont des sorciers. Ma première expérience, je l'avoue a commencé sous bonne augure avec un japonais spécialiste des vins de Sauternes. L'analyse par spectrométrie de masse des composés volatiles du vin était la version robotisée de notre sens: l'art de la dégustation. Ce fut une des rares occasions qui me fut donnée pour expérimentalement affirmer que l'homme supplante aisément la machine... Ma formation scientifique m'a ensuite permis de me déplacer vers des contrées plus exotiques (que ma très non moins célèbre université parisienne pour son amiante) : La Norvège. Petit étudiant Erasmus, j'ai rencontré la vie de labo à travers une superbe équipe, très humaine ou le champion de ski: André Aamodt pouvait à lui seul déplacer un lab meeting ou une expérience de biacore. Le dépaysement ou 'culture shock', les petites rivières et ses truites, saumons, le vélo sur neige et le rugby sur sable ont crée ces souvenirs inoubliables. Plein d enthousiasme, le virus du voyage m'amène par la suite à Grenoble ou j'intègre l'équipe de Michel Vivaudou. Cette année fut pleine de bons moments scientifiques mais aussi musicoculinosportifs; les Grands ducs en Chartreuse, le Beaujolais, la torture du patch clamp, les naufrages en ski à Chamrousse me confortèrent dans mes aspirations lointaines. La collocation avec gens bien ouverts, du monde en transit, des idées qui s'échangent, le bonheur. Pourquoi partir ou quitter ce tableau ?

Je ne le sais pas vraiment encore, goût de la nouveauté, curiosité, stupidité…..On en arrive donc à la thèse, aux antipodes, et un nouveau pays La

Nouvelle-Zélande. Petit paradis sur terre oublié du reste du monde lorsque la coupe du monde rugby s'éteint. La première année je l'avoue fut très dure car il faut tout recréer, des amis, une vie, agiter ses cheveux dans tous les sens pour mieux avancer. Chose assez étrange ou bien logique le monde scientifique s'est alors éclipsé pour des valeurs plus simples, plus humaines. J'ai appris qu'un métier ne pouvait pas lier les hommes, seules les passions y parviennent. Ce chemin bien sinueux fut des plus plaisant et original. Je tiens pour ce fait remercier toutes les personnes comme Corinnie du groupe Daisy chain Halo pour les drumming sessions, la médecine traditionnelles Maori. J'ai vraiment apprécié rencontrer ces personnes pour les enseignements et les joies qu'elles m'ont procurées. Deidre, la petite fée de Raglan avec son caractère et gentillesse bien trempée. Elle pourrait à elle seule résumer ce qu'est la Nouvelle-Zélande. Julien l'homme orchestre aux milles facettes, pêcheur de sensation sur la Waipa river qui a agit comme un accélérateur sur ma façon de penser (même s'il me reste beaucoup de travail de ce coté). Mason, Tony et Droid, et leur maison du bonheur. Rochelle pour m'avoir supportée aussi longtemps. Drew et sa passion pour la musique. Merci de m'avoir donné la chance de pouvoir connaître le monde de la radio et ses acteurs à travers le Womad et le Sunsplash. Dr. Reiner Eschenbruch et sa femme Janis pour leur amour du vin et leur conception de la vie qui fait bien défaut à certains, au grand dam de leur maigre breuvage. Des vendanges entre amis et nombreuses dégustations qui rendent ces instants magiques. Les habitants de Raglan, comme Tony le pêcheur, Ross, les Cornerstones roots, la communauté du town hall et ses danses allemandes. Aileen et Alex, les habitants de Karuna falls et toutes les personnes que nous avons rencontrées au cours de nos expériences en tant que woofers. L'équipe de rugby de Marist et ses coachs légendaires, les matchs de rugby à sept 'façon Samoa' avec Feelo, Soss, Johnny, Massy, Houlo, Alex, et les autres.

Merci à tous, j'espère que j'aurai l'occasion de revenir dans ces îles magnifiques pour passer de bons moments avec vous.

J'aimerais aussi remercier très chaleureusement toutes les aimables personnes qui ont pues m'héberger à Grenoble à maintes reprises lors de mes passages lorsque j'en avais besoin. Mention spécial pour Thibault Geoui qui à préféré démarrer sa

carrière d'oenologue chez Quiagen, Fabien Boursate, la légende des ravioles savoyardes, Greg et Julien pour les kitchen parties, Carlos fervent défenseur de l'huile d'olive made in Jaen, Carole pour sa marmotte empaillée, Morgan pour son amour des Chocapics au Comté, Nathalia pour son amitié et Suzel pour me ramener toujours au bon endroit……
Merci à tous

On aurait pu arrêter ces quelques lignes sur ces notes teintées de nostalgie et de joie mais comme je l'ai souligné la science a un coté pourrie, en ceci qu'elle concentre une multitude d'égocentriques au trop plein d'ambition qui paradoxalement est sans commune mesure avec leur valeur morale. Ces ambassadeurs du néant écrasent tout devant eux, confiance, idéal et humanisme et sont en quelque sorte le phylloxera de la Vie. J'ai au cours de ces 5 années perdu un temps inconsidérable à me battre contre cette idéologie que je qualifierais de 'malbouffe'. Bien sur, on peut sourire de cette naïveté et faire abstraction de ce monde qui est voué à perte mais le changement est possible. Pour l'illustrer j'aimerais citer deux films qui m'ont vraiment interpellés au cours de ces derniers mois. Premièrement : 'There will be blood', ou la ressemblance avec certains des acteurs et le plus haut responsable de ce travail est fort troublante. La fin justifie les moyens. On défend des valeurs et non pas sa petite place de nanti. Un deuxième long métrage: 'Louise Michel' pour l'humour si décapant et si réaliste de la vie d'en bas.

Idéaliste je le resterai et le combat ne fait que commencer. C'est pourquoi ce travail de thèse signera la fin de ma très courte carrière en science mais peut-être une nouvelle vie.

David

Preface

The work presented in this thesis has been performed in a collaborative joint work between mainly two main groups: RM Daniel's group from the University of Waikato which is the 'biochemist half' of the collaboration and J. Smith's group from the University of Heidelberg, which is the 'physics half' of the team. Prof .J Finney and Dr. M Tehei also played an important role in this team as biophysicists, expert in neutron scattering techniques. Some of the materials of this PhD have therefore not been processed by myself, the neutron scattering chapter includes a strong input from Dr. M Tehei who has processed the data. Nevertheless, I have interpreted and written the remaining material concerning this section.

Chapter 5 has the form of an article. I have detailed my input in the introduction but my contribution was restricted to the sample preparation for the neutron experiment. I have also written the purification protocol and help in the discussion of the first versions of the manuscript. I have no input in the data processing.

Chapter 7 deals with the crystallographic structure of the complex between the enzyme dihydrofolate reductase and the inhibitor: methotrexate. I have collaborated with Dr Suzanna Teixeira (Dlab, Grenoble) to perform this work. She has helped me in all aspects of the project. However, the goal of this study was for me to gain an expertise in X-ray crystallography as I planned to join an X-ray crystallography laboratory. I have therefore obtained crystals by myself and assisted actively in the data collection (ESRF, Grenoble). The data processing was made under Dr. Teixeira supervision but I have processed the whole set of data.

Table of contents

PAGE

List of figures

List of Tables

List of Abbreviations and or Symbols

Å	Ångstrom (10^{-10} m); unit of distance
APS	Ammonium Persulfate
CS	Conformational Substate
CD	Circular dichroism
$^{\circ}$C	Celsius degree; unit of temperature
D	Deuterium (^{2}H)
DHFR	Dihydrofolate reductase
DIE	Deuterium Isotope Effect
D_2O	Deuterium oxide
H	Hydrogen (^{1}H)
HEPES	4-(2-HydroxyEthyl)Piperazine-1-EthaneSulfonic acid
H_2F	7, 8-Dihydrofolate
H_4F	5, 6, 7, 8-Tetrahydrofolate
ILL	Institut Laue Langevin
ITC	Isothermal Titration Calorimetry
K	Kelvin
kDa	kiloDalton (10^3 Da); unit of mass
K_i	Inhibition constant
K_m	Michaelis constant
λ	Lambda; Wavelength of Electromagnetic Radiation
M	Molar; unit of concentration

µ	Mu; Micro, as in µm, or micrometer (10^{-4}m)
MALDI	Matrix-Assisted Laser Desorption Ionization
MD	Molecular Dynamics
mg	milligram (10^{-3} gram (g)); unit of mass
mL	milliliter (10^{-3} liter (l)); unit of volume
mm	millimetre (10^{-3} meter (m)); unit of distance
MS	Mass Spectrometry
MTX	Methotrexate (Amethopterin)
M_w	Molecular weight
N	Stoichiometry
NACs	Near Attack Conformers
$NADP^+$	Nicotinamide Adenine Dinucleotide Phosphate, Oxidized Form
NADPH	Nicotinamide Adenine Dinucleotide Phosphate, Reduced Form
NMR	Nuclear Magnetic Resonance
PEG	Polyethylene Glycol
RMSD	root mean square deviation
RT	Room temperature
s	second; unit of time
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
TOF	Time Of Flight
TS	Transition State
V_{max}	Maximum velocity
ΔH	Enthalpy

ΔS	Entropy

1. Protein dynamics and function: A moving story

1.1. Molecular flexibility and protein function

Proteins are the tools of the living; they participate or perform all the reactions within the cell. They are noticeably important in cell signaling, cell adhesion or immune responses. Because their role is very versatile, they are classified in families. One particular class of proteins, namely enzyme, often called “life catalysts” lie at the interface between the energies of the living and protein function. These biocatalysts enhance chemical reaction over 5 to 10 orders of magnitude by lowering the activation barrier. They ultimately ease reactions to occur on a life time scale but paradoxically to their biological relevance, the origin of their prodigious power is unknown. In this respect, a considerable effort and theories were developed to unravel the origin of enzymatic catalysis but none succeed to establish a general consensus [1]. The concept of transition state or activated complex first developed in the 1940s was reformulated by Pauling: ‘The activation barrier must be lowered to speed the reaction’[2-4]. According to the postulate, enzymatic catalysis must pass through a critical and fleeting state: the transition state corresponding to the maximum energy state along the reaction coordinate. Although, this postulate originally focuses on the chemical reaction itself, it also strongly suggests the need of a structural flexibility in the architecture of the enzyme. In other words, the enzyme might also experiences structural changes in order to guide the reactant in the proper orientation for the reaction to occur. Indirectly, this statement was introducing the premise of a new era in enzymology: conformational functionality [5].

The earliest indication of this structural flexibility came from crystallographic investigations, which was at first glance surprising since the crystallization process implies a structural homogeneity. Indeed, when the first two structures were solved by X-ray crystallography, lysozyme and ribonuclease respectively, small conformational changes were found between the structure in the presence and in the absence of substrate [6-8]. Daniel Koshland (who passed away in July 2007) was very surprised that these small changes were discarded and not being addressed as

evidence of a new theory: induced fit (1958) [9]. This particular mechanistic model supposes that the protein structure changes upon ligand binding. This view was therefore prompting the first challenging points to the traditional paradigm, 'one sequence, one structure, one function'. The magnitude of conformational diversity observed in proteins was just emerging and was later confirmed by several techniques such as Nuclear Magnetic Resonance (NMR) [5]. Nowadays with many structures from NMR and X-ray crystallography, structural changes are widely accepted as an intrinsic property of proteins. Consequently, there is good evidence that most, if not all proteins require molecular flexibility for function. These fluctuations range from side chain motions to loop or sub-domain movement or even tertiary structure rearrangement [10]. A growing and important collection of molecular motions is depicted in a specific database: The Database of Macromolecular Movements (http://molmovdb.org/) which could be considered as the dynamic counterpart of the Protein data bank (PDB) [11, 12]. The movements are identified and classified according to their nature and geometry. Overall, the most intriguing consequence of conformational diversity is that it provides a structural mechanism for protein function. A fine-tuning is operating between the architecture and the chemical process. However, if the structures of the reaction intermediates are sometimes well defined, the frequencies of the conformational changes and their biological importance remain largely unknown.

Pre-steady state kinetics and NMR studies have shed light on this fundamental question [13, 14]. They have given a picture of both kinetic and structural details of the catalysis mechanism of various enzymes such as cyclophilin A [15]. In most cases, a direct relationship can be established between the turnover number and the frequencies of specific motions. In other words, milli to micro second motions are the 'key stones' in protein activity since most enzyme catalysed reactions occur on these specific timescales [16]. Nonetheless, if these studies have contributed to improve our understanding of protein function, they have assumed that 'functional dynamics' must be restricted to this particular time window. Though this assumption can be applied to many proteins, some proteins do not exhibit such dynamics changes to perform their biological function but on the other hand, we cannot consider them as

rigid elements. Indeed, the hierarchy of protein dynamics spans over at least 15 orders of magnitude, some proteins fluctuate with speeds ranging from femto-seconds to minutes (table 1.). Clearly, the various classes of motions extending over multiple timescales throughout the protein might potentially play a role in protein activity and/or stability. However, as reviewed by Daniel *et al*, the role of dynamics in enzymatic activity is very versatile: different protein functions may have different dynamical dependencies. Conformational dynamics are likely to be dependent on the nature of enzyme and its catalytic mechanism [17]. For instance, quantum/classical molecular dynamics with site directed mutagenesis approaches have correlated femto-pico second timescale conformational changes with kinetic motions in the DHFR kinetic cycle (millisecond) [18]: suggesting that motions can be coupled to each other in a specific network throughout the protein. It also suggests that fast conformational fluctuation might be channelled into productive events for protein function, which is intriguing since their frequencies do not match the kinetically relevant motions [19]. On the other hand, protein stability is also dependent on dynamics, in the sense that native proteins are flexible within much narrower limits than their denatured form [20]. This has given rise to the paradigm that activity and stability are inversely related by internal dynamics; i.e.

Activity ∝ Dynamics (flexibility) ∝ 1/Stability

Consequently, proteins need to be sufficiently flexible to be catalytically active, but not so flexible that they readily denature. However, the nature of the dynamic requirements for activity and stability and in particular whether they are the same are unclear. Protein dynamics and how dynamical fluctuations are coupled to each others is still a paradigm of complexity. Various methods have been used to address this problem representing a gap in our understanding of protein function. A number of concepts have been advanced to rationalise their precise role but the basis of this inter-relationship still remains uncertain...

Motions	Amplitude (nm)	Approximate time scale Log (s)
Electronic rearrangement	0.001 to 0.01	-15 to -14
Atomic vibrations	0.001 to 0.01	-14 to -9
Loop motions	0.05	-11 to -6
Domain motions	0.1 to 0.5	-11 to -6
Allosteric transitions	0.1 to 0.5	-6 to 0
Folding and unfolding transitions	0.5 to 1	-3 to 1

Table 1. The timescale and amplitude of various dynamic events that can occur in proteins [21].

1.2. Fast motions for slow action? Fast swings for slow steps?

In the literature devoted to protein dynamics fast protein dynamics lie at the border of the spectrum with time scale ranging from the nanosecond to the femto second. The role of these motions that constantly fluctuate along the protein backbone and side chains is still a matter of controversy. A number of studies have clearly identified these motions as ‘random’ vibrations not necessary for protein function [22-24]. On the other hand, in the recent years, these ultra fast fluctuations have gained popularity in view of enzymology and protein evolution [5, 19, 25-28]. In fact, there is much evidence indicating, that many fluctuations can be coupled together to greatly enhance the catalytic power of enzyme or participate in protein stability [19, 29].

Are these motions simply inherent properties of proteins or have they been harnessed to enhance the functional efficiency?

A number of concepts have also emerged to rationalise the role of fast protein dynamics in terms of protein action. If proteins have a hierarchical spatial organization, which is naturally mapped onto the spectrum of their motions, we would in principle assume that the protein conformation adopted must be functionally relevant. NMR was used to test this determinist hypothesis in order to depict the 'true' number of alternative conformations of a single protein in solution. This number was exceeding largely what was expected to be the catalytically active conformations [30]. Interestingly, these findings have confirmed the postulate of Landsteiner and Pauling in 1930, suggesting that immunoglobulins can exist as an ensemble of isomers with different structures but with a similar free energy. The Monod-Wyman-Changeux (MWC, 1965), model of allosteric regulation was actually reformulating this concept of conformational diversity in view of enzymology [31]. The simple rigid scaffold of a protein was replaced by a more modern view: a dynamic equilibrium of pre-existing conformations. This idea although considered outlandish for decades is now elegantly presented in energy diagrams such as folding funnels found in Dill's theory [32]. This 'new view' of protein as inferred for instance by James *et al* depicts proteins using a rugged energy landscape [5] (figure 1).

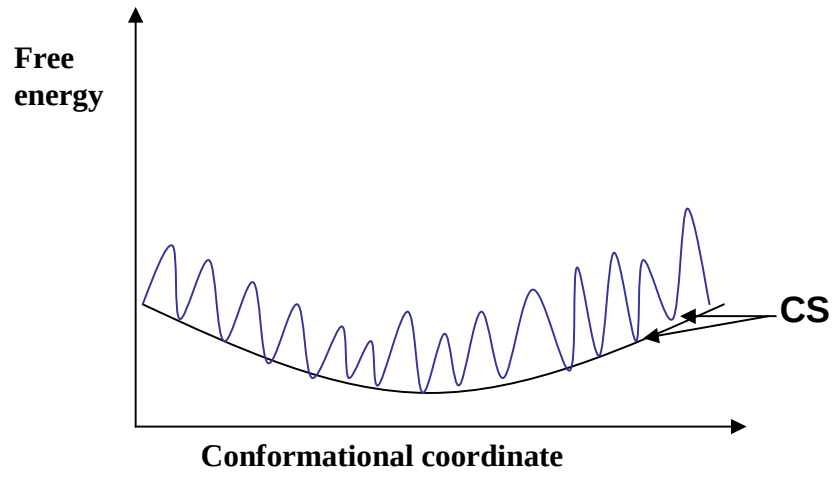


Figure 1. Two dimensional energy landscape of a protein**.** The energy minima correspond to the different conformational substates (CS) explored by the protein.

This theory was originally formulated by the Frauenfelder's group 20 years ago [33]. Indeed, energy landscapes are defined as the potential energy of the system as a function of all coordinates. The native folded form of a protein can assume a very large number of nearly iso-energetic conformations or conformational substates. The hierarchy is organised in tiers which are shifted by the presence a specific ligand [29, 33-35]. Energy landscapes are therefore important since they help to characterize the different possible structures, but also because they determine the dynamics of the system. More recently, a number of studies have concurred to build an unifying picture, connecting: protein structure, energy landscape, dynamics and enzyme catalysis. Free energy landscapes can be viewed as three-dimensional catalytic pathway [29, 36]. Enzymes reactions are suggested to operate through "catalytic networks" involving multiple energy minima (reaction intermediates) and unfavourable energy states (transition states) (figure 2). Thus, this new view is stochastic rather than deterministic and emphasises growing evidence for 'non-catalytically active' conformations [19]. In other words, enzymes must be considered as existing in an ensemble of coupled conformational states, working in concert to achieve their remarkable efficiency [37].

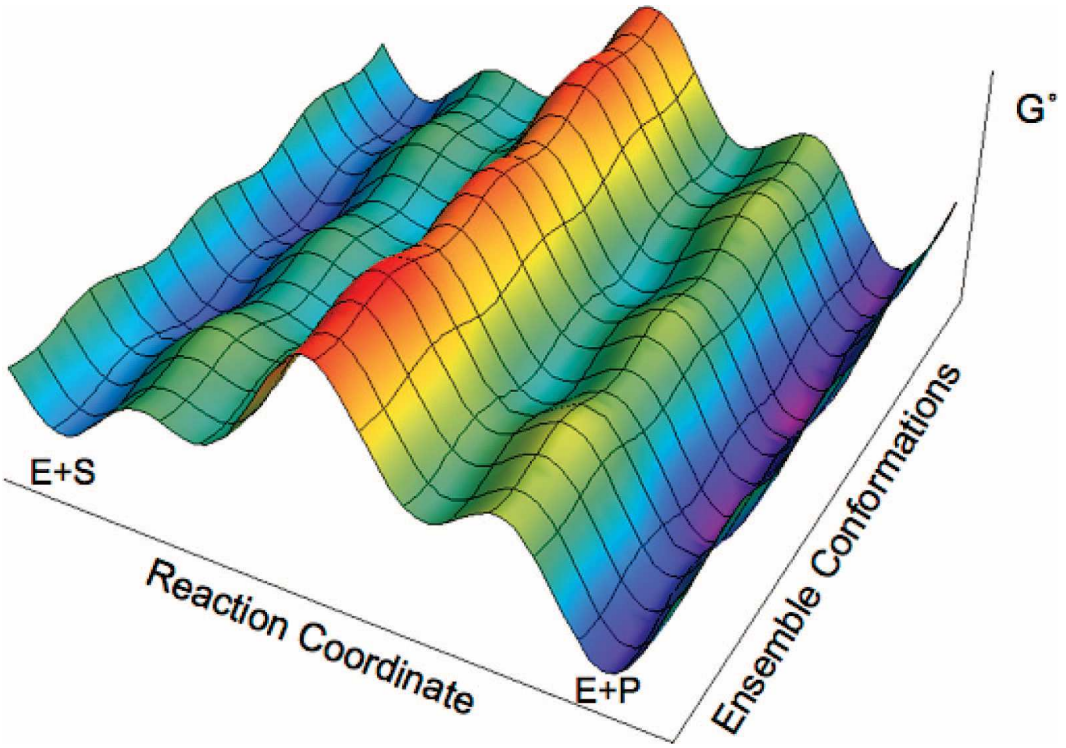


Figure 2. Schematic three-dimensional representation of the standard free-energy landscape for the catalytic network of an enzyme reaction**.** G° is the free energy of the system, E+S represents the enzyme and its substrate and E+P the enzyme and products respectively. This figure illustrates the multiple populations of conformations, intermediates (minima) and transitions states (maxima) that may facilitate the enzyme catalyzed reaction. This figure was created by Sarah Janes Edward and reproduced from reference [36].

However, we still do not know how the different conformational substates inter-converse. Fast protein dynamics might promote the vibrational modes necessary for the jumps between the various energy levels or conformational substates (CS). Intuitively, these 'promoting motions' could be coupled to slower motions relevant for protein function [38]. In this respect, fast protein dynamics are believed to serve as lubricant by enhancing the conformational space sampled by the protein [17]. Computational approaches such as molecular dynamic (MD) simulation have contributed to bring some insights in this framework [39]. The calculations and consequently the timescales accessible from simulation are matching the motion frequencies (ps-ns). These approaches have naturally provided an opportunity to examine if fast motions were required in energy barrier crossing, which are in principle localized and of high frequency. In view of enzymology, fleeting states such

as Transition State (TS) or Near Active Conformers (NACs) might require such motions to progress along the reaction coordinate [40, 41]. By definition, NACs are a ground state of the Michaelis complex which differ slightly in the geometry of the TS within 0.4 Å and 15º [42]. The changes in going from the enzyme NAC to enzyme TS are therefore very subtle and may therefore link backbone fluctuations among the conformational substates population to their functional relevance. A recent MD study has examined the negative effect of a site specific mutation in the network of coupled motions over different time-scales[18]. A clear connection between backbones fluctuations on the ns-ps timescale with the kinetic rate of DHFR [43] was established. It seems also that the efficiency of hydrogen tunnelling might depend on fast motions. This phenomenon by which a particle 'digs' in the reaction barrier because of its wavelike property is exquisitely sensitive to the distance or donor and acceptor molecule and temperature [44]. The tunnelling effect decreases respectively with temperature. It is therefore tempting to assume that these thermally excited fluctuations are involved in the hydride transfer reaction. Consequently, if these views turn out to be correct, ps to ns motions will be in particular responsible for the crossing of activation barriers in protein function. In addition, these fast motions would determine the shape of the free-energy landscape of the catalysed reaction [18, 45].

Now the research in these fields has to address these finding on issues relating to protein allostery, enzyme evolution and drug design. The implication of fast protein dynamics in fundamental sciences might profoundly affect our understanding of protein structures.

1.3. The dihydrofolate reductase enzyme: A prime model

1.3.1. Presentation

DNA is one of the major macromolecules in the cell, it supports the genetic information necessary for the cell to produce proteins. The DHFR enzyme, which is found in all living organisms is of prime importance in this path. It catalyses the

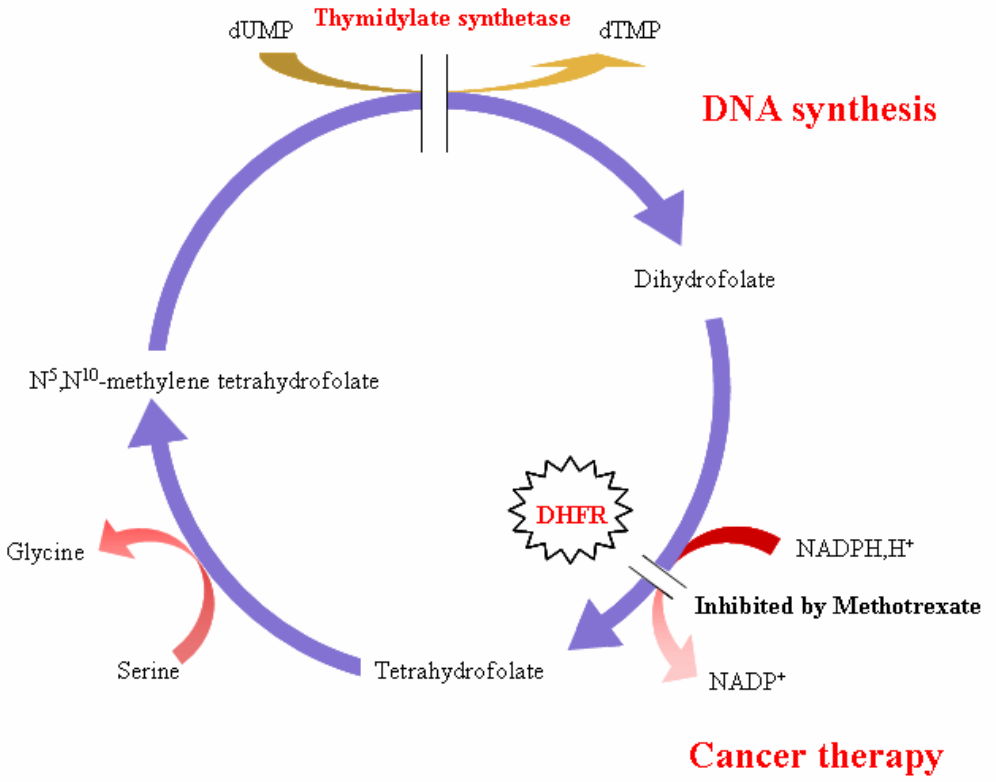


Figure 3. Schematic representation of the folate pathway**.**

reduction of dihydrofolate into tetrahydrofolate and therefore participates in normal folate biosynthesis. Consequently, DHFR is a key-stone in nucleic acids and protein biosynthesis (figure 3). The biological relevance of this particular enzyme was first indirectly highlighted by cancer therapy strategies, which were using inhibition of the folate pathway to slow and block the growth of tumors. The discovery in 1957 of the dihydrofolate reductase enzyme by Futterman was the starting point of an avalanche of studies not only dedicated to its medical application but also in fundamental research field [46, 47]. This wealth of information on the DHFR enzyme was taking

advantage of the important number of anti-cancer strategies inhibiting directly the DHFR enzyme. For instance, the development of drugs such as methotrexate, which is known as a DHFR stoichiometric inhibitor [48], is still one of the most potent treatments in certain forms of cancer [49]. The success of this enzyme as a prime target in cancer therapy is still not questioned; the recent development drug Permetrexed for lung and oesophageal cancer reflects its relevance in drug design [50]. Nowadays, this success has expanded beyond the pharmaceutical world, extending to a number of research fields, which encompass not only applied science but also fundamental research areas such as enzymology and computational science. The structure analysis of this small enzyme has revealed an interesting architecture: the structure is made of two domains: the adenosine binding domain and the major subdomain which are organized in an α/β fold [51]. The β secondary structures forms a relatively rigid skeleton counterbalance by an extensive number of loop which accounts about 40% of the major subdomain. The loops are believed to control the catalytic power of the enzyme by acting as a dynamic gate [51]. This unique interrelationship has therefore prompted a growing interest in protein function, particularly in enzymology where protein motions i.e. protein dynamics are implicit in enzyme action but not completely understood. Overall, the DHFR biological relevance in cancer therapy and this unique structure/function relationship have promoted the DHFR enzyme as a prime model for an unprecedented number of research projects.

1.3.2. Kinetic mechanism

The dihydrofolate reductase first discovered by Futterman in 1957 belongs to an ubiquitous pathway found in all living organism. It plays a key role in the normal folate pathway enabling nucleic acid and protein synthesis. DHFR catalyses the reduction of 7,8-dihydrofolate (H_2F) to 5,6,7,8-tetrahydrofolate (H_4F) using the reduced form of nicotinamide dinucleotide phosphate (NADPH) as a cofactor (figure 4). In many organisms, the reduction of folate to H_2F is also catalysed, although less efficiently.

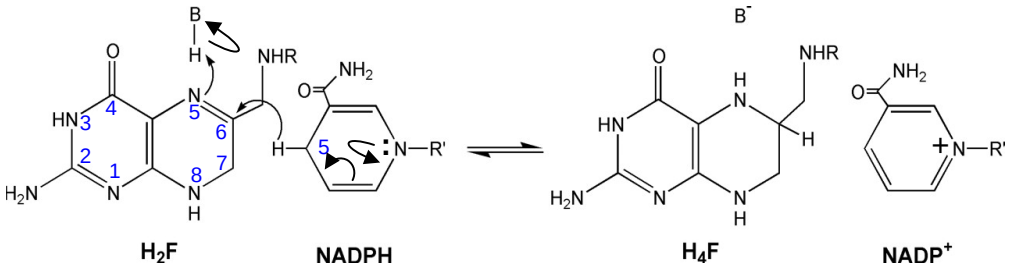


Figure 4. Schematic representation of the reduction of dihydrofolate in tetrahydrofolate using NADPH as a cofactor by the enzyme DHFR**.** This material is based upon the review of Benkovic *et al* [45].

The reaction occurs by stereospecific hydride transfer of the pro-R hydrogen of NADPH to the C6 atom of the pterin ring of H_2F with concomitant protonation at N5 [45]. However, the protonation mechanism is still a matter of debate. Indeed, for hybrid transfer to proceed, the substrate must be protonated at the N5 of the pteridin ring but at physiological pH, only a small proportion of the substrate is protonated [52]. Hence, numerous experimental and theoretical studies have attempted to rationalise a possible mechanistic scheme [47]. However, the complex contribution of amino acids distal or proximal to the active site, long-lived water molecules and the effect of coenzyme binding still hindered the answer. The main question is how the enzyme provides an architecture that complements the chemical reaction resulting in the correct ionisation of H_2F for hydride transfer to occur. Not surprisingly, site directed mutagenesis and sequence alignment has shed the light on a strictly conserved carboxy residue in the active site (Aspartate 27 in *E. coli*, Aspartate 26 in *L. casei*, Glutamate 30 in human DHFR). The interaction of this acidic residue and the pteridin fragment was proposed to drive the substrate protonation. In addition, a keto-enol tautomerism mechanism, involving long-lived water molecules in the vicinity of the active site, was proposed to participate actively in the mechanism. However, the position of the water molecules obtained by X-ray crystallography did not agree with the model, which finally raised more questions than it answered.

Consequently, in spite of the abundant data, the mechanism of substrate protonation is not fully understood [53]. On the other hand, pre-steady and steady state kinetic analyses have given a detailed map of *E. coli* DHFR catalytic scheme Benkovic *et al* [45]. The enzyme is cycling between five kinetically observable intermediates with a preferred pathway (figure 5). The cofactor NADPH is believed to initiate the catalysis by forming the binary complex E:NADPH. The subsequent binding of H_2F leads to the fleeting Michaelis complex E:NADPH:H_2F, and to the products E:NADP+:H_4F, E:H_4F and E:NADPH:H_4F. In this complex mechanism, the association and dissociation of the ligand are of primary importance.

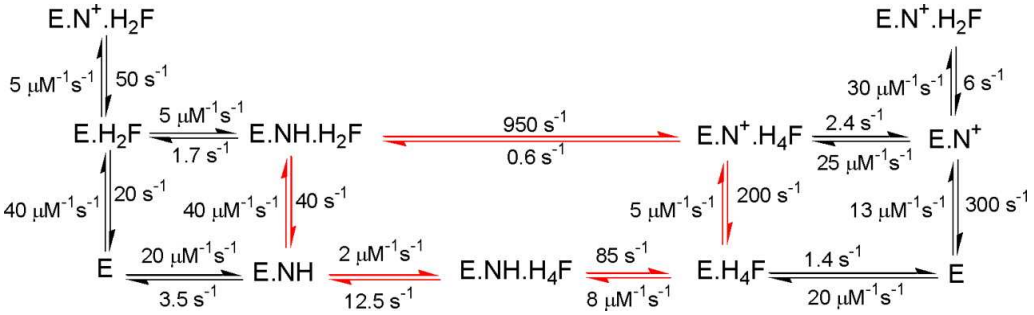


Figure 5. Kinetic scheme for DHFR catalysis at 25°C. E is DHFR, NH is NADPH; N+ is NADP+;H_2F is dihydrofolate,H_4F is tetrahydrofolate. The preferred pathway during multiple turnovers is highlighted in red [45].

It actively modulates the overall rate of the process along a preferred pathway. The hydride transfer rate, which is the key stage of the catalytic process, is one of the most rapid stages in the catalytic cycle whereas the limiting step is associated with the rebinding of a new NADPH molecule. Interestingly, the reaction product H_4F is not released after hydride transfer. A new NADPH molecule must bind the binary complex, E:H_4F, for H_4F release. A synergistic coordination is therefore achieved between product release and cofactor/ substrate binding which plays a key role in the regulation of the whole pathway.

1.3.3. Structure

The pharmaceutical and biological relevance of DHFR have led to a number of structural studies. Various techniques such as X-ray crystallography, NMR, and neutron crystallography have contributed to a detailed map of the architecture of the enzyme in various ligated states. In particular the work of Sawaya and Kraut has depicted the reaction intermediates of the kinetic cycle with more than 24 different isomorphous crystals [51]. The corresponding number of *E. coli* structures found in the Brookhaven Protein Databank (PDB) now exceeds 80. In fact, *E. coli* DHFRs complexed with small molecules such as drugs have been very popular over the last 20 years. In addition, despite a low sequence homology between DHFRs isolated from various sources (less than 30 %), DHFRs display a high degree of structural similarity. Sequence alignment for DHFRs shows a series of strictly conserved residues not only at the active site but also in the distal part of the enzyme. Consequently, this wealth of structural studies has brought *E. coli* DHFR a role as a prime 'model' and the main reference against which other DHFRs sources are compared.

Overall, all chromosomally encoded DHFRs are small (≈18-25 kDa for most enzyme species isolated from various sources) single domain proteins organized in two subdomains, the adenosine-binding domain and the loop domain also named the major subdomain [14, 51]. These two structural motifs are organized in a α/β structure consisting of a central eight stranded β-sheet (composed of β-strands A-H) and four flanking α-helices (designated αB, αC, αE, and αF) connected by the loops. The β-sheet structure is suggested to form a rigid skeleton balancing the flexibility from the loop sub-domain [51]. The adenosine binding subdomain (residues 38-88 in *E. coli* DHFR) which is composed of two α-helices is the smaller of the two subdomains and provides the binding site for NADPH moiety. On the other hand, the major subdomain consists of approximately 100 residues from the N and C termini and is dominated by a set of three loops that surround the active site. Interestingly, in terms of sequence length, these loops makes up approximately 40 %-50 % of the major subdomain; hence it is sometimes called the loop subdomain. This subdomain is made of three loops: Met20 loop (residues 9-24), F-G loop (residues 116-132), and

G-H loop (residues 142-150) in *E.coli* DHFR [51]. The Met20 and F-G loops are also referred to in the literature as loop 1 and loop 2, respectively.

At the interface of the two domains, we find a hinge region, which plays a role in the access to the active site through bending motions. The adenosine binding domain moves relative to the major domain upon binding of various ligands, resulting in closure of the active site cleft. The active site cleft lies in the hydrophobic pocket surrounded by the α-helix B, the central β –sheets (a, e and b), and the loop I. The nicotinamide fragment of the cofactor resides in a deep pocket in the vicinity of the substrate whereas the rest of the molecule is positioned at the protein surface. Similarly, the pteridin ring of the substrate resides in the centre of the protein with its *p*-aminobenzoylglutamine part at the protein surface (figure 6). This structural arrangement provides the environment for hydride transfer to occur. The Met20 loop is often depicted in the literature in respect with the catalytic pathway with two distinct conformations, closed and occluded and four in the crystalline state open, closed, occluded and disordered (figure 9). The conformation of the active site loops depends on the ligand bound in the substrate and cofactor binding site. The loop adopts the closed conformation when it packs against the nicotinamide ring of the cofactor bound in the active site. On the other hand the occluded conformation is prompted when substrate site is occupied. The amplitude of variation between closed and occluded conformation is up to 9Å. The Met20 loop is therefore an active gate controlling the access of the active site.

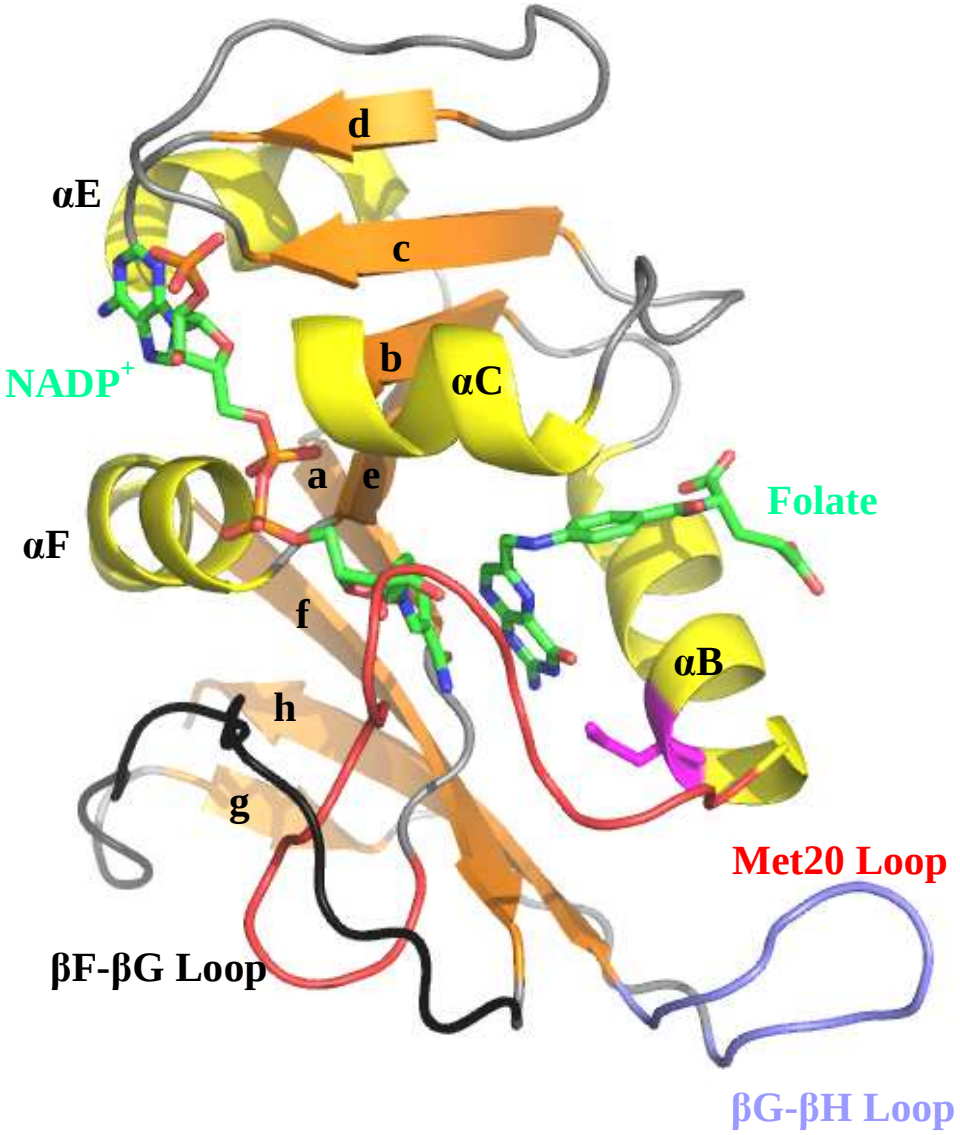


Figure 6 Cartoon representation of the tertiary structure of E. *coli* DHFR bound to the substrate (folate) and the cofactor (NADP+). Secondary structural elements, α-helices in yellow and β-strands in orange are denoted by capital Latin letters (B, C, E and F) and lower-case Latin letters (a-g) respectively. Major loop regions are labeled. The active site Asp 27 is labeled in purple. The adenosine-binding loop (residues 62-70) is at the top of the structure, linking the c and d β-strands together. The coordinate were taken from the X-ray structure of the corresponding ternary complex (Brookaven PDB code 1RX2). This figure was made using PyMOL [54, 55]

1.3.4. *Bacillus stearothermophilus* DHFR structure

Since our laboratory (Thermophile unit) is interested in the interrelationship between activity/stability and protein function under extreme temperature conditions, we have used a DHFR from thermophilic organism: *Bacillus stearothermophilus* (*Bs* DHFR*)*.

The X-ray structure of the free form of *Bs* DHFR (1ZDR) was first determined by Klinman *et al* [56]. This structure is the first monomeric DHFR structure from a moderate thermophilic organism. The structure was resolved at a 2 Ångstrom resolution and exhibits the common α/β fold organised in two subdomain found in all chromosomal DHFRs. Relative to the other known structures of DHFRs, the sequence of *Bs* DHFR shares the greatest sequence identity and similarity with the DHFR from *E. coli* (38 and 58 %, respectively). In comparison, DHFRs from *Lactobacillus casei* or *Thermotoga maritima* share 35 and 23 % of sequence identity and 56 and 40 % of similarity respectively (figure 7) [56]. The superimposition of the two structures from *E. coli* and *Bs* show a very little deviation in the protein backbone (figure 8). The average RMSD/residue was found to be 4.0 and 4.7 Å for the main and side chains respectively (table 2). The folding pattern is extremely well conserved between the various evolved thermal sources.

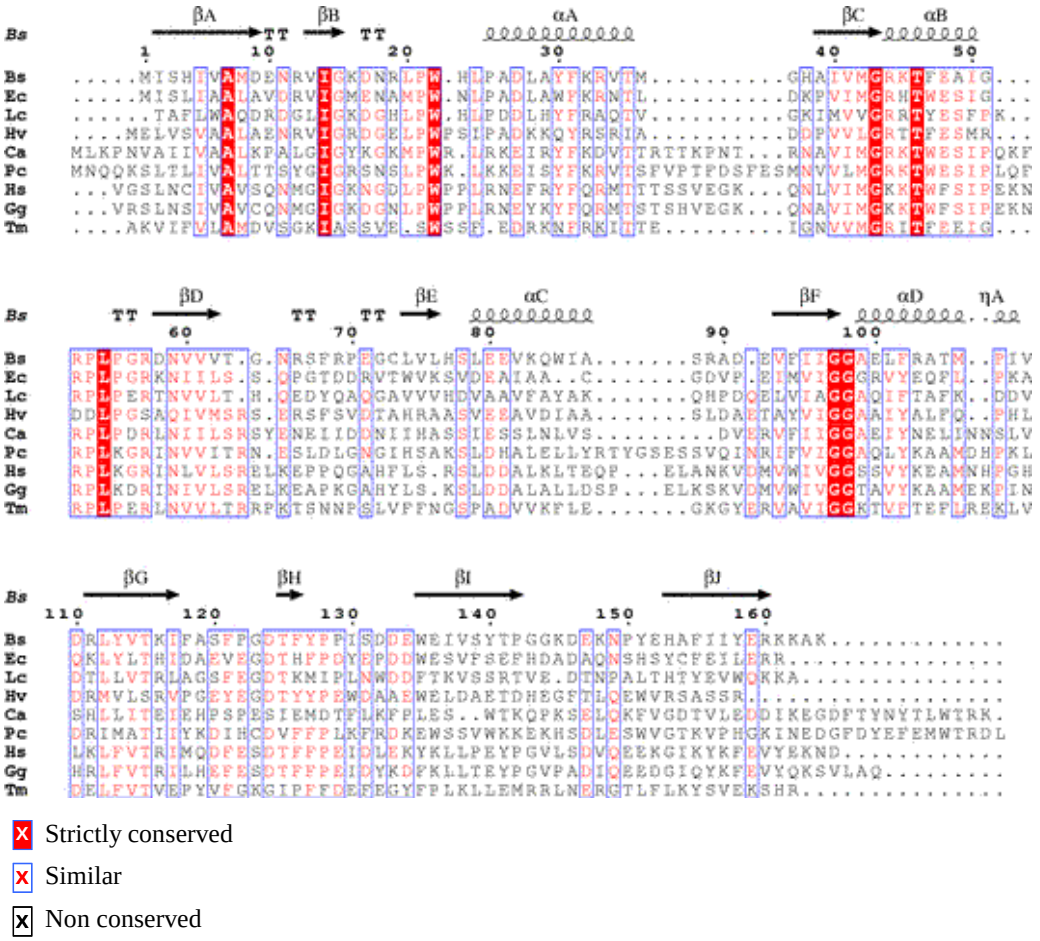


Figure 7. Structure-based sequence alignment [57] of representative DHFR structures: *Bs, B. stearothermophilus;Ec, E. coli; Tm, T. maritima; Ca, C. albicans; Pc, P .carinii; Hs, H. sapiens; Gg, G. gallus; Hv, H. volcanii; and Lc, L. casei.* Secondary structures found in *Bs* DHFR are indicated on the top line. The sequence alignment was constructed using the software Multalin [58].

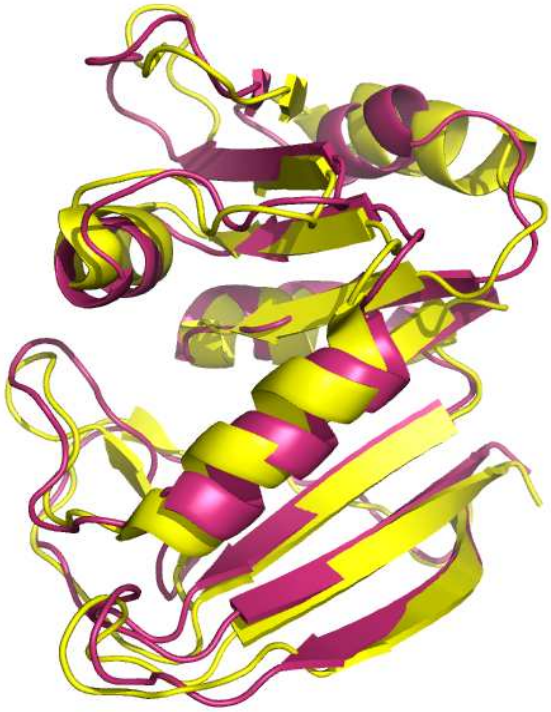


Figure 8. Overlay of three-dimensional structure of *Bs* DHFR (1ZDR, magenta) and *E. coli* DHFR (1DRA yellow). This figure was made using the program Superpose [59] available in the CCP4 package (Collaborative Computational Project, Number 4) [60].

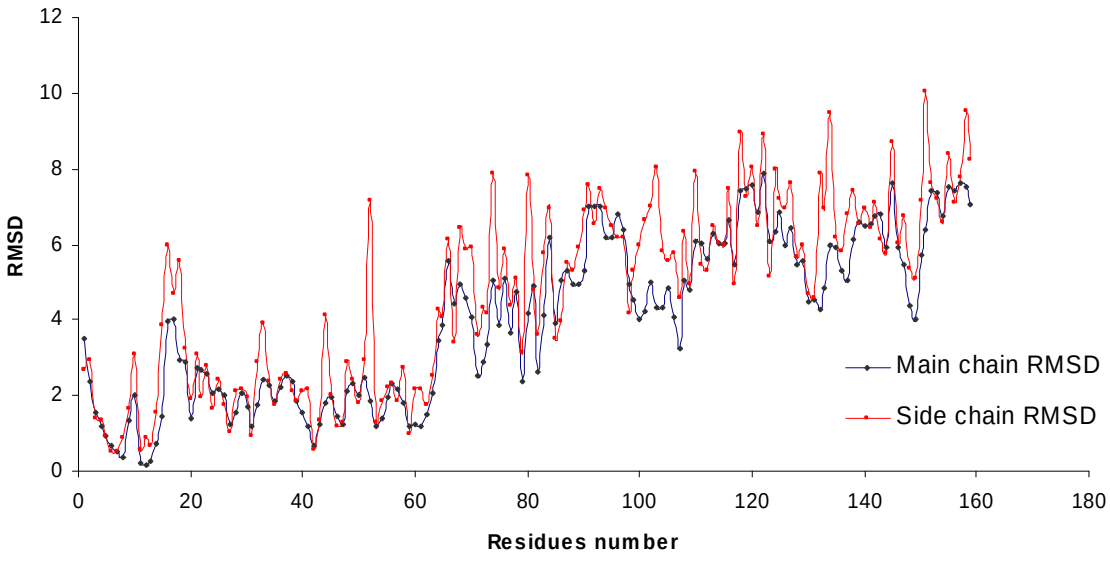


Table 2. RMSD values of the main chain (blue) and side chain (red) between *Bs* DHFR and *E. coli* DHFR. The graph was obtained using the program Superpose [59].

1.3.5. Motions and DHFR catalysis: a synergistic mechanism

In this picture of catalysis, which takes advantage of the X-ray and NMR models, the structure of all the kinetic intermediates or models of the intermediates have been solved. The various intermediates have been compiled in a movie to highlight the dynamical key elements in view of the kinetic cycle (http://www.chem-faculty.ucsd.edu/kraut/dhfr.html). This study has unambiguously revealed an essential role for substantial backbone and side chains motions in cofactor and substrate binding and in the catalytic cycle. Almost 40 % of the major subdomain is formed by loops (Met20, F-G and G-H) which potentially implies a high structural flexibility. In particular, the Met20 loop is one of the major sites of conformational change with the substrate binding pocket. The closed, occluded, and opened conformations correlate to the various binding states along the reaction cycle (figure 9). The loops movements are believed to assist the proper alignment of the reactive groups involved in the chemical reaction [47]. The timescale and amplitude of backbone motions depend primarily on the conformation of the active site Met20 loop. For instance conformational changes on the micro-millisecond time-scale appear to be related to the structural changes in the occluded-closed transition [61]. This is of particular interest since these motions are comparable to the timescale of product release (12 s^{-1}), which is the rate-limiting step under conditions of steady state turnover [45]. In addition, in the closed conformation, picoseconds-nanoseconds timescale motions become attenuated [62]. The various conformations adopted by the Met20 loop have therefore specific fluctuations modes over a wide range of timescales, from pico to milliseconds. In this respect, the Met20 loop serves as a dynamic gate controlling the active site accessibility. These finding are prompting strong evidence of a synergistic mechanism between specific motions and catalysis. Boehr *et al* have elegantly deepened our understanding on these so called ‘promoting’ dynamics or motions in the DHFR catalytic cycle [37]. Using NMR relaxation dispersion they have detected a conformational exchange on a micro to millisecond time scale between the ground-state structure and one or two excited states that resemble the ground state of the preceding and/or the following intermediate in the

catalytic cycle. They suggest that there is an active cycling conformational selection during DHFR catalysis. It provides solid evidence that the whole kinetic pathway is governed by the dynamics of the conformational fluctuations between the ground and excited states of all kinetics intermediates. Consequently, certain enzymes such as DHFR combine flexibility and plasticity to achieve their function. The coordination of the fluctuations is therefore expected to have a long range impact on the performance and the regulation of the enzyme.

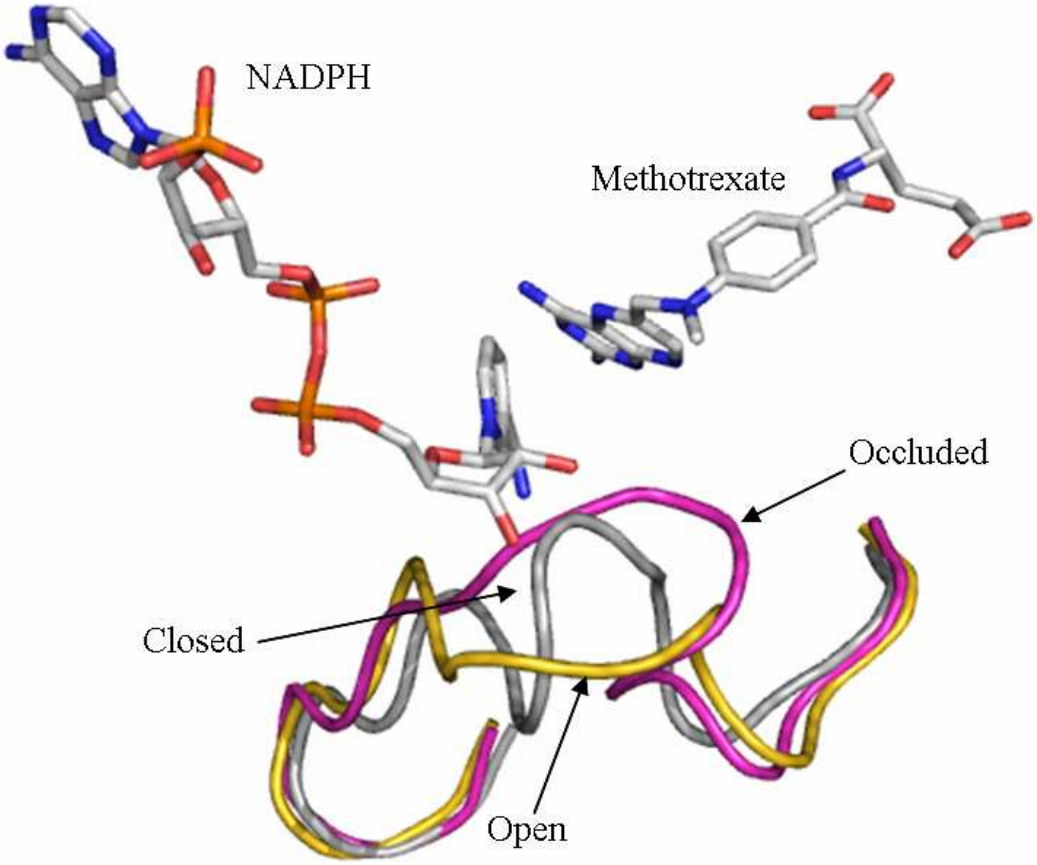


Figure 9. Three main types of the Met20 loop (residues 9-24): open, closed, occluded. The picture is based on the atomic coordinates of the X-ray structures of the complexes of E. coli DHFR with NADPH (open, gold), with NADPH and MTX (closed, grey), and with H2F (occluded, purple) (PDB codes are 1RH3, 1RA1 and 1RX5, respectively). This figure was built using PyMOL [54] based on the reference [63].

1.3.6. The transition state analog Methotrexate

DHFR is a key enzyme in normal folate synthesis. Folate is an essential precursor metabolite required for many one-carbon transfer reactions involved in purines, pyrimidines synthesis, and amino acids [47]. This particular biochemical pathway has therefore gained over the years an important potential for inhibiting rapidly dividing or invasive cells. Sidney Farber and coworkers at the Children's Hospital Bostonhad used this fact to cure cancers. In the 1940s, they were the first to observe the effects of aminopterin, a folate antagonist, on cancer remission. The therapeutic potential of the folic acid biosynthesis pathway was just emerging. The importance and recognition of this work came in 1988 when Hitchings, Elion and Black shared the Physiology and medecine Nobel prize for their joint work on this important biosynthetic pathway (http://www.en.wikipedia.org/wiki/Methotrexate). Nowadays, the purine and pyrimidine biosynthesis pathways have triggered an avalanche of studies which have eventually led to the development of therapeutic agents such as trimethoprim, pyrimethamine, methotrexate [49]. These so-called antifolate drugs compete with dihydrofolate for the substrate binding site in DHFR. Particularly pertinent to our discussion is the impact of MTX on the catalytic properties of DHFR. MTX is a folate analog, only differing at position 4 with respect to folate competing reversibly with folate. It was developed by Yellapragada Subbarao in 1948 and is now used to treat million of patients suffering with both malignant and autoimmune diseases.

To understand the physiological and pharmacologic effects of MTX in the cell, we shall see MTX inhibition as a combination of at least two distinct effects: a partial depletion of reduced folate substrates but also in a direct enzymatic inhibition of various key enzymes involved in the pathway such as DHFR, and thymidylate synthase (Ts) [64]. Although MTX is not in the strictest sense a specific inhibitor of DHFR, this enzyme has become its prime target [65]. Indeed, overall drug effectiveness depends significantly of the strength and selectivity of the inhibitor. In this respect the interaction between MTX and DHFR has become a perfect example of a high affinity interaction. The dissociation constant of methotrexate for *E. coli* DHFR is about one thousand-fold that of folate (K_d= 10^{-10} Mol l^{-1}) [65]. MTX has

become the 'DHFR stoichiometric inhibitor'. However one fundamental question remains unsettled:

Why does MTX bind so tightly to DHFR?

It is undoubtedly due to the structural features shared with H_2F. Most potent inhibitor of chromosomal DHFR contain N1-2,4-diamino heterocycles. MTX and H_2F differ noticeably at position 4, where the 4-oxo group of DHF is replaced with a 4-amino group (please, see figure 10). Numerous crystallographic studies on MTX DHFR complexes have revealed a specific geometry of binding [14, 51, 66]. The MTX binding site is well defined, a loop that forms the active site, residues 9-24 in *Lactobacillus. casei* closes over MTX when bound to DHFR [51, 66]. On the other hand, the interaction of small molecules with DHFR was depicted extensively by NMR studies. Interestingly, the pteridin ring of MTX is inverted (180^o) in the active with respect the natural substrate H_2F [67, 68]. This flipped orientation is believed to alter the H-bond network favoring a direct interaction between two conserved backbone carbonyl oxygen from the highly conserved residu Asp 27 and the 4-amino group from MTX (figure 11) [69]. In addition, on going from the free to the bound state, the pKa of MTX (complexed with *L. casei* DHFR) marked an increased, from 5.3 to 10.5 [70]. It further favors the H-bond formation between the conserved active site carboxylate group which is negatively charged and the protonated N1 from MTX. This specific architecture contributes actively to the stabilization of the MTX heterocycle in the enzyme active site [51]. The total energy of binding of MTX to DHFR involves also a significant contribution from hydrophobic interactions [71]. The pteridine and benzoyl rings of MTX interact with lipophilic amino acids present in the active site of the enzyme. For instance, the key residues ensuring such interactions are known in *L. casei* DHFR: Leu 27, Phe 30, Phe 49 and Leu 54 and contribute to a stabilizing energy not lower than 10.5 kJ mol^{-1}[64].

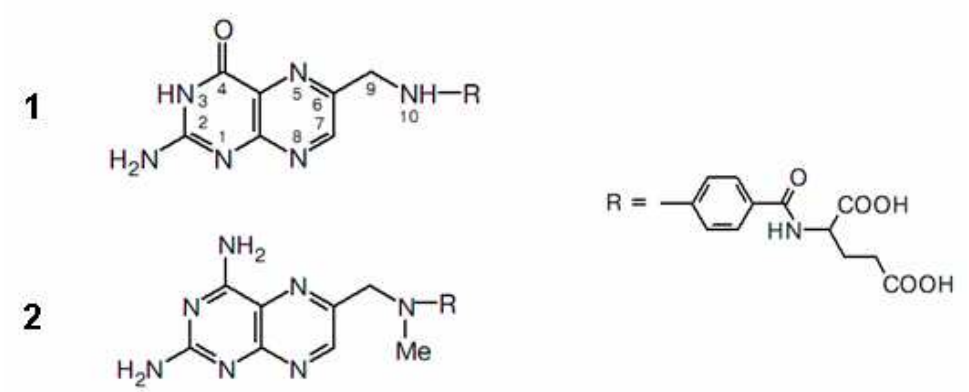


Figure 10. Structure of Dihydrofolate (1) and Methotrexate (2).

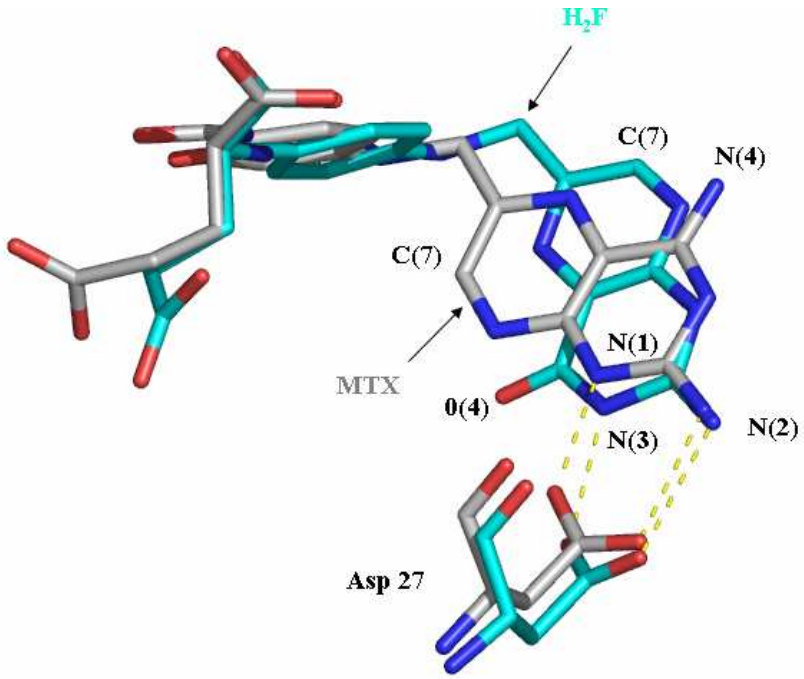


Figure 11. The arrangement of the pteridin ring of MTX (grey) and H_2F (light blue) with respect of the carboxy group of the aspartic residu 27 (Asp 27). Selected atoms are numbered. The atomic coordinate were taken from the structures of the complexes *E. coli* DHFR-MTX NADPH (PDB code 1RX3) and *E. coli* DHFR H_2F $NADP^+$ (1RX2) and were superimposed by the protein backbone. The figure was made using PyMOL [54] .

However, even if it is now clear that these structural factors contribute to the extremely high affinity of MTX for the DHFR, it does fully describe the unique properties of the inhibitor toward DHFR. Interestingly, the dynamic characteristics of the interaction between MTX and *E. coli* DHFR were studied by stopped-flow kinetics and single molecule fluorescence [72]. The binding of MTX is a complex process involving multiple enzyme conformations. In particular, a conformational change was attributed to the opening and closing of the Met20 loop and later confirmed by X-ray and neutron crystallography studies [51, 67, 73]. This conformational diversity may underlie the great potency of MTX toward DHFR catalysis. One other striking property is the cooperative binding of NADPH when MTX is bound to the enzyme [74, 75]. The affinity constant in the binary complex NADPH-DHFR is lowered by 100 times when the inhibitor is bound. This observation supports interestingly the idea that MTX can favor particular conformational state of the enzyme and therefore lowered free energies. Thus, the DHFR-MTX complex would represent a favorable energy state of the enzyme whereas the transition state marked by the association of NADPH-H_2F-DHFR would represent the maximal height in the energy profile of the reaction [63, 75]. This discrepancy in the binding constant between the two complexes, DHFR MTX NADPH and DHFR H_2F NADPH might provide the evidence for a synergistic arrangement of MTX with enzyme structure, function and dynamics.

1.4. How to study fast protein dynamics: Theory of neutron scattering

1.4.1. Basic properties of the neutron

A number of experimental and theoretical techniques can be used to assess protein dynamics. However, the key parameter to choose a particular method is invariably dependent of the space and time window of interest. Using a schematic energy diagram (figure 12), it is clear that the scope of protein dynamics can not be measured by a single method.

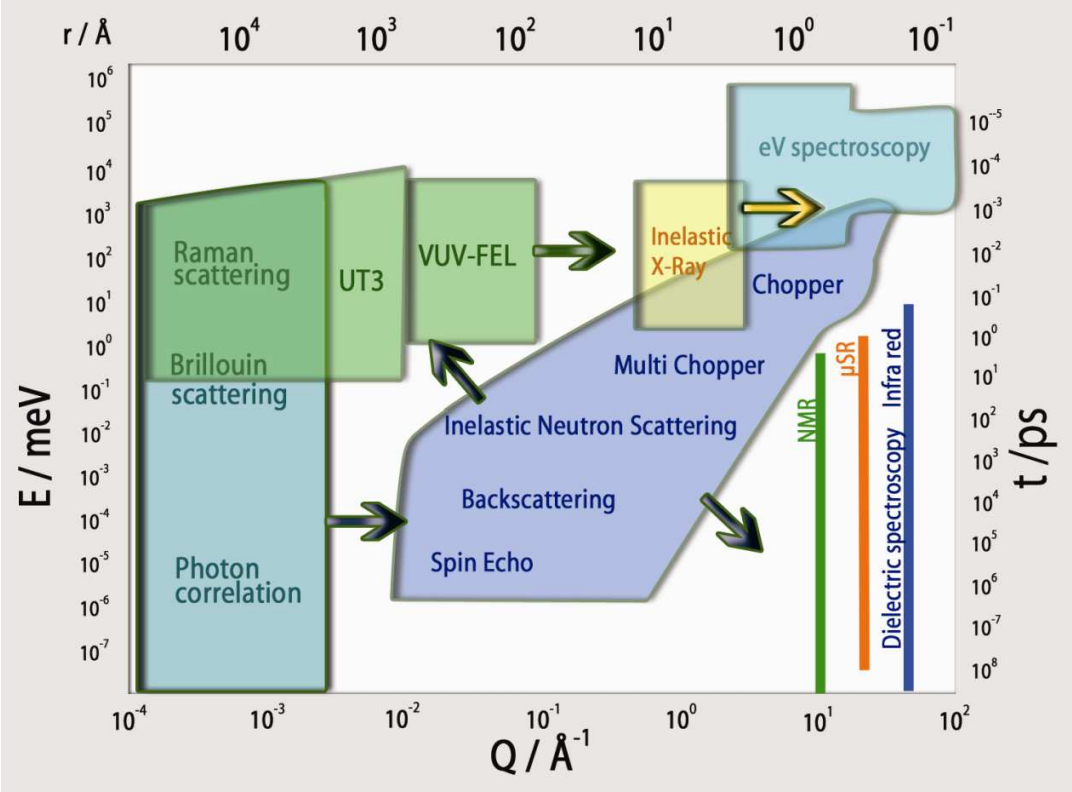


Figure 12. Comparison of neutrons with other biophysical methods: time and space ranges accessible to different techniques. The different types of neutron spectrophotometers are denoted by chopper, backscattering, multi chopper, spin echo. Adapted from www.neutron-eu.net.

From MD to NMR or fluorescence, several orders of magnitude are covered and sometimes overlap. However in most cases, the different methods used to probe protein dynamics gives complementary information based on the radiation-sample interactions. To a certain extent the method of choice depends of the advantages and disadvantages which can be obtained. Our goal on the present work was to focus on fast protein dynamics which occurs on the ps-ns timescale with a space window in the Ångstrom range. Early experiment in this field used Mossbauer spectroscopy [76] but as illustrated, NMR, and neutron scattering, can also provide significant insights in these regions.

1.4.2. A powerful tool for structural biology: neutron scattering

Neutron spectroscopy was developed after the Second World War, in the late fifties, through essentially E. Fermi. and later on by C.G Shull and Brockhouse who obtained the Nobel prize in 1994 for their pioneering work. However, the establishment of neutron spectroscopy in biology is surprisingly recent, the first experiences on DNA fibres were performed by Dahlborg and Rupprecht in 1971 and further on membrane proteins in 1972 [77]. Nowadays, neutron spectroscopy has gained in popularity with the development of innovative instruments dedicated to biological sciences [78]. Techniques such as small angle neutron scattering have reinforced the strength of neutrons in biology [79, 80]. The opening of new research centres with more powerful sources has also contributed to the rapid development of neutron methods in sciences.

The unique features of neutron spectroscopy reside in physical properties of these subatomic particles. Neutrons were first discovered in 1932 by J Chadwick in the nucleus of atoms. Neutrons are either produced in nuclear reactor through the fission of uranium-235 or in spallation sources when an accelerated proton beam hits a heavy metal target. After moderation, the neutron beam obtained has an energy between 0.2 and 500 meV (1 meV=1,602.10^{-22} J) with wavelength ranging from 2 nm to 0.04 nm. The most striking properties of neutrons are their relatively high mass: 1,675.10^{-27} kg (i.e. 1836 times the mass of an electron), the absence of electric charge, and their spin: 1/2 [81].

The energy of a neutron is related to its speed by the following equation:

$$E = \frac{1}{2}mv^2 \qquad \textbf{Equation 1}$$

As all subatomic particles, neutrons obey to the laws of quantum mechanics, which means they behave like waves as well as particles, therefore we can alternatively use a planar wave of wave vector $\vec{k}$ with a wavelength λ to describe neutrons. The module of the wave vector is equal to

$$\left|\vec{k}\right| = 2\pi/\lambda \qquad \textbf{Equation 2}$$

Their energy can be therefore described with the following relations

$$E = \frac{\hbar^2 k^2}{2m} = \frac{h^2}{2m\lambda^2} \qquad \textbf{Equation 3}$$

With h being the Planck constant

$$\hbar = \frac{h}{2\pi} = 1{,}055.10^{-34}\, J.s \qquad \textbf{Equation 4}$$

A neutron of energy of 25 meV, at a temperature of 293 K, has an associated wavelength: λ of ~ 1.8 Å with a velocity of ~ 2.2 km/sec. These values clearly illustrate the strength of neutron in biology, their wavelength (about tenth of nanometers) is comparable to interatomic distances found in proteins and their associated energy (meV) are very similar to the forces that maintain biological structure such as hydrogen bonding, electrostatic interaction, van der Waals or ones arising from the hydrophobic effect [82]. These forces are therefore weak and are in the range of a few meV. Neutron spectroscopy therefore matches uniquely the amplitude and energies associated with atomic motions found in biological materials. Other important advantage of the neutron method is that it can characterize the dynamics of any type of biological sample, which not need not to be crystalline [83]. Small Angle Neutron Scattering takes advantage of this property and allows the exploration of ‘soft’ matter structures with length between 0.5 and 500 nm [79, 84].

As C.G Shull and Brockhouse summed up elegantly in their Nobel citation, neutrons enable us to know 'where atoms are and what atoms do'.

1.4.3. Interaction between neutrons and matter

Neutrons interact with atoms of an element in a manner that does not seem correlated with the atomic number of the element (table 3). It varies from one isotope to another. Unlike X-rays, neutrons interact with the nucleus of atoms with an order of 10^{-15} m. The nucleus can therefore be considered a point-like scatterer, since the neutron wavelength is five orders of magnitude larger, implying that the waves scattered are isotropic [81]. The scattering of neutron by a single nucleus is described in terms of a cross section σ measured in barns (1 barn=10^{-28} square meter). It is equivalent to the effective area presented by the nucleus to the passing neutron. The strength of the interaction is denoted by the scattering length: b and related to the cross section $\sigma = 4\pi b^2$ [85].

Atomic number Z	**Element**	**Scattering length (10^{-14} m)**		**Scattering length X-ray* (10^{-14} m)**
		Coherent Diffusion b_{coh}	**Incoherent Diffusion** b_{inc}	
1	H D	-0.374 0.667	2.530 0.404	0.282
6	C C^{13}	-0.665 0.619	0.005 -0.05	1.69
7	N	0.937	0.198	7.05
8	O	0.581	0.004	7.33

Table 3. Neutron scattering lengths for studying organic molecules [86]. b_{coh} is the coherent scattering length and b_{inc} is the incoherent scattering length.

This isotopic difference is of particular interest for instance for hydrogen and deuterium which both interact weakly with X-rays but have neutron scattering lengths, b_{inc} and b_{coh} that are relatively large and quite different. P. Schoenborn exquisitely highlighted the strength of the neutron techniques in biology, 'The strength of neutron scattering remains principally in the manipulation of scattering density through hydrogen and deuterium' [87]. The difference between the scattering lengths of hydrogen and deuterium is the basis of an isotopic-labelling technique, called contrast matching.

1.4.4. Scattering law

As displayed pictorially in the so-called scattering triangle (figure 13), a neutron scattering experiment consists to measure the exchange of momentum (Q) and energy ($\hbar\omega$) between the neutron and the sample. The scattering can be either elastic when there is no change in the energy ($\hbar\omega$=0) of the incident neutron or inelastic when the magnitude of the neutron wave vector $\mathbf{k}_1$ change ($\hbar\omega \neq 0$).

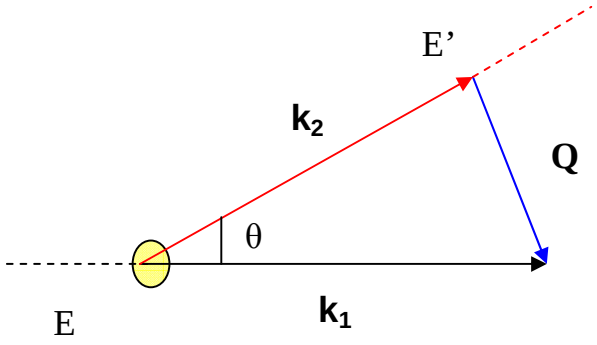


Figure 13. Schematic representation of a scattering event (scattering triangle).

The process can be describe using the following relationships

$$\mathbf{Q} = \mathbf{k_1} - \mathbf{k_2}$$ Momentum transfer

$$\hbar\omega = E - E'$$ Energy transfer

The incoming neutron wave is characterized by a wavevector k_1 and energy E. During the scattering event between, some neutrons are scattered in a certain solid angle $d\Omega$, in the direction of the wavevector k_2 with a final energy E' along the diffusion angle θ. This quantity is called the double differential cross section

$$\frac{d^2\sigma}{d\Omega d\omega} \qquad \textbf{Equation 5}$$

It can be shown that for a system of N identical atoms of equal scattering length b:

$$\frac{d^2\sigma}{d\Omega d\omega} = N\frac{k_2}{k_1}b^2 S(Q,\omega) \qquad \textbf{Equation 6}$$

Where S(Q,ω) is called the neutron scattering function.

$$S(Q,\omega) = \frac{1}{2\pi\hbar N}\sum_{i,j}\int_{-\infty}^{+\infty}\left\langle \exp[-iQ.r_i(0)].\exp[iQ.r_j(t)]\right\rangle e^{-i\omega t}dt$$

Equation 7

The Fourier transform of S(Q, ω) equation introduce the intermediate time dependent function S (Q, t).

$$S(Q,\omega) = \int_{-\infty}^{+\infty}\exp(i\omega t))S(Q,t)dt \qquad \textbf{Equation 8}$$

The intermediate scattering function S(Q, t) can be written in terms of time-dependent correlations between the position of pairs of atoms and the autocorrelation of the same atom in the sample [81].

$$S(Q,t) = \sum_{i,j}^{N,N} b_i b_j \left\langle \exp[iQ.r_i(0)].\exp[-iQ.r_j(t)]\right\rangle + \sum_{i}^{N}(b_i inc)^2\left\langle \exp[iQ.r_i(0)].\exp[-iQ.r_i(t)]\right\rangle$$

Equation 9

The intermediate scattering function denotes the position of the atom i at time t, $r_i(t)$, in a system of N atoms with i=1,….,N.

The brackets reflect the average of all enclosed states.

1.4.5. Instruments

As implied by figure 12, neutron spectroscopy used a number of instruments and strategies to explore matter.

Spin echo is primarily used to study slow relaxation phenomena in polymers, glasses or magnetic materials [81]. It exploits one neutron's intrinsic angular momentum or spin to access slow relaxation processes (~10^1ns). This kind of spectrometer allows studies of dynamic processes in macromolecular systems that are relevant to, polymer [88] and biological sciences [89]. Neutron diffractometers and small angle neutron scattering instruments essentially count neutrons as a function of the scattering angle, and display few particularities compared to X-ray diffraction [87]. D11 and D22 at the ILL, Grenoble, are neutron diffractometers where small angle measurements are performed. They are often used to resolve the structure of large scale molecules (10 to 1000 nm) such as macromolecular complexes or polymers [90].

Some instruments such as neutron spectrophotometers measure not only the scattering angle but also the neutron's energy. Neutron detectors are devices that can 'count' neutrons but not resolve their energies. The incident and scattered energies must be known, so the beam must first be 'prepared' for the measurement by selecting a particular wavelength or energy. After, the energy of the scattered neutron is measured. Different methods exist to determine neutron energies, the main being: the time of flight and the backscattering (diffraction off crystals).

The time of flight technique uses choppers (rotating disks with slits) to select neutrons depending on their velocities. A first chopper with a slit will sends a pulse of neutrons at different speeds. A second chopper turning at a certain speed will only let through neutrons of a certain wavelength. Knowing the time of flight of the neutrons and the distance between the sample and the detectors, it is possible to calculate the neutron outgoing energy. Examples of these types of spectrometers are IN5 and IN6 at the ILL, Grenoble. On the other hand, crystals can be used to select a particular wavelength and therefore neutron's energy. The so-called backscattering spectrometers use this method. Indeed, it is based on the Bragg's law, which is extremely wavelength selective when the scattering angle is equal to 180°. To select a

specific wavelength the monochromating crystal is simply heated. IN13, IN10 and IN16 at the ILL, Grenoble used this strategy. Since IN13 has been used in the present study, a detailed description will be given but more information can be easily obtained at (http://www.ill.fr/YellowBookCDrom).

1.4.6. Dynamic studies using neutron scattering

Equation 9 is fundamental in neutron spectroscopy. Neutrons are somehow 'schizophrenic' particles, behaving either as a coherent and incoherent wavelike function. The first term found in the intermediate scattering function S (Q, t) is the intermediate coherent scattering function, which describes the relative position of all the atoms present in the system. This type of scattering depends on the relative distances between the constituent atoms and thus gives information about the structure of the materials. In the next part of the equation, a second type of scattering: incoherent scattering is depicted. The neutron wave interacts independently with each nucleus in the sample so that the scattered waves from different nuclei do not interfere. This scattering gives information about the interaction of a neutron wave with the same atom but at different positions and different times, thus providing information about atom diffusive motions involved in dynamic studies [87].

Macromolecular dynamics are essentially followed by incoherent neutron scattering (INS). The scattering intensity can be described with the relation:

$$S_{inc}(Q,\omega) = \int_{-\infty}^{+\infty} \exp(-i\omega t) \sum_i (b_i inc)^2 \left\langle \exp[iQ.r_i(0)].\exp[-iQ.r_i(t)] \right\rangle dt$$

Equation 10

This relation denotes the motions of the same atom over time. Since, $S_{inc}(Q, \omega)$ is proportional to the square of b_{inc}, the incoherent signal is highly dominated by the scattering of hydrogen atoms reflect the dynamics. Protons are therefore an exquisite probe to investigate dynamic motions of the groups to which they are attached. This is of particular interest for biological materials such as protein which contains predominantly hydrogen atoms uniformly distributed in the structure. In the

frequency range examined, pico to nano-second, protein dynamics can be therefore studied with incoherent neutron scattering. A schematic picture of the theoretical spectrum obtained from INS is given in figure 14. It depicts the three main components: elastic, quasi-elastic and inelastic scattering. These correspond to the energy resolved when neutrons interact with atoms. Different cases are possible. A peak centred at zero energy transfer over all Q values signs an elastic scattering. The momentum may change but its energy is conserved ($\hbar\omega$=0). For an ideal spectrophotometer, the elastic peak is a Dirac function, however in practise, the form and the width of this peak represents the instrument resolution. At higher temperatures, atoms begin to move, the elastic peak begins to diminish in height as a function of the wavevector Q. Two other types of scattering are now observed when ($\hbar\omega \neq 0$): quasi-elastic and inelastic centered on $\hbar\omega$=0 and $\hbar\omega \neq 0$ respectively. Quasi-elastic broadening is rich in information. It reflects how diffusive processes such as translations or rotations occur in the sample. On the other hand, inelastic scattering arises from energy transitions from the high frequency vibration of individual atoms.

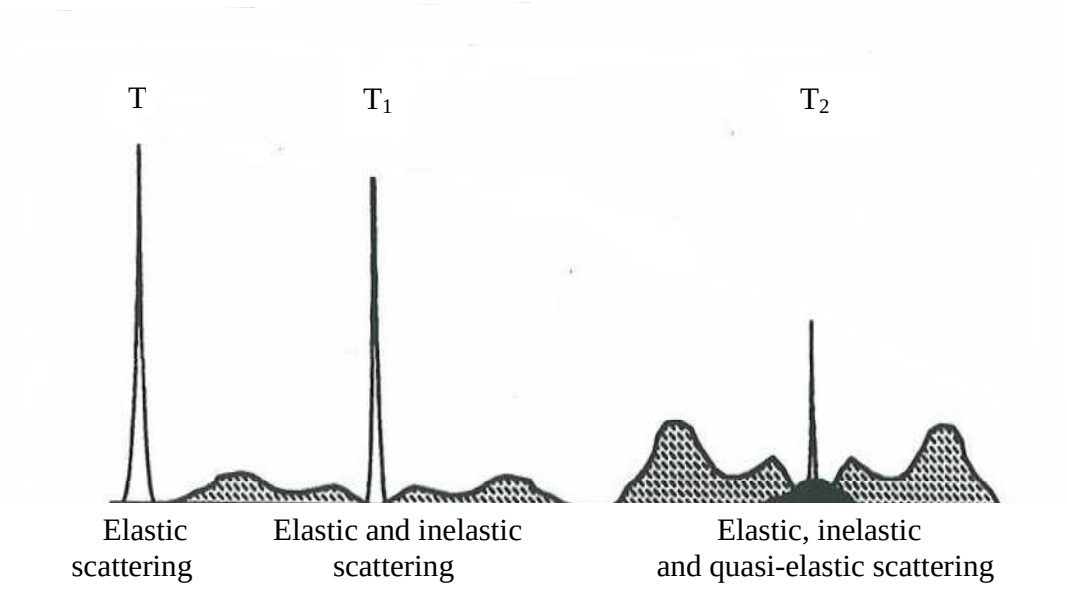


Figure 14. Scattering dependency over temperature with T>T1>T2. Adapted from [91].

Hence, the variation of the intensity of the elastic peak as a function of temperature reflects various motions modes within the sample. However, what can be gleaned from these three types of scattering?

Different models are used to interpret and rationalise the spectra. However, it is worth noting, that the weight given to these scattering modes depends primarily on the energy resolution of the instrument. The energy and Q-values accessible to a spectrophotometer correspond in a reciprocal fashion to a window in space and time. For instance, a backscattering and a time-of flight spectrophotometer have very different time and space dependencies. Typically, vibrational modes are analysed within an energy range from 20 to 125 meV whereas rotational and translational motions occurs at lower energy levels from 10^{-2} meV to 30 meV. The selection of the time and space window and therefore the instrument is crucial to follow in a satisfactory manner a specific type of motions. In addition, there are number of models to describe and interpret macromolecular dynamics [85, 92]. Since, it was not the aim of this work to focus on this issue, only one particular model is described in chapter two.

1.5. Aim and objectives

The aim of this research is to increase the understanding of the role of fast protein dynamics in enzyme activity and stability, and specifically to test the hypothesis that enzyme activity and stability are inversely related by their internal dynamics.

Activity ∝ Dynamics (flexibility) ∝ 1/Stability

The specific objectives if we redefine more narrowly the interrelationship between enzyme activity stability and dynamics are to improve our fundamental understanding of protein dynamics to protein function. The idea that enzymes might have evolved not just structurally but also dynamically to optimize their function in intriguing but not proven [93]. We seek an improvement in this fundamental issue.

To address this problem, we have embarked a multidisciplinary project using the technique of neutron scattering as the main probe to depict fast protein dynamics on the ps-ns time scales. Our strategy was to perturb simultaneously activity, stability and dynamics and to cross-correlate these findings. We have performed this work using a particular enzyme, the Dihydrofolate Reductase (DHFR) from a moderate thermophilic organism, *bacillus stearothermophilus* (*Bs* DHFR). We have also taken advantage of the pharmacological relevance of DHFR in cancer therapy for the extensive structural information and the number of specific inhibitors available. The anti-cancer drug methotrexate was of particular interest since it could impact not only on activity, but potentially also on stability, structure and protein dynamics. We have therefore used MTX as an informative probe in this network of pairs of properties and examined dynamics/stability/structure relationships. The thermodynamic analysis combines both, circular dichroism and isothermal calorimetry in order to assess the effect of MTX on DHFR thermal stability and its binding ‘signature’ respectively. Classical kinetic measurements were performed to examine the potency of the inhibitor on DHFR catalysis. The cross correlation of the various results is therefore expected to provide the framework for shedding the light on our current paradigm.

Fast proteins dynamics and its correlation to activity and stability

2. Effect of Methotrexate on DHFR fast protein dynamics: a neutron scattering study

2.1. Introduction

The dependence of the catalytic action of DHFR on dynamics has already been demonstrated using a combination of techniques such as NMR and steady state kinetics [37, 45]. Motions concomitant with the rate of catalysis have been often highlighted as the driving force, but little is known of the role of faster motions occurring on the ps-ns timescales. In this context, it is important to distinguish which motions are required for enzyme function and conceptually for protein function.

The aim of my work is to further the understanding of the role of fast protein dynamics in the activity and stability of this specific system. A key element in protein function is molecular flexibility without which enzymes for example cannot function or fold correctly [93]. A key point in enzymology is the binding of specific ligands which can selectively (or not) modify the dynamical behavior of such biocatalysts [5]. We have taken advantage of the tight binding inhibitor: methotrexate to examine the interrelationship between dynamics/activity and stability. In our dynamical map, MTX is an informative ligand in this network of properties since it can impact not only on activity and stability but potentially on all these parameters.

As presented in the introduction, neutron scattering is particularly well-suited to study protein dynamics on the pico and nano-seconds time-scale. We have therefore examined the effect of the anti-cancer drug MTX on the dynamical behavior of the enzyme DHFR and correlated these findings with activity and stability measurements under the same conditions.

2.2. Technique overview: Neutron scattering

As mentioned in the introduction, there are a few ways to characterize the dynamical properties of biological materials using the technique of neutron scattering. A relatively fast and simple way is to measure the elastic intensity as a function of temperature. During the experiment the energy of the scattered neutron is only

recorded at the elastic peak. If the atomic displacements within the sample are confined within the space-time window of the spectrophotometer, their displacements can be described by the Gaussian approximation. In this case the measured elastic intensity (at zero energy transfer) is related to the mean square displacement as shown in equation 10 (a full development is given in Gabel *et al* [94]).

.

A simple way to describe the incoherent scattering function S_{inc} is possible.

$$S_{inc}(Q, \omega \pm \Delta\omega) = A_0 \exp\left(\frac{-\langle u^2 \rangle Q^2}{6}\right) \qquad \textbf{Equation 11}$$

where $<u^2>$ is the mean square displacement

The expression is valid for small Q values, where $<u^2>Q^2$ is smaller than 2.

The mean square displacement is therefore calculated from the slope of the natural logarithm of the elastic intensity as a function of Q^2.

$$\langle u^2 \rangle = -6 \frac{d \ln(S_{inc}(Q, \omega = 0))}{dQ^2} \qquad \textbf{Equation 12}$$

It is to be noted that different authors use either a factor 1/6 or 1/3 in the previous expression. If a motion is a fluctuation around equilibrium position, the $<u^2>$ values is given by three times the slope. On the other hand, the factor 6 derives to an extent from molecular dynamics calculations which are though to better approximate this factor. The physical reason also derives from the fact that the displacement is not necessarily around an equilibrium position.

It is possible to extract, from the way the logarithm of the elastic intensity decreases as a function of Q^2, a root mean square displacement (RMSD). The RMSD corresponds to the mean fluctuation of an atom around its equilibrium position in the space and time window defined by the instrument resolution. The RMSD describes the overall flexibility of the system, it is often noted as $<u^2>$ and given in $Å^2$. An independent parameter, postulated by Zaccai *et al* [95], is also commonly used: the

mean resilience <k>. This force constant introduces in a more quantative way the 'stiffness' of biological molecules. This parameter is proportional to the inverse of the slope of the mean square displacement as a function of temperature. The mean resilience is given in $N.m^{-1}$. The higher the slope, the smaller the force constant and the softer the environment can be considered to be, and inversely. An interesting way to describe the flexibility, resilience and energy potential landscape was first proposed by Doster *et al* [92]. In this model, the RMSD can be assumed as the conformational space that can be explored by the atoms in single energy well and the resilience as the energy barrier between the various conformational substates. In the present work, the model of force constants was used to describe macromolecular dynamics.

2.3. The backscattering spectrometer: IN13

The backscattering spectrometer IN13 at the ILL was used during this work (figure 15). The principle of backscattering is an inverse spectroscopy where the energy of the scattered neutrons remains fixed and only the incident particle energy is varied. The energy of the incident neutrons is defined by the backscattering conditions by the Bragg relation:

$$\lambda = 2d\sin\theta \qquad \textbf{Equation 13}$$

The energy resolution, which is one of the fundamental instrument characteristics, is given by the differentiation of the Bragg equation:

$$\frac{\Delta E}{E} = \frac{2\Delta\lambda}{\lambda} = 2\cot\theta + \frac{2\Delta d}{d} \qquad \textbf{Equation 14}$$

The backscattering condition on IN13 permits a very good energy resolution [96] as can be seen by the latter equation and takes its value when $\theta = 0$. The energy resolution depends on the quality of the crystal and takes crystal imperfections and impurities into account ($\Delta d/d \sim 5*10^{-5}$). The factor θ is the angular divergence of the beam and depends on the wavelength and the monochromator crystal. The last factor, cot θ, shows the advantage of the backscattering instrument. The best energy

resolution is achieved when the neutron are backscattered from the monchromator when θ is close to π/2.

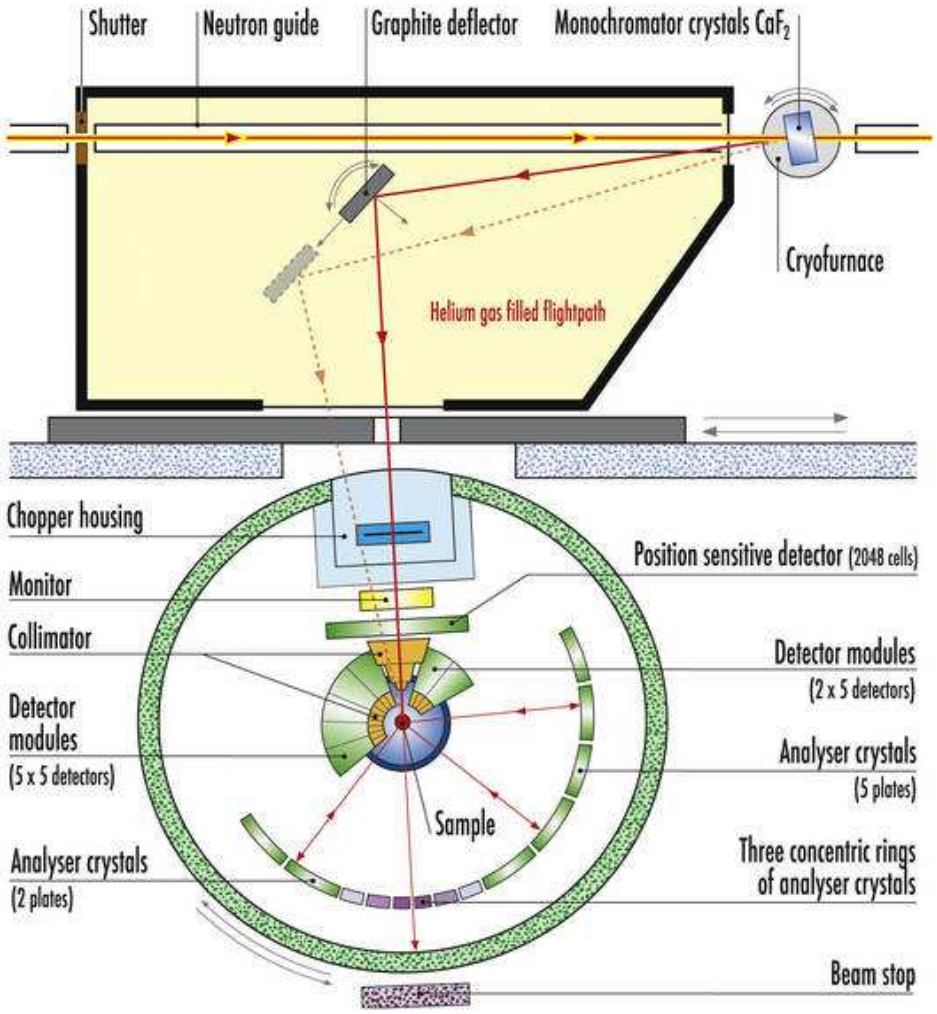


Figure 15. Thermal neutron backscattering spectrometer IN13 layout

When neutrons have left the monochromator they are reflected on a deflector. Then, they pass through a chopper and a monitor before reaching the sample. Scattered neutrons are first analyzed by another set of crystals, which reflect only neutrons fulfilling the Bragg condition into the ^{3}He detectors (situated on the opposite side of the sample). The basic principle of a neutron detector is to convert the neutral neutron

to a charged particle. There are a number of neutron detectors [97] but the most common gas is ^{3}He. The underlying equation is given below:

$$^3He+n \longrightarrow {}^3H+{}^1H+0.765\ MeV \qquad \textbf{Equation 15}$$

The energy released by the reaction between ^{3}He gas and the neutron (n) produces a current. This current, which is proportional to the number of neutron hitting the detector, gets record by the electronics. The incident wavelength is varied via heating and cooling of the monochromator. Indeed, the lattice of the monchromator changes upon heating and cooling, which results in neutron of different wavelength being reflected. The role of the choppers is to phase the electronic circuit of the ^{3}He detectors in a way that only backscattered neutrons are counted and neutrons directly scattered into the detector are ignored. The monitor counts a fraction of the incident beam and is used for normalisation of the different measurements.

2.4. Material and methods

2.4.1. Chemicals

Deuterium oxide (D_2O, 99.9 % and 98 %) was purchased from Minipul, Norell Inc. (Landisville, NJ., USA). Methotrexate (MTX) was purchased from Sigma-Aldrich Inc (St.Louis, MO., USA). Reagents and medium components for the purification and the analysis of *Bs* DHFR were purchased from Sigma-Aldrich Inc (St.Louis, MO., USA) and Merck KgaA (Darmstadt, Germany).

2.4.2. Over expression and purification of recombinant *Bacillus stearothermophilus* DHFR

Bacillus stearothermophilus DHFR clone was a kind gift of Judith Klinman, (University of California at Berkeley, US). The recombinant protein was purified from transformant *E. coli* cells (BL21/DE3) bearing the plasmid pET-21a-(+) encoding for the DHFR gene. The *E. coli* clone was grown at 37 °C in Luria broth media, containing 100 μg/ml ampicillin for the initial inoculum. The cells were grown in a bulk growth media pH 7 containing 100 μg/ml ampicillin as a batch-fed culture in a 8.5 litres fermenter. Expression of *Bs* DHFR was induced by addition of

IPTG at a final concentration of 1 mM at the end of the first logarithmic growth phase. The cells were then harvested generally after 11-15 hours of growth and concentrated by hollow fiber filtration. The cells were collected by centrifugation at 8500 rpm for 30 min, washed in 0.9 % NaCl, recentrifuged, and then resuspended in 40 mM Hepes pH 6.8 before being lysed by sonication (5 cycles of 5 minutes each). The cell debris were removed by centrifugation at 8500 rpm for 30 min and the resulting supernatant was heat treated at 55 °C for 20 min and centrifuged at 10000 rpm for 20 min. The protein solution was then loaded on a SP-Sepharose XK 50 cation exchange column (Amersham Biosciences) pre-equilibrated in 40mM Hepes at pH 6.8 and then eluted with a linear gradient of 0-500 mM NaCl. Fraction showing DHFR activity were pooled and then concentrated by ultrafiltration using an Amicon concentrator (YM-10 membrane). The concentrated fractions were then dialyzed extensively (molecular cut off: 6000-8000), freeze-dried and stored at 4 °C until use. The purity was assessed by SDS-PAGE (Sodium-Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) and size exclusion HPLC (High Performance Liquid Chromatography) analysis.

2.4.3. SDS-PAGE gel

SDS-PAGE gels were prepared according to Laemmli [98]. The gels were constituted of a 15 % polyacrylamide separating gel layered with a 5 % stacking gel. Gel preparation was performed with a concentrated 30 % (w/v) acrylamide solution containing 3 % methylene bis-acrylamide (ProtoGel™, National Diagnostics). All solutions are stored at 4 °C. The electrophoresis was performed in a Mini-Protean III dual slab cell (Bio-Rad Laboratories, Watford, UK). The separating gel was first gently poured in the casting module. Isopropanol was added on the top to avoid the formation of air bubbles. After complete polymerization (20 to 30 min), isopropanol was removed and the stacking gel was layered on the top of the separating gel. After polymerization of the stacking gel, the wells were washed with the cathodic buffer.

Samples were prepared by dissolving 10 µl of sample with 10 µl of double concentrated sample loading buffer (125 mM Tris-HCl, pH 6.8, 4 % SDS, 10 % β-mercaptoethanol, 0.4 % bromophenol blue and 20 % glycerol). The mixture was

heated for 2 min at 95 °C and centrifuge 1 min at 13200 rpm. The samples and molecular weight standard (Roche) are then applied to the gel (typically 10 µl per well) with a Hamilton micrometer syringe. Electrophoresis was performed at room temperature using a voltage stepped procedure: the voltage was kept constant at 80 volts until the samples have reached the separating gel. The voltage was then, raised to 120 volts for approximately 20 min until the tracking dye reached the bottom of the gel. Immediately after ending electrophoresis, the proteins embedded in the gel were stained by Coomassie brilliant blue R-250 (Roche) for 30 min at room temperature under gentle shaking. Destaining was applied for 10 min at room temperature and stopped with water when the bands were well defined. Finally, the gels were photographed and conserved if necessary in a solution of 10 % glycerol, 5 % acetic acid in 100 mM Tris-HCl pH 8.8 and dried under vacuum between sheets of transparent foil.

Separating gels (15 %) :

43% (v/v)	Protogel™
358 mM	Tris-HCL, pH 8.8
0.1 % (w/v)	SDS
0.1 % (w/v)	TEMED
0.1 % (w/v)	APS

Stacking gel (5 %):

18.2 % (v/v)	Protogel™
220 mM	Tris-HCL, pH 6.8
1.8 % (w/v)	SDS
0.1 % (w/v)	TEMED
0.1 % (w/v)	APS

Running buffer:

25 mM	Tris base
250 mM	Glycine

0.2 % (w/v) SDS
pH adjusted to 8.3

Staining solution (Coomassie):

0.1 % (w/v) Coomassie Brilliant Blue R-250
450 ml (v/v) Methanol
100 ml (v/v) Acetic acid
Adjusted to 1 l with distilled water

Destaining solution:

450 ml Methanol
150 ml Acetic acid
Adjusted to 1 l with distilled water

2.4.4. Neutron scattering sample preparation

To minimize the contribution of hydrogen scattering, the labile hydrogen atoms from the protein and/or the ligand have to be exchanged with deuterium oxide (D_2O). For the ligand MTX, a single deuterium exchange of 4 hours at 6-7 °C was carried out before freeze-drying. As MTX is highly insoluble in acidic condition, the 'pH' of the solution was adjusted to pD = 7.5, using the relationship $pD= pH_{meter}+0.4$ unit with ammonium bicarbonate 100 mM.

The purified protein was dissolved in D_2O (10 mg/ml), gently stirred 12 hours at 4 °C and then freeze-dried. The operation was repeated two more times with higher grade D_2O (purity 99.9 %), freeze-dried, and then stored at 4 °C until use. Activity was checked and less than 10 % of the original activity was lost during the process. The DHFR-MTX complex was prepared in a similar manner. Since MTX is a tight (K_d value is approximately 10^{-9} M) and specific inhibitor of DHFR with a binding stoichiometry of 1:1, we saturated the enzyme preparation with a 2/1 mole ratio (MTX to respect to DHFR) during the final deuterium exchange. After lyophilisation,

both enzyme preparations (100 mg) were dried over phosphorus pentoxide. To reach a higher hydration level, the proteins were equilibrated in a desiccator over a saturated solution of NaBr (which gives a relative humidity of 50 % at 20 °C). After exposure at this relative humidity (for two days), the D_2O content was shown to be 20 %, (i.e. 25 mg of D_2O per 100 mg of dry weight protein). At this hydration level the sample holder was quickly sealed and stored at 4 °C until use. The NaBr salt was rotary evaporated before use to near dryness in order to remove as much as possible any remaining water molecules and then dissolved in high grade D_2O (99.9 %).

2.4.5. Neutron scattering and Macromolecular Dynamics

Results described in this work are from experiments performed on the backscattering spectrometer IN13 at the Institute Laue Langevin Grenoble, France (information on the Institute and the instrument is available on the web at (www.ill.fr). The instrument allows the examination of atomic motions in the space and time window of about 1 Å in 0.1 ns. In this space-time window and according to a Gaussian approximation, the incoherent elastic scattered intensity can be analysed as [39]:

$$I(Q,0 \pm \Delta\omega) = A_0 \exp\left(\frac{-\langle u^2 \rangle Q^2}{6}\right) \qquad \textbf{Equation 16}$$

where Q is $4\pi\sin\theta/\lambda$, 2θ is the scattering angle and λ the incident neutron wavelength, $\langle u^2 \rangle$ values include all contributions to motions in the accessible space and time windows, from vibrational fluctuations (usually expressed as a Debye-Waller factor) as well as from diffusive motions. The validity of the Gaussian approximation for the mean square fluctuation $\langle u^2 \rangle$ and its analogy to the Guinier formalism for small angle scattering by particles in solution has been discussed by Réat *et al*. [99] and more recently by Gabel *et al* [100]. In the Guinier formalism a radius of gyration Rg^2 of a particles in solution is calculated [101]. The particle equivalent is the volume swept out by a single proton during the time scale of the experiment (~ 100 ps). The analogy holds if the motion is localized well within the space-time window defined

by the Q and energy transfer ranges, respectively. The Guinier approximation is valid if $\sqrt{R_g^2 * Q^2} \approx 1$. Following our definition of $\langle u^2 \rangle$, $Rg^2 = 1/2 * \langle u^2 \rangle$. As a consequence, the Gaussian approximation is valid in the domain where $\sqrt{< u^2 > * Q^2} \approx \sqrt{2}$. Data were normalized, at each Q-value, by the lowest temperature (20 K) data point (figure 16). The mean square fluctuations $\langle u^2 \rangle$ at a given temperature T were calculated according to the Gaussian approximation as:

$$\ln[I(Q,0 \pm \Delta\omega)] = K + A * Q^2 \qquad \textbf{Equation 17}$$

The mean square fluctuations were therefore calculated as:

$$\langle u^2 \rangle = -6A \qquad \textbf{Equation 18}$$

Elastic incoherent scattering data were collected in a scattering vector range of 0.52 $Å^{-1} \leq Q \leq 1.58$ $Å^{-1}$. The $\langle u^2 \rangle$ values were then plotted as a function of absolute temperature T (figure. 17). An effective mean force constant <k'>, defining mean resilience, can be calculated from the derivative of $\langle u^2 \rangle$ plotted versus temperature, T [95, 102] (figure. 17):

$$\langle k' \rangle = \frac{0.00276}{\frac{d\langle u^2 \rangle}{dt}} \qquad \textbf{Equation 19}$$

The numerical constants are chosen to express <k'> in Newtons/metre (N/m) when $\langle u^2 \rangle$ is in $Å^2$ and T in Kelvin.

2.5. Results

Our experiments were performed on the IN13 backscattering spectrometer which allows the study of protein backbone motions in a typical space-time window in the range of ~1 Å in 0.1 ns. The elastic incoherent scattering intensities obtained on IN13 for a range of temperature are shown in figure 16. The fitted straight lines were

used to plot $<u^2>$ versus temperature for the free and MTX complexed enzyme (figure 17). The mean square displacement $<u^2>$ attributed to both enzymes increases with temperature with a constant slope between 50 and 250 K with a sharp change above 220 K. This deviation has been previously identified as the dynamical transition. It is generally attributed to the onset of thermally excited inharmonic motions [103], but maybe due to the speeding of such motions until they are observable within the instrumental timescale, rather than their onset [96]. In the harmonic region T=50-230 K, the temperature dependence of $<u^2>$ measured from the uncomplexed state behaves differently in comparison to the complexed state. The presence of MTX seems to systemically lower the mean square displacement amplitude by 15 % with a maximum up to 50 % around the dynamical transition. It is clear, that MTX significantly alters the structural flexibility of the protein by confining the atoms in a narrow energy well. Additionally, the mean environmental force $<k'>$, calculated from the slope of the straight-line part of the scan indicates a two fold increase $<k'>$= 3.3 N/m and $<k'>$= 1.4 N/m in presence of MTX for the complexed and uncomplexed enzyme respectively (inset, figure 17). However, above the dynamical transition, in contrast with the flexibility, which holds a significant difference over the anharmonic regime i.e lower values in the presence of MTX, we did not observe any significant effect of MTX on the resilience of the system. It remains essentially unchanged from 250 K to 320 K for both samples.

This is at first glance surprising, however because neutron scattering experiment depends of the time and space resolution given by the instrument, other energy conformations of the ensemble might be hindered. Indeed, there might be additional contribution from exited states not accessible to the technique. In addition results must be interpreted with the greatest care because the conformational landscape of macromolecules is, to some extent temperature dependent. For example, ‘cryo structures’ (below the dynamical transition) may correspond predominantly to a structural state weakly populated at room temperature [104].

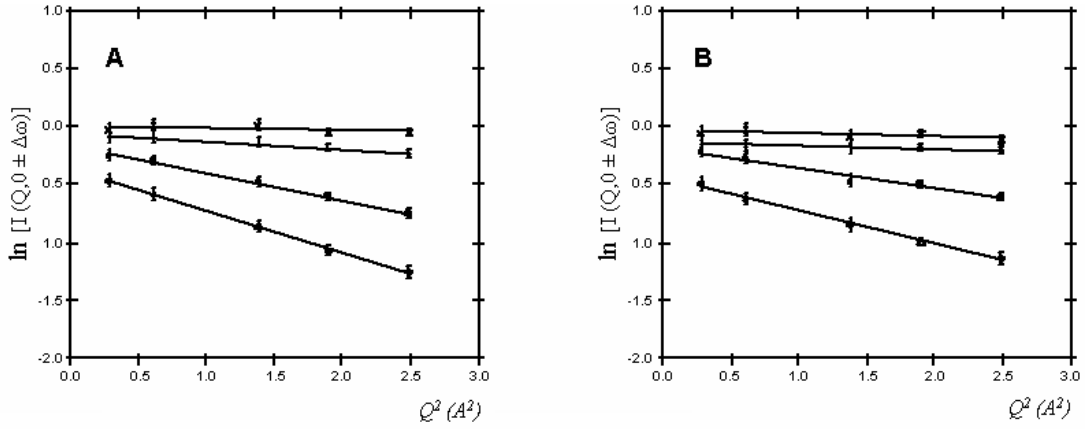


Figure 16. Variation of ln [I (*Q*, *0±Δω*)] as a function of Q^2 for the un-complexed DHFR (A) and complexed DHFR with methotrexate (B) at 100 K (×), 230 K (❍), 270 K (□) and 320 K (✧), from which mean mean square fluctuations $\langle u^2 \rangle$ were calculated.

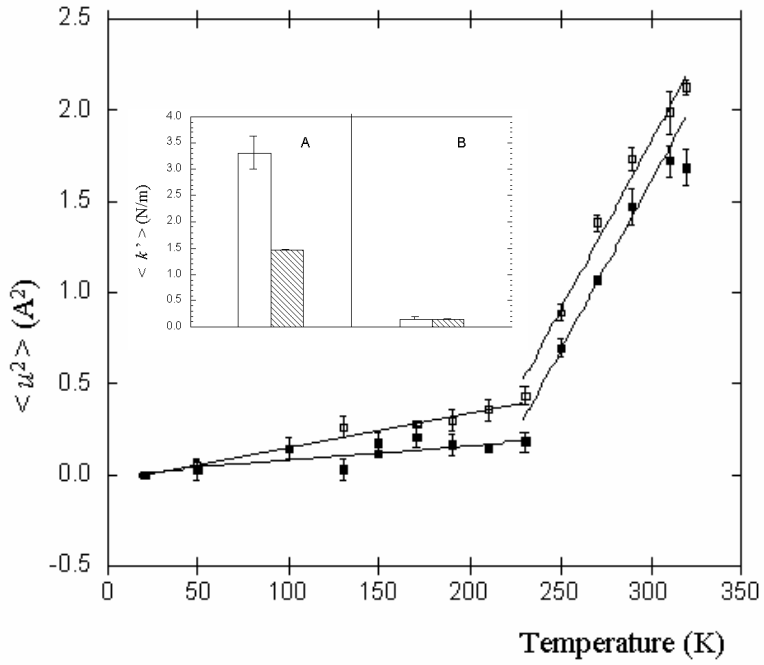


Figure 17. Mean square fluctuation $\langle u^2 \rangle$ were plotted against absolute temperature for the un-complexed DHFR (❒) and complexed DHFR with methotrexate (■). In the inset mean macromolecular force constant values < k' > measured for the MTX complexed DHFR (blank) and the un-complexed DHFR (hatched) are plotted versus below (A) and above (B) the dynamical temperature transition (230 K).

2.6.Discussion

In conclusion, the data obtained by elastic neutron scattering in the harmonic region shows that the free DHFR presents a larger $<u^2>$ and lower $<k>$ compared to the MTX-DHFR complex. Using the theory of energy landscape and assuming that the protein can execute harmonic motions along each local minimum from the energy landscape, the binding of MTX must be accompanied by the reduction of the conformational freedom. As elegantly suggested by Zaccai *et al* $<u^2>$ and $<k>$ can be easily represented with an energy well [105]. Following this graphical interpretation of macromolecular dynamics, the energy barrier between the various conformational substates must be increased and the relative flexibility which is reflected by the width of the energy well must be somehow decreased. It clearly shows that the relative plasticity of the energy landscape is modulated by the precise binding event of MTX. An alternative way to present these results is to suggest that the various conformational substates 'melt' to favor particular conformations (figure 18). The number of CS is reduced and the shape of the energy well is remodeled. Possibly, the well might harbor a funnel shape i.e. deeper with a restricted space in presence of MTX.

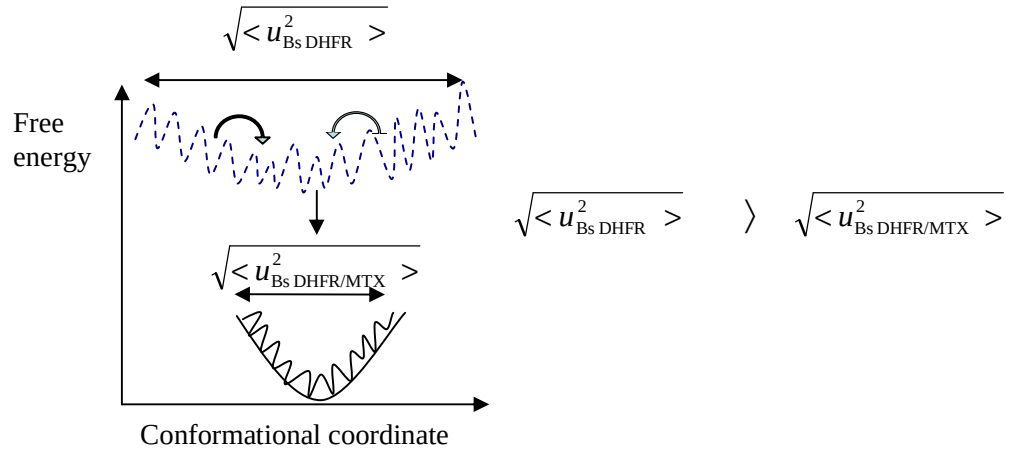


Figure 18. Schematic representation of the possible rearrangement of the free energy landscape of DHFR upon MTX binding.

Interestingly, it is therefore possible to observe such dynamical ‘groups’ and their properties with neutron scattering. For instance, the pseudo force <*k*>, i.e. protein resilience is often associated with structural rigidity and is one of the observed properties of the new conformational substates. It has been observed that under thermal denaturation MTX acts as a protective agent [106]. A recent study using atomic force microscopy has also correlated an increase in the mechanical stability of DHFR upon the binding of MTX [107]. It is therefore very tempting to suggest that MTX can modify the energy landscape of *Bs* DHFR by guiding the protein toward a particular folded state which is less susceptible to unfolding.

Relatively few studies have investigated the effect of ligand binding on internal protein dynamic using the technique of neutron scattering [102, 108]. Using this particular technique, most studies have focused on the effect of hydration and/or unspecific molecules such as sugars on internal proteins dynamics [17, 109]. The first lines of evidence investigating the effect of a small, specific non-covalently bound molecule were brought by Balog *et al* (2006) [110]. Performing an inelastic approach on the Time of Flight (TOF) spectrophotometer IN6, they have associated a softening effect, entropically driven upon the binding of MTX/NADPH on the *E. coli* DHFR internal dynamics. These findings are therefore in opposition with our results which tend to demonstrate a stiffening effect on both the flexibility and resilience. There are a few possible reasons for this discrepancy. The instrument resolution is of particular importance. We recall that TOF instrument such as the IN6 spectrophotometer detects motions over a wide energy shell (i.e. 10 ps-100 ps) but with a limited space window which is appropriated for pure vibrational analysis (i.e. backbone motions). On the other hand, backscattering spectrophotometers such as IN13 monitor motions on the pico-second time scale (100 ps) over a relatively wide space window which matches perfectly side chains motions [96]. From X-ray structures and particularly during the refinement process, it well known that amino acid side chains present a larger number of accessible conformations or rotamers. On the other hand, structural constraints do not allow this plasticity for backbone atoms. In other words, the potential changes observed in neutron spectroscopy might be more significant when observing side chains than backbone nuclei. It is therefore not surprising that IN13 and IN6 reveal

very different motional frequencies and modes but the question remains the same; which ones are the more relevant to unravel the interrelationship between protein dynamics and protein function?

A specific asset derived from the observation of side chains is highlighted by the technique of site directed mutagenesis. Beyond the simple deletion of key amino acids, the single replacement of a side chain can have a deleterious effect on protein function. Briefly, backbone nuclei can be depicted as the protein skeleton whereas side chains hold the chemical reactivity necessary for protein to function. However, even if the correlation between fast protein dynamics and function is not completely understood, the observation of side chain nuclei instead of backbone motions might in the first place be a better approximation to encompass such inter-correlation. In the future, it will be possible to depict the relevance of side chain motions and their contribution in internal dynamics with the increasing development of specific labeling (personal discussion with Dr. M. Haertlein, Deuteration laboratory, ILL, Grenoble).

3. Binding energetics and effect of Methotrexate on dihydrofolate reductase activity

3.1. Introduction

Proteins interact with a variety of ligand to perform a myriad of function in the cell. Through these interactions, they control a wide range of biological processes including protein synthesis, receptor signaling or signal transduction. A crucial part of this process is the modulation of protein dynamics upon ligand binding. Numerous studies have observed a structural functionality along preferred enzymatic pathway but the modulation of protein motions upon ligand binding remained elusive.

An understanding of how enzymes distinguish between closely related ligands on the molecular level might shed the light on this interrelationship. To gain insight into the factors that are most important for enzyme-ligand interactions and dynamics, the binding of closely related molecules having distinct effect on enzyme activity has been examined. This chapter will assess the inhibitory mode mediated by the substrate analog MTX and its binding energetic toward DHFR. We will rationalize these findings against the parameters obtained with the natural substrate H_2F.

Different techniques can be used to fulfill these goals. Here, Isothermal Titration Calorimetry (ITC) was used to compare the thermodynamic signature of these two structurally related molecules, and steady state kinetics were used to characterize the MTX inhibitory pattern.

3.2.Steady state kinetics

3.2.1. Technique overview

3.2.1.1.The Michaelis-Menten equation

Enzyme kinetics is a conceptual tool that allows us to interpret quantitatively and measure enzyme activity under variant parameters such as substrate concentration, temperature, pH, etc. It explains how enzymes work and behave in living organisms. In 1913 Leonor Michaelis and Maud Menten suggested that an enzyme catalysed reaction can be described by the following scheme

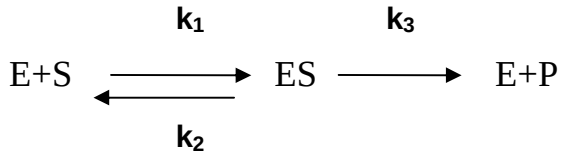


Diagram 1

where E is the enzyme, S the substrate, ES the complex enzyme substrate, P the product and $k_{1,2,3}$ the rate constants. The overall reaction rate can be described using the Michaelis-Menten's postulate. This postulate makes the assumption that enzymatic reaction can be divided into two discrete parts, the binding of the enzyme and the substrate (ES formation and breakdown) and the catalytic step concurring in product formation. The concept of steady state corresponds to a particular case where the concentration of the ES complex is constant. Mathematically, the reaction rate V can be measured with the universally known Michaelis-Menten equation:

$$V = V_{max} \times \frac{[S]}{[S] + Km} \qquad \textbf{Equation 20}$$

with $K_m = (k_2+k_3)/k_1$

Note that this equation introduces a number of parameters such as K_m and $V_{max.}$ These kinetic parameters are critical to attempt to rationalized enzyme action. The

new constant K_m has the form of a dissociation constant. Indeed, it is an apparent equilibrium constant which measures the affinity of the enzyme for the substrate.

The V_{max} is the maximum velocity when the enzyme is saturated by the substrate. The meaning of V_{max} is also apparent from inspection of the Michaelis-Menten equation. At very high [S], the quantity of:

$$\lim_{[S]\to\infty} \frac{[S]}{[S] + Km} \approx 1 \qquad \textbf{Equation 21}$$

And thus, $V=V_{max}$ so the reaction rate V is independent of the substrate concentration. For a more detailed description and complete derivation of the Michaelis-Menten equation see [111]. We should also see that a number of methods are available to determine these valuable parameters (direct linear, Lineweaver-Burk as seen in figure 19, computer least-square fit method, etc). They all transform the Michaelis-Menten equation and gives alternative statistical criteria to estimate the goodness value [111].

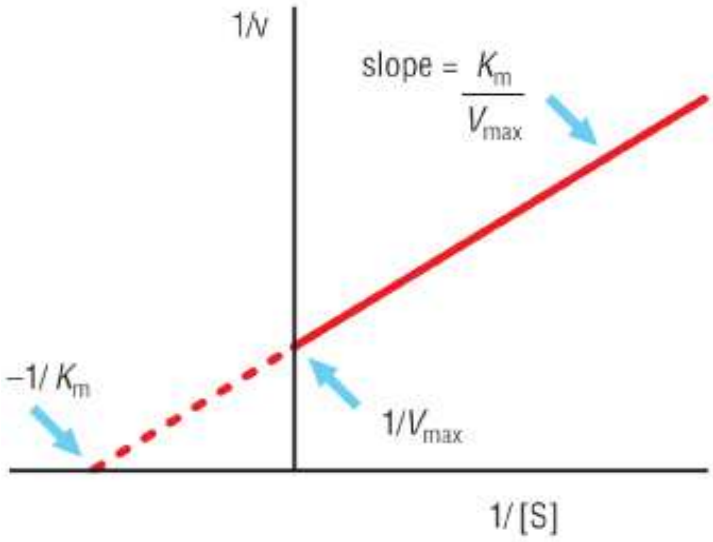


Figure 19. Lineweaver-Burk plot (or double reciprocal plot) of a typical enzyme catalysed reaction. The y-intercept of the graph is equivalent to the inverse of V_{max} and the x-intercept of the graph represents $-1/K_m$.

3.2.1.2.Competitive inhibition

Inhibitors can interact with an enzyme in different ways. In this respect a number of inhibition models have been described and extensively studied using enzyme kinetics. Accordingly, inhibitors are often divided into three classes, competitive, mixed (non-competitive), or uncompetitive [111]. If a reversible inhibitor can bind to the enzyme active site in place of the substrate, it is described as a "competitive inhibitor." In pure competitive inhibition, the inhibitor is assumed to bind to the free enzyme but not to the enzyme-substrate (ES) complex. The binding is described as shown below:

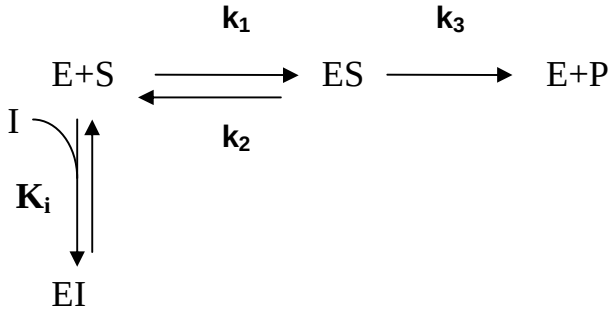


Diagram 2

here K_i is the dissociation constant for the enzyme inhibitor complex (EI). EI does not react to form E + P, and the enzyme is unable to bind both S and I at the same time. In this specific case, the mathematical treatment derived from the Michaelis-Menten equation is modified and the Michaelis-Menten equation for competitive inhibition is now:

$$V = V\max \times \frac{[S]}{[S] + Km\left(1 + \frac{[I]}{Ki}\right)}$$

Equation 22

$$\text{With } K_m^{app} = K_m \times \left(1 + \frac{[I]}{Ki}\right)$$

where [I] is the concentration of the inhibitor and K_i the inhibition dissociation constant. The bracket denotes the concentration of the enclosed molecule. K_m^{app} is the apparent K_m for the substrate when the inhibitor I is present. As inferred previously, there are several graphical methods for detecting and analyzing competitive inhibition. For instance the Lineweaver-Burk equation [112] for competitive inhibition is:

$$\frac{1}{V} = \frac{Km}{V_{max}} \times \frac{1}{[S]}\left(1 + \frac{[I]}{Ki}\right) + \frac{1}{V_{max}}$$ **Equation 23**

This unweighted fit double reciprocal plot is the least appropriate to use for determination of kinetic values but it is the most employed method in the peer reviewed literature since it is one of the better methods for display, and convenient to differentiate graphically the different inhibitory patterns. A secondary plot is therefore necessary to obtain K_i, K_m and V_{max} and their standard deviations (a more detailed description is given in the following section).

3.2.1.3.Slow binding inhibition

Slow-binding inhibitors are compounds that inhibit their target in a time-dependent manner. Slow binding inhibition is a phenomenon in which equilibrium between enzyme, inhibitor, and the enzyme-inhibitor complex is established slowly on a time scale of seconds to minutes. In the literature, three main mechanisms are proposed [113], the first one is very similar to the scheme presented in diagram 2. It assumes that the formation of the enzyme-inhibitor complex is a single slow step relative to the Michaelis complex formation. This mechanism will be referred as mechanism A. The second mechanism is more general and includes an extra step. It assumes the rapid formation of the EI complex but it suggests a slow and favorable isomerization of this EI complex to an EI* complex with an equilibrium constant K_{i*} (diagram 3). This mechanism will be referred as mechanism B.

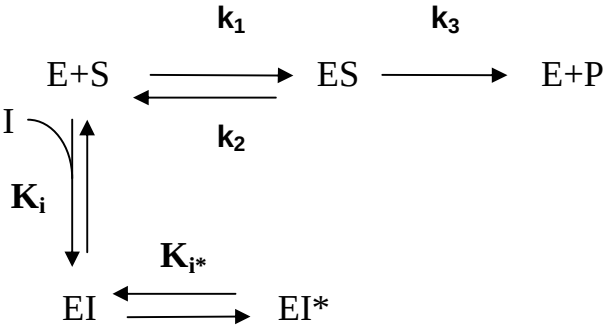


Diagram 3

In a third variant, the enzyme itself must slowly isomerize before it can bind to the inhibitor. These mechanisms can be distinguished by the careful analysis of the reaction progress curves. In the presence of a classical inhibitor, product formation remains constant over the reaction unless substrate depletion. With slow binding inhibitors, progress curves show clearly two phases, an initial burst of reaction which is relatively linear with an initial velocity V_0 and a second phase corresponding to a steady state with a velocity V_S (with often $V_0 \rangle V_S$) [114]. Classically, the initial rate (V_0) and steady state velocity (Vs) can be estimated graphically from the slope given by the tangent at time zero and at steady state. Such time-dependent inhibitions were observed between MTX and various sources of DHFR such as *E. coli* [65], *L. casei* [115], *and N. gonorrhoea* [48]. We have in this respect approach the operating mechanism and work out the parameters relative to the inhibition mediated by MTX.

3.2.2. Material and methods

Bacillus stearothermophilus dihydrofolate reductase activity was measured continuously by following the decrease in absorbance at 340 nm using a molar extinction coefficient of ε_{340}= 12300 M^{-1} L cm^{-1} [116]. The standard assay mixtures contains sodium phosphate buffer 100 mM at pH 8, at a fixed concentration of NADPH (100 µM) at variant concentration of H_2F (3 µM-200 µM). The final enzyme concentration was typically ~ 0.3 nM. The reaction was initiated by the addition of 1 µL of enzyme to the reaction mixture (469 µL), pre-incubated 2 min at a temperature of 30 °C. Assays were run for 60 seconds using a thermospectronic helios γ-spectrophotometer interfaced with the Vision™ 32 software (version 1.25, Unicam Ltd.). Reaction progress curves were recorded in triplicate and repeated if the data deviated by more than 10 %. The Michaelis-Menten parameters: Km, V_{max}, were determined by using the Enzpack software version 3.0 (Biosoft, Cambridge, U.K) with the appropriate kinetic model. This software uses a least-square fit regression analysis to fit the experimental data to the selected kinetic model.

The inhibition constant: K_i for MTX was determined using essentially the same protocol as described above with the specified variable H_2F concentrations at different fixed concentration of the inhibitor: 5, 20, 50 nM. The enzyme was pre-incubated 2 min at a temperature of 30 °C in the reaction mixture containing, MTX, NADPH. The reaction was initiated by the addition of H_2F. To avoid substrate depletion, we followed the linearity of the progress curves in the absence of the inhibitor. Data collection and further analysis were performed using the software Vision™ 32 (version 1.25, Unicam Ltd.). Reaction progress curves were recorded in triplicate and repeated if the data deviate by more than 10 %. The inhibition constant K_i, was determined graphically and computer assisted by using the Enzpack software version 3.0 (Biosoft, Cambridge, U.K) with a competitive model using a least-square fit regression analysis.

3.2.3. Results

Kinetic profile: *Bs* DHFR is a ternary complex mechanism enzyme. Previous kinetics studies have indicated an ordered kinetic mechanism with NADPH binding first to the enzyme [56]. The reaction progress curves (substrate breakdown Vs substrate concentration) in absence of the inhibitor were linear. From their slopes representing the initial rates V_0, the two kinetic parameters V_{max} and K_m were evaluated using one of the linearized form of the Michaelis-Menten equation (Lineweaver-Burk) [112]. From the double reciprocal plots, the V_{max} value was 0.178 ± 0.022 μmole/min. The Michaelis constant for the natural substrate H_2F was found to be 7.5 ± 1.4 μM at 30 °C. This value is in good agreement with Kim *et al* who have determined a K_m of the same order of magnitude. K_m=1.5 and 12.5 μM at 20 °C and 60 °C respectively [56].

Inhibitor binding profile: When DHFR activity was continuously assayed after the addition of the enzyme to assay mixture containing MTX, NADPH and H_2F, the resulting progress curves displayed a time dependent decrease in the reaction rate before substrate depletion. This suggests the slow establishment of an equilibrium between the enzyme, inhibitor and or enzyme-inhibitor complex. We also observed a

time dependent inhibition when the reaction was initiated by the addition of NADPH. These observations clearly outline the slow acting inhibition of MTX and consequently, MTX can be classified as a slow binder. The first step of our analysis was to find out the appropriate inhibition mechanism for data processing.

As inferred previously, the possible mechanisms can be discriminated rapidly from the inspection of the initial rate and steady state rate. Under our experimental conditions, the initial velocities V_0 were dependent on MTX concentration; their rates were decreasing with increasing concentration of MTX. Moreover, a similar trend was observed with the steady state velocity. In consequence, the diagnostic plots were in favor of mechanism B. The determination of the equilibrium constant K_i was possible using the initial rate of the reaction as given by the relationship:

$$\frac{1}{V_0} = \frac{K_m}{V_{max}} \times \frac{1}{[S]}\left(1 + \frac{[I]}{K_i}\right) + \frac{1}{V_{max}} \qquad \textbf{Equation 24}$$

Graphically, the primary plot 1/Velocity versus $1/[H_2F]$ (figure 20) shows the common characteristic of a classical competitive inhibition. The V_{max} values are almost equivalent, 0.166 μM min^{-1} and 0.178 μM min^{-1} in the presence and absence of MTX respectively whereas the apparent K_m increases with inhibitor concentration. An apparent K_m of 14.3 μM was estimated for [MTX] =10 nM. The secondary plot or the so-called Dixon plot [117], as shown in figure 21, gives an estimated K_i value of approximately 11 nM which in good agreement with the computer least square analysis which yielded K_i of 10.9 ± 1.6 nM. The MTX concentration is therefore very close to that of enzyme (0.3 nM) therefore MTX behaves as a tight inhibitor. It also interesting to note that the inhibitor concentration is not reduced upon enzyme binding ([I]>>[E]).

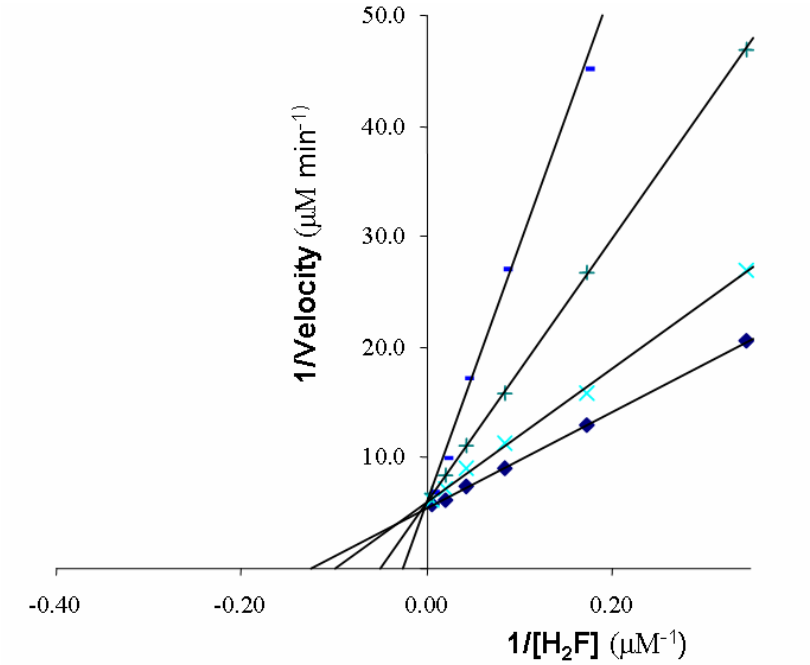


Figure 20. Double reciprocal plot of the competitive inhibition of the *Bs* DHFR by methotrexate. H_2F concentration was varied between 3-200 μM while MTX was at different fixed concentrations: (◆) 0 nM; (X) 5 nM; (+) 20 nM; (-) 50 nM. The straight lines are the best fit value.

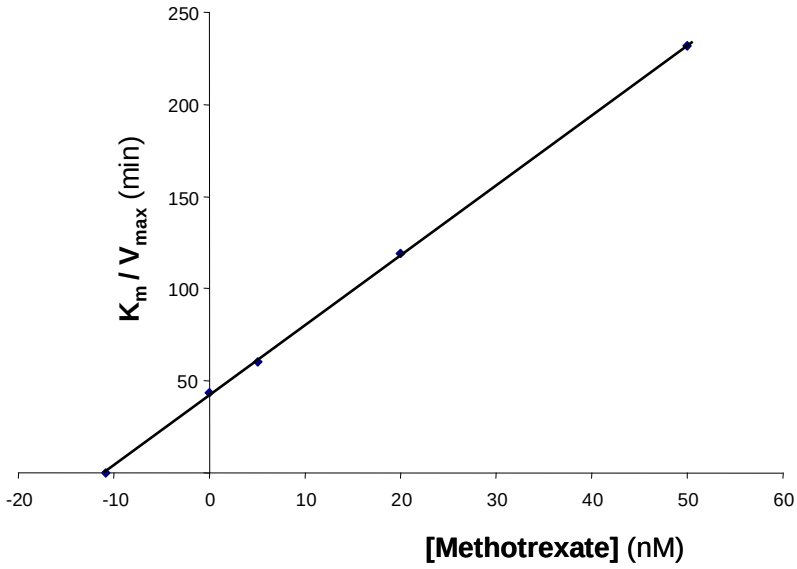


Figure 21. Replot of the slope from the double reciprocal plots from figure 20 versus methotrexate concentration**.**

3.2.4. Discussion

As often described in the literature from many DHFRs sources, MTX is a slow, tight and competitive inhibitor [118]. Our experimental data are consistent with this inhibition pattern with respect to H_2F. The inhibition constant which is an apparent dissociation constant is three orders of magnitude lower than the natural substrate H_2F. However MTX is not an exceptionally tight binding inhibitor of *Bs* DHFR. The K_i value is substantially greater than various bacterial and mammalian enzymes: *Homo sapiens* K_i = 6.1 pM [119], *Lactobacillus leichmannii* K_i = 7 pM [120]. However, as inferred by Baccanari *et al* a number of studies have over or under-estimated numerous K_i values [118]. They noticeably neglect the enzyme ligand equilibrium time dependency or the partial depletion of the free inhibitor when [I]~[E] and [E]>K_i [121]. An illustration of such discrepancy can be found in enzyme database Brenda (http://www.brenda-enzymes.info/): The DHFRs K_i values yielded from the same organism: *Pneumocystis carinii* are over three orders of magnitude, 16 pM and 20 fM. Consequently, it is not possible to compare in the strictest sense the overall MTX potency with previous studies if the assay conditions and data treatment are not similar [122].

Another point of interest came from the inhibitory pattern. The kinetic data presented here clearly show that MTX is an effective inhibitor of *Bs* DHFR but not the origin of its power and mechanism. A complete kinetic analysis for the calculation of the kinetic constants would certainly help to draw the complete mechanism of the slow-binding inhibition of *Bs* DHFR by MTX. For instance, the second equilibrium constant K_i* relative to the isomerizaton of the enzyme may have been evaluated. However the experimental acquisition of V_0 is not straightforward since it is difficult to measure the initial slope of the progress. Indeed to follow such events, a rapid experimental technique is sometimes required with dead times in the milli-second time range. Moreover, the analysis must exploit the entire progress curves as suggested in [122] with multiple non-linear regressions [123, 124]. The mathematical treatment and the powerful computer packages they involved (such as DynaFit: www.biokin.com) [125] were beyond the scope of the work proposed in this chapter. It is maybe frustrating for a kineticist who knows what is hidden in the

‘black box’ but to rationalize the enormous power of MTX such analysis is needed. This issue is particularly relevant in drug design. The differences in binding affinities are currently used to target preferentially either mammalian, prokaryotic or parasite cells [126]. The design of new anti-folate drugs such as MTX in cancer chemotherapy, bacterial or parasitic infections still requires such efforts.

3.3.Isothermal titration calorimetry

3.3.1. Technique overview

The aim of this chapter is to determine the binding affinity and the thermodynamics parameters associated with the interaction between the enzyme *Bs* DHFR and the anti-cancer drug MTX.

The recognition of a specific ligand by a protein comes from a delicate balance between various factors such as the entropy and enthalpy. The overall change in enthalpy and/or entropy is a sum of several contributions that can be either favorable or unfavorable. For instance, high specificity can be achieved by the formation of optimal contacts such as hydrogen bonds, Van der Waals contacts, electrostatic interactions, pKa changes [127] or even conformational changes [128]. Due to the many possible contributions to the observed changes in the thermodynamic parameters upon binding, assigning the roles of each can be difficult. Nevertheless, this knowledge is crucial if we want to understand how specific is a protein toward its partner.

Binding affinities are usually determined by using various techniques: UV-visible spectroscopy, surface plasmon resonance, ITC, etc. Unlike other techniques, ITC allows the complete determination of the thermodynamics signature of the interaction between two molecules [129]. The binding association constant K_a, the stoichiometry N and the binding enthalpy ΔH_b are available in a single experiment [130]. The free energy ΔG_b and entropy ΔS_b of binding are calculated from the association binding constant K_a using the equations [131].

$$\Delta G_b = -RT\ln(K_a)$$
$$\Delta G_b = \Delta H_b - T\Delta S_b.$$

Equations 25

ITC is a differential technique that measures directly the binding equilibrium by determining the heat evolved on association of a ligand with its binding partner [129]. For a typical titration, the macromolecule is placed into the sample cell and

small injections of the ligand are made until all of the binding sites on the macromolecules are saturated. The heat released from each injection is recorded and integrated yielding a typical sigmoidal curve that can be fitted to a binding model [132]. The amount of heat evolved (Q) on addition of ligand can be represented by the equation:

$$Q = V_0 \Delta H_b [M]_t K_a [L] \times \frac{1}{1 + K_a [L]} \qquad \textbf{Equation 26}$$

where V_0 is the volume of the cell, ΔH_b is the enthalpy of binding per mole of ligand, $[M]_t$ is the total macromolecule concentration including bound and free fraction, K_a is the association binding constant and [L] is the free ligand concentration.

In this respect, this technique has been successfully used in drug design [133]. Indeed, ITC is particularly useful in the optimization of drugs binding affinities since the various forces involved in the equilibrium can be dissected [134]. We have therefore used this calorimetry technique for its unique features to characterize our enzyme-MTX complex and depict the most important driving forces embedded in the equilibrium. ITC can be used to obtain accurate measurements of the binding affinities for protein ligand interactions with binding constants between 10^3 and 10^8 M^{-1} [131]. A common problem with molecules that have a very high affinity (that is, nanomolar range or lower) is the difficulty of determining accurately the binding affinity. ITC has an upper limit set with a maximum of 10^8-10^9 M^{-1} for the binding constant. This restriction is a serious drawback since the reported MTX binding affinities (K_a) toward DHFR approach the nanomolar or sub-namolar range [119, 135, 136]. To circumvent this problem, we have performed a competitive titration as described by Sigurskjold [137] and implemented by Velazquez-Campoy *et al* [131]. In this experimental approach, the goal is to lower the affinity of the high affinity ligand to a level that can be measured. Three titrations are needed. First a titration with the weak ligand (H_2F) in order to characterize its binding thermodynamics, second, a titration with the high affinity inhibitor (MTX) to measure its binding enthalpy and third the displacement titration involving a mixture of both ligands. This method has been validated and used successfully to determine the association

constant between inhibitors, peptides, carbohydrate and nucleotides with proteins [131, 133, 137].

3.3.2. Material and methods

3.3.2.1.Sample preparation

Proper sample preparation is essential for successful ITC testing. In particular a minimal number of requirements need to be fulfilled to insure an accurate estimate of stoichiometry (N), enthalpy of binding ($\Delta H_{b)}$, and binding association (K_a). Beyond the technical requirements implied by the machine itself, temperature, stirring speed, reference power, particular care has to be given to the macromolecule solutions. For all experiments, the protein was dialyzed overnight with the same buffer used for the ligand solution to avoid any mismatch between solvent compositions. We have used a 100 mM sodium phosphate buffer pH 8, 200 mM NaCl, 0.2 mM Tris [2-carboxyethyl] phosphine (TCEP) for the ligands and the protein. This buffer was selected for the optimal stability and solubility of the reactants. The samples were filtered and degased prior use.

The protein, H_2F and MTX concentrations were assessed by UV spectroscopy. Molar absorption coefficients of 25440 L M^{-1} cm^{-1} at 280 nm, 28000 L M^{-1} cm^{-1} at 282 nm and 22100 L M^{-1} cm^{-1} at 302 nm were used respectively.

The thermodynamics binding parameters of MTX to DHFR were measured using a Microcal VP-ITC instrument (Microcal. Inc). The setting up of the machine was identical for all three experiments. The measurements were done at 25 °C with a reference power of 10 μCal/sec with high feedback gain.

3.3.2.2.Direct titration: MTX and H_2F binding.

The titration of the high inhibitor affinity MTX consisted of 28 injections of 10 μl each of a 200 μM stock solution of MTX. MTX was injected into the sample cell (V= 1.42 ml) containing 33 μM of *Bs* DHFR (the protein:ligand ratio used was 1:6). The injections were made over a period of 10 seconds with a 5-min interval between subsequent injections. The sample cell was stirred at 286 rpm.

The calorimetric titration of the natural substrate: H_2F was determined with 28 injections of 10 µl each of a 500 µM stock solution of H_2F. The sample cell (V= 1.42 ml) contained 76 µM of *Bs* DHFR (the protein:ligand ratio used was 1:7). The injections were made over a period of 10 seconds with a 4 min interval. The sample cell was stirred at 307 rpm.

3.3.2.3.Displacement titration experiment

The displacement titration was set as previously described. The pre-bound complex contains 48 µM *Bs* DHFR with 500 µM H_2F (weak ligand: the protein:ligand ratio used was 1:10). The solution in the syringe contains 800 µM MTX (high affinity ligand: the protein: ligand ratio used was 1:15).A series of 20 injections over a period 10 s with 4 min interval were made to assess the binding thermodynamics.

3.3.2.4.Data analysis

The binding thermodynamics parameters: Stoichiometry (N), association constant (Ka) and the binding enthalpy (ΔHb) were determined using the Origin version 5.0 software from MicroCal interfaced to a Gateway PC for data acquisition and analysis. Direct titration of H_2F and MTX were performed using a single site model. Heat dilutions and baseline correction were carefully subtracted from the raw data before analysis. The tail of the titration often includes heats from ligand dilution and buffer-buffer interaction, I have performed a titration of the ligand against the buffer to determine this heat of dilution, then get the average and subtract from all data points.

For the competitive binding, the mathematics of this method have been presented by Wang and Sigurskjold and implemented by a number of groups [137, 138]. In these experiments, the observed affinity and enthalpy are made up of two contributions, the energetics for the displacement of the primary ligand H_2F and the binding energetics for the secondary ligand MTX. The observed binding affinity (K_a^{obs}) is given as:

$$K_a^{obs} = K_x \frac{1}{1 + K_s[S]}$$ **Equation 26**

where K_x is the binding constant for the strong ligand, K_s is the binding constant for the weak ligand and [S] the concentration of the free weak ligand in the calorimeter cell. The observed binding enthalpy also contains the contribution from both the primary and secondary ligand

$$\Delta H_{obs} = \Delta H_x - \Delta H_s \frac{K_s[S]}{1 + K_s[S]}$$ **Equation 27**

Using the software package from Microcal, the analysis of the displacement titration was performed using a competitive binding model. The binding energetics of H_2F with equation 27 were used to determine the binding energetics of MTX. The data correction was performed differently as mentioned previously. The endothermic peaks at the end of the titration were average from the stable saturated baseline and substracted from all the data points.

3.3.3. Results

As previously stated, due to the extremely high affinity of MTX, the binding thermodynamic parameters were measured in a competition experiment using H_2F as a weak ligand. Figure 22 illustrates the three titrations necessary for the complete description of the binding signature.

Panel (a) from figure 22 shows the direct titration of MTX against the enzyme. As inferred by the graph, the characteristic sigmoid shape obtained generally from the binding isotherm is not observed [131]. The transition between the non-saturated and saturated states is very sharp and approaches nearly linearity. Indeed, if the region of the binding isotherm in which the transition takes place contains less than two points, the slope of the inflection (which determines the binding constant) can not be reasonably determined. Consequently, the equivalence point can not be identified accurately which necessarily impairs the accurate determination of the binding

constant. Nevertheless, using a single site model, the dissociation constant (K_d) which is the inverse of K_a value was found to be 1.3 nM with a 1/1 stoichiometry (N= 1.1). It clearly demonstrates that the free enzyme binds tightly and selectively MTX at a single site. However, this value reflects only a lower limit for the binding affinity since we are beyond the limit of the ITC resolution. On the other hand as described by Sigurskjold *et al* [137], this restriction did not impact a precise measurement of the binding enthalpy: ΔH_b= -16.8 kcal mol^{-1}.

The two other experiments consist of the titration of the weak ligand (H_2F) alone and secondly the displacement reaction (DHFR- H_2F against MTX). The titration of the natural substrate H_2F with DHFR shows the common binding isotherm. The transition and therefore the equivalent point are easily fitted with a single site model. The binding stoichiometry gives a 1/1 ratio (N= 0.9) with a binding constant of 1.8 μM. This value is very close to the Michaelis constant determined by steady state kinetics K_m= 7.5 μM. The binding enthalpy was found to be favorable ΔH_b= -12.1 kcal mol^{-1} and characterized by unfavorable positive entropy change ΔS_b= -14.3 cal $K^{-1}mol^{-1}$. The most notable feature of this experiment is that the heat effect associated with binding is exothermic. The calculated binding Gibbs energy ΔG_b= -7.8 kcal mol^{-1} is therefore enthalpy driven.

On the other hand the displacement titration can be followed with the heat binding isotherm shows in panel (c). The plot has sufficient curvature to allow the determination of the apparent binding affinity. Under the conditions of this experiment and using a competitive model, the binding constant was K_d= 2.9 nM which is surprising since the value was expected to be lower than 1.3 nM as inferred previously. However, the error is quite large ± 5 nM so it can explain such discrepancy. It also yielded a negative enthalpy which seems to be underestimated ΔH_b= -11.9 kcal mol^{-1} with an almost favorable entropy term ΔS_b= -0.7 cal $K^{-1}mol^{-1}$. A second mechanism must contribute to this excessively low value of enthalpy and change in entropy. Structural studies have depicted conformational changes in the DFHR structure upon ligand binding [51]. However, the competitive binding model does not take into account conformational reorganization. This is a serious drawback since as inferred by Horn *et al*, conformational heterogeneity can drastically

modulates the thermodynamic signature and particularly the enthalpy and entropy values [128]. The dynamic nature of proteins and the change in the position of atoms could therefore contribute to a significant level in the binding thermodynamics. This particular thermodynamic signature could reflect a less synergistic mechanism when H_2F is bound prior to DHFR. The absence of linear portion in the first injections could also derive from this contribution. Unfortunately, it is not possible to determine the various factors involved in this specific behavior. Another contribution which is scarcely approached in the literature is the statistical analysis. Data fitting using a specific model is relatively easy and straightforward with the Microcal package. For instance, the stoichiometry values are given with a relative error of 1 % which may be far too good for an ITC experiment. On the other end the K_d values are less accurate, the relative errors are over 150 % in all tested cases. Therefore, it might be harzadous to give a precise statistical meaning of ΔS_b since it is calculated from the enthalpy and binding affinities using equation 25. It would in principle include the statistical error from both, K_d and ΔH_b. In fact, the statistical deviations from table 4 only reflects how good the experimental values agree with the analysis model. What would be probably more useful and relevant will imply the repetition of each experiments, at least twice to cross correlate the results. Nevertheless, we shall see that an ITC experiment is time demanding and only two or three titrations can be made during a day work (if everything goes well). Thus, it impairs such analysis.

Ligand	**K_d (μM)**	**ΔH_b (kcal mol^{-1})**	**ΔS_b (cal K^{-1} mol^{-1})**	**N**
MTX	$1.3\ 10^{-3} \pm 2.1\ 10^{-3}$	-16.8± 0.4	-15.6	1.14±0.01
H_2F	1.8± 3.3	-12.1± 0.1	-14.3	0.90±0.01
MTX into *Bs* DHFR-H_2F	$2.9\ 10^{-3} \pm 5\ 10^{-3}$	-11.9± 0.4	-0.7	0.96±0.01

Table 4. Thermodynamics of binding of MTX, H_2F to *Bs* DHFR**.**

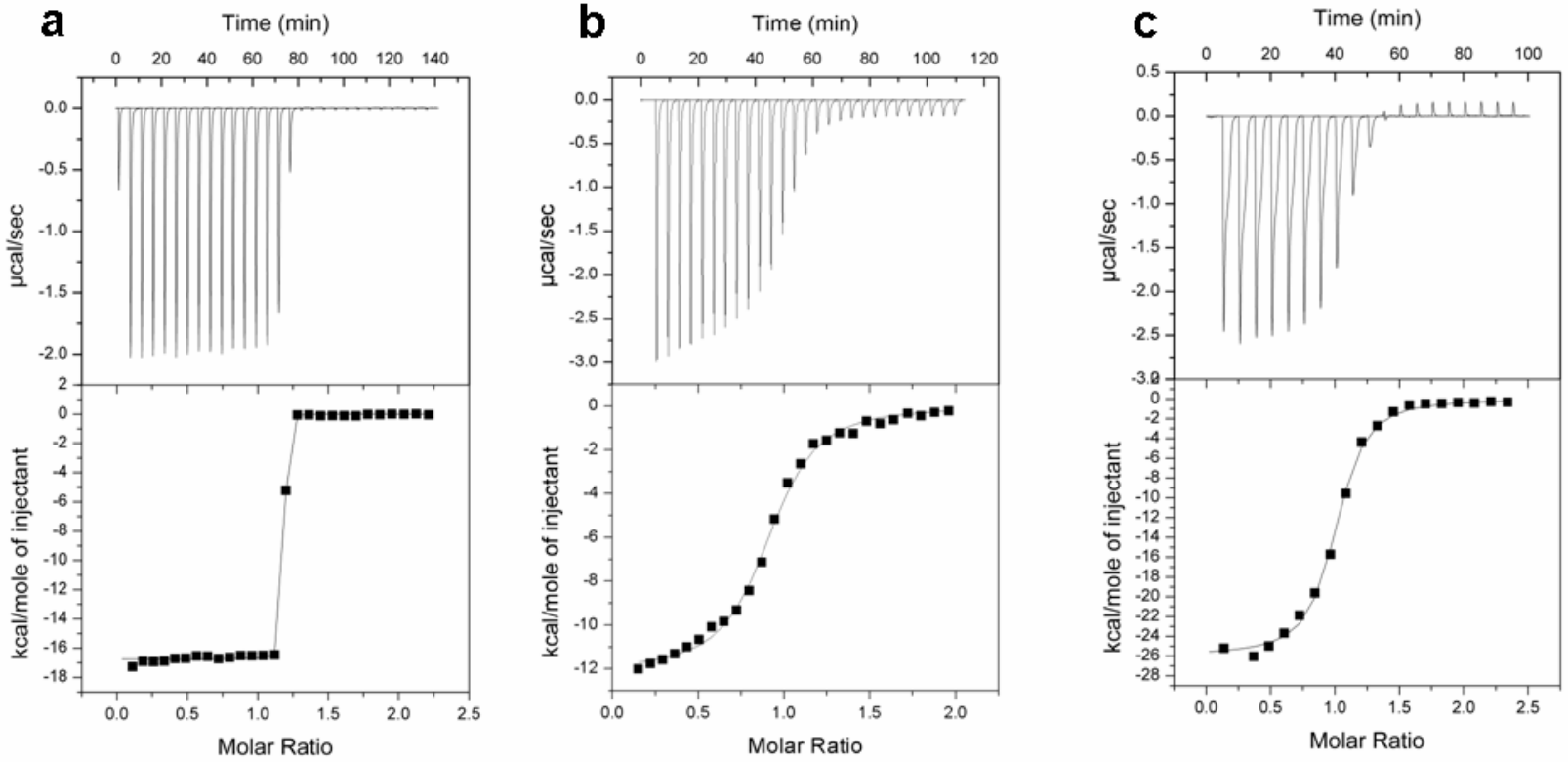


Figure 22. ITC determination of the binding thermodynamics of MTX to the *Bs* DHFR. Calorimetric titrations of *Bs* DHFR with the inhibitor MTX (a), the natural substrate H_2F (b) and with MTX in presence of H_2F (displacement titration) (c).

3.3.4. Conclusions and discussion

The most notable feature of this experiment is that, as expected, MTX binds to the enzyme with a higher affinity than H_2F (by three orders of magnitude). In addition, the thermodynamic profile for both ligands is very similar qualitatively. The binding is enthalpically favorable ($\Delta H<0$) and entropically unfavorable ($\Delta S<0$). The reduced affinity for H_2F is due to a significantly reduced favorable binding enthalpy from ΔH_b= -16.8 kcal mol^{-1} for MTX to ΔH_b= -12.1 kcal mol^{-1} for H_2F ($\Delta\Delta H_b$= -4.7 kcal/mol). There is some enthalpy-entropy compensation in that the entropic contribution to the binding of H_2F is less unfavorable than MTX. However the difference is not large enough to counter the magnitude of the favorable enthalpy loss. At the thermodynamic level, the binding enthalpy, primarily reflects the strength of the interaction of the ligand with the targeted protein (e.g Van der Waals, hydrogen bonds, ionic bridge, etc.) relative to those existing with the solvent [139]. A potential source of this more favorable accommodation in the case of MTX could reflect the better ability of MTX to create additional bonds strengthening its interaction with the enzyme. Using the R67 dihydrofolate reductase *Jackson et al* [140] have pointed out that the enthalpic signal arising from folate and antifolate drugs primarily originates from the *p*-aminobenzoylglutamate fragments and not the dihydrobiopterin group. In addition, the reduced and keto form (N3 protonated) of folate were found to be the preferred binding states. It further supports a strong role for O4 and/ or N3 atoms in the interaction. These finding agree well with nearly all available biochemical and structural data from *E. coli* and *L. casei* DHFR. From these DHFR sources, a particular ionic interaction between the Oδ2 of the highly conserved aspartate 27 and the N1 atom of MTX has been identified as the main driving force [14, 63, 66, 67, 73, 141, 142]. It clearly indicates that the replacement of the 4-oxo group of H_2F by an amino group has a dramatic effect on the binding energetics. It induces a unique binding geometry in the enzyme active site. This finding involves the positioning of pairwise interactions or counter charges such as that an increase in ligand order or rigidity occurs. However, these results are in disagreement with NMR [143] and X-ray [144] studies which suggest that the MTX high binding energy is due to an increase in entropy of the system. These studies noticeably suggest that the binding

energetics associated with folate and its analogs comes from a disordered segment of the molecule. The relative disorder presents in the molecule would help the positioning of the negatively charged tail (*p*-aminobenzoylglutamate). It should be emphasized that the entropic term depends primarily on two main forces. An unfavorable term, which originates from the loss of conformational degrees of freedom from the ligand and from some residues in the protein. Secondly, a favorable term, which comes from the release of water molecules upon binding or by a less efficient burial of hydrophobic groups from the solvent (solvation effect) [145]. Unfortunately, water molecule maps of the free and the various complexed state of the enzyme are not available. A recent crystallographic neutron structure has initiated such research but more it needed for a complete description of the system [73]. On the other hand a number of studies have investigated the conformational dependency of the DHFR-MTX complex [14, 67, 146]. Recent H/D exchange studies coupled with Mass Spectroscopy (MS) experiments have reported a decrease in structural fluctuation upon ligand binding [147]. This effect might shed the light on the unfavorable character of the entropy term. The conformational space explore by the protein atoms is reduced. However our data set can not explain why the entropy term differs between H_2F and MTX. The $\Delta\Delta S_b$= 1.3 cal/K/mol is very low and can not be used to draw a direct conclusion.

Consequently, the main difference between MTX and H_2F is its higher affinity toward the enzyme and this affinity is achieved by the combination of a strongly favorable binding enthalpy and unfavorable entropy. Overall, inhibitors such as MTX, displaying favorable binding enthalpies do not require extreme binding entropies to achieve high binding affinities. A similar trend was found with HIV-1 protease inhibitors such as ritonavir and KNI-64 [145]. These findings can be useful in drug design since drug screening could be easily achieved and improved by introducing flexible segment in the molecules without lowering the affinity toward their respective target.

4. Effect of methotrexate on the thermal stability of *Bs* DHFR

4.1. Introduction

Binding of an effector molecule is often the primary step in biological reactions. For instance, an enzyme must bind to a substrate to initiate its catalytic reaction. This association between an enzyme and a ligand is also often accompanied with substantial changes in enzyme properties [37]. Such conformational changes have been extensively studied and have given rise to an important number of concepts. The 'induced fit' model suggests the need for functional plasticity for the enzyme active site to accommodate the substrate [9]. Since the three-dimensional architecture of a protein is often described using a unique model, i.e. static, the spatial rearrangement upon ligand binding is very interesting since it must involve a dynamic process. It is therefore very tempting to correlate specific movement such as side chain rotation or loop reorientation to a binding event. Consequently, if protein conformational changes and ligand binding have gained a paramount importance in understanding enzyme kinetics and dynamics, it is also relevant in protein stability. For example the binding of a cofactor has been directly correlated with an increase of protein stability and folding kinetics [148]. Undoubtedly, there is a very subtle and narrow boundary between ligand binding and the dynamic processes involved in the folding or unfolding pathways. The ligand could be in fact an initiating signal for the the onset of original protein properties. Now the question is to understand how the nature of the ligand correlates with structural changes and or protein stability.

A number of structural techniques are available to unravel this question, such as X-ray crystallography, NMR or circular dichroism. In the first place, we have assessed the effect of MTX and H_2F on the DHFR structure and its thermal stability using the well known spectroscopic technique of Circular Dichroism (CD).

4.2. Technique overview: circular dichroism

As with many biological molecules, proteins possess chiral centers. This confers to the protein an optical activity. This property is exploited by spectroscopic methods such as CD [149, 150]. Briefly CD is defined as non equivalent absorption of the left-handed and right-handed circularly polarized light. When asymmetric molecules interact with the beam, they may absorb left circularly polarized (LCP) and right circularly polarized light (RCP) to different extents (figure 23). The result of this discrepancy is termed circular dichroism and can be measured in terms of absorbance:

$$\Delta A = A_L - A_R \qquad \textbf{Equation 28}$$

where A_L and A_R are the absorbance of the left and right circularly polarized light.
By applying the Beer's law, it can be expressed as:

$$\Delta A = (\varepsilon_L - \varepsilon_R) * C * l \qquad \textbf{Equation 29}$$

Where ε_L and ε_R are the molar extinction coefficient for LCP and RCP light, C the molar concentration (in M)and l the path length (in cm).
Then we can define the molar circular dichroism $\Delta\varepsilon$:

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \qquad \textbf{Equation 30}$$

As shown in figure 23, the ellipticity is defined as the angle θ whose tangent is the ratio of the minor and the major axis of the ellipse.
The ellipticity of polarization is equal to:

$$\tan\theta = \frac{E_R - E_L}{E_R + E_L} \qquad \textbf{Equation 31}$$

where E_R and E_L are the magnitude of the electric field vectors of the right-and left-circularly polarized light, respectively. Generally, the circular dichroism effect is small, tan θ can be therefore approximated as θ. The ellipticity can be now written as:

$$\theta = \Delta A\left(\frac{\ln 10}{4}\right)\left(\frac{180}{\pi}\right) \qquad \textbf{Equation 32}$$

The linear dependence of the solute concentration C and pathlength l is removed and the molar ellipticity becomes:

$$[\theta] = \frac{100\theta}{Cl} \qquad \textbf{Equation 33}$$

Then combining the last two equations with Beer's law, the molar ellipticity becomes

$$[\theta] = 100\Delta\varepsilon\left(\frac{\ln 10}{4}\right)\left(\frac{180}{\pi}\right) \qquad \textbf{Equation 34}$$

For illustration of the CD phenomena see this interesting link (http://www.enzim.hu/~szia/cddemo/).

In proteins, CD signal arises from optically active groups such as backbone amide bonds, disulfide bonds, aromatic amino acids such as Phenylalanine (Phe), Tryptophan (Trp), and Tyrosine (Tyr). In secondary structure conformations, the backbone and hence the amide bond chromophores are arranged in regular organized patterns. In the far ultraviolet region (<250nm), CD spectroscopy is extremely sensitive to these patterns and each conformation gives rise to characteristic spectral features. For example, α-helical proteins have a negative peak at 222 nm and 208 nm and a positive peak at 193 nm. Proteins with well defined anti parallel β-pleated sheets (β-helices) present a negative peak at 218 nm and a positive peak at 195 nm. On the other hand disordered proteins present a small ellipticity above 210 nm and a negative band near 195 nm [149]. A typical CD spectrum from a globular protein is

therefore the sum of all these contributions (see fig 24). Therefore, the CD spectrum c(λ) of a protein can be reconstructed as the linear combination of the base spectra multiplied by the abundance of the respective structure elements:

$$c(\lambda) = \sum_{i=1}^{n} fi\,bi(\lambda)$$ **Equation 35**

Here, *n* is the number of secondary structure components, and *fi* is the fraction of structure *i* in the protein.

This direct relationship between protein secondary structure and circular dichroism spectra means that CD is an excellent spectroscopic technique for studying the conformation changes due to temperature, mutations, heat, denaturant or binding interactions [150]. Although CD is not giving detailed information on a residue basis such as NMR or X-ray crystallography, the real power of CD resides in the analysis of structural changes in a protein upon some perturbations. This chapter focuses on the use of CD to follow the structural change of the native state of the enzyme upon MTX binding but also its thermal stability when it is bound to specific ligand such as H_2F and MTX.

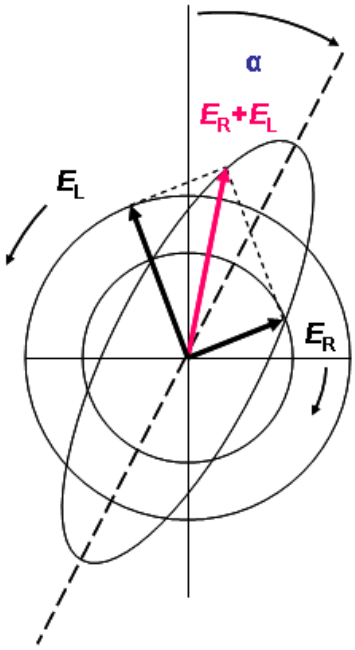


Figure 23. Elliptical polarized light is composed of unequal contributions of right (E_R) and left (E_L) circular polarized light.

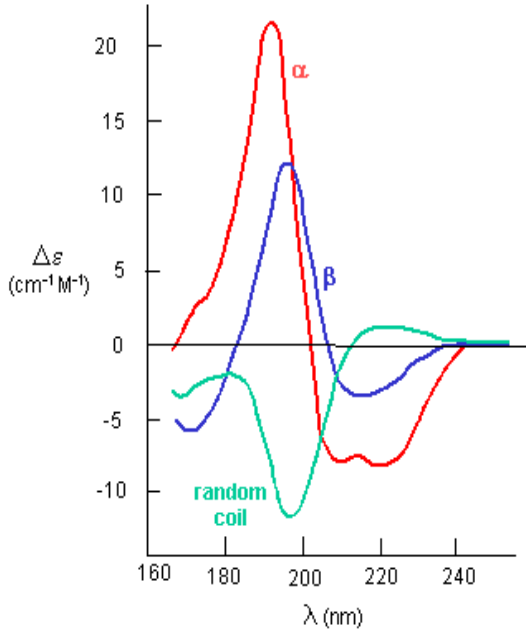


Figure 24. CD spectra of polypeptides and proteins with representative secondary structure such as α-helices (α) β-sheets (β) or random coils.

4.3.Material and Methods

4.3.1. Sample preparation.

Bs DHFR was dissolved in potassium phosphate buffer 10 mM, pH 8, degassed, and centrifuged at 14000 rpm for 5 min to remove any insoluble particles from the solution. Protein concentration was estimated by UV spectroscopy using an absorption coefficient ε_{280} of 25440 L/Mol/cm. Since MTX is a tight binder, it was added at a 1/1 molar ratio with respect to the enzyme. H_2F was used at a 10 fold

excess with respect to the enzyme. The concentrations of H_2F and MTX were estimated by UV spectroscopy using ε_{282} 28000 and ε_{258} 23250 L/Mol/cm respectively. Their respective contribution (below 5% of the overall signal) was subtracted from the CD spectrum.

4.3.2. Thermal unfolding at a single wavelength.

CD experiments were carried out in a Jasco J-810 with a Jasco PTC-348WI peltier effect temperature controller. Quartz cell with a 1 cm path length with a stir bar were used for thermal unfolding. Samples were monitored for helical content by circular dichroism at 222 nm at an enzyme concentration of 7 μM in potassium phosphate buffer of 10 mM pH 7. Thermal denaturation was performed at scan rate of 1 and 2°C/min with a data pitch of 0.2 nm and a bandwidth of 4 nm; denaturation behaviour did not vary with the ramp at these specific rates.

4.3.3. Secondary structure estimation and thermal unfolding

Far-ultraviolet (far-UV) spectra were monitored using a 0.1 cm quartz cell at wavelength between 190 and 250 nm. The final spectrum was the summation of five accumulated spectra taken at a scan rate of 100 nm/min every five degrees. The temperature increment was 1°C/min.

4.3.4. Data processing

Thermal denaturation was analyzed using the program Sigma plot 10 from Systat, Inc to calculate all $T_{1/2}$ values. The relationship between the CD signal measured at a temperature T, y (T), and the proportion of the unfolded state, %U (T), is given by the following equation:

$$\%U(T)= [y(T)-yN(T)]/ [yU(T)-yN(T)] \qquad \textbf{Equation 36}$$

$y_N(T)$ and $y_U(T)$ are the linear dependency of T of the native and unfolded protein in order to take in account the pre and post dependencies of the CD signal on the final

spectrum [148]. Because the reaction was not reversible, the thermodynamic parameters were not determined using the fitting procedure.

Secondary structure estimation was analyzed using the deconvolution software SOMCD [151]. This method is based on a self organizing neural network. The reference spectra consists of 24 CD spectra taken from Yang *et al* [152].

4.4. Results

Far-UV CD spectra of a protein generally reflects its secondary structure composition [149]. The analysis of the CD spectra using the SOMCD software has given an estimate of the secondary structures content in the enzyme *Bs* DHFR in presence or in absence of MTX. As shown in figure 25, the two spectra are overlaid. The common features between the spectra are a large negative ellipticity below 205 nm with broad minima between 205-220 nm. The CD signal obtained from the *Bs* DHFR-MTX complex presents two minima at approximately 207 and 218 nm whereas the free enzyme only shows one minima centered between 208 and 210 nm. Table 5 summarized the characteristics of these spectra and those predicted from the X-ray structure of the free enzyme. A large contribution of the secondary structure elements in *Bs* DHFR is described by random and β-turn structures which contribute to almost 80 % of the overall secondary structure. Other

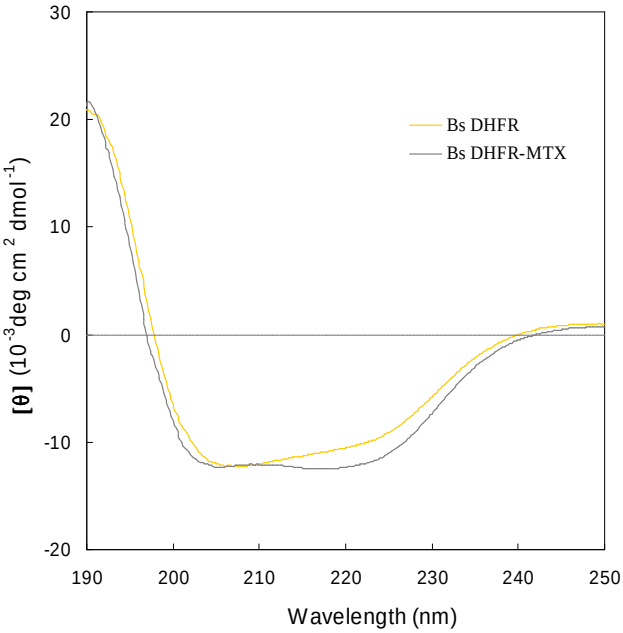


Figure 25. Far UV spectra of unbound *Bs* DHFR (orange line) and MTX complexed *Bs* DHFR (grey line).

secondary structures: α-helix and β-structures account for ~ 17 % and ~ 35-40 % respectively. These findings agree well with the predicted value from X-ray crystallography [56]. Although the two spectra graphically differ in the 215-225 nm regions, the additional minimum observed at 222 nm in presence of MTX, typical of α-helix, seems to have a limited impact on the overall secondary structure distribution. Their estimation reveals only a slight decrease in β-sheets in favor of more random and β-turn conformational elements but according to the relative error, it seems that these differences are not significant.

	Secondary structure		
Sample	α-helix (%)	β-sheet (%)	Other* (%)
Bs DHFR	17.4± 6.4	41.4± 8	41.2± 15.2
Bs DHFR-MTX	17.6± 5.1	34.6± 9.4	47.9± 11.3
DSSP prediction *Bs* DHFR	22	34	Not given

Table 5. Secondary structure of *Bs* DHFR estimated from experimental circular dichroism data and predicted from X-ray structure using the DSSP model (database of secondary structure assignments) [153].

Figure 26 illustrates the melting of secondary structures over temperature increase. For the non-complexed enzyme, as the temperature increases, the intensity of the CD spectrum progressively decreases. The loss of CD signal in the far-UV region started at a temperature equal to or over 55°C with the strongest transition occurring between 60 and 65°C. The overall shape of the graph remains essentially unchanged below 65°C. Above this temperature, the initial minimum centered at 218 nm is shifted to ~ 208 nm. Finally, heating at 80°C caused a total loss of ellipticity and absorbance, suggesting aggregation and precipitation of the protein. On the other hand, the spectra of the binary complex between *Bs* DHFR and MTX denotes a very different behavior. The temperature effect on the CD signal is relatively low below 65°C with no change in the minima position (207 and 218 nm). We clearly observed a red shift of the initial minima observed at 207 nm to approximately 220 nm. An extremely sharp transition reinforcing this observation occurred between 80 and 85

°C. Heating to 90°C leads to almost a complete loss of CD signal over the observed region.

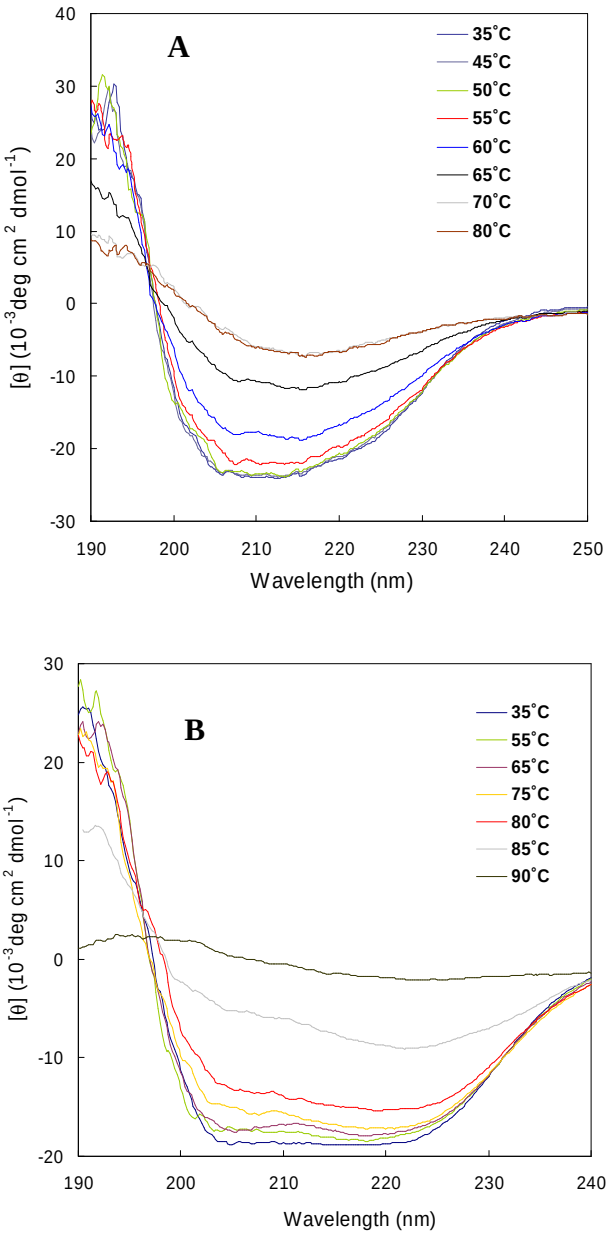


Figure 26. Temperature dependence of the far ultraviolet CD spectra of *Bs* DHFR (A) and MTX-*Bs* DHFR (B). The temperatures are indicated in the inset.

We also attempted to fractionate the loss of secondary structure elements using the SOMCD method [151]. It seems that the CD changes with or without MTX were consistent with loss of most protein's β-sheet structure concomitant with an increase in random structures. However, we have also noted a significant increase in α-helix. It might indicate the limit of the analysis especially at high temperature where the protein is actually 'seen' as a pure α-helix rich protein. The model used could weight external factors from the unfolding process in the conformational analysis misleading the data interpretation.

Because the loss of CD signal was relatively well marked at 222 nm, we have examined the thermal behavior of the enzyme and its binary complex with MTX and H_2F at this single wavelength (figure 27).

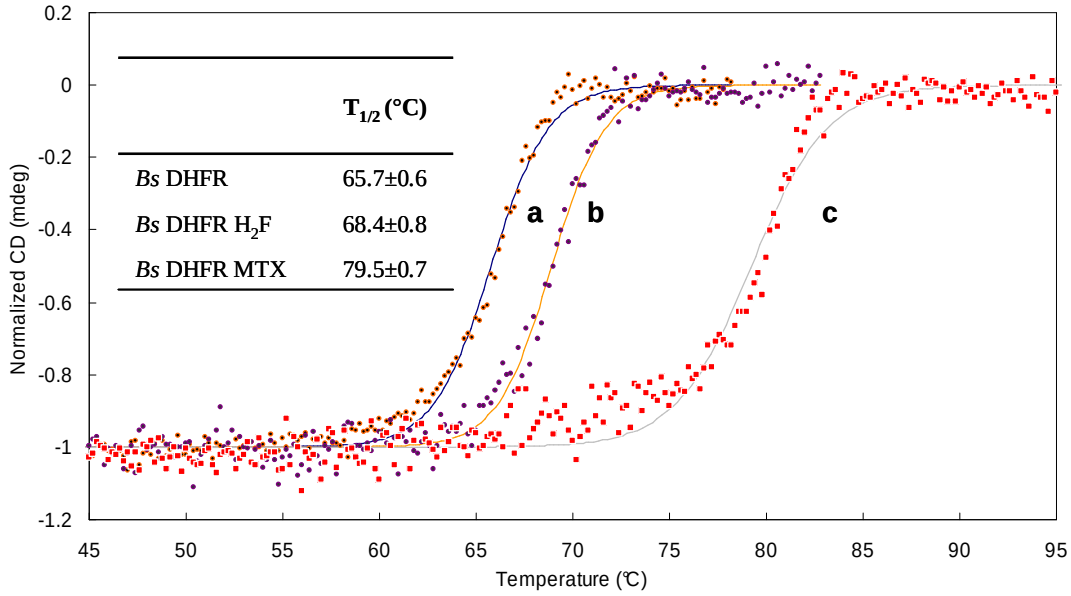


Figure 27. Temperature dependence of the CD at 222 nm of *Bs* DHFR (a), *Bs* DHFR-H_2F (b) and *Bs* DHFR-MTX (c). The dots are the raw data and the solid are the best fit. In the inset, the $T_{1/2}$ values are listed.

Thermal denaturation of the *Bs* DHFR resulted in the progressive and irreversible changes in the CD spectra as shown in figure 27. Indeed, once the initial melt was complete, the sample was brought back to 30 degrees to verify the folding was reversible. The spectra before and after could not be superimposed indicating that the protein can not refold once the temperature is reduced. This inability of the protein to completely renature is the result of either precipitation or aggregation of the unfolded form [154]. Only if the melting is fully reversible, the melting temperature is directly related to conformational stability and the thermodynamics of protein folding can be extracted from the data. In our case, it is more appropriate to depict a $T_{1/2}$ since the melting temperature reflects primarily the kinetics of aggregation and the solubility of the unfolded form as well as the intrinsic conformational stability. Heat denaturation of *E. coli* DHFR gave a similar pattern and was found to be irreversible in [155, 156]. Under the experimental conditions, *Bs* DHFR was denatured at an average temperature $T_{1/2}$ of 65.7°C. The presence of the natural substrate H_2F results in a significant increase of the $T_{1/2}$ by 2.7°C with no change in the unfolding progress curve. In marked contrast to the heat induced effect on the enzyme and its binary complex with H_2F, the presence of MTX has a drastic effect on the $T_{1/2}$. The $T_{1/2}$ value was increased by 13.8 °C reaching the value of 79.5 °C. This later finding corroborates well with the heat denaturation temperatures obtained in binary complexes with MTX between *E. coli* and *b*DHFR (Bovine liver DHFR) in comparison with the free enzyme. The stabilization effect was as much as 14.2 ˚C and 13.3 ˚C for *E. coli* and *b*DHFR respectively. Interestingly, there was also a slight shift from the fitted curve and the raw data between 70 and 75°C and a steeper transition in the $T_{1/2}$ region in comparison with the free enzyme and its binary complex with H_2F. It possibly denotes a cooperative effect [157, 158] in the unfolding process.

4.5. Discussion

Comparing the secondary structure in the free enzyme and its binary complex with MTX, we observed a notable change in the CD signal centered around 220 nm, when MTX binds. As inferred by Hood *et al* it could reflect some reorganization of secondary structure on MTX binding but it could be also due to far-UV transition of

MTX itself [159]. However the secondary structure analysis did not yield a significant difference between the two folds. Standard error observed for spectroscopic techniques such as CD is probably beyond the calculated difference between the uncomplexed and MTX complexed enzyme. The validity and limitations of the deconvolution packages such as SOMCD is discussed in [151]. Overall, the technique, gives accuracies up to 95 % for α-helical content but drop to 50 or 75% for β structures or mixed structures. Other CD deconvolution packages were used such as DICROPROT [160] and *k2d* [161] but they gave very similar results without bringing further information. This ambiguity is noticeably introduced by the nature of the protein reference set used in the statistical analysis or in the strength of the mathematical model to take in account only the conformational factors in the spectroscopic analysis. CD is a low resolution technique, only a high resolution technique such as X-ray crystallography could measure the possible differences.

Similarly, the dissection of the melting of secondary structures upon temperature increase did not yield solid evidence in the hierarchy of the unfolding process. Theoretical studies using MD on *E. coli* DHFR have suggested that α-helical structures unfold prior to the β 'core' [162]. Our results disagree with those finding but we clearly observed a very distinctive behavior between the two temperature scans. The red shift observed in the presence of MTX might underlie a different unfolding pathway. It could reflect the presence of alternative structural intermediates along the unfolding landscape. Additionally, the single wavelength denaturation performed on the MTX binary complex seems to emphasize this hypothesis. The shoulder prior to the thermal transition might indicate a possible non-cooperative effect in the unfolding process. At this stage, only, very flexible regions of the protein are unfolding or partially unfolding leading to a heterogeneous population of folded structures [163, 164]. On the other hand when the temperature approaches $T_{1/2}$, the cooperativity which is measured qualitatively by the width and shape of the unfolding transition is steeper in comparison with the free enzyme or with its binary complex with H_2F. A calorimetry study also reported an asymmetry of the melting curve, a shoulder in the low temperature part of the heat induced transition was observed for the binary of *E. coli* DHFR and MTX [165]. It reflects the melting of at least two

regions of the protein at different temperatures but also it may be indicative of the unfolding of more compact or well folded structure. The presence of an intermediate in the unfolding pathway has been previously reported for *E. coli* DHFR [166]. Interestingly, the kinetic cycle also involved well defined intermediates. MTX could therefore bind, induce and stabilize preferentially particular enzyme conformations but what are the fundamentals behind the MTX stabilization effect?

A ligand such as MTX may have stabilized DHFR against thermal unfolding over a variety of mechanisms. The increase thermal stability is the result of a combination of effects such as ligand-protein specific interactions and or conformational changes. One hypothesis explaining the high temperature stability of the MTX-*Bs* DHFR complex suggests that there are more stable interactions than in the free enzyme. The natural substrate H_2F binds to DHFR at the same site as MTX, however, the stabilization effect is not as effective as when MTX binds. The difference in their binding constants, and other studies highlight the presence of different interactions between the enzyme and its inhibitor MTX, compared to its natural substrate [73]. These interactions could contribute directly by resisting to the unfolding process. Using various complexed form of *E. coli* and *b*DHFR with anti-folate compounds (such as TMP or MTX) and/or the natural cofactor NADPH, Sasso *et al* suggested that the main driving force in the *E. coli* DHFR structure stabilization originated from hydrophobic contacts [165]. A recent study using single molecule atomic force microscopy further supports this hypothesis [167]. The DHFR mechanical stability was found to be increased in the presence of MTX but the origins of this increased mechanical stability was not discussed. As previously demonstrated, the binding strength of MTX is three orders of magnitude higher than H_2F. This discrepancy might be one of the key factors in the unique properties of stabilization. However, we can also speculate that thermal fluctuations are involved. Indeed, thermal unfolding is mainly induced by the enhanced thermal fluctuations overcoming the stabilisating forces such as hydrogen bonds and Van der Waals interactions [168]. Using hydrogen/deuterium exchange Yamamoto *et al* found that the binding of ligands such as NADPH, H_2F or MTX reduce DHFR structural fluctuations [147]. They also point out that MTX was exerting the most significant

effect among the later ligands. In addition according to the X-ray structure, MTX binding causes a 'tightening' of *E. coli* enzyme-a general decrease in the distances within the molecule [146]. It is therefore very tempting to conclude that the reduction of in the amplitude of fluctuations are directly correlated to the binding of MTX and change in thermal stability. However, the role of fluctuations is not fully understood.

In conclusion, in the present study we have observed a puissant stabilization effect of MTX toward the thermal unfolding of the enzyme *Bs* DHFR. Noteworthy, we should also see that thermal stability depends on multiple forces. The sum of these forces necessarily implies a global effect. Interestingly the binding of MTX is a local event. Since CD as mainly all spectroscopic technique gives an average view of the studied system, the binding of MTX must have been propagated far away from its binding site to affect the overall stability of the protein. A potential chain of connectivity could be also assigned upon this particular binding event as inferred by Benkovic *et al* [45]. The pre-existing equilibrium hypothesis or Monod-Wyman-Changeux model (MWC) have also 'propagated' this idea in the field of enzymology and protein function 40 years ago [31]. The binding of MTX have therefore a deep and complex impact on the dynamics and structure of the enzyme. It might underlie a synergistic mechanism that could imply cooperativity.

5. Protein dynamics and stability: the distribution of atomic fluctuations in thermophilic and mesophilic dihydrofolate reductase derived using elastic incoherent neutron scattering

5.1.Introduction

The structures of enzymes reflect two tendencies that appear opposed. On the one hand, they must fold into compact, stable structures; on the other hand, they must be flexible to bind or accommodate a ligand. It appears that there is a trade off between enzyme stability and enzyme function [17]. Indeed, it has given rise to a paradigm, that activity and stability are inversely related by internal dynamics; i.e.

Activity ∝ Dynamics (flexibility) ∝ 1/Stability

Consequently, enzymes need to be sufficiently flexible to be catalytically active, but not so flexible that they readily denature. According to that, thermophilic proteins are believed to be less flexible, more stable and less active than their mesophilic counterpart at room temperature. However, the nature of the dynamic requirements for activity and stability and in particular whether they are the same are unclear. In this chapter which has been recently published, the relationship between thermal adaptation and fast protein dynamics (pico-second timescale) is explored by neutron scattering means. My input in this collaborative work was to provide both enzymes (*E .coli* DHFR and *Bs* DHFR) in 100 mg quantities. I tested the activity of both enzymes. I prepared the neutron sample, i.e.: bringing the protein to the appropriate hydration threshold and taking care of the sample holder. I also assisted in the neutron scattering experiment.

Protein Dynamics and Stability: The Distribution of Atomic Fluctuations in Thermophilic and Mesophilic Dihydrofolate Reductase Derived Using Elastic Incoherent Neutron Scattering

Lars Meinhold,* David Clement,†‡ Moeava Tehei,†‡ Roy Daniel,† John L. Finney,¶ and Jeremy C. Smith*§
*Heidelberg University, Interdisciplinary Center for Scientific Computing, Computational Molecular Biophysics, Heidelberg, Germany; †Department of Biological Sciences, University of Waikato, Hamilton, New Zealand; ‡Institut Laue-Langevin, Grenoble, France; ¶Department of Physics and Astronomy, University College London, London, Great Britain; and §University of Tennessee/Oak Ridge National Laboratory Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee

ABSTRACT The temperature dependence of the dynamics of mesophilic and thermophilic dihydrofolate reductase is examined using elastic incoherent neutron scattering. It is demonstrated that the distribution of atomic displacement amplitudes can be derived from the elastic scattering data by assuming a (Weibull) functional form that resembles distributions seen in molecular dynamics simulations. The thermophilic enzyme has a significantly broader distribution than its mesophilic counterpart. Furthermore, although the rate of increase with temperature of the atomic mean-square displacements extracted from the dynamic structure factor is found to be comparable for both enzymes, the amplitudes are found to be slightly larger for the thermophilic enzyme. Therefore, these results imply that the thermophilic enzyme is the more flexible of the two.

INTRODUCTION

Protein function is commonly understood to depend both on the three-dimensional structure and the dynamics of the polypeptide chain. Further, it has been proposed that increased structural stability of proteins arises from increased rigidity, while increased flexibility may favor higher activity (1–3).

Proteins extracted from mesophilic and thermophilic organisms are interesting subjects for studying the relationships between protein structural stability, dynamics, and function (4,5). A structural comparison between mesophilic and thermophilic protein homologs has revealed that different protein families employ different structural mechanisms to adapt to higher temperatures, with the only systematic rule being an increase in the number of ion pairs with increasing growth temperature (6). This suggests that dynamics may play an important role in thermal stability.

Thermophilic enzymes, which are stable and catalytically active at higher temperatures than their mesophilic counterparts, have therefore been hypothesized to have higher rigidity and correspondingly lower activity than their mesophilic counterparts (7–10). According to this "corresponding state" hypothesis, at moderate temperature the thermophilic protein is less flexible than its mesophilic counterpart but both proteins exhibit the same flexibility when compared at their respective optimal growth temperature. However, some questions have been raised regarding the inverse relationship between activity and stability, as mediated by dynamics (2). For example, and in contrast to the above-mentioned studies, a higher structural flexibility on the picosecond timescale has been measured for a thermostable α-amylase as compared with its mesophilic counterpart (11,12). Moreover, a study at moderate temperature of the millisecond-timescale flexibility of rubredoxin from a hyperthermophile organism has provided no evidence that enhanced conformational rigidity underlies thermal stability (13). These seemingly contrasting findings underline the question raised above whether indeed dynamics plays a key role in the thermal adaptation of proteins, and whether this dynamics may be timescale-dependent.

The protein studied here, dihydrofolate reductase (DHFR), is an enzyme important for cell growth. The structure and function of DHFR are well characterized, and DHFR from *Escherichia coli* (*Ec*) has become an important model for investigating the relationship between protein dynamics and catalytic function (14). Here, the dynamics of DHFR extracted from mesophilic *Ec* and from thermophilic *Geobacillus stearothermophilus* (*Bs*) are studied over a physiological temperature range (of *Ec*) on the sub-nanosecond timescale. Although motions on this timescale do not comprise the full range required for enzymatic function they are indicative of global flexibility. DHFR from *Ec* and *Bs* show closely similar overall and secondary structures as is shown in Fig. 1, *a* and *b* (15,16). However, the x-ray crystallographic B (or temperature) factors, which are indicative of equilibrium structural flexibility and are shown in Fig. 1 *c*, suggest that *Bs*DHFR is, on average, more flexible than its mesophilic *Ec* counterpart.

The pico- to nanosecond timescale dynamics present in proteins can be determined using incoherent neutron scattering (INS) (17). INS has been extensively used to study the

Submitted September 19, 2007, and accepted for publication January 18, 2008.

Address reprint requests to Lars Meinhold, E-mail: lars.meinhold@caltech.edu.

Lars Meinhold's present address is California Institute of Technology, Physical Biology Center for Ultrafast Science and Technology, 1200 East California Boulevard, Pasadena, CA 91125.

Editor: Helmut Grubmuller.

0006-3495/08/06/4812/07 $2.00

doi: 10.1529/biophysj.107.121418

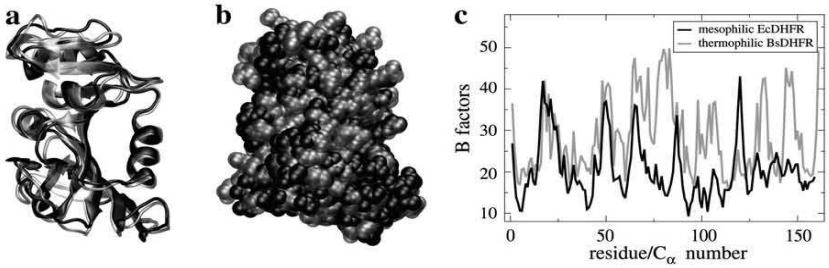


FIGURE 1 Structural comparison between mesophilic *Ec*DHFR (*solid representation*) and thermophilic *Bs*DHFR (*shaded representation*), (*a*) in cartoon and (*b*) in space-filling van der Waals representation. (*c*) The C_α-atom crystallographic B factors are plotted for both enzymes. The *Ec* and *Bs* structural coordinates and B factors were obtained from the Protein Data Bank (43) accession codes 2ANQ (16) and 1ZDR (15), respectively. Both structures were determined at the same temperature, 100 K.

dynamics of proteins and, in particular, to characterize the temperature-dependent change in inferred dynamics that is often referred to as the dynamical transition (18). Much of the dynamical transition work has involved the examination of the elastic incoherent neutron scattering (EINS) from which the average atomic mean-square displacements (MSD) can be derived (19). The physical models used to interpret the experimental EINS data have been extensively tested using molecular dynamics simulations. It has been shown that dynamical inhomogeneity in a protein contributes significantly to EINS (20–22). Furthermore, diffusive protein motions and the finite energy resolution of the spectrometer also influence EINS (23–25). However, a method still commonly used to extract atomic fluctuations from experimental EINS data assumes that all atoms have the same, i.e., an average, fluctuation amplitude. To avoid using this oversimplified description, in this report a model based on a distribution function for the atomic fluctuations is proposed and used to analyze the experimental EINS data.

METHODS

Sample preparation

Chemicals

Deuterium oxide (D_2O, 99.9% and 98%) was purchased from Minipul (Norell, Landisville, NJ). Reagents and medium components for the purification and the analysis of variant DHFRs were purchased from Sigma-Aldrich (St. Louis, MO) and Merck KgaA (Darmstadt, Germany).

Overexpression and purification of recombinant Ec *and* Bs*DHFR*

Recombinant variants of *Ec*DHFR (provided by Carston R. Wagner, University of Minnesota) and *Bs*DHFR (provided by Judith Klinman, University of California at Berkeley) were purified from *Ec* cells (BL21 (DE3)) bearing the plasmid encoding for the DHFR genes, pTZwt1-3 and pET-21a, respectively. The *Ec*DHFR variant was purified by a one-step procedure (26), using methotrexate affinity chromatography (Sigma, St. Louis, MO). The *Bs*DHFR enzyme was first partially purified by heat denaturation (incubation 20 min at 55°C) and then subjected to anion exchange chromatography instead of affinity chromatography (27). Afterwards, a final ultrafiltration step (Amicon concentrator, YM-10 membrane; Amicon Plastics, Houston, TX) was carried out. Finally, the enzymes were lyophilized and kept at 4°C. Their purity was assessed by sodium-dodecyl-sulfate gel electrophoresis.

Neutron scattering sample preparation

The purified protein was dissolved in D_2O (purity 98%) and gently stirred at room temperature overnight to replace the labile hydrogen atoms by deuterium and then freeze-dried. The operation was repeated two more times with higher grade D_2O (purity 99.9%) and the sample was then freeze-dried and stored at 4°C until use. The dry enzyme (115 mg) was mixed into D_2O (345 mg) as a homogenous highly-concentrated solution (300 mg protein/ml) in which protein translational and rotational diffusion is likely to be considerably lower than in a dilute solution. The samples were then sealed in a flat aluminum sample holder (dimension $0.4 \times 30 \times 50$ mm^3).

Neutron scattering, data acquisition, and processing

The neutron scattering experiments were performed on the IN13 backscattering spectrometer of the Institut Laue-Langevin (ILL) in Grenoble, France. This spectrometer is sensitive to the q-range $0.3\ \text{Å}^{-1} \leq q \leq 5.5\ \text{Å}^{-1}$, with an energy resolution of 8 μeV corresponding to observable motions on the timescale of 40–100 ps or faster.

Sample containers were mounted on a cryostat and cooled to 280 K at a rate of -5 K min^{-1}. Scattering data were taken at 280 K and at intervals of 5 K to 305 K (heating rate $+5$ K min^{-1}). At each temperature, the scattering intensity was integrated for 4 h (280 and 285 K) or 5 h (290–305 K) to ensure sufficient statistics. The weights of the sample containers were measured before and after the scattering experiment to ensure that no sample was lost during the experiment; no loss was detected.

The raw data were corrected for scattering of the empty sample container and pure solvent (D_2O), detector response (by using a standard vanadium sample), and for self-absorption events (by using the transmission of the sample) using the softwares Capri and Elascan provided by the ILL for IN13 to obtain $S_{\text{inc}}(q, 0;T)$ at various temperatures T.

Analysis of neutron scattering data

INS provides information on the self-correlations of atomic motions (28). Due to their large incoherent scattering cross-section, the scattering from hydrogens (^{1}H) dominates the EINS from the present samples. For a Gaussian scatterer, the elastic incoherent scattering is given by (29)

$$S_{\text{inc}}(q, \omega = 0) = A\exp\left(-\frac{1}{6}\langle \Delta \mathbf{r}^2 \rangle q^2\right), \qquad (1)$$

where q is the momentum transfer of the scattered neutron, $\langle \Delta \mathbf{r}^2 \rangle(t) = \langle [\mathbf{R}(t) - \mathbf{R}(0)]^2 \rangle$ is the time-dependent mean-square displacement (MSD) of the scatterer on the timescale of the instrument, and A is a constant amplitude. Note that the time-dependent MSD is related to the static thermal atomic mean-square position fluctuation by $\langle u^2 \rangle = (1/2)\lim_{t\to\infty} \langle \Delta \mathbf{r}^2 \rangle(t)$.

In a commonly-used method to extract the temperature-dependent atomic MSD from $S_{\text{inc}}(q, 0)$, use is made of Eq. 1 and linear regressions are performed on $\log S_{\text{inc}}(q, 0)$ plotted against q^2. However, as is shown in Fig. 2, the

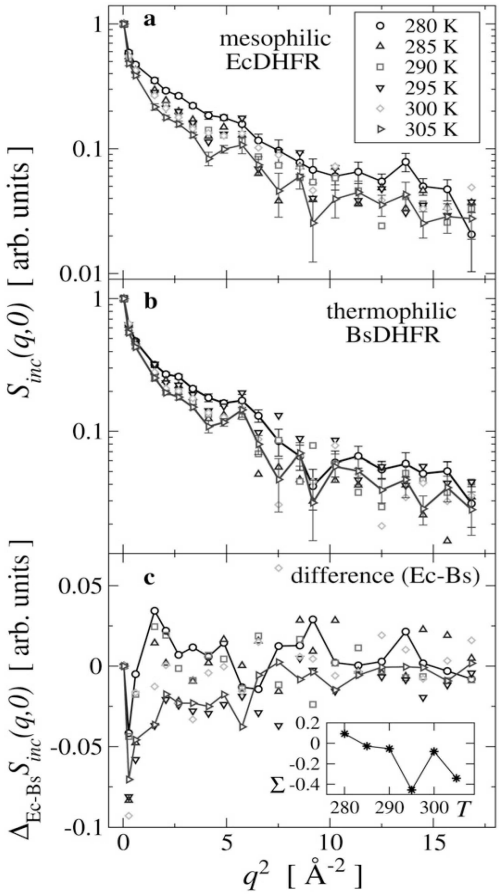


FIGURE 2 Elastic scattering intensity, $S_{\mathrm{inc}}(q, 0)$ measured at various temperatures for (*a*) *Ec*DHFR and (*b*) *Bs*DHFR. For clarity, representative error bars are shown only for the 280 K and 305 K data. Connecting lines are drawn for convenience and the vertical axes are logarithmic. (*c*) The difference, $\Delta_{\mathrm{Ec-Bs}}S_{\mathrm{inc}}(q, 0) = S^{\mathrm{Ec}}_{\mathrm{inc}}(q, 0) - S^{\mathrm{Bs}}_{\mathrm{inc}}(q, 0)$, for each temperature; the inset shows the integrated difference, $\Sigma(T) = \Sigma_{\mathrm{q}}\Delta_{\mathrm{Ec-Bs}}S_{\mathrm{inc}}(q, 0)$ plotted against temperature, *T*.

experimental log $S_{\mathrm{inc}}(q, 0)$ data is not linear over the full q^2-range. This nonlinearity, which has also been reported in previous studies (18,30–32), may, in principle, be due to anharmonic motion and/or the presence of dynamical inhomogeneity, i.e., a distribution of MSD amplitudes. However, it has been demonstrated using molecular dynamics simulations that dynamic inhomogeneity is the major contributor (20,22,33,34). Furthermore, due to their large abundance in proteins, the (rotational) dynamics of methyl groups has recently been identified as contributing significantly to this dynamical inhomogeneity (32,35–38).

Whereas Eq. 1 is not applicable for anharmonic motions, in the case of dynamical inhomogeneity and assuming that the scattering from individual atoms can be described by Eq. 1, the observed $S_{\mathrm{inc}}(q, 0)$ is given by the sum

$$S_{\mathrm{inc}}(q, 0; \{A_{\mathrm{i}}, \Delta T_{\mathrm{i}}^2\}) = \sum_{\mathrm{i}=1}^{\mathrm{N_G}} A_{\mathrm{i}} \exp\left(-\frac{1}{6}\langle \Delta \mathbf{r}_{\mathrm{i}}^2 \rangle q^2\right), \qquad (2)$$

where N_{G} is the number of distinct populations of Gaussian scatterers with MSD $\langle \Delta \mathbf{r}_{\mathrm{i}}^2 \rangle$. A version of Eq. 2 has been used in the literature (30,31,39). to perform independent linear regressions to distinct q-regions. For instance, in Engler et al. (39). the authors used Eq. 2 to fit experimental data taken from Doster et al. (18) with $N_{\mathrm{G}} = 3$, but supplied the weights A_{i}. In doing so, the question arises of which number N_{G} should be used to yield a physically meaningful description of the dynamics present while not overfitting the data. Here, a generalization of Eq. 2 is proposed by using the continuum limit,

$$S_{\mathrm{inc}}(q, 0; A, \alpha) = A \int_0^\infty \mathrm{d}\langle \Delta \mathbf{r}^2 \rangle \, \rho(\langle \Delta \mathbf{r}^2 \rangle; \alpha) \exp\left[-\frac{1}{6}\langle \Delta \mathbf{r}^2 \rangle q^2\right], \qquad (3)$$

where $\rho(\langle \Delta \mathbf{r}^2 \rangle;\alpha)$ is the distribution function of the MSD-amplitudes with a set of parameters α. In a recent molecular dynamics study it has been shown that the non-Gaussian behavior of the EISF of globular proteins can be well described using Eq. 3 (22).

A priori, the functional form of $\rho(\langle \Delta \mathbf{r}^2 \rangle)$ is not known. However, for a given system, $\rho(\langle \Delta \mathbf{r}^2 \rangle)$ can be directly obtained from molecular dynamics simulation. In Fig. 3, $\rho(\langle \Delta \mathbf{r}^2 \rangle)$ is shown derived from an MD simulation of a globular protein (40). The quantity $\rho(\langle \Delta \mathbf{r}^2 \rangle)$ strongly increases at small values of $\langle \Delta \mathbf{r}^2 \rangle$, has a single maximum at $\langle \Delta \mathbf{r}^2 \rangle \approx 0.5$ Å^2, and then decreases with a tail to zero for larger $\langle \Delta T^2 \rangle$. Besides this shape description, any analytical function for $\rho(\langle \Delta \mathbf{r}^2 \rangle;\alpha$ must fulfill two other prerequisites. First, since $\langle \Delta T^2 \rangle < 0$ is unphysical, $\rho(\Delta \mathbf{r}^2 < 0) \equiv 0$, thus precluding the use of a Gaussian distribution. Second, the number of parameters α should be small enough to allow meaningful fitting and interpretation. Here, a Weibull distribution was chosen as the functional form, given by (41)

$$\rho(\Delta \mathbf{r}; \alpha, \beta) = \frac{\alpha}{\beta}\left(\frac{\Delta \mathbf{r}}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{\Delta \mathbf{r}}{\beta}\right)^{\alpha}\right]. \qquad (4)$$

The parameters α and β determine the shape and the scale of the distribution, respectively. As an example, a fit to the simulated MSD data is also shown in Fig. 3. Although the height of the peak and the length of the tail in $\rho(\langle \Delta \mathbf{r}^2 \rangle)$ are underestimated by the Weibull distribution, the general shape is reproduced. Equations 3 and 4 were used to fit the experimental $S_{\mathrm{inc}}(q, 0)$ in a least-squares sense,

$$\min_{\mathrm{A},\alpha,\beta} \sum_{\mathrm{q}} [S^{\mathrm{exp}}_{\mathrm{inc}}(q, 0) - S_{\mathrm{inc}}(q, 0; A, \alpha, \beta)]^2. \qquad (5)$$

The average root mean-square displacement $\mu_{\Delta \mathbf{r}} = \sqrt{\langle \Delta \mathbf{r}^2 \rangle}$ is then readily calculated from the distribution parameters (41),

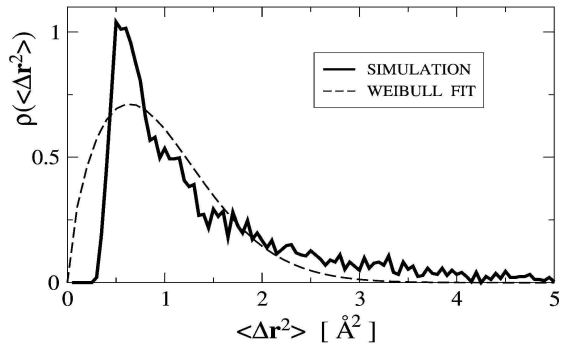


FIGURE 3 Distribution of atomic mean-square displacements from a molecular dynamics simulation of crystalline Staphylococcal nuclease calculated from a 1-ns trajectory and with $\Delta t = 40$ ps corresponding to the IN13 energy/time resolution. Simulation details are described elsewhere (40). The simulation data is fitted using a Weibull distribution, Eq. 4, with the parameters $\alpha = 1.68$ and $\beta = 1.09$.

$$\mu_{\Delta r} = \beta\Gamma\left(1+\frac{1}{\alpha}\right), \quad (6)$$

where $\Gamma(\cdot)$ denotes the γ-function. To obtain an estimate of the error in $\mu_{\Delta r}$, the fit was performed on 100 subsets of data points, randomly chosen from the full q-range with weights proportional to the inverse of their statistical error. Finally, note that for the limit $\alpha \to \infty$, the Weibull distribution Eq. 4 converges toward the Dirac distribution, such that the monodisperse Gaussian model is retrieved: With $\Gamma(1) = 1$, the limits for the mean and variance, b^{-2}, become $\lim_{\alpha\to\infty} \mu_{\Delta r} = \beta\Gamma(1) = \beta$ and $\lim_{\alpha\to\infty} \sigma^2 = \lim_{\alpha\to\infty} \beta^2$,

$$\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)\right] = 0,$$

respectively. Vanishing variance and normalization 1 are properties of the Dirac distribution, and it follows that $\lim_{\alpha\to\infty} \rho(\Delta\mathbf{r}; \alpha, \beta) = \delta(\Delta\mathbf{r} - \beta)$.

RESULTS

In Fig. 2 are plotted log $S_{inc}(q, 0)$ against q^2 for mesophilic and thermophilic DHFR at all temperatures studied (280–305 K) and over the full q-range. For comparison, for each enzyme, the data are normalized such that for the lowest q-value $S_{inc}(q^2 = 0.038\ \text{Å}^{-2}, 0) = 1$ at all temperatures. With increasing temperature, the average slope in the scattering intensity increases for both samples, indicating an increase with temperature in the structural flexibility. The difference in scattering intensity from the *Ec* and *Bs* samples at all temperatures is plotted against q^2 in Fig. 2 *c*. The figure shows a significant difference in low-q scattering ($q^2 \lesssim 6\ \text{Å}^{-2}$), whereas the differences at larger q are somewhat smaller. The inset to Fig. 2 *c* shows that the integrated difference depends on temperature.

For both DHFR samples and all temperatures, log $S_{inc}(q, 0)$ versus q^2 clearly deviates from linearity (Fig. 2, *a* and *b*), indicating the presence of anharmonic dynamics and/or dynamical inhomogeneity. Here, the analysis is performed assuming the validity of the Gaussian approximation but explicitly considering dynamical inhomogeneity, modeled by a Weibull distribution for the atomic displacements. The following analysis utilizes only the elastic scattering data to study differences in the intramolecular flexibility between the two enzymes. In principle, translational and rotational diffusive whole-molecule motions are also present and will contribute to the elastic intensity in the back-scattering regime (19,42). However, the proteins *Ec* and *Bs*DHFR have very similar mass (18.0 kDa and 18.7 kDa, respectively) and three-dimensional structure/shape (Fig. 1, *a* and *b*), which determine the whole-molecule diffusive dynamics. Therefore, differences in the elastic scattering can, to a good approximation, be attributed to differences in the protein-internal dynamics.

A realistic description of the elastic scattering within the framework of the Gaussian approximation is given by Eq. 3. An example of the analysis of experimental data using the Weibull model, i.e., Eqs. 3–5, is presented in Fig. 4 and found to reproduce $S_{inc}(q, 0)$ reasonably well over the full q-range.

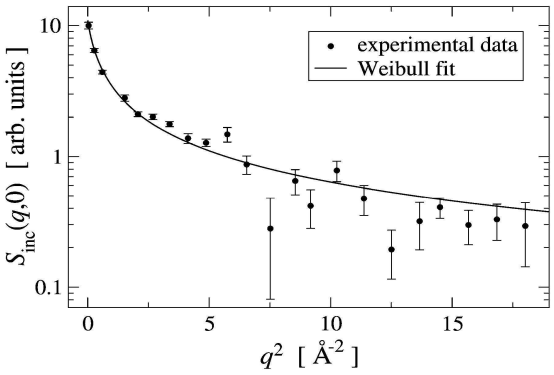


FIGURE 4 Example fit of the Weibull model, Eqs. 3–5, to the experimental elastic scattering data for thermophilic *Bs*DHFR at 300 K over the full q-range. The vertical axis is logarithmic.

In particular, the Weibull model provides an excellent fit to the data for $q^2 < 6\ \text{Å}^{-2}$, where log $S_{inc}(q, 0)$ versus q^2 is strongly nonlinear and the commonly-used analysis method using only Eq. 1 is inadequate.

Using Eqs. 3–5, the average RMS-displacements, $\mu_{\Delta r}(T)$ were determined for both mesophilic *Ec* and thermophilic *Bs*DHFR and are shown in Fig. 5. For both enzymes, $\mu_{\Delta r}(T)$ significantly, and roughly linearly, increases with increasing temperature, with an approximately equal rate of increase. The *Bs*DHFR data point at 305 K appears anomalous. If this point is not considered, the slope of $\mu_{\Delta r}(T)$ is the same for *Ec* and *Bs*DHFR, being 0.036 ± 0.005 Å K^{-1} and 0.037 ± 0.004 Å K^{-1}, respectively.

At a given temperature, $\mu_{\Delta r}(T)$ is somewhat larger for the thermophilic *Bs*DHFR, implying that the thermophilic enzyme is more flexible than its mesophilic counterpart. However, $\mu_{\Delta r}(T)$ provides only an average, i.e., an overall figure that relates to the protein flexibility. In the following, therefore, the utility of the Weibull model is demonstrated by

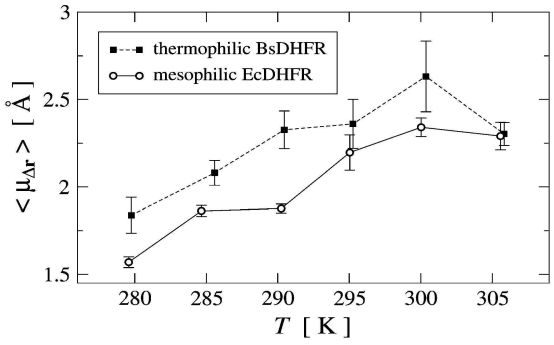


FIGURE 5 Temperature dependence of the average RMS-displacements $\mu_{\Delta r}$ obtained by fitting the Weibull model, Eqs. 3–5, to the experimental EISF data. Error bars denote the standard deviation for 100 fits to data subsets as described in the text. Lines connecting data points are drawn for convenience.

directly visualizing the estimated distributions of atomic fluctuation amplitudes.

In Fig. 6 are shown the temperature-dependent distributions, $\rho(\Delta\mathbf{r})$ of displacement amplitudes for both enzymes. The temperature dependence of $\rho(\Delta\mathbf{r})$ is similar for both enzymes. For both systems with increasing temperature the distribution $\rho(\Delta\mathbf{r})$ becomes broader and the maximum shifted to larger displacement amplitudes. However, this behavior is significantly stronger for the thermophilic *Bs*DHFR, for which the distribution $\rho_{\mathrm{Bs}}(\Delta\mathbf{r})$ is significantly broader than $\rho_{\mathrm{Ec}}(\Delta\mathbf{r})$. Furthermore the shift of the distribution maximum is larger for the thermophilic *Bs*DHFR: the maxima for $\rho_{\mathrm{Bs}}(\Delta\mathbf{r})$ and $\rho_{\mathrm{Ec}}(\Delta\mathbf{r})$ are at 1.6 Å and 1.5 Å, respectively, at 280 K but 2.1 Å and 1.8 Å at 300 K.

Fig. 6 also shows the temperature dependence of the Weibull fit parameters α and β. The shape parameter α is similar for both enzymes and decreases with increasing temperature for $T < 300$ K, indicating a longer tail in the distribution, ρ. The average root mean-square displacement, $\mu_{\Delta\mathbf{r}}$, is only slightly affected by the variation of α in this parameter range. However, $\mu_{\Delta\mathbf{r}}$ is directly proportional to the scale parameter β. The value β is significantly larger for the thermophilic enzyme, indicating that ρ_{Bs} is broader than ρ_{Ec}. For both enzymes, β increases with increasing temperature. Thus, while the distributions of atomic fluctuations have approximately the same shape (determined by α) for both enzymes, the scale or width (determined by β) is larger for the thermophilic protein. With increasing temperature, the distributions for both enzymes become longer-tailed, reflecting that large-scale atomic fluctuations become more likely.

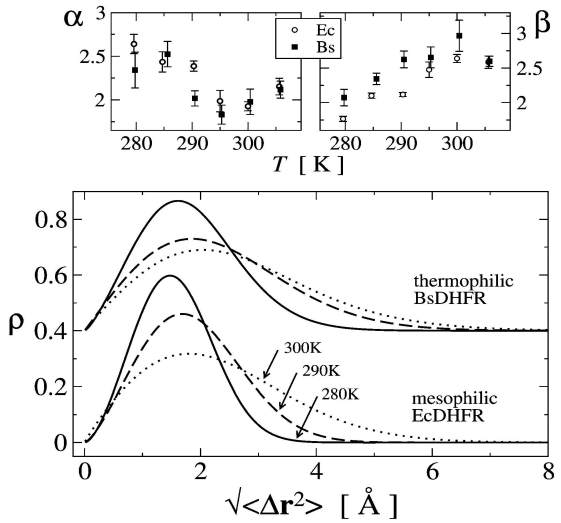


FIGURE 6 (*Top*) Fit parameters α and β for the Weibull model plotted against temperature. (*Bottom*) Average Weibull-distributions of RMS-displacements for mesophilic *Ec*DHFR and thermophilic *Bs*DHFR at 280 K (*solid line*), 290 K (*dashed line*), and 300 K (*dotted line*). For convenience, the profiles for *Bs*DHFR are vertically shifted by 0.4.

DISCUSSION AND CONCLUSION

The sub-nanosecond dynamics of mesophilic and thermophilic dihydrofolate reductase are studied here using elastic incoherent neutron scattering. The scattering data were analyzed using a model based on the Gaussian approximation (quasi-harmonic dynamics) that explicitly incorporates dynamical inhomogeneity using a distribution of atomic displacement amplitudes. Here, a Weibull function was used to model this distribution and the experimental scattering data were well reproduced over the full accessible q-range. In comparison, the commonly-used analysis method based on only one average fluctuation amplitude for all atoms, Eq. 1, typically needs to be restricted to certain q-ranges and, in particular, is incapable of reproducing $S_{\mathrm{inc}}(q, 0)$ for the present samples at small q.

For mesophilic *E. coli* DHFR the average RMS-displacements, $\mu_{\Delta\mathbf{r}}(T)$, increase from 1.6 Å at 280 K to 2.3 Å at 305 K. Over the temperature range studied here, $\mu_{\Delta\mathbf{r}}(T)$ of *Ec*DHFR increases approximately linearly with temperature. For thermophilic *G. stearothermophilus* DHFR $\mu_{\Delta\mathbf{r}}(T)$ increases from 1.8 Å at 280 K to 2.6 Å at 300 K. The decrease of $\mu_{\Delta\mathbf{r}}(T)$ in the last temperature step at 305 K appears anomalous and may be erroneous. However, considering the relatively large uncertainty for the 300 K $\mu_{\Delta\mathbf{r}}$-value, a flattening of $\mu_{\Delta\mathbf{r}}(T)$ for $T \lesssim$ 295 K is also compatible with the experimental data.

A particular advantage of this analysis method, i.e., the Weibull model, is that the distribution of atomic fluctuation amplitudes can be estimated. Although the Weibull model has only two adjustable parameters, the combination of a power law and an exponential function confers versatility on the distribution profile. Furthermore, a comparison between the results for $\mu_{\Delta\mathbf{r}}(T)$ and $\rho(T)$ in Figs. 5 and 6 shows that, due to the tail in the distribution ρ, $\mu_{\Delta\mathbf{r}}$ is generally larger than $\Delta\mathbf{r}_{\mathrm{max}}$, i.e., the position of the maximum in ρ. This also illustrates the difficulty of using only one single value (or moment) to characterize the distribution of atomic fluctuation amplitudes. For a strongly skewed ρ, the average fluctuation amplitude can be significantly different from the value where ρ is maximal. The ambiguity in distinguishing the contributions of protein-internal and whole-molecule dynamics to the elastic scattering remains but can, in principle, be alleviated by exploiting the quasielastic scattering (42). This suggests an analysis method combining distribution functions for the amplitudes of internal fluctuations with rigid-body displacements for whole-molecule translation and rotation.

A further finding concerns the shape of the distribution of atomic displacement amplitudes. These results indicate that this is broader for the thermophilic enzyme. The significance of this is that it suggests that a larger proportion of atoms in the thermophilic enzyme fluctuate with high amplitude. For example, in the distributions at 290 K, 17% of the atoms in the thermophilic enzyme fluctuate with $\Delta\mathbf{r} > 3.5$ Å whereas this value is 4% for the mesophilic species. One can speculate that this highly-mobile fraction might involve the relatively

nonstructured loops of the protein, thus preserving a relatively rigid functional core at higher temperatures. This hypothesis is supported by the observation that the offset between the $\mu_{\Delta r}(T)$-slopes for *Ec*DHFR and *Bs*DHFR in Fig. 5 is $\approx$7 K, whereas their optimal growth temperatures differ by $\approx$15 K. Testing such hypotheses will become possible with specific deuteration and facilitated by the coming on line of next-generation neutron sources, such as the Spallation Neutron Source at Oak Ridge National Laboratory.

Finally, the biological relevance of the results for mesophilic and thermophilic protein dynamics is addressed. The importance of protein rigidity for structural stability has been discussed previously (1–3), but the question arises as to whether rigidity should refer to smaller displacements or a smaller change in displacements with increasing temperature. This report finds the increase in flexibility with increasing temperature to be similar for both enzymes, whereas the fluctuation amplitudes are found to be slightly larger for the thermophilic enzyme. This suggests that thermophilic *Bs*DHFR is intrinsically more flexible than its mesophilic counterpart *Ec*DHFR. The greater flexibility of the thermophilic enzyme may permit the larger fluctuation amplitudes at higher temperatures to be more easily accommodated within the native structure. Further studies on different proteins will be required to ascertain whether this is a general characteristic of mesophilic and thermophilic counterpart proteins.

We thank Dr. V. Kurkal-Siebert for helpful discussions. We also thank an anonymous referee for pointing out the relationship between the Weibull and Dirac distributions.

We gratefully acknowledge supply of sample material by Drs. C. R. Wagner and J. Klinman, help from Drs. F. Natali and M. Bée at IN13, and support from the ILL.

REFERENCES

1. Daniel, R. M. 1986. Protein structure, folding and design. *In* UCLA Symposia on Molecular and Cellular Biology, New Series, Vol. 39. D. L. Oxender, editor. A. R. Liss, New York.
2. Jaenicke, R. 2000. Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *Proc. Natl. Acad. Sci. USA.* 97:2926–2964.
3. Arnold, F. H., P. L. Wintrode, K. Miyazaki, and A. Gershenson. 2001. How enzymes adapt: lessons from directed evolution. *Trends Biochem. Sci.* 26:100–106.
4. Scandurra, R., V. Consalvi, R. Chiaraluce, L. Politi, and P. C. Engel. 1998. Protein thermostability in extremophiles. *Biochimie.* 80:933–941.
5. Fields, P. A. 2001. Review: Protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. Physiol. A.* 129:417–431.
6. Szilagyi, A., and P. Zavodszky. 2000. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comparative survey. *Structure.* 8:493–504.
7. Varley, P. G., and R. H. Pain. 1991. Relation between stability, dynamics and enzyme activity in 3-phosphoglycerate kinases from yeast and *Thermus thermophilus*. *J. Mol. Biol.* 220:531–538.
8. Kohen, A., and J. P. Klinman. 2000. Protein flexibility correlates with degree of hydrogen tunneling in thermophilic and mesophilic alcohol dehydrogenases. *J. Am. Chem. Soc.* 122:10738–10739.
9. Zavodszky, P., J. Kardos, A. Svingor, and G. A. Petsko. 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. USA.* 95:7406–7411.
10. Wolf-Watz, M., V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E. Z. Eisenmesser, and D. Kern. 2004. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struct. Mol. Biol.* 11:945–949.
11. Fitter, J., and J. Heberle. 2000. Structural equilibrium fluctuations in mesophilic and thermophilic α-amylase. *Biophys. J.* 79:1629–1636.
12. Fitter, J., R. Herrmann, N. A. Dencher, A. Blume, and T. Hauss. 2001. Activity and stability of a thermostable α-amylase compared to its mesophilic homologue: mechanisms of thermal adaptation. *Biochemistry.* 40:10723–10731.
13. Hernandez, G., F. E. Jenney, Jr., M. W. W. Adams, and D. M. LeMaster. 2000. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proc. Natl. Acad. Sci. USA.* 97:3166–3170.
14. Schnell, J. R., H. J. Dyson, and P. E. Wright. 2004. Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.* 33:119–140.
15. Kim, H. S., S. M. Damo, S. Y. Lee, D. Wemmer, and J. P. Klinman. 2005. Structure and hydride transfer mechanism of a moderate thermophilic dihydrofolate reductase from *Bacillus stearothermophilus* and comparison to its mesophilic and hyperthermophilic homologues. *Biochemistry.* 44:11428–11439.
16. Summerfield, R. L., D. M. Daigle, S. Mayer, D. Mallik, D. W. Hughes, S. G. Jackson, M. Sulek, M. G. Organ, E. D. Brown, and M. S. Junop. 2006. A 2.13 Å structure of *E. coli* dihydrofolate reductase bound to a novel competitive inhibitor reveals a new binding surface involving the M20 loop region. *J. Med. Chem.* 49:6977–6986.
17. Smith, J. C. 1991. Protein dynamics: comparison of simulations with inelastic neutron scattering experiments. *Q. Rev. Biophys.* 24:227–291.
18. Doster, W., S. Cusack, and W. Petry. 1989. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature.* 337:754–756.
19. Gabel, F., D. Bicout, U. Lehnert, M. Tehei, M. Weik, and G. Zaccaï. 2002. Protein dynamics studied by neutron scattering. *Q. Rev. Biophys.* 35:327–367.
20. Hayward, J. A., and J. C. Smith. 2002. Temperature dependence of protein dynamics: computer simulation analysis of neutron scattering properties. *Biophys. J.* 82:1216–1225.
21. Becker, T., J. A. Hayward, J. L. Finney, R. M. Daniel, and J. C. Smith. 2004. Neutron frequency windows and the protein dynamical transition. *Biophys. J.* 87:1–9.
22. Tokuhisa, A., Y. Joti, H. Nakagawa, A. Kitao, and M. Kataoka. 2007. Non-Gaussian behavior of elastic incoherent neutron scattering profiles of proteins studied by molecular dynamics simulation. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 75:041912.
23. Gabel, F. 2005. Protein dynamics in solution and powder measured by incoherent elastic neutron scattering: the influence of Q-range and energy resolution. *Eur. Biophys. J.* 34:1–12.
24. Gabel, F., and M. C. Bellissent-Funel. 2007. C-phycocyanin hydration water dynamics in the presence of trehalose: an incoherent elastic neutron scattering study at different energy resolutions. *Biophys. J.* 92: 4054–4063.
25. Kneller, G. 2007. Estimating the influence of finite instrumental resolution on elastic neutron scattering intensities from proteins. *J. Chem. Phys.* 126:125107.
26. Baccanari, D. P., D. Averett, C. Briggs, and J. Burchall. 1977. *Escherichia coli* dihydrofolate reductase: isolation and characterization of two isozymes. *Biochemistry.* 16:3566–3572.
27. Wilquet, V., J. A. Gaspar, M. van de Lande, M. V. de Casteele, C. Legrain, E. M. Meiering, and N. Glansdorff. 1998. Purification and characterization of recombinant *Thermotoga maritima* dihydrofolate reductase. *Eur. J. Biochem.* 255:628–637.
28. Bée, M. 1988. Quasielastic Neutron Scattering. Adam Hilger, Bristol and Philadelphia.

6. Effect of perdeuteration on the thermal stability of *Bs* DHFR

6.1. Introduction

Sample preparation, especially for biological molecules can be the bottle neck of a neutron scattering study. To overcome this problem, a number of strategies have been explored. The ability to deuterate selectively biological molecules has facilitated the use of neuron scattering techniques in biology [80]. For instance the use of contrast variation based on the dramatic scattering difference between hydrogen and deuterium atoms has a major impact on the range of applications of small angle neutron scattering [169]. The solution conformation of macromolecular complexes such as viruses [170] or enzymes [171] have elegantly highlighted the power of perdeuteration in structural biology. Consequently, the number of dedicated laboratories producing specifically labelled proteins for NMR, SANS or incoherent neutron scattering is growing fast. Although the applicability of deuteration is unquestionable, the effect of deuteration on living organism is nevertheless poorly understood. A deuterium nucleus contains one neutron and one proton whereas the hydrogen atom does not contain any neutrons. Deuterium behaves similarly to ordinary hydrogen, but there are differences (due to the large mass rest of deuterium atoms) which are larger than the isotopic differences in any other element (see table 6) [172]. The effect of exchanging hydrogens by deuterons is further magnified in biological systems, which are very sensitive to small changes in the solvent properties of water. For instance, at high concentration, mice, rats and dogs (over 25% of body weight) suffer from acute neurological symptoms, liver hyperplasia, anemia, and eventually death [173]. On a molecular level, the effects of H/D exchange is more tricky. As a solvent D_2O increases stability of proteins [174-176], on the other hand perdeuteration does not affected thermal stability of stable proteins whereas proteins with low denaturation temperatures are profoundly destabilized by perdeuteration [177]. In the case of gluthatione S-transferase, the destabilisation effect of side chain deuterons outweighs the stabilisating effect of the core hydrogen bonding [177]. The stability of perdeuterated proteins seems therefore to depend on a number of factors, including steric crowding of the transition state, in which hydrogens have been

exchanged especially in hydrophobic regions [178]. Perdeuteration can also modify protein function, it lowers antifreeze activity of glycoproteins [179]. On the other hand enzymic activities have been reported to be positively [177], negatively (this study) or remain unaffected [180, 181]. Protein structure is often not modified by such isotopic replacement but the differences are very subtle and hard to discern [182]. However, if there is a recent growing interest in protein labelling for structural biology there are too few examples on the effect of perdeuteration (substitution of non exchangeable carbon-bound H for D atoms) on protein structure and function in the literature [181].

Property	**D_2O**	**H_2O**
Melting point ($^{\circ}$C)	3.82	0.00
Boiling point ($^{\circ}$C)	101.72	100.0
Density (20 $^{\circ}$C, g/mL)	1.1056	0.9982
Viscosity (20 $^{\circ}$C, centipoise)	1.250	1.005
Surface tension (25°C, µN-cm)	719.3	719.7
Heat of fusion (cal/mol)	1515	1436
Heat of vaporization (cal/mol)	10864	10515

Table 6. Some physical properties of deuterium oxide and water (from [173]).

This chapter will assess the consequence of H/D exchange in the functional and thermal behaviour of the enzyme *Bs* DHFR. As hydrogens can be replaced in proteins by two ways, it is important to distinguish two sets of effects in this study. The first is a 'solvent isotope effect' that results from the change of solvent from H_2O to D_2O. Only exchangeable hydrogens bound to sulfur, nitrogen or oxygen atoms (relatively weak for C atoms) will be readily replaced by deuterium in a reversible (and time dependent) manner when the solvent is deutrium oxide. This solvent isotope effect, often called the 'primary isotope effect' [183] will principally affect hydrogen bonding in proteins and their hydration shell since D_2O is more structured than H_2O. Secondly the 'Deuterium Isotope Effect' (DIE) resulting from the ability of deuterium to replace H with D in biological molecules at non exchangeable positions. The replacement of non exchangeable hydrogen in protein is obtained by biosynthesis

in fully deuterated media and is called perdeuteration. The DIE is usually considered only in terms of D linkage to C atoms since these linkages are essentially irreversible. These hydrogens are essentially bound to the carbon atoms of the protein backbone and also to the side-chains of amino-acids. This secondary isotope effect [184] will be studied in both H_2O and D_2O solvent.

To follow the effect of perdeuteration and the solvent isotope effect on the thermal stability and activity of *Bs* DHFR, we have used a novel and original model: The Equilibrium model developed at the University of Waikato and University of Bath [185, 186].

6.2.Technique overview: The 'Equilibrium' model

Rates of all reactions, including those catalysed by enzymes, rise with increase in temperature in accordance with the Arrhenius equation [187].

$$K = Ae^{\frac{-\Delta G^{\ddagger}}{RT}} \qquad \textbf{Equation 37}$$

where K is the kinetic rate constant for the reaction, A is the Arrhenius constant, also known as the frequency factor, $\Delta G^{\ddagger}$ is the standard free energy of activation (kJ M^{-1}) which depends on entropic and enthalpic factors, R is the gas law constant and T is the absolute temperature. In general, it would be preferable to use enzymes at high temperatures in order to make use of this increased rate of reaction. However, enzymes are proteins and undergo essentially irreversible denaturation (i.e. conformational alteration entailing a loss of biological activity) at temperatures above those to which they are ordinarily exposed in their natural environment. The effect of temperature has therefore a profound effect on enzyme activity and is often described by two well-established thermal parameters: the free energy of activation $\Delta G^{\ddagger}_{cat}$, which describes the effect of temperature on the catalytic rate constant and $\Delta G^{\ddagger}_{inact}$, which describes the effect of temperature on the thermal inactivation rate constant. The so-called 'classical model' assumes a simple two states equilibrium (see diagram 4) between an active and a thermally-denatured state X [188].

$$E_{act} \longrightarrow X$$ Diagram 4

The variation in enzyme activity with temperature and time of assay can be described as in equation 38

$$V_{max} = k_{cat} \,.\, [E_0] \,.\, e^{-K_{inact}\,.}$$ **Equation 38**

Where V_{max} is the maximum enzyme velocity; k_{cat} is the catalytic constant of the enzyme; k_{inact} is the thermal inactivation constant; $[E_0]$ is the total enzyme concentration; and t is the assay duration.

The variation of the two rate constants in equation 38 with temperature is given by equation 39 and 40.

$$k_{cat} = \frac{k_B T}{h} * e^{-\left(\frac{\Delta G^{\ddagger}_{cat}}{RT}\right)}$$ **Equation 39**

and

$$k_{inact} = \frac{k_B T}{h} * e^{-\left(\frac{\Delta G^{\ddagger}_{inact}}{RT}\right)}$$ **Equation 40**

Where k_B is the Boltzmann's constant; R is the gas constant; T is the absolute temperature; and *h* is the Planck's constant. $\Delta G^{\ddagger}_{cat}$ and $\Delta G^{\ddagger}_{inact}$ are the activation energy of the catalysed reaction and thermal inactivation process respectively. However, this model presents some anomalies as it does not account for the observed temperature-dependent behaviour of enzymes. For instance, at least one study indicates that enzymes at elevated temperature can become less active over timescales which seem too short to involve denaturation [189]. This behaviour means that the observed loss of activity can not be directly associated with an irreversible denaturation process. Indeed, the common graphical temperature optimum: Topt when the enzyme activity (V) is plot versus temperature T : V= f(T) (figure 28) is actually a time dependent parameter, which means experimentally, the longer the assay duration is, lower is Topt and vice versa. Consequently, Topt is not an intrinsic

enzyme property but a mixture of at least two effects: the effect of temperature and time on both enzyme activity and stability.

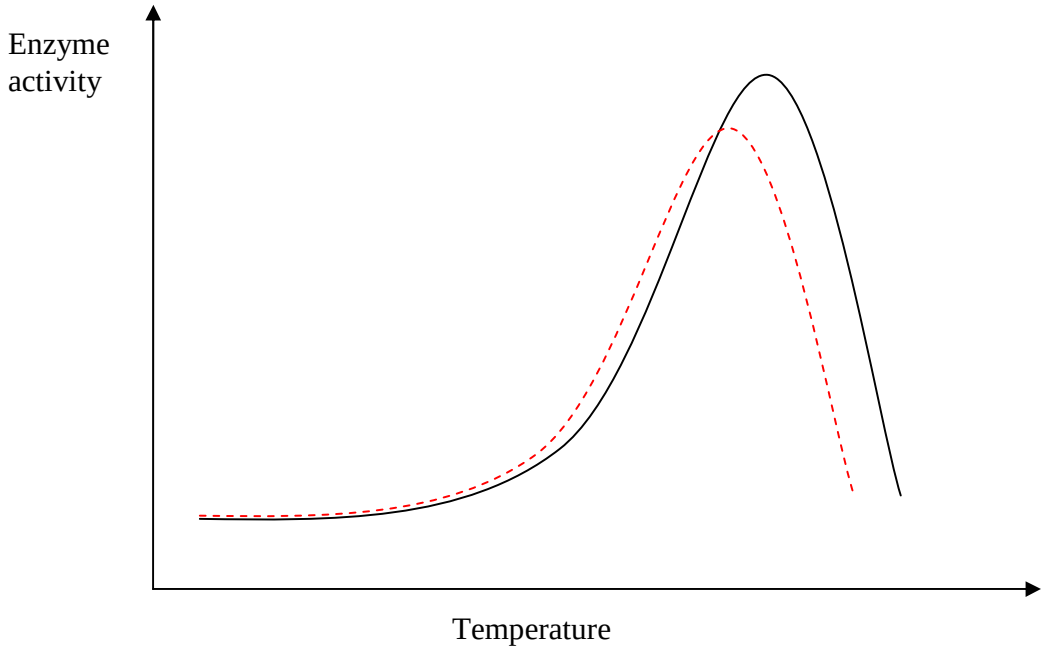


Figure 28. A schematic diagram showing the effect of the temperature on the activity of an enzyme catalysed reaction. ------ Long incubation time, —— short incubation time. Note that the apparent temperature optimum varies with the incubation time

To describe a true temperature optimum, a new model has been developed and validated [185, 186]. This model namely the 'Equilibrium model' describes a new mechanism by which enzyme loses activity at high temperatures. This model includes an extra step in the mechanism where an inactive form of the enzyme E_{inact} undergoes in a reversible equilibrium with an active form of the enzyme E_{act} (very similar to that found in protein unfolding model). According to the postulate (see below), this newly inactived form of the enzyme will irreversibly go to the denatured state X [190, 191].

$$E_{act} \underset{}{\overset{K_{eq}}{\rightleftharpoons}} E_{inact} \xrightarrow{K_{inact}} X$$

Diagram 5

The mathematical description of the reaction velocity introduces a new component: $[E_{act}]$ and becomes dependent of the equilibrium between the inactive and active form of the enzyme. It can be written as followed:

$$V_{max} = k_{cat} \cdot [E_{act}] \qquad \textbf{Equation 41}$$

The concentration of E_{act} and E_{inact} at any time is defined by:

$$[E_{act}] = \frac{[E_0] - [X]}{1 + K_{eq}} \qquad \textbf{Equation 42}$$

Where K_{eq} is the equilibrium constant between the active and inactive form of the enzyme

$$K_{eq} = \frac{[E_{act}]}{[E_{inact}]} \qquad \textbf{Equation 43}$$

The variation of K_{eq} with temperature is given by equation 38:

$$\ln(K_{eq}) = \frac{\Delta H_{eq}}{R}\left(\frac{1}{T_{eq}} - \frac{1}{T}\right) \qquad \textbf{Equation 44}$$

Where ΔH_{eq}: is the enthalpic change associated with the conversion of an active to an inactive enzyme, K_{eq}: Equilibirum constant, R: gas constant, T_{eq}: equilibrium temperature. This equilibrium includes both reversibility and time-dependent enzyme inactivation and can be described by an equilibrium constant K_{eq}. K_{eq} is characterized in terms of temperature dependence by the enthalpy term ΔH_{eq} and a third thermal parameter T_{eq} which is the temperature at which the concentration E_{act} and E_{inact} are equal (see equation 44). T_{eq} can be considered as the thermal equivalent of Km. In addition, we can yield a zero time temperature optimum Topt, independent of the assay duration. Since T_{eq} arises from K_{eq}, it is independent of enzyme thermostability. In other words, T_{eq} reflects a more intrinsic enzyme property albeit it is not directly measurable by experimental means. The overall dependence of enzyme activity on temperature with time is defined by equation 45.

$$Vmax=\frac{\frac{K_B T.e^{\left(-\frac{\Delta G^{\ddagger}cat}{RT}\right)}.E_0.e^{\left(-\frac{K_B T.e^{-\left(\frac{\Delta G^{\ddagger}inact}{RT}\right)}.e^{\left(\frac{\Delta Heq\left(\frac{1}{Teq}-\frac{1}{T}\right)}{R}\right)}}{h\left(1+e^{\left(\frac{\Delta Heq\left(\frac{1}{Teq}-\frac{1}{T}\right)}{R}\right)}\right)}.t\right)}}{h\left(1+e^{\left(\frac{\Delta Heq\left(\frac{1}{Teq}-\frac{1}{T}\right)}{R}\right)}\right)}$$

Equation 45

This equation is therefore used to derive the parameters of interest; $\Delta G^{\ddagger}_{cat}$, $\Delta G^{\ddagger}_{inact}$, T_{eq}, and ΔH_{eq} using a data set using thousand of points. Figure 29 shows the main graphical difference for both models, the classical model does not present a temperature optimum at zero time (a) in contrast with (b), which actually shows an activity optimum at zero time. Secondly, (b) gives a strong evidence of the inactivation equilibrium has taken place since half of the activity is lost at ~350 K, without any contribution deriving from thermal denaturation (zero time).

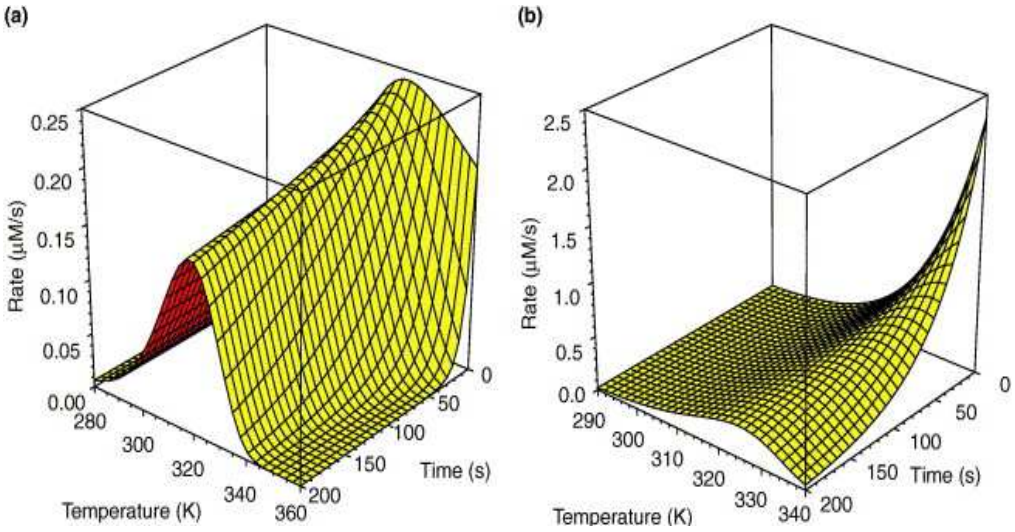


Figure 29. Comparison of the **(a)** Equilibrium and **(b)** Classical models of temperature dependence of enzyme activity**.** The data were simulated using parameter values: (a) $\Delta G^{\ddagger}_{inact}$=95 kJ mol^{-1}, $\Delta G^{\ddagger}_{cat}$=75 kJ mol^{-1}, T_{eq}=320 K; (b) $\Delta G^{\ddagger}_{inact}$=95 kJ mol^{-1}, $\Delta G^{\ddagger}_{cat}$=75 kJ mol^{-1}. Note the absence of a temperature optimum at zero time in the Classical model [192].

If currently, there is no evidence for the molecular basis of such a reversible equilibrium, experimental observations suggested that such interconversion must take place on a ms scale or faster [193]. To date, this recently proposed and experimentally validated model provides a full description of the temperature effect on enzyme catalytic activity for all 40 or so of the enzymes to which it has been applied and for which V_{max} values can be obtained. Furthermore the equilibrium model has shown its usefulness in enzyme eurythermalism and thermophily [193]. For a more detailed description see [194, 195].

6.3. Material and Methods

6.3.1. Enzyme perdeuteration

For protein expression in perdeurated media, the plasmid encoding the DHFR gene was subcloned in a pET-24a (Novagen) vector that confers kanamycin resistance. Perdeuterated *Bs* DHFR was expressed in *E. coli* BL21 (DE3) competent cells. Cells were grown in minimal media ENFORS: 6.86 g/L $(NH_4)SO_4$, 1.56 g/L KH_2PO_4, 6.48 g/L $Na_2HPO_4.2H_2O$, 0.49 g/L diammonium hydrogen citrate, 0.25 g/L $MgSO_4.7H_2O$, 0.5 mg/L $CaCl_2.2H_2O$, 16.7 g/L$FeCl_3.6H_2O$, 0.18 g/L $ZnSO_4.7H_2O$,

0.16g/L $CuSO_4$. $5H_2O$, 0.15g/L $MnSO_4.4H_2O$, 0.18g/L $CaCl_2.6H_2O$, 5 g/l glycerol, 40mg/L Kanamycin. To minimise hydrogen contamination, mineral salts were dried out in a rotary evaporator (Heidoph) at 333 K and labile protons were exchanged for deuterons by dissolving in pure D_2O (99.8 %) and re-dried. Perdeuterated d_8-glycerol (Euriso-Top, France) was used as a carbon source. Since D_2O is toxic for most organisms at high concentration, *E. coli* cells were therefore adapted to fully deuterated medium in three steps [196]. Firstly cells were plated on Luria Bertani solid medium containing Kanamycin (40 mg/mL). A colony was selected and plated on a solid hydrogenated minimal medium. After overnight growth, cells were plated on a partially deuterated medium (v/v 15 % D_2O). After 24-36 hours a colony was selected and plated on a v/v 30 % D_2O medium. This operation was repeated until reaching a value of v/v 100 % D_2O. Liquid deuterated media minimal medium was then inoculated with the adapted cells. Typically, 1500 mL of deuterated medium was inoculated with 100 mL preculture of pre-adapted cells in a 3 L fermentor (labfors, Infors). During the batch and fed-batch phases the pH was adjusted to 6.9 (by addition of NaOD) and the temperature to 303 K. Gas-flow rate was 0.5 L/min. Stirring was adjusted to ensure a dissolved oxygen tension (DOT) of 30 %. The fed-batch phase was initiated when the optical density at 600 nm reached 6.0. D_8-glycerol was added to the culture to keep the growth rate stable during fermentation. When OD_{600nm} reached 12, DHFR overexpression was induced by the addition of 1 mM IPTG and incubation continued for 24 hours. Cells were then harvested, washed in 100 mM NaCl, 10 mM HEPES (pH 6.8) and stored at 193 K. Purification of perdeuterated *Bs* DHFR was undertaken as described previously [56]. Protein purity was assessed by SDS-PAGE and molecular weight was determined by MALDI-TOF spectrometry.

6.3.2. Determination of protein concentration

Protein concentration was estimated using the Buiret reaction [197]. 0.5 mL of the the unknown protein solution concentration is mixed with 2.5 mL of Buiret reagent. The mix is let to stand for at least 15 min and no more than an hour. Absorbance is read at 540 nm and the protein concentration is determined against the protein standart curve. Samples are made in duplicate. The protein standart curve is

made by varying concentrations of stock standard bovine serum albumine solution from 1 to 10 mg/mL. The Buiret reagent consists of 0.96 g of $CuSO_4$ (5 H_2O); 4.89 g of sodium potassium tartrate; 300 mL of 10 % sodium hydroxide. The final solution is brought to 1 L with distilled water.

6.3.3. Mass spectrometry: MALDI-TOF

MALDI-TOF (Matrix-assisted Laser Desorption Ionization Time-of-Flight) analyses were performed using a Brucker Autoflex MALDI-TOF from Daltonik (Bremen, Germany) in linear mode. Data for several 2 ns pulses of the 337 nm nitrogen laser were averaged for each spectrum. An acceleration voltage of 25 kV and a laser intensity of of approximately 10 % greater than the threshold were used. A peptide calibration standard mixture (Brucker) of insulin (5734 Da), ubiquitin (8566 Da), cytochrome C (12361 Da) and myoglobin (16952 Da) was used for external mass-axis calibration. A saturated solution of sinapinic acid in acetonitrile/water (70:30, v/v) containing 0.1 % TFA (Trifluoroacetic Acid) was used as a matrix solution. The protein (0.5 mg/mL) was mixed with the matrix solution at 1:10 (sample: matrix, v/v) ratio. The sample/matrix solution (1 µL) was then deposited on the target plate and air-dried at room temperature. Samples were ionized with 100 laser shots and spectra were obtained by summing at least 500 laser shots. Peptide masses were submitted to an analysis programme: Flex-analysis from Brucker Daltonik (Bremen, Germany) and molecular weights were determined using a Savitzky-Golay filter.

6.3.4. Enzyme assays

Bs DHFR activity was measured continuously by following the decrease in absorbance at 365 nm (ε_{365}= 16700 M^{-1} L cm^{-1}) associated with the oxidation of NADPH and the reduction of dihydrofolate. Standard assay mixtures used sodium phosphate buffer 100 mM at pH 8 or pD 8 (using the relationship pD= pH +0.4), 800 µM NADPH, H, 500 µM H_2F and *Bs* DHFR at a concentration of typically ~3 nM. These substrates concentrations were maintained at ~ 10 times K_m to ensure V_{max} or near V_{max} values were obtained. The reaction was initiated by the addition of 5 µL of enzyme to a cuvette containing 472.5 µL of buffer, 10 µL NADPH, H and 12.5 µL

H_2F which has been preincubated 5 min at the appropriate temperature. Assays were run for 150 s

To analyse the solvent isotope effect, phosphate salts were prior exchanged three times in D_2O before use by rotoevaporation. The substrate and cofactor were also deuterium exchanged at pD 8 and then lyophilized in presence of 1 mM DTT in a 100 mM ammonium bicarbonate buffer at pD 7.9. A detailed account of the experimental procedures needed to determine the Equilibrium model parameters is described in [198].

6.3.5. Data collection

The data were collected on a thermospectronic helios γ spectrophotometer equipped with a thermospectronic™ single cell peltier effect cuvette holder. This system was networked to a computer installed with Vision™ 32 (version 1.25, Unicam Ltd.) software including the vision enhanced rate programme: the time interval was set so that an absorbance reading was collected every second. Enzymatic assays were performed at a variety of temperatures, from 25 °C to 85 °C with 5 °C increments. Reaction progress curves were recorded in triplicate. The reaction rate for each assay was determined using Vision32™ Rate Programme and were repeated if the data deviated by more than 10 %.

6.3.6. Data analysis

Using the values for $\Delta G^{\ddagger}_{cat}$ (80 kJ.mol^{-1}), $\Delta G^{\ddagger}_{inact}$ (95 kJ.mol^{-1}) ΔH_{eq} (100 kJ.mol^{-1}) and T_{eq} (320 K) as initial parameter estimates [191], the experimental data were fitted to the Equilibrium model using a standalone Matlab® (Version 7.1.0.246 (R14) Service Pack 3, the Mathworks, Inc.) application, enabling the facile deviation of the equilibrium model parameters from a Microsoft® Office excel file of experimental progress curves (product concentration versus time).

The standard deviation was not included in the tables as it refers to the fit of the data to the model. Based on the variation between the individual triplicate rates from which the parameters are derived for all the enzymes we have assayed thus far, we have found that the experimental errors in the determination of $\Delta G^{\ddagger}_{cat}$, $\Delta G^{\ddagger}_{inact}$ and

T_{eq} are less than 0.5 %, and less than 6 % in the determination of ΔH_{eq}. For more details see [198].

6.4. Results

Below, are the results from four sets of experimental data aiming at understanding the effect of perdeuteration on the thermal parameters of *Bs* DHFR and also the solvent isotope effect induced by heavy water (table 7). I have only plotted one graph since the shape was very similar in all cases (figure 30).

Enzyme	**Buffer**	**Topt °C**	**Teq °C**	**$\Delta G^{\ddagger}_{cat}$ *kJ mol^{-1}***	**$\Delta G^{\ddagger}_{inact}$ *kJ mol^{-1}***	**ΔH_{eq} *kJ mol^{-1}***
Bs DHFR	H_2O*	64.2	53.3	66.5	97.9	91.9
Bs DHFR	D_2O**	69.4	53.6	66.8	99.3	86.2
dBs DHFR	H_2O*	70.5	55.5	67.9	96.9	90.0
dBs DHFR	D_2O**	68.4	55	68.4	98.7	91.3

Table 7. Summary of experimentally determined thermodynamic parameters.*sodium phosphate buffer 100mM pH 8 in H20, ** deuterated sodium phosphate buffer 100mM pD 8 in D_2O

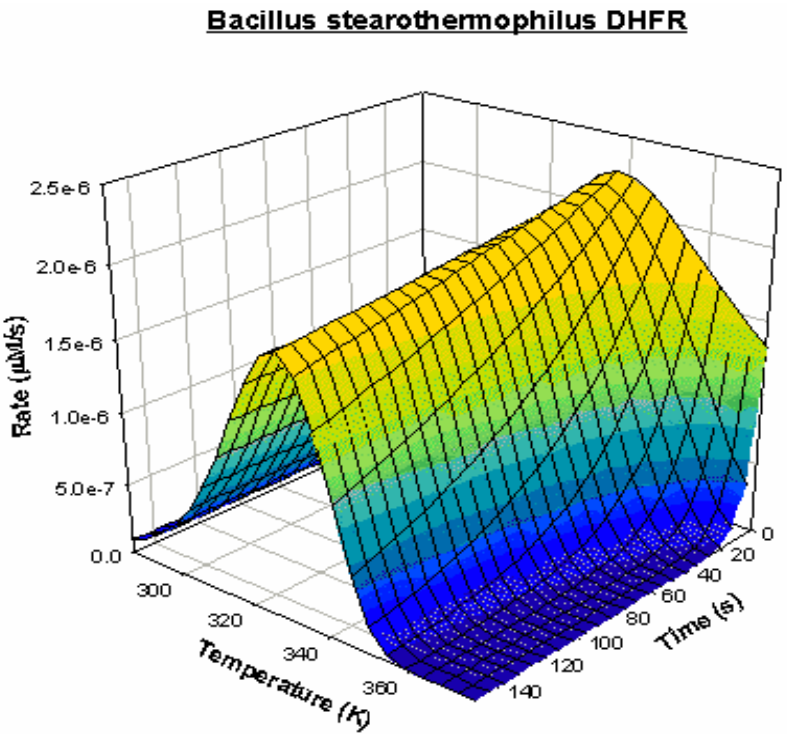


Figure 30. 3D plot of temperature dependence of *Bacillus stearothermophilus* dihydrofolate reductase activity. This plot was created as described in 'Methods' from experimental data to fit the equilibrium model.

6.4.1. $T_{eq,}$ Topt

The hydrogenated *Bs* DHFR presents a T_{eq} and Topt values of 53.3° C and 64.2 ° C respectively. A previous study on *Bs* DHFR using the same methodology reported very similar values for T_{eq} 53.9 ° C and T_{opt} 63 ° C [199, 200]. In addition Klinman *et al* [56] have reported optimum activity values peaking at 75 °C at pH 7 and 65 °C at pH 9. These findings corroborate well our results and also emphasize the thermal origin of the source organism. *G. stearothermophilus* is a moderate thermophilic organism with an optimal growth temperature of 55 °C. The close value of T_{eq} and the growth temperature of *G. stearothermophilus* supports the idea by which T_{eq} is an excellent parameter to predict the thermal environment of organisms as proposed in [200].

T_{eq} shows a dependency by solvent substitution, the temperature shift is about 1.5 to 2° C but is unaffected upon solvent substitution. We recall that T_{eq} reflects the reversible equilibrium between the inactive and active form of the enzyme. An

increase means a higher 'buffering effect' in the interconversion E_{act} to E_{inact}. As perdeuteration consists of exchanging covalently bound hydrogens to deuterium, it implies that only the interactions affected by deuterons are responsible for such shift. Additionally, it is interesting to note that solvent substitution does not significantly affect T_{eq}. The H_2O/D_2O substitution primarily affects exchangeable hydrogens found thorough the protein structure and solvent. Consequently, these observations suggest that the 'hydrogen-exchangeable' network, a prerequisite in normal enzyme structure, did not modify the equilibrium between the active and inactivated complex. On the other hand, Topt behaves differently over perdeuteration. The Topt value is higher upon perdeuteration ≈ 5 °C, but in the mean time, this effect is annealed upon solvent substitution. The values remain constant around a mean value of 68-69 °C. This compensation over solvent substitution might underlie a complex interaction between the effect of temperature on activity and stability.

6.4.2. k_{cat}, $\Delta G^{\ddagger}_{cat}$

	H_2O buffer*			D_2O buffer**		
Enzyme	$\Delta G^{\ddagger}_{cat}$ (kJ mol^{-1})	k_{cat} (s^{-1})	k_{cat}^{D}/k_{cat}^{H}	$\Delta G^{\ddagger}_{cat}$ (kJ mol^{-1})	k_{cat} (s^{-1})	k_{cat}^{D}/k_{cat}^{H}
Bs DHFR	66.5	117	0.59	66.8	105	0.55
dBs DHFR	67.9	69		68.4	58	

Table 8. Summary of experimentally determined $\Delta G^{\ddagger}_{cat}$ and k_{cat}
* Sodium phosphate buffer 100 mM pH 8 in H_2O
** Deuterated sodium phosphate buffer 100 mM pD 8 in D_2O

The perdeuteration systemically lowers the turn over number of the enzyme whether a deuterated buffer is used or not. The k_{cat}^{D}/k_{cat}^{H} ratio between the light and heavy enzyme is almost constant ~0.5-0.6 at all tested temperatures (see Appendix). Perdeuteration changes the temperature dependence of the rate constants as shown by $\Delta G^{\ddagger}_{cat}$ values. Additionally, solvent substitution impacts clearly in this discrepancy. It accounts for a decrease of 10 to 15 % in the k_{cat} values for both enzyme preparations. These data together with those from the literature, indicate that perdeuteration can affect negatively the rate of catalysis [201-203]. However, this effect is relatively low

in comparison of DIE effect which accounts usually over tenfolds difference in rate [204, 205].

6.4.3. $\Delta G^{\ddagger}_{inact}$, $t_{1/2}$,

	H_2O buffer*		D_2O buffer**	
Enzyme	$\Delta G^{\ddagger}_{inact}$ (kJ mol^{-1})	$t_{1/2}$ (s)	$\Delta G^{\ddagger}_{inact}$ (kJ mol^{-1})	$t_{1/2}$ (s)
Bs DHFR	97.9	713	99.3	1189
dBs DHFR	96.9	480	98.7	930

Table 9. Summary of experimentally determined $\Delta G^{\ddagger}_{inact}$ and $t_{½}$ at 323 K. * sodium phosphate buffer 100 mM pH 8 in H_2O, ** Deuterated sodium phosphate buffer 100 mM pD 8 in D_2O

The half-life ($t_{1/2}$) of an enzyme is the time it takes for the activity to reduce to a half of the original activity. The deuterated enzyme shows a lower inactivation energy which means a lower $t_{1/2}$ both in heavy and light water. The perdeuterated enzyme is almost ≈ 1.5 fold less stable than its protiated equivalent. Additionally, using a deuterated buffer we observed an increase in $t_{1/2}$ of 1.7 and 1.9 for the hydrogenated and perdeuterated enzyme respectively. The observed effect is very similar for both enzymes and reveals a clear stabilisating effect over solvent substitution.

6.4.4. ΔH_{eq}

ΔH_{eq} gives the enthalpy change from the active to the inactive form of the enzyme. It can be considered as a measure of the sensitivity of the enzyme catalytic activity to temperature. Due to the very small amplitude of the differences and the relatively high error introduce by the experimental sampling (5%), the values presented here between perdeuterated and protiated enzyme and over solvent substitution are not considered to be statistically significant. Nevertheless ΔH_{eq} is in a quantitative fashion a measure of the enzyme's ability to function over a broad range of temperature: eurythermalism [200]. The application of the model to 22 enzymes [200] has allowed to scale up ΔH_{eq}. *Bs* DHFR and alkaline phosphatase from *Bos taurus* present the smallest ΔH_{eq} over the various enzymes studied. It means that *Bs*

DHFR will function at relatively high activity over a broad range of temperature, i.e. behaves in a eurythermal manner. Conversely, the perdeuration has no effect on this thermal property.

6.5. Discussion

The first goal of this study was to characterize the effect of perdeuteration on *Bs* DHFR kinetic and temperature dependency prior neutron crystallography. We have obtained significant changes upon deuteration (which are commonly believed to be small) using an original model: the equilibrium model.

We have observed a lower catalysis efficiency: k_{cat} upon perdeuteration, about half of the activity is lost in both heavy and light water buffer. A similar reduction of the rate of catalysis was reported by Rokop *et al* using a deuterated alkaline phosphatase [206]. The reasons for this decrease in activity might be widespread since many factors contribute to enzyme activity. For instance, perdeuteration was reported to change significantly the pKa of catalytic residues [207]. Thus, it can led to a completely new reaction site motif and modify the recognition of substrate, bond formation and so on. Other studies using gluthatione *S*-tranferase reported an increase of 25 % of the catalytic rate of the enzyme when deuterium labelled [177]. Cytchrome P450 [181] or Human Arginase I [180] exhibit identical activity upon perdeuteration. It seems that perdeuteration can affect rates of catalysis either positively or negatively. Consequently, the effect of perdeuteration on protein and enzyme activity is very versatile. The structural and physicochemical requirement implied by the reaction pathway might be the key stone in this interdependency but need to be tested over a large number of enzymes with various kinetic mechanisms.

T_{eq} is higher upon perdeuteration and it is suggested to be an active site linked phenomena [185, 192]. If it is clear that T_{eq} is independent of $\Delta G^{\ddagger}_{cat}$ and $\Delta G^{\ddagger}_{inact}$, it is interesting to note that the observed change in k_{cat} is accompanied with a change in T_{eq}: higher T_{eq} with lower k_{cat}. It might be a coincidence but the inter-relationship between these two parameters is intriguing but it has not been discussed since the molecular basis of the model is not known. k_{cat} measures the overall catalytic power of the enzyme and if it is highly dependent of the active site, distal portions of the enzyme might also participate in the catalysis. On the other hand T_{eq} examines

possible conformational changes at the active site and more precisely a subtle equilibrium between an active and inactive form of the enzyme. A simple hypothesis would be: A higher T_{eq} might reflect a higher resilience (rigidity) of the system enabling a higher buffering effect between E_{act} and E_{inact}. However, the cost of this stiffening effect would result in a reduced k_{cat}. Interestingly T_{eq} remains unchanged over solvent substitution. One explanation is that the effect is too small to be measured (less than 5%) or the hydration and water network has little influence on the equilibrium. Stabilizing agent such as glycerol has no effect on T_{eq}. Glycerol protective mechanism is often attributed to a 'preferential hydration' of the protein without modifying its conformation [208]. Overall it seems that D_2O and glycerol enhance the stability of the whole protein rather at the active site. If T_{eq} is an active site parameter then it is normal that the value remains constant over solvent isotope substitution. Interestingly T_{eq} is related to the fundamental thermodynamic parameters the enthalpy ΔH_{eq} and entropy ΔS_{eq} as follow:

$$T_{eq} = \frac{\Delta H_{eq}}{\Delta S_{eq}} \qquad \textbf{Equation 46}$$

with $$\Delta G_{eq} = -RT\ln K_{eq} = \Delta H_{eq} - T_{eq}\Delta S_{eq} \qquad \textbf{Equation 47}$$

Equations 46 and 47 show unambiguously that an entropy and enthalpy compensation mechanism can explain why T_{eq} remains constant over solvent substitution. Such enthalpy-entropy compensation have been observed in many systems [209]. It has been suggested that hydration was probably responsible for the observed consequence. However to fully understand this phenomenon, the dissection of the energetics behind T_{eq} is necessary. Important details of the inactivation process are certainly impaired. The thermal stability of a protein depends generally on the balance of several factors, including ionic force, pH, steric crowding of the transition state, etc...[82]. In the present study, the measured ΔG_{inact} reflects the irreversible thermal inactivation of the inter-conversion of E_{inact} to X. The perdeuterated enzyme shows systemically a smaller $\Delta G^{\ddagger}_{inact}$. This clearly demonstrates a destabilization effect due to perdeuteration. Deuterated proteins are generally less stable against thermal unfolding [181, 206, 210]. The exchange of the covalently bound hydrogens by deuterium reduces essentially the hydrophobic interactions between amino acids

[210, 211]. Hence, hydrophobic forces are believed to be the key stones in protein stability [178]. A reduction of these stabilisating forces would invariably decrease the overall stability of the protein. The destabilizing effect may be caused by steric crowding. Hattori *et al* [211] suggested that C-D bonds have a lower steric requirement compared to the C-H bonds. This implies a reduction of the bond vibration amplitude which diminishes the steric constraints for forming the activated complex and therefore diminishes the activation energy for the thermal unfolding of the protein to occur. With the present model we can not probe in the strictest sense if the protein is globally unfolded or tends to form aggregates. However it has been argued than intermolecular contacts between side chains are more pronounced in perdeutrated proteins. The side chains become more hydrophobic and less bulky which facilitate aggregation [212]. Non-specific or inappropriate interactions would result in a whole destabilizing effect of the ensemble. In contrast, the solvent substitution outweighs the destabilizing effect of perdeuteration. It is well known that protein in D_2O buffer show a higher thermal stability than in H_2O [105, 174-176]. This is curious at first glance since proteins have evolved in protonated environment but the explanation of this apparent contradiction is: Substitution of H_2O by D_2O affects only the exchangeable hydrogens which form the hydrogen bonds in proteins. Hydrogen bonds are strengthened by deuteration cause of their larger mass rest that leads to lower vibration frequencies, respectively lower bond energy for the donator-D bond (peptide N-D) compared to the donator-H bond (peptide N-H). This facilitates the bonding to the acceptor atom (peptide C-O). Consequently, it is predicted that hydrophobic bonds involving non polar side chains of protein will be slightly stronger in D_2O than H_2O [178]. For proteins, it leads to stronger hydration bonds a lower solubility of apolar group which favours hydrophobic interactions [105]. Accordingly, the native structure of the protein will be better maintained in D_2O than H_2O. To understand the energetics behind this stabilisating effect, a thermodynamic analysis is required. It has been undertaken in a number of studies using calorimetry [213], enzyme kinetics [174] or analytical centrifugation [178]. As deuterium bonds are somewhat stronger than hydrogen bonds, it will not be surprising to find larger enthalpy values for the inactivation process in D_2O than H_2O. Accordingly if the

enthalpy changes, the entropy term is also expected to be changed too. Conversely, as described by the following equilibrium

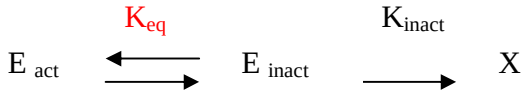


Diagram 5

The critical step at which E_{inact} undergoes the irreversible denaturation to state X must be energetically more demanding than E_{act} to E_{inact}. The dissection of the energy involved in both conversions would undoubtedly reinforce the relevance of the equilibrium not only in enzymology but overall in protein thermal stability. Biophysical means such as calorimetry, circular dichroism or fluorescence would certainly help to initiate the first step.

In conclusion, the Equilibrium model has been successfully applied in this study. It has yielded a number of significant differences in the thermal and functional behaviour of the deuterium labelled and non labelled DHFR. As presented in the introduction, the major application of perdeuterated proteins is 'structural'. It is actually a tool aimed at resolving atomic structure by neutron crystallography, SANS or NMR. This particular application may have led to a common idea that perdeuteration is not invasive as the differences in terms of structures are very subtle and difficult to draw (depending on the resolution shell). However, differences exist and are not easy to predict. The methodology used during this particular study allows and required the presence of the substrates and or products (i.e., physiological conditions) and may have also magnified the effect. Our data with those from the literature indicate that, the physicochemical properties between protonated and perdeuterated proteins can be quite different [177, 181]. If in general it is not easy to draw a common pattern, the thermal behaviour and the solvent isotope effect observed agreed well with a number of studies [174, 181, 210, 211, 213, 214]. However the effect on enzyme activity was quite surprising, being large in comparison with other studies which tend to favour a limited effect on protein activity [177, 181, 207]. These findings might indicate either a specific case with *Bs* DHFR or ability of the model to emphasize small differences which are thought to be small. The effect of perdeuteration and or solvent substitution is maybe more complex or magnified when the effect of temperature is investigated in parallel.

It is also interesting to note the lack of evidence on the molecular basis of the equilibrium between E_{act} and E_{inact}. Although there is no currently no evidence of connecting protein dynamics to the molecular basis of the equilibrium, there is a strong correlation between protein stability and protein dynamics. Thus we might expect a change in protein dynamics maybe localised at the active site when temperature is raised or approaches T_{eq}. The debate of how global and active site protein dynamics participate in protein function is still unravelled. Interestingly a recent study using α-chymotrypsin has concluded conformational dynamics at the active site have a distinct connection with the enzyme functionality [215]. It is very tempting to include T_{eq} in this scheme. T_{eq} is considered as the thermal equivalent of K_m so it is intrinsically a measure of the energetics involve in the transition state. Teq could be an informative probe of active site dynamics reflecting their connection to temperature (i.e. activation energy). This hypothesis needs to be tested but actually the bottleneck is coming from the timescale and space covered by the instruments. In this quest, there may be role for perdeuterated proteins as an indirect informative probe. Selective deuteration of active site residues may shed the light if T_{eq} is an active site phenomenum without the requirement of fast acquisition methods such as stop flow in combination with laser temperature jump devices [216]. Techniques using the kinetic isotope effect (induced by the labelling of at least one of the reactant in enzyme reaction) have been extremely powerful to detail enzyme mechanisms [217]. The perspectives of heavy water and deuterated compounds are therefore widespread and also extended to physiological, pharmacological and also drug design [173].

7. Preparation of samples for neutron crystallography and protein perdeuteration

7.1. Introduction

One of the subsidary goals of this thesis is to rationalize the tremendous power of MTX. Here the proposal is to use X-ray crystallography and neutron crystallography in combination with the ITC data to obtain a comprehensive picture of the energetic and architecture that govern the MTX-DHFR complex. This chapter will summarize the actions that have been taken to prepare samples for X-ray and neutron experiments, including efforts to obtain crystals and to effectively fully deuterate the enzyme *Bs* DHFR.

Although this chapter does not provide the full description of the atomic structure of the MTX-DHFR complex (refined structure), a number of preliminary results have been obtained. Technical parts such as protein perdeuteration, crystal genesis are described. The reason why the X-ray structure has not been entirely refined is primarily financial. This work as the neutron experiment, planned originally at the end of this year, 2008, have been therefore postponed. I have done this work with the help of Dr Suzanna Teixeira, ILL, Grenoble. She has introduced me to all the aspects of protein crystallization and data collection at the ESRF, Grenoble. I have grown and successfully obtained crystals. The structure refinement was made by myself under her supervision and kindly advice.

The catalytic mechanism and ligand binding of DHFR is still a matter of controversy. Indeed previous evidence from X-ray crystallography, NMR, UV/VIS difference, and fluorescence spectra suggest that the N1 atom of MTX is protonated and, thus, positively charged when bound to DHFR [67, 70, 218-220]. The binding of MTX in the DHFR active site is such that the N1 atom is within 2.6-2.7 Å of the Oδ2 of Asp27. It has been suggested that the orientation of the pteridin ring is flipped over a 180° rotation about the C6-C9 bond as compared to the folate and H_2F [218]. The close proximity of N1 (MTX) when bound to Asp27 carboxylate group (DHFR) strongly suggests an electrostatic interaction, possibly a salt bridge (please see figures 9-11 for additional details on the DHFR/MTX complex and interaction). In contrast a

recent study using molecular dynamics suggests that Asp27 is protonated whereas the N1 of MTX is not [221]. The interaction is suggested to be neutral dipole-dipole rather than ionic. In order to resolve the present controversy that represents a gap in our understanding of DHFR's catalytic mechanism and ligand binding, a number of structural studies have been initiated. Structural insights on an atomic scale can be explored with a number of techniques such as X-ray crystallography. However, hydrogen atoms interact weakly with X-rays, causing ambiguities in their accurate positioning upon electron density maps. On the other hand neutron crystallography is well suited for resolving the position of hydrogen atoms in macromolecules [222]. It can resolve the position of hydrogen atoms, especially the deuterium isotope at a moderate resolution (at 2.5-2.0 Å and beyond) [222]. Nevertheless, neutron crystallography has a number of disadvantages. The refinement of the structure and phasing need a model which is often determined by X-ray crystallography. The first step is therefore to obtain the X-ray structure and then move to the neutron experiment. In addition, because of the weak flux generated by spallation or reactor sources, crystal volume needs to be sufficiently high to get sufficient signal to shorten the time needed for the data collection. For instance, crystal volume of 1 mm^3 is often a pre-requisite for a neutron structure and the data collection is generally operated over 15 days (LADI, ILL, Grenoble). Another disadvantage of neutron crystallography is that hydrogen nuclei have a strong incoherent component. Hydrogen atoms appear as negative peaks in neutron density maps and impair the signal by 'cancelling' positive neutron scattering from the atoms to which they are covalently bound like carbon or nitrogen. This problem can be overcome and the signal to noise ratio improved by exchanging hydrogen atoms (H; M= 1.0063 Da) by deuterium (D; M= 2.0126 Da). This can be done by soaking or growing the crystal in D_2O based buffer. In this case, only chemically exchangeable hydrogens (such as hydrogen bound to amino, hydroxyl and carboxylate groups) would be replaced by deuterons. Another way is to express the protein in *E. coli* cells grown in D_2O based media. This latter process is called perdeuteration and has the advantage to replace all the hydrogens by deuterons even at chemically non-exchangeable positions such as hydrogens bound to aliphatic or aromatic carbons. H/D exchange in proteins

significantly reduces the background of neutron data collection, due to deuterium positive and coherent scattering behaviour.

7.2. Technique overview: X-ray crystallography

X-ray crystallography is a method used to resolve the three-dimensional structure of a molecule or complex of molecules [223]. This is done through the interpretation of the characteristic diffraction pattern obtained when a crystal of such a compound is irradiated with X-rays of a certain wavelength, under specific experimental conditions.

X-ray crystallography has proven its significance with the determination of keystone mechanisms in enzymology, chemistry and physics [224]. The structure of a macromolecule contains vital informations on its function. There is a tight relationship between structure and function, at many levels, in all living organism. We can therefore understand or predict the properties of an increasing number of compounds with this technique. The range of application has increased tremendously with instrument development and new sources opening around the world.

The following pages will introduce to the main aspects of the theory of X-ray scattering. Additionally the description of the methods used in crystal genesis, data collection and refinement process will be given.

7.2.1. Crystallization theory

7.2.1.1. Protein crystallization

Obtaining crystals is currently the bottleneck in protein structure determination by X-ray crystallography [225]. The most frequently used crystallization methods are the vapour diffusion and the dialysis technique. The first technique is usually aimed at defining the crystallization conditions, and the screening can be carried out relatively rapidly. With the development of nano-drop robots, hundreds of conditions can be tested within a day. The second technique using dialysis buttons is less frequent in lab routine but it is very useful when large crystals are needed for neutron experiments. We have used both techniques to obtain large, well diffracting crystals.

Crystallization is not a trivial and fully understood process. It is still a semi empirical process often described as ‘an art’ in the literature [226]. Crystallization is a fine equilibrium between a fluid and solid phase of a given material. It depends and can be influenced by a high number of interdependent variables difficult to control such as temperature, ionic strength, pressure, impurities, etc. The interplay between these various ‘actors’ make crystallization very complex and hard to predict. In contrast, the fundaments behind the process of crystallization can be simply illustrated by figure 31.

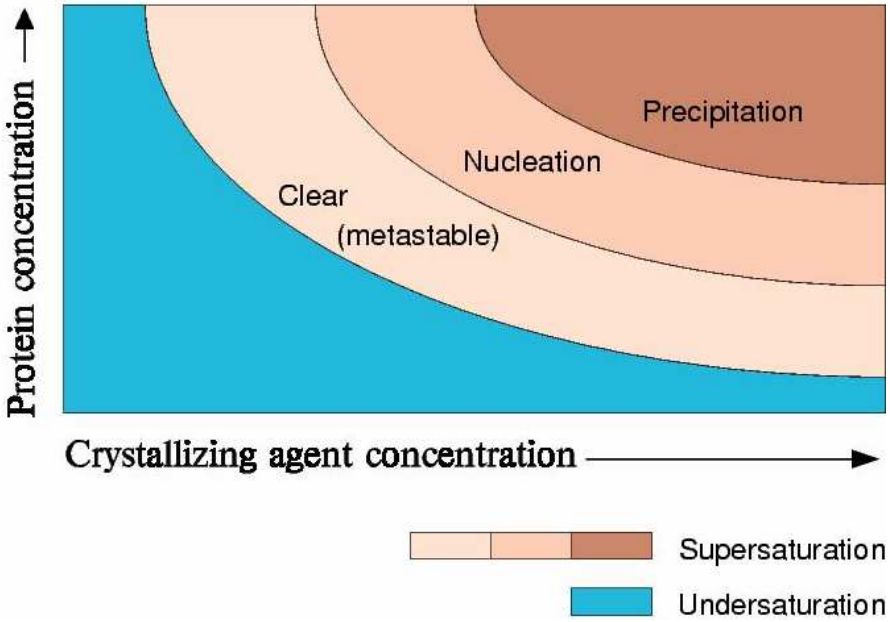


Figure 31. A schematic diagram of a typical solubility curve (from www-structmed.cimr.cam.ac.uk).

This schematical picture shows a typical solubility curve for a macromolecule to be crystallized. Nucleation is achieved with either increasing the molecule or precipitant concentration. The plot can be divided into four main phases: Above the solubility curve there is the so called metastable zone, where supersaturation is achieved. Nucleation does not occur here, as an energy barrier needs to be overcome for protein/protein contact to occur (figure 32). At higher concentrations (either protein or crystallizing agent) the nucleation zone begins, and therefore the crystallization process. Ideally, when a few nuclei are present, crystals start to grow and the concentration of the protein in solution starts to decrease, reaching the

metastable zone. If the concentrations are increased, the protein will start to precipitate and eventually lead to phase separation.

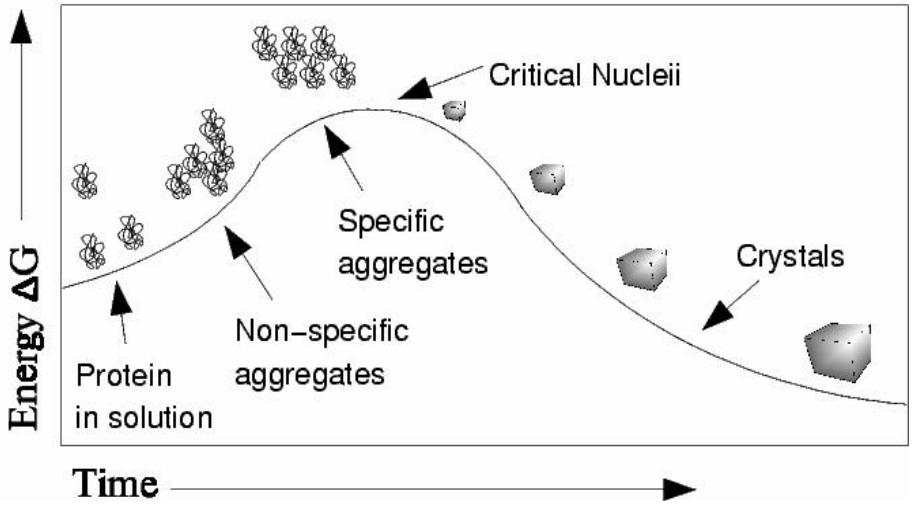


Figure 32. Energy diagram for crystallization (from www-structmed.cimr.cam.ac.uk).

Crystallographers have found ways to rationalize this particular work: statistical approaches based on the experimentally resolved structures [227] or specific screening methods. Crystallization usually begins by screening over a wide range of conditions until preliminary conditions are determined. The refinement of the primary conditions will operate as a fine tuning, enabling the production of optimal crystals. This process can take from days, to months or even years. Several techniques have been developed to crystallize proteins with a better control of the variable at stake, reproducibility, efficiency, and cost in terms efforts and samples. By far the most popular are the vapour diffusion methods which basically concentrate the sample slowly through vapor diffusion. The two most well known variants of this technique are the hanging drop and the sitting drop. The macromolecule is usually placed in a drop containing a buffer and one or more salts. The drop is equilibrated against a reservoir containing one or more molecules such as PEG (Polyethylene Glycol), MPD (methyl 1-2, 4-pentanediol) These act as precipitating agents and the spontaneous diffusion of water vapour from the drop to the reservoir will concentrate the macromolecule in the drop, which ultimately will form nucleation sites. These methods are easy to set up, relatively cheap and alow the fast screening of a large

number of conditions. However, they usually require a large amount of sample. The representation of the apparatus needed for these operations is given by figure 33.

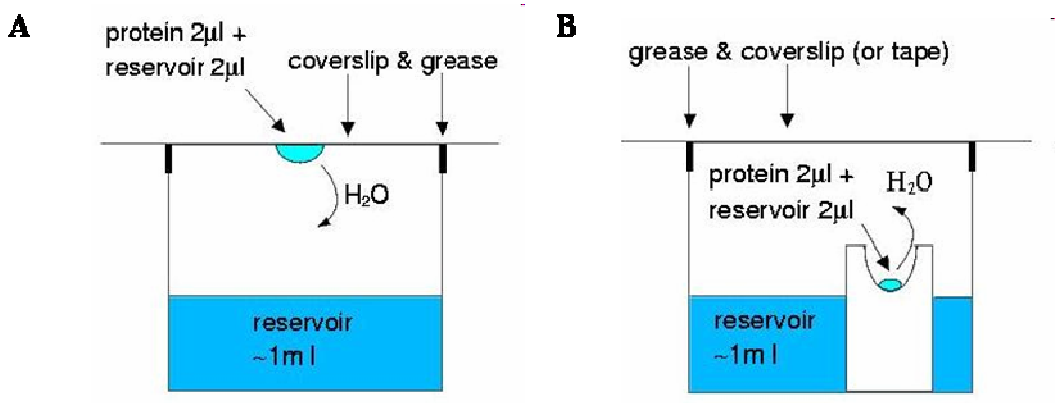


Figure 33. Schematical representation of the two variants of the vapour diffusion method. (A) corresponds to the hanging drop method whereas (B) illustrates the sitting drop method (from www.kenividilaseris.wordpress.com).

Others strategies make the use of different techniques such as the batch method, liquid-liquid diffusion methods, dialysis methods. The batch method brings the macromolecule to the nucleation zone when the system moves slowly from the supersaturation zone to the metatasble zone. This method requires a relatively small amount of the macromolecule and other reagents but remains less successful than vapour diffusion. In liquid-liquid diffusion methods the macromolecule is slowly mixed with the precipitating solution by diffusion. When capillaries are used, gels create barriers between solutions thus slowing down the process. This usually creates a concentration gradient itself intiating the nucleation process. This method requires a considerable amount of material but produces large crystals that can be easily and directly tested for diffraction. Dialysis methods are somewhat interesting as they allow a slow change of the precipitating solution. This method is often used in dialysis buttons sealed with a permeable membrane with a defined cut off. The advantages of this technique are of particular importance when the macromolecule is not stable if sudden changes are made in the crystallization buffer. It could also allow the diffusion of any molecules at any stage of the crystallization. This is particularly useful when one wants to add a specific ligand to an enzyme or add a cryo protecting solution with minimal damage to the crystal. In addition because of the large volume

of the buttons (50 to 200 µL) large crystals can be obtained which is an asset for neutron crystallography.

Besides these common techniques, a number of factors have to be taken into account for the optimization of the crystals. For instance, the drop size, the temperature, the speed of diffusion, etc, could have a strong influence on the results. If a ‘good looking’ or promising crystal is obtained the next step is to test their diffracting pattern.

7.2.2. X-ray scattering and crystals

A crystal is a periodic solid with an internal ordered structure, made of the repetition of building locks called a ‘unit cell’ in three dimensions [223]. The unit cell can be described by the length of its edges (a, b and c) and the angles between them (by convention: α is the angle between b and c, β is the angle between a and c, γ is the angle a and b). A crystal structure can be made of primitive or multiple unit cells but certain conventions have been established to guide the choice. The unit cell should be of the smallest for which, the vector coincides with direction of the highest symmetry. A symmetry operation is an operation that can be performed either physically or imaginatively that results in no change in the appearance of an object. For crystals the symmetry is internal and if each unit cell can be represented by a single point, the result is called a crystal lattice, which is an imaginary three-dimensional arrangement of points such that the view in a given direction from each point in this lattice is identical to the view in the same direction from any other lattice point. The convolution of the structural motif and the crystal lattice will detail the entire crystal structure.

The defining property of a crystal is its inherent symmetry, by which we mean that under certain operations the crystal remains unchanged. Depending on the type of unit cell, there are seven possible crystal systems. Crystal systems are a grouping of crystal structures according to the axial system used to describe their

lattice. Each crystal system consists of a set of three axes in a particular geometrical arrangement. The simplest and most symmetric, the cubic (or isometric) system, has the symmetry of a cube, that is, it exhibits four threefold rotational axes oriented at 109.5 degrees (the tetrahedral angle) with respect to each other. The other six systems, in order of decreasing symmetry, are hexagonal, tetragonal, rhombohedral (also known as trigonal), orthorhombic, monoclinic and triclinic. Combining the various possible lattice centerings with the different crystal systems defines the Bravais lattices. They describe the geometric arrangement of the lattice points, and thereby the translational symmetry of the crystal. In three dimensions, there are 14 unique Bravais lattices which are distinct from one another in the translational symmetry they contain. Symmetry operations in a crystal are not limited to translation, other types of symmetry are found: rotation, reflection and inversion. The crystallographic point group or crystal class define mathematically the different group of symmetry enumerated earlier. There are 32 possible crystal classes, classified into one of the seven crystal systems. The result of combining the 14 Bravais lattices with the 32 point groups and translational symmetry (glide and screw axes) define the space group of the crystal structure. There are 230 distinct space groups but for proteins this number is lower, as only L amino-acids occur naturally and because of the chirality of α-helices. In fact, any symmetry operation that would change the handedness of the structure is not allowed (mirror planes, inversion centers). The same thing can be said about DNA with its helical structure. This leaves 65 possible space groups for biological molecules [223].

7.3. Diffraction

7.3.1. Bragg's Law

The scattering of electrons, X-rays and neutrons by crystals can be described as a reflection of the beams at planes of atoms (lattice planes). If the incident plane wave hits the crystal at an arbitrary angle, the interference of the reflected waves can be either destructive or constructive. For a constructive interference to occur, the path difference between the two incident and the scattered waves, which is $2d\sin\theta$, has to be an integer number of wavelengths ($n\lambda$). For this case, the Bragg law then gives the relation between interplanar distance d and diffraction angle θ:

$$n\lambda = 2d\sin\theta \qquad \textbf{Equation 48}$$

When X-rays strike a crystal they will be diffracted when and only when, this equation is satisfied (figure 34).

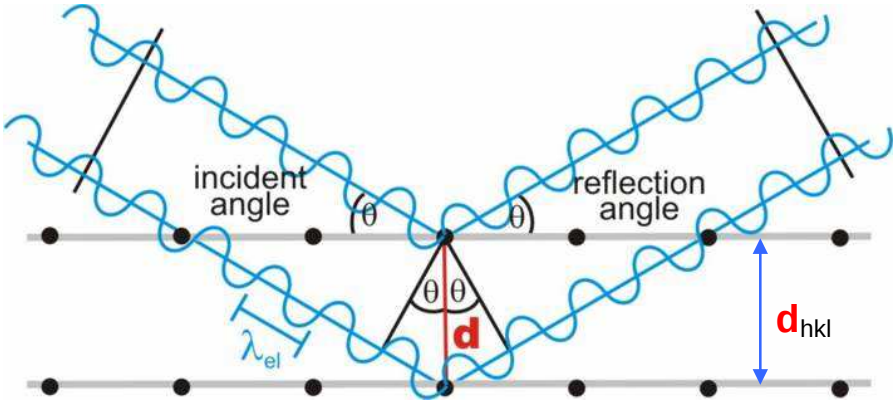


Figure 34. Schematic representation of a crystal lattice (black dots correspond to the lattice point) with a distance d_{hkl} between the lattice plane. θ is the angle of reflection. (from www.microscopy.ethz.ch/bragg.htm).

7.3.2. Notions of reciprocal lattice

The crystal lattice describes the internal crystal structure but does not provide information on the detailed arrangement of atoms within a unit cell. The main question is: How do we know the relationship between the crystal and the diffraction pattern that we will obtain from it?

In this case it is often easier to think of the diffraction experiment with respect to a reciprocal lattice rather than the crystal lattice plane. The concept of reciprocal lattice was proposed by Josiah Willard Gibbs in 1884 [228] and introduced in crystallography by Ewald in 1921 [229]. The reciprocal lattice is constructed from the spacing between the crystal lattice planes d_{hkl}. In the reciprocal lattice the point hkl is drawn at a distance $1/d_{hkl}$ from the origin and in the direction of the perpendicular between the lattice planes. The concept of reciprocity has already appeared intuitively in Bragg's law. It is clear that sinθ is proportional to 1/d. It is therefore easier to convert 1/d to a reciprocal d* that will have a direct relationship to sinθ. For this

reason it is more convenient to describe scattering of X-rays by a crystal in an imaginary lattice called the reciprocal lattice. The reciprocal lattice is constructed on the a*, b*, c* axis derived from the (direct) crystal lattice axis so that a* is perpendicular to b and c, b* is perpendicular to a and c and c* is perpendicular to b and a.

7.3.3. The Ewald sphere

The Ewald sphere is a geometrical description of diffraction which encompasses both the Bragg's law and the Laue equations. It was initially proposed by P.P. Ewald [229] In the Ewald construction (please see figure 35) a circle with a

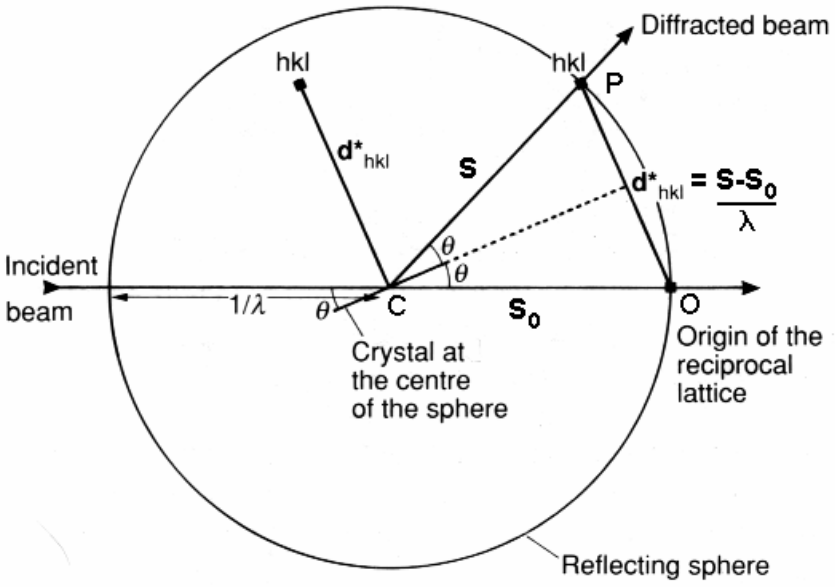


Figure 35. The construction of an Ewald circle, radius 1/λ, with the crystal at the center C. S_o and S are the incident wave vector and scattered wave vector respectively.

radius proportional to 1/λ with the crystal as is drawn. The Ewald construction provides a geometrical relationship between the orientation of the crystal and the direction of the X-ray beams diffracted by it. The advantage of this description is that it allows the determination of which Bragg reflections will be observed knowing the orientation of the crystal with respect to the incident beam. As the crystal is rotated, so is its crystal lattice and thus also the reciprocal lattice. If during the rotation of the crystal a reciprocal lattice point (hkl) intersects the sphere, Bragg's law is satisfied.

The result is a reflection in the direction S_o, with values of h, k, l corresponding both to the values of the reciprocal lattice point and for the crystal lattice planes.

7.3.4. Structure factors and electron density: the phase problem

Previously we have seen how to obtain a diffraction pattern. The question arising now is: how do we get from the diffraction data to a model of the actual protein structure?

The X-ray diffraction pattern of a crystal is made of a multitude of spots with various intensities. If we measure the intensities and the position of reflections, we would in principle determine the geometry of the crystal lattice (driven by the spatial arrangement of the spots) and the arrangement of atoms within one unit cell (given by the intensities). With the symmetry operations implied by the direct lattice, the entire crystal structure will be solved with only one cell. The intensities recorded on an X-ray detector during data-collection can be expressed as:

$$I = K|F|^2 (Lp)(Abs) \qquad \textbf{Equation 49}$$

Where F is the structure factor, K is a scale vector (necessary to bring all reflections into the same scale), Lp is a geometric correction (combination of geometric corrections-shape, orientation and composition of the crystal), Abs is a correction due to absorption effects. This intensity is proportional to the square of the structure factor amplitude, $|F_{hkl}|$. A structure factor describes the way in which an atom or group of atoms scatter X-rays. The structure factor $F_{|hkl|}$ for a reflection *h,k,l* itself is a complex number derived as follows :

$$F_{hkl} = \sum_{j=1}^{n} f_j . \exp\left[2\pi . i\left(hx_j + kx_j + lx_j\right)\right] \qquad \textbf{Equation 50}$$

f_j is the scattering factor of atom *j*, and *h, k, l* the three indices of the corresponding reflection. Of particular importance is the exponential which appears as a complex number with *x,y,z* the fractional coordinates of each atom in the summation. There is an imaginary component that applies to the phase component: α,

of the structure factor. The real importance of this missing component is signified in the electron density (ρ) equation, which is described as a Fourier series:

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l |F_{hkl}|.\exp\ -2\pi.i(hx+kx+lx-\alpha_{hkl}) \qquad \textbf{Equation 51}$$

Where |F*hkl*| is the amplitude of the structure factor, V is the volume of the unit cell, *i* is the imaginary part of the complex number and α_{hkl} is the phase of the reflection hkl. As seen in equation 45, the electron density is expressed as a function of structure factors. We are transforming an inverse space into a real or direct space (the electron density at a real point x,y,z in space). This transformation is accurate and in principle complete but several parameters need to be determined to calculate the real structure. Particularly, we need the complex structure factors $F_{(hkl)}$ but we know only their magnitude $|F_{hkl}|$. Indeed, we know only the absolute value (magnitude) of the complex structure factor F_{hkl} but not its phase, α_{hkl}. To fully characterize the electron density, the phase angle of each reflection α_{hkl} must be known. However, it can not be measured directly: this called the 'phase problem' in crystallography. A number of methods have been established to solve this problem and will be described next.

7.3.5. Methods to solve the structure

Because we can not measure the phase directly, we have to deduce it indirectly. There are a number approaches following four general methods (that can be used either separately or combined): direct methods, Molecular Replacement (MR), Isomorphous Replacement (IR) and Multiwavelength Anomalous Dispersion (MAD). Selection between different techniques and between different programs is a very important step for protein structure solution. Many structures were never been solved due to mistakes in structure solution procedure, which in fact ruins all protein crystallography experiments.

7.3.5.1.Direct methods

The argument that phase information is lost in the diffraction process is not totally correct. The phase problem can be solved directly from the experimental data:

$|F_{hkl}|$ values. The direct methods basis [230] exploits two conditions. Firstly, the scattering density must be a positive real number (or zero) as the probability to find an electron at a position (x,y,z) can not be negative. Secondly, the electron density maps should have high values near atomic position and zero elsewhere else. If these two conditions are satisfied then the correct value of $\alpha_{(hkl)}$ would appear in the Fourier series. Herb Hauptman and Jerome Karle won the Nobel Price in chemistry (1985) owing to their pioneering and outstanding contributions on direct methods. At present, direct methods are the preferred method for phasing crystals of small molecules having up to 1000 atoms in the asymmetric unit. However, they are generally not appropriate for larger molecules such as proteins [230, 231].

7.3.5.2.Molecular replacement

This is by far the most popular method to solve protein structure and the basic ideas were described by Michael G. Rossman and David M. Blow in 1962 [230, 232]. The principle is quite simple, as you need to place the model structure in the correct orientation and position in the unknown unit cell. To orient a molecule you need to specify three rotation angles and to place it in the unit cell you need to specify three translational parameters. The molecular replacement problem is therefore a 6-dimensional problem but to minimize time and power it is separated in two independent searches. So, it became two three-dimensional problems. A rotation function can be computed to find the three rotation angles, and then the oriented model can be placed in the cell with a 3D translation function. Molecular replacement (MR) commonly used a structurally related model (or portion) of the unknown structure in the crystal. As a rule of thumb, MR requires a model that shares at least 40 % sequence identity with the unknown structure. It becomes progressively more difficult as the model becomes less complete or shares less sequence identity. The first step in MR is to calculate the Patterson map for both, the models and the unknown structure. The Patterson function was introduced by A. Lindo Patterson in 1934 [233]. It is a Fourier synthesis that does not require phase information. It is usually defined as:

$$P(uvw) = \frac{1}{V}\sum_{hkl}\left|F_{(hkl)}\right|^2 \cos[2\pi(hu + kv + lw)] \qquad \textbf{Equation 52}$$

Where (u v w) is a vector from a position x, y, z to another x + u, y + v, z + w. The Patterson map gives information on the orientation of a molecule of known structure in the unit cell. It is therefore very useful for positioning the unknown structure in the new unit cell. First a rotation function is used to assess the degree of agreement or coincidence between the two Patterson function from each model, i.e.if peaks overlap in the two maps. Once the orientation of the molecule is determined it is necessary to position the molecule with respect of the crystallographic axes. A translation function is therefore used to find the translation necessary to move the correctly oriented molecule in the unit cell. The solutions are found when the differences between the observed and calculated positions are minimal.

There are a number of different MR methods, all differing in the rotation and translation function used. For instance AMoRe (automated Molecular replacement package, [234]) screens the superposition of the Patterson maps in a particular volume of the unit cell. However AMoRe requires a high homology between the two structures. It also often fails when there are multiple copies of the model inside the unit cell. To overcome this problem, alternative MR methods using different maximum likelihood algorithms were introduced. MR programs such as Refmac [235], MOLREP [236] or Phaser [237]. Maximum likelihood is a statistical method used for fitting a mathematical model to some data. In crystallography it means that the likelihood of the model given the data is defined as the probability of the data given the model. It asserts that the best model on the evidence of the data is the one that explains what has in fact been observed with the highest probability [238]. The introduction of maximum likelihood estimators into the methods of refinement, experimental phasing and, with Phaser, has substantially increased success rates for structure solution [237]. However, in case of DNA, MR methods show their limitations. Elongated molecule or stacked structures are non trivial and are usually approximated with errors.

7.3.5.3.Isomorphous replacement

This method derives the relative phase by comparing the intensities of corresponding Bragg reflections from two or more isomorphous crystals (the crystal possesses the atom arrangement but differs in the atomic number of one or more

atoms). The derivatives are prepared by soaking the crystal of a sample to be analyzed with a heavy atom solution or co-crystallization with the heavy atom. A dataset of the native structure is collected along with at least one heavy-atom derivative. The interpretation of the Patterson difference map reveals the heavy atom's location in the unit cell. This allows both the amplitude and the phase of the atom to be determined. Since the structure factor of the heavy atom derivative (Fph) of the crystal is the vector sum of the lone heavy atom (Fh) and the native crystal (Fp), the phase of the native Fp and Fph vectors can be solved geometrically.

$$F_{ph} = F_p + F_h \qquad \textbf{Equation 53}$$

At least two isomorphous derivatives must be evaluated but the more derivatives available the more accurate the phase determination will be. Multiple Isomorphous replacements are among the most successful methods for phasing.

7.3.5.4.Multiwavelenght Anomalous Dispersion (MAD)

The MAD technique is conceptually similar to the isomorphous replacement methods. It uses only one crystal, which contains anomalous scatterer atoms (Se, Hg, Fe, Pt, U). Anomalous dispersion will occur when the refractive index of the radiation changes when absorbed. So by changing the wavelength of the X-rays, you can change the degree to which the anomalous scatterers perturb the diffraction pattern and therefore there will be a phase change. This phase shift in anomalous scattering leads to a breakdown of Friedel's law ($I_{hkl}= I_{-h-k-l}.\alpha_{hkl}= -\alpha_{hkl}$) which is shown in the complex scattering factor of the anomalous atom (f_{anom}).

$$f_{anom=} f_j + \Delta f_j^{'} + i\Delta f_j^{''} \qquad \textbf{Equation 54}$$

Where $i=\sqrt{-1}$, f_j is the normal scattering factor of atom j and $\Delta f_j^{'}$, $i\Delta f$ are the perturbed scattering factor of the anomalous scatterer j. They will vary with the wavelength of the incident radiation. The Δ contribution of the anomalous scatterer will be determined by collecting data at several wavelengths, two near the absorption edge of the anomalous element in the crystal and one further away. From this, you

can obtain phase information analogous to that obtained from MIR. The MAD technique has become very popular in protein crystallography with the development of tuneable synchrotron sources and protein engeenering (selenomethionine labeling).

7.3.5.5.Fitting and refinement

Whatever the method chosen to obtain the phase and therefore the structure, now the second phase is to build an optimal model from the measured data. Indeed the phase information is usually not accurate, so is the electron density map. As a result, the initial model contains a lot of errors and needs to be modified. The first phase is the fitting process which involves a number of numerical and graphical steps testing the overall agreement between the experimental and calculated data. Basically, it uses software such as Coot [239] or O [240]. First, the initial model is modified to fit the electron density maps. Differential maps such as: (F_{obs}-F_{calc}) and ($2F_{obs}$-F_{calc}) (see equations 49 and 50) are usually compared to graphically fit the first model.

$$\rho_{obs}(xyz) - \rho_{calc}(xyz) \approx \sum_{hkl} (F_{obs} - F_{calc}) \exp(i\phi_{calc}) \exp(2\pi i(hx + ky + lz))$$

Equation 55

Where Φ is the phase F_{obs} and F_{calc} are the observed and calculated structure factors.
The residual map defined by equation 49 indicates how much different is the model compared to the observed structure factors. For instance, a positive peak would definitively imply a lack of atoms in the model and on the other hand a negative peak will sign an atom misplacement in the model. A second map is used to judge the imperfections of the model, it is a linear combination of the electron density map and the previous differential map defined by equation 50.

$$2\rho_{obs}(xyz) - \rho_{calc}(xyz) \approx \sum_{hkl} (2F_{obs} - F_{calc}) \exp(i\phi_{calc}) \exp(2\pi i(hx + ky + lz))$$

Equation 56

An atomic model can never be perfect, but it can be improved a great deal by a process called: refinement, in which the atomic model is adjusted to improve the agreement with the measured data. Basically, the strategy uses the addition of "observations" in the form of restraints, or constraint of the model by reducing the number of parameters. Typical restraints include bond lengths, bond angles and van

der Waals contact distances. They derive from high resolution models and use softwares such as Refmac or SHELX [237, 241]. On the other hand, constraints aimed at reducing the number of parameters. For instance, constraining the model can be made by fixing the torsion angle, as rigid body, or the number of copies present in the asymmetric unit.

7.3.5.6. Validation tools

To validate the model, we need to know the overall accuracy of the model. There are a few common validation tools and other coefficients to guide the search and hopefully find the best solution. The solutions are then compared through the use of correlation coefficient ($CC(F_{obs}, F_{calc})$ defined as:

$$CC(F_{obs}, F_{calc}) = \frac{\left\langle F_{obs} F_{calc} - \left\langle F_{obs} \right\rangle \left\langle F_{calc} \right\rangle \right\rangle}{\sqrt{\left\langle \left\langle F_{obs}{}^2 - \left\langle F_{obs} \right\rangle^2 \right\rangle \left\langle F_{calc}{}^2 - \left\langle F_{calc} \right\rangle^2 \right\rangle \right\rangle}} \qquad \textbf{Equation 57}$$

The correlation coefficient between F_{obs} and F_{calc} must be over 0.9. The correlation coefficient is a measure of the similarity between the structure factors calculated directly from the structural model and one calculated from experimental data. Another popular tool is the R-factor (R_{fact}). Several types of R-factors have been defined [242, 243], all based in the simple principle of comparing measured or theoretical values and calculated variables. It gives a crude measure of the correctness of the structure and data quality. It is defined as:

$$R_{fact} = \frac{\sum_{hkl} \left| F_{obs} - F_{calc} \right|}{\sum_{hkl} F_{obs}} \qquad \textbf{Equation 58}$$

The R-factor for macromolecules should be equal or below 20 % depending on the data quality. Imprecise structures have higher R-factors (over 25 %). At moderate resolutions, it is possible to overfit the data. This problem can be circumvented by introducing an analog of the R-factor, the R_{free}. The R_{free} is based on a test set consisting of a small percentage (usually ~5-10%) of reflections excluded from a

structure refinement.separate a subset of the data (usually around 5%) [244]. It is defined as:

$$R_{free} = \frac{\sum_{free} \left| F_{obs} - F_{calc} \right|}{\sum_{hkl} F_{obs}} \qquad \textbf{Equation 59}$$

If R_{free} drops, then the model must really have improved because there is no pressure to overfit R_{free}. R_{free} is very relevant to keep the model objective in respect with the experimental data. R_{free} values for macromolecules are usually around 25 % to 30 %.

Analysing the quality of the model also involves other types of tools. Those validation coefficients used are model based. For instance, the well known Ramachandran plot [245, 246] for protein is often used to compare average or behaviour pattern from a protein database with the new structure. A typical example, is the program Procheck [247] which checks the stereochemical quality of a protein structure by examining torsion angles between residues and/or side chains. Other methods such as Profile from Eisenberg's group [248] detects if the model does not present anomalies. It basically peforms a 'threading'along the peptide chain to find polar or hydrophobic regions. If a hydrophobic region is solvent exposed, this will result in an unfavourable environments for amino acids and therefore the model must contain errors.

In conclusion the success of an atomic model is often judged through a number of crystallographic indices such as R_{fact}, R_{free}, or whether it satisfies validation methods. However if the accuracy of the model has greatly improved over the last 30 years, the structure determination from other physical methods such as NMR often disagree with X-ray models. There are still a number of biases conferred for instance by the crystallization process, data collection or modelling. At present, improvement must be aimed at converging those methods for structure determination.

7.3.6. Data collection

Data collection is the next step after crystallization. X-ray data collection is a time consuming procedure and is very expensive. A correct prediction of the strategy of data collection is crucial. The crystal needs to be placed in front of a radiation source and a very important part of this process is the geometry of the experiment. Correct alignment of the crystal with controlled rotation allow the completeness on the final data set, improve signal/noise ratio and reduce radiation damage [223].

7.3.6.1.Crystal mounting

Protein Crystallography deals sometimes with crystals about 0.1 mm in size. It must be mounted for data collection on a goniometer. This apparatus allows the crystal to be rotated around different axes without moving away from the beam. Several devices have been engineered to obtain a maximum of degree of freedom. In the past, it was common to mount a crystal in a glass capillary but if this method is still used for room temperature experiment, it needs considerable practice and delicate handling of the crystal. Nowadays, the most common methods use loops in cryo conditions. The crystal is maintained in a cryo-protecting liquor inside the loop and then flash cooled in a stream of nitrogen gas at temperature of 100-120 K.The advantages of the technique are numerous; including reduction of radiation damage to the crystal, and improvement of the resolution shell. Cryo-crystallography has therefore tremendously enhanced the quality and amount of results possible [249]. More recently automatic mounting devices are available for high throughput crystallography. This expands the number of collections that can be made in a relatively short period of time.

7.3.6.2.X-ray sources

Discovered by Röntgen in 1895, X-rays are radiation on the 10 to 0.001 nanometers range of wavelengths (energies in the range 120 eV to 120 keV). X-rays have many well-known applications in medicine (radiography, radiotherapy) but they can also be used to reveal important information in structural sciences since their wavelengths are in the same order of magnitude of bond lengths. X-rays are produced from heavy elements when their electrons make transitions between the lower atomic

energy levels. This occurs when a beam of electrons with sufficient energy knocks out an electron atom's inner K shell. This vacancy will be quickly filled by electrons dropping down from higher levels, emitting X-rays with a sharply defined energy. There are different ways to produce X-rays, some processes used X-ray tubes, rotating anode and other synchrotrons. X-ray tubes or rotating anodes typically involves bombarding a metal target with high speed electrons which have been accelerated by tens to hundreds of kilovolts of potential. The X-ray radiation produces in these cases can be used in crystallography but their intensity limits the resolution shell to the Ångstrom range. For instance, the copper anode which is often used in the studies of macromolecules present a typical wavelength defined by CuK_{α}= 1.5418 Å. The achievable resolution is not sufficient for high resolution structural studies. These X-ray sources have been challenged by synchrotron facilities (please see figure 36). Synchrotron radiation was first observed in 1947 [250] and began to be used as a research tool in the mid-1960s. Synchrotrons are based on the principle that an accelerated electron emits radiation in proportion to its energy. Indeed, when an electron is accelerated close to the light velocity, it emits photons and X-rays.

Figure 36. Photo of the European synchrotron radiation facility situated in the Grenoble, France. The ESRF is a third generation synchrotron and one of the three most powerful synchrotrons in the world.

In synchrotron facilities, electrons are emitted by an electron gun and are first accelerated in a linear accelerator (linac). Afterwards they are transmitted to a circular accelerator (booster synchrotron) where they are accelerated to reach an energy level

of 6 billion electron-volts (6 GeV). These high-energy electrons are then injected into a large storage ring (844 metres in circumference at the ESRF) where they circulate in a vacuum environment. The storage ring is made by a series of magnets separated by straight sections. The bending magnets maintain the electron in their trajectory (circular orbit) and as the electrons are deflected through the magnetic field created by the magnets. They give off electromagnetic radiation that leaves the storage ring

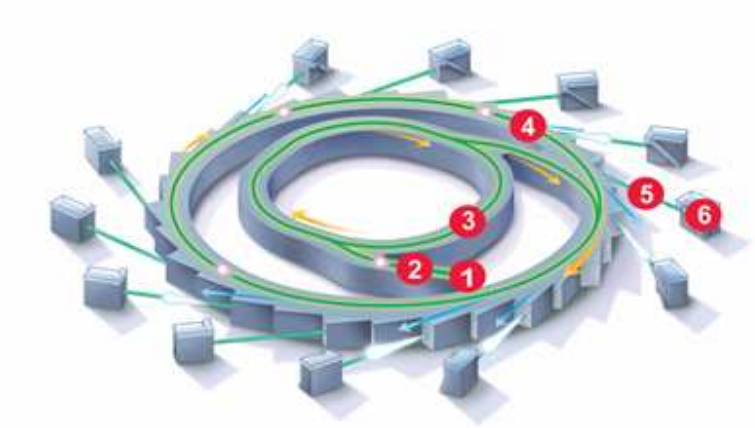


Figure 37. General diagram of a synchrotron. 1 is the electron gun, 2 is the linac, 3 is the booster ring, 4 is the storage ring, 5 the beamline, 6 end station. The synchrotron emits a "synchrotron radiation", especially X-rays; these are sent into the various beamlines (the straight lines branching out of the synchrotron). Each beamline contains scientific instruments, experiments etc. and receives an intense beam of radiation (from http://www.synchrotron.vic.gov.au).

through beam ports in a direction tangential to the ring where experimental hutches have been designed (see figure 37). Because electrons loose energy during their rotation, other type of magnets are found in the straight section of the storage ring. These insertion device magnets create a periodic acceleration. There are two main

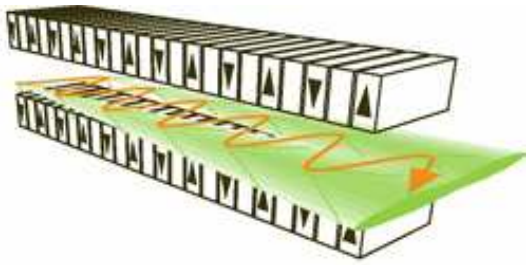


Figure 38. Schematic representation of a wiggler (from http://www.synchrotron.vic.gov.au).

classes of insertion devices namely: wigglers (figure 38) and undulators (figure 39). Wigglers deflect the particle beam from side to side in a ‘wiggling’ way. The radiation emitted is a broad beam of incoherent synchrotron light and can feed several beamlines at the same time. As the cones of light can superimpose on each other the intensity is increased at each bend. Undulators are a multipole wiggler but with a moderate magnetic fields and a large number of poles closer together. In this respect, the poles produce less deflection of the electron beam. This results in a narrow beam of coherent synchrotron light, with certain frequencies amplified by up to 10,000 times.

Figure 39. Schematic representation of an undulator (from http://www.synchrotron.vic.gov.au).

The advantages of synchrotron radiations are multiple:

High brightness: synchrotron light is extremely intense (hundreds of thousands of times more intense than that from conventional x-ray tubes) and highly collimated.

Wide energy spectrum: synchrotron light is emitted with energies ranging from infrared light to hard x-rays.

Tunable: it is possible to obtain an intense beam of any selected wavelength.

Highly polarised: the synchrotron emits highly polarised radiation, which can be linear, circular or elliptical.

Emitted in very short pulses: pulses emitted are typically less than a nano-second (a billionth of a second), enabling time-resolved studies.

Nowadays, synchrotron facilities are in the so called third generation. The main advances made from the first and second generation synchrotron came essentially from bending magnets and insertion devices. Third generation synchrotrons aim at optimising the intensity that can be obtained from insertion devices and uses the three different types of light sources (bending magnets, multipole wigglers and undulators). These third generation synchrotrons have been optimized not only in their architecture but also in terms of sample facility, wet lab at the ESRF, automated crystal loading (see millennium upgrade program at the ESRF), etc. A fourth generation synchrotron exists with intensities four to nine orders of magnitude higher than the current ones but the relevance to crystallography is questionable [251].

7.3.6.3.Detectors

A diffraction pattern presents thousands or even millions of intensities for separated reflections and there is no such thing as a perfect detector: each design has its virtues and problems. Diffractometers have become obsolete due to their poor resolution and sensitivity to radiation damage [252], area detectors have gained popularity in X-ray crystallography. There are four common types of area detectors:

Film-based detectors
Image plate
Charged Coupled Device (CCD)
Multi-wire detectors

Film based detectors used photographic film but even if the resolution is very high, they have become obsolete due to their cost and time consuming efforts. Image plates have replaced them. Image plates are made of a photostimulable phosphor matrix. When the plate is exposed to X-rays the image is temporarily stored in the phosphor coating. After exposure the image is read out by the laser head which scans across the plate radially as the plate rotates. The plate may be immediately erased and re-used. There are different image plate designs depending on the size, the shape, the laser used, etc. Image plates present a major disadvantage since the immediate processing of the diffraction spots is not possible. The introduction of gas-filled

ionization chambers (multi-wire detectors) or image intensifiers (video based or Charged Coupled devices-CCD) have overcome this problem. A CCD detector is composed of a phosphor screen (up to thousand mm^2) that converts X-rays into visible light. The phosphor screen is usually coupled to the CCD with a tapered optical fiber. A CCD chip collects the electron charge image. This electric charge is read out and digitized in order to transfer to the computer. The pixel size is about 10-50 µm with readout times of less than 1 s. CCD detectors generally show high signal/noise ratio and a good dynamic range. They are currently the "standard" detector technology in macromolecular crystallography. However they are expensive, too small and require correction for image distortions. Improved CCD array systems are being made: better phosphors, reduction of electronic noise, increased multiplexing of parallel readout (to increase speed), and increased numbers of modules (to increase size).

7.4. Material and methods

7.4.1. Protein perdeuteration, purification and enzyme assay

The same procedure as described in chapter 2.4.2 was applied to obtain hydrogenated *Bs* DHFR and 6.3.1 for fully perdeuterated *Bs* DHFR. Specific activities of *Bacillus stearothermophilus* dihydrofolate reductase were measured continuously by following the decrease in absorbance at 340 nm using a molar extinction coefficient of ε_{340}= 12300 M^{-1} L cm^{-1} [116]. The standard assay mixtures contain sodium phosphate buffer 100 mM at pH 8, at a fixed concentration of NADPH (100 µM) and H_2F (50 µM). The final enzyme concentration was typically ~ 0.3 nM. The reaction was initiated by the addition of 1 µL of enzyme to the reaction mixture (469 µL), pre-incubated 2 min at a temperature of 30 °C. Assays were run for 60 seconds using a thermospectronic helios γ-spectrophotometer interfaced with the Vision™ 32 software (version 1.25, Unicam Ltd.). Reaction progress curves were recorded in triplicate and repeated if the data deviate by more than 10 %. One unit of enzyme is defined as that amount of the enzyme that catalyzes the conversion of 1 micro mole of substrate per minute under the previous assay conditions.

7.4.2. Protein crystallization

The MTX-*Bs* DHFR binary complex was crystallized using the hanging drop method vapor diffusion method at 291 K and an automatic system for crystallization using a robot to set up 96-well crystallization plates. Initial crystallization trials were performed using 0.2 M Tris (pH 7.4) and 1.75 M ammonium sulfate. As MTX is not soluble at high concentrations, the protein and the drug were incubated for several minutes at relatively low concentration 1mg/ml with a protein/MTX ratio of 1/2 at room temperature. After complexation the protein concentration was increased up to 15 mg/ml with a Centricon YM10 membrane filtration device (Millipore, MA, USA). Crystals were grown by mixing equal volumes (10 μL) of the protein/ MTX complex and reservoir buffer. Crystal appeared in two days and grew to full size within a week. Before X-ray exposure, a cryoprotecting solution was made by adding glycerol to the mother liquor to a final concentration of 20 %. Crystals were mounted onto loops and frozen in liquid nitrogen prior to data collection.

7.4.3. X-ray data collection and structure refinement

Diffraction data were collected on the ID14-EH1 beamline of the ESRF, Grenoble, France which is equipped with a two crystal monochromator and a Quantum 210 CCD detector. A wavelength of 0.934 Å was used together with oscillations of 0.5 ˚. Crystal to detector distance was set at 1024 mm. Crystals were cryocooled in a stream of nitrogen to minimize radiation damage (100 K). A total of 360 frames were recorded. Data were indexed, integrated and scaled with Mosflm software package on the basis of a centered orthorhombic space group I_{222} with the following dimension a= 123.82 Å, b= 156.23 Å, c= 179.53 Å with α =β =γ = 90˚. Molecular replacement was performed using the maximum likelihood software Phaser. A molecular solution was found using *Bs* DHFR (PDB entry 1ZDR) as a search model, with non conserved residues mutated to alanine. There are 8 molecules in the asymmetric unit corresponding to a solvent content of 57.6 %. Molecular packing of alternative solutions was checked manually using the program Coot based on $2F_{obs}-F_{calc}$ and $F_{obs}-F_{calc}$ electron density maps. Solvent building was done using the program ARP/wARP [253] from the CCP4 package. 20 ARP/wARP refinement

cycles were run in total. Atoms were removed if below 1.0 sigma in $2F_{obs}-F_{calc}$ map and atoms added if above 3 sigmas in $F_{obs}-F_{calc}$ map.

7.5. Results

7.5.1. Protein perdeuteration

7.5.1.1. Production

The aim of this work was to produce reliably and economically a fully deuterated form of *Bs* DHFR. In order to achieve this goal the preparation of deuterium resistant bacterial strain is necessary and performed by a multi-stage adaptation process by plating on a solid deuterated minimal media and repeated liquid deuterium cell culture (with increasing D_2O content). When the strain is adapted to the growth in deuterated media with glycerol as the sole source of carbon, the second step is to obtained a high cell density culture (HCDC). The 'batch fed batch' technique is a suitable method to control the growth rate of the adapted strain and to minimize inhibition from the by products of the HCDC and obtain a large amount of recombinant protein. The growth rate is generally lower than in hydrogenated media but the final yield in terms of biomass will be essentially the same. HCDCs are generally obtained in bioreactor also called fermenters. These present a number of advantages, noticeably in their ability to control simultaneously pH, oxygen pressure, temperature, feeding. In figure 40, the fermentation profile of the *E. coli* cells expressing the *Bs* DHFR gene is given. The different phases of the 'batch fed batch' culture are delimited and the respective parameters such as OD_{600}, PO_2, pH, are shown. The inoculation of the culture media marks the beginning of the batch culture. A short latency phase is observed with a diminishing PO_2, lower pH, and increased stirring. During this phase the cells are using the glycerol present in the initial media and will reach their maximum growth rate (exponential). OD_{600} follows an exponential law. The second phase, 'fed batch', is initiated when the PO_2 and stirring reach a critical value. It indicates a change in the metabolism of the cell, the carbon source glycerol becomes limiting, and the cells start to exploit other compounds such as lactate or acetate. To avoid a stationary phase, glycerol is added to the culture media during the 'fed batch'. During this step the growth rate is controlled and remains linear. This step is sub-optimal but limits the formation of toxic by products.

The last phase induced by the addition of IPTG initiates the protein expression. The glycerol feeding is stopped as the protein production is not greatly limited by the energy demands of the cells. The biomass does not increase significantly which shows that cell division has ceased. From 50 grams of d_8 glycerol (99 % deuterated glycerol) 45 grams of cell were obtained. Perdeuterated *Bs* DHFR was readily expressed, yielding approximately 8 mg of protein per gram of cell paste. In comparison, the unlabeled enzyme is produced at 10 mg of protein per gram of cell paste. The two-step purification protocol consisting of: heat precipitation and cation exchange chromatography yielded typically 150 mg of pure protein/L of culture. As implied by by table 10, the protein purification factor is low ~ 4 which indicates a relatively high level of enzyme in the cell supernatant compared with other proteins. The purified protein is pure as assessed by SDS-PAGE electrophoresis (a single band of approximately 19 kDa in lane 1, figure 41).

7.5.1.2.Characterization by MALDI-TOF

MALDI-TOF mass spectra indicate (see figure 42) that the molecular weight (M_w) of unlabeled protein is 18697 Da and perdeuterated protein 19740 Da in H_2O (figure 43). For reference, the theoretical molecular weight of unlabeled *Bs* DHFR is 18695 Da [56] and that of perdeuterated *Bs* DHFR is 19724 Da (including only non-exchangeable). The theoretical number of exchangeable hydrogens is 1315

The extent of deuteration can be determined using the following equation:

$$\%\ Deuteration = \frac{\left(M_W\ Per_{Exp} - M_W\ Hydr_{Exp}\right)}{\left(M_W\ Per_{Theo} - M_W Hydr_{Theo}\right)} \times 100 \qquad \textbf{Equation 59}$$

Where $M_W\ Per_{Exp}$ is the experimental molecular weight of the perdeuterated DHFR, $Hydr_{Exp}$ is the experimental molecular weight of the hydrogenated DHFR, $M_W\ Per_{Theo}$ is the theoretical molecular weight of the perdeuterated DHFR and $M_W\ Hydr_{Theo}$ is the theoretical molecular weight of the hydrogenated DHFR. From these data, the extent of perdeuteration is 101 % indicating that not all exchangeable deuterons are substituted in the protein sample when soaked in H_2O. A similar observation was made by Di Costanzo *et al* [180] where the deuterium enrichment exceeded the calculated value (105 %).

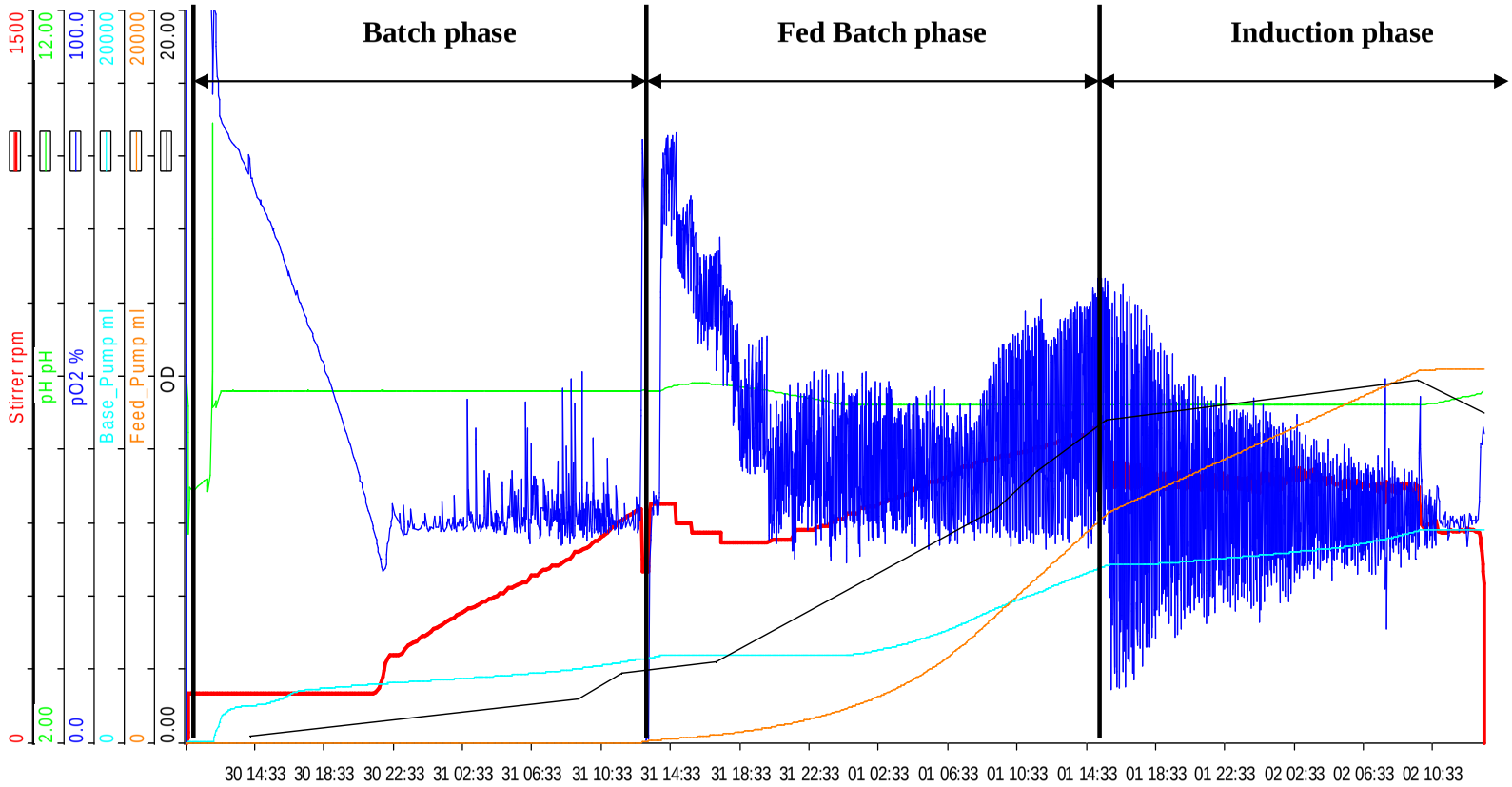


Figure 40. Fermentation profile of *E. coli* cells in deuterated Enfors minimal media. The pattern is decomposed in three distinct phases: The batch phase where the growth rate is maximum; the fed batch phase which marks the addition of glycerol to the growth media and finally the induction phase where the protein is over-expressed. The X axis corresponds to time.

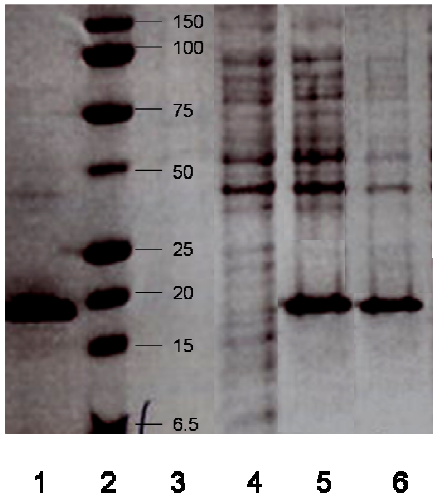


Figure 41. SDS-PAGE gel of *Bs* DHFR purification. (1) is the collected fractions of pure *Bs* DHFR, (2) and (3) is the molecular weight ladder and its correspondence in kDa. (4) is a cell extract before induction, (5) is a cell extract after induction and (6) is the heat treated supernatant.

	Protein			Enzyme			
Step	Fraction volume (mL)	Protein concentration (mg/mL)	Total (mg)	Specific activity (units/mg)	Total (units)	Yield (%)	Purification factor (fold)
Cell extract	200	7.5	1500	1.2×10^6	1.9×10^9	100	1
Heat denaturation	200	4.5	900	1.9×10^6	1.7×10^9	90	1.5
Anion exchange	100	2.8	280	4.6×10^6	1.3×10^9	68	3.8

Table 10. *Bs* DHFR purification table summary. The cell mass used was 30 grams. One unit of enzyme is defined as that amount of the enzyme that catalyzes the conversion of 1 micro mole of substrate per minute under our assay conditions.

The back exchange between hydrogen and deuterium is dynamic process, largely favorable, occurring spontaneously in a well-hydrogenated environment. However, some deuterons are so deeply embedded in the protein structure that their solvent access is very limited. To test this possibility pre-incubation of the perdeuterated protein in mild denaturing conditions would certainly help to exchange those deuterons [177]. Nevertheless, our data indicate that the protein is deuterated at an appropriate level and therefore suitable for crystallization and X-ray analysis.

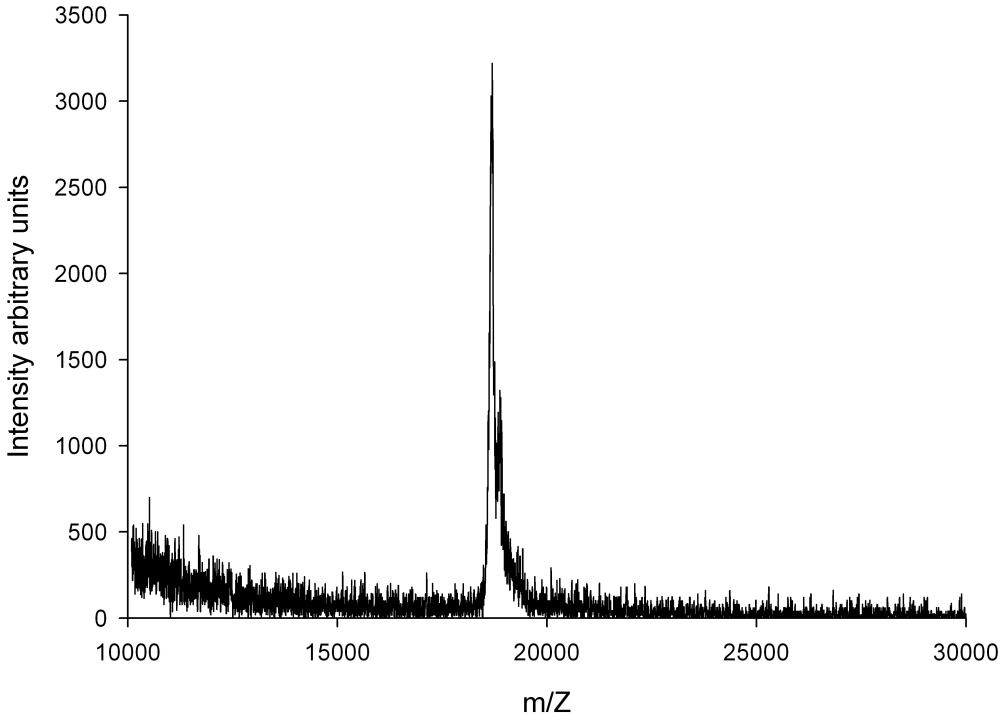


Figure 42. MALDI mass spectra of hydrogenated *Bs* DHFR. The main peak corresponds to a *m/Z* of 18697.

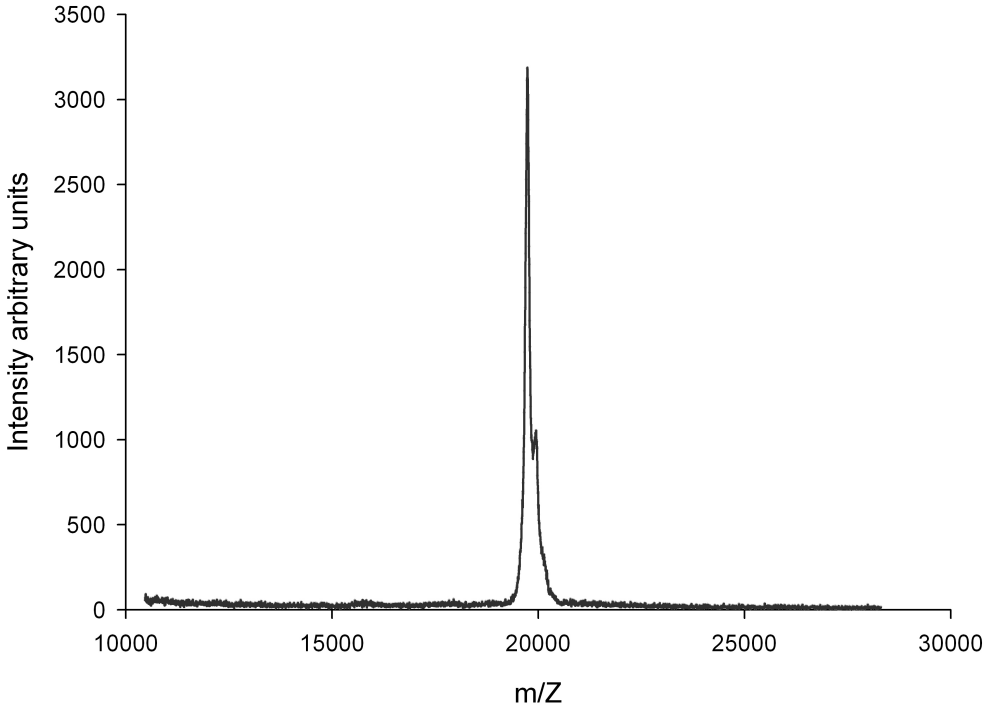


Figure 43. MALDI mass spectra of perdeuterated *Bs* DHFR. The main peak corresponds to a *m/Z* of 19740.

7.5.2. Crystallization

A previous crystallographic study has determined the structure of the free enzyme [56]. We have therefore crystallized the MTX-*Bs* DHFR complex over relatively similar conditions (figure 44). Crystal obtained in A and B were relatively large, for instance, using the conditions described in (B), the approximate volume is 0.4x0.7x0.2=0.06 mm^3. However we found that the most appropriate crystals for crystal diffraction were obtained using a solution containing 0.1 M HEPES; pH = 7.0 ; 1.6 M $(NH_4)_2SO_4$. These crystals were the most resistant to the cryo-protecting process. Indeed, the crystals in (A) and (B) tend to crack when the cryo-protecting solution was added to the mother liquor. This is a relatively common consequence when the crystal volume is large. A second problem comes from the presence of $(NH_4)_2SO_4$ in the precipitating solution. When flashed cooled, $(NH_4)_2SO_4$ salt crystals grow rapidly impairing the data collection. To circumvent these problems, we have therefore worked with the relatively small crystals shown in (C).

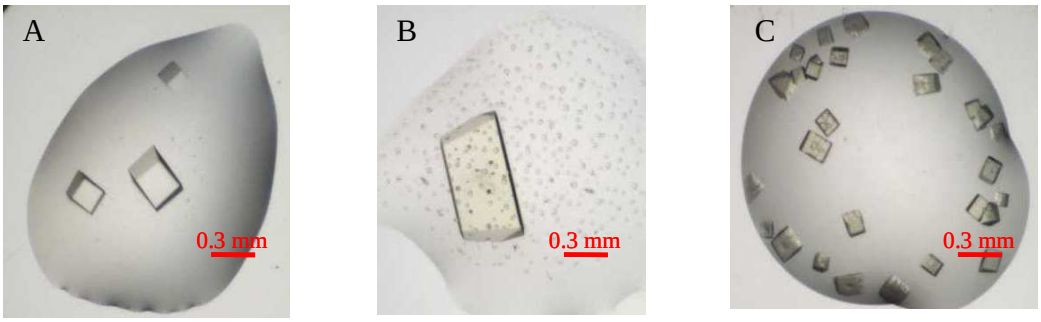


Figure 44. Crystals of *Bs* DHFR. Crystals shown in (A) were obtained using 0.1 M NaCl ; 0.1 M HEPES ; pH = 7.5 ; 1.6 M $(NH_4)_2SO_4$, in (B) the precipitating solution used was 1.0 M $(NH_4)_2SO_4$; 0.1 M HEPES ; pH = 7.0 ; 0.5 %w/v PEG 8000 and crystals presented in (C) were grown using 0.1 M HEPES ; pH = 7.0 ; 1.6 M $(NH_4)_2SO_4$. The bar gives the scale in mm.

7.5.3. Preliminary X-ray structure

7.5.3.1.Data processing

After X-ray exposure, we have obtained via the CCD detector a diffraction pattern (figure 45). This map, made of hundred of spots, contains all the information to solve the structure of the MTX-*Bs* DHFR complex. The current diffraction pattern

show well defined spot, not overlapping. Mosaicity is a useful indice to examine the degree of long-range order of the unit cell within a crystal. The mosaicity is 0.434 which indicates a well behaved crystal (mosaicity is generally between 0.2 for well diffracting crystals to 1.5 for badly behaving crystals such as thin plates, or damage by cryo cooling). The resolution range achievable according to this image is expected to be not less than 2.5 Å since the spots are going beyond the most outward circle corresponding to a resolution shell of 2.5 Å. The set of native data is therefore promising for structure determination. However, this two-dimensional map needs to be converted into a three dimensional model. The strategy used during data collection was based on the rotation oscillation method. As a result the full data set consists of 360 separate images taken at different orientations of the crystal.

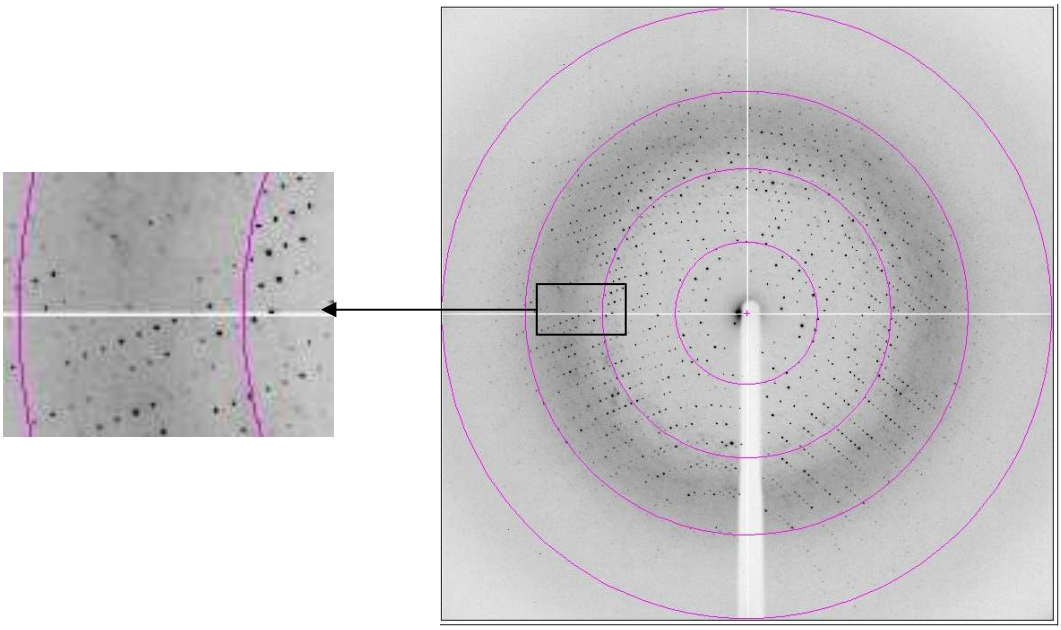


Figure 45. Diffraction pattern of MTX-*Bs* DHFR crystal on the beamline. The purple circles give the relative resolution shells: 9.9, 4.9, 3.3, 2.5 Å (from the most inner circle to outer shell respectively).

The goal is now to convert these raw data into a list of Bragg reflections with measured intensities. This process called data processing can be broken down into three sections:

(1) Data Indexing : determination of crystal orientation, cell parameters, possible space group

(2) Integrating the data: define the reflection list from the image sets

(3) Scaling and merging the data

We have performed these steps using MOSFLM [254] and SCALA [255] . Integrating the data show that the space group of crystallization was I_{222} (centred orthorhombic) with the following cell parameters a= 123.82 Å, b=156.23 Å, c= 179.53 Å with α=β=γ=90˚. The Matthew's coefficient was calculated with the following relationship:

$$V_M = \frac{V_{unitcell}}{MW \times Z \times N} \qquad \textbf{Equation 60}$$

Where V_M is the Matthew's coefficient, M_W the molecular weight of *Bs* DHFR (18695 Da), Z the number of asymmetric units per unit cell (8 for space group I_{222}) and N is the number of molecule per asymmetric unit. V_M gives the most probable number of molecules in the asymmetric unit and usually falls into the range of 1.6-2.5 $Å^3$ [256]. Using a cell volume of 3472899.9 $Å^3$ the number of molecules per asymmetric unit was 10 corresponding to a V_M of 2.90 $Å^3$ Da^{-1} with a solvent percentage of 57.6 %.

After indexing and integration of the intensities from the 360 images, the data set was merged and scaled together. The summary of the data is given in table 11. The data corresponds to 158956 unique reflections. The resolution limits range from the highest shell: 1.66 to 65.94 Å to the lower shell 7.44 to 65.94 **Å**. Probably the best criterion to describe the quality of the structure is the resolution limit of the data. In order to define this limit a number of factors have been defined, such as the ratio of recorded intensity and its standard deviation I/σ(I), the completeness, R-factors and data redundancy. As presented earlier in 7.3.5.6, R-factors are the traditional overall measures of quality. They measure the relative agreement between equivalent observations. There are a number of R-factors. Merging R factor arises from the averaging of multiple measurement of reflection of the same (hkl) and of symmetry-related reflections. R_{merge} is defined as:

$$R_{merge}(I) = \frac{\sum_{hkl}\sum_{i}\left|I_i(hkl) - \overline{I(hkl)}\right|}{\sum_{hkl}\sum_{i} I_i(hkl)} \qquad \textbf{Equation 61}$$

Note that R_{sym} is often used in literature instead of R_{merge}. R_{sym} is a subset of R_{merge} since it is measure of how well symmetry related to reflection merge.

Another calculator than can be helpful in data quality assessment is the so-called precision-indicating merging R factor $R_{p.i.m.}$ [257]. It describes the precision of the average measurement.

$$R_{p.i.m.} = \frac{\sum_{hkl} [1/(N-1)]^{1/2} \sum_{i} \left| I_i(hkl) - I(\overline{hkl}) \right|}{\sum_{hkl} \sum_{i} I_i(hkl)} \qquad \textbf{Equation 62}$$

$R_{p.i.m}$ gets better (smaller) with increasing multiplicity. It estimates the precision of the merged <I>.

R factor: R_{meas} [258] is the subsitute of R_{merge} and describes the precision of the individual measurements independent on how often a given reflection has been measured.

$$R_{meas} = \frac{\sum_{hkl} [N/(N-1)]^{1/2} \sum_{i} \left| I_i(hkl) - I(\overline{hkl}) \right|}{\sum_{hkl} \sum_{i} I_i(hkl)} \qquad \textbf{Equation 63}$$

A second quantity is the completeness. The completeness of a data set at a given orientation is defined as the ratio of the number of unique reflections measured on the total of unique reflection. In practise completeness should be equal to, or over 95 % for the electron density calculation [225]. The inner shell completeness of our data set is 99.7 % for a resolution range of 7.44 and 65.94 Å. For the highest resolution shell, 1.66-7.44 Å, it is 78.4% and 8.9 % at 1.66-1.71 Å. For these later values, the number of unique reflection is too low and therefore a truncation of data is needed for accurate model building. Symmetry related reflections are usually more accurate in defining the overall quality of the data set. The information given by the different geometrical conditions is preferred over repeated measurement of the same reflection. The factor used is called redundancy or multiplicity of intensity measurement. The overall multiplicity 5.6 and 6.5 for the overall and inner shell respectively, but falls to 1.8 for the outer shell. If the first two values are well

behaved, the third value is too low and corroborates the finding enumerated earlier. The building of the electron density should exclude the highest resolution shell. Within table 12, we are able to assess in a satisfactory way data quality according to the value of R-factors, multiplicity, and completeness. D_{min} is defined as the minima distance at which two features in the electron density map can be resolved. The completeness is dropping with D_{min} below 1.92 Å. A decrease in completeness clearly demonstrates a deterioration of the model parameters [257, 259]. It indicates that the data must be truncated before this threshold.

	Overall	**InnerShell**	**OuterShell**
Low resolution limit (Å)	65.94	65.94	1.71
High resolution limit (Å)	1.66	7.44	1.66
R_{merge}	0.108	0.043	1.831
R_{meas} (within I+/I-)	0.119	0.046	2.554
R_{meas} (all I+/I-)	0.119	0.046	2.554
$R_{p.i.m.}$ (within I+/I-)	0.048	0.018	1.773
$R_{p.i.m.}$ (all I+ & I-)	0.048	0.018	1.773
Fractional partial bias	-0.034	-0.024	-0.387
Total number of observations	897723	15993	2358
Total number unique	158956	2445	1317
Mean ((I)/σ(I))	9.0	37.9	0.3
Completeness (%)	78.4	99.7	8.9
Multiplicity	5.6	6.5	1.8

Table 11. Summary of data statistic processed in space group I_{222} (centred orthorhombic).

D_{min} (Å)	Number of reflection	Completeness (%)	multiplicity	R_{meas}	R_{sym}	$R_{p.i.m.}$
7.44	15993	99.7	6.5	0.046	0.043	0.018
5.26	30998	100.0	7.2	0.061	0.056	0.023
4.30	40374	100.0	7.4	0.064	0.060	0.023
3.72	47745	100.0	7.4	0.068	0.063	0.025
3.33	54389	100.0	7.5	0.076	0.071	0.028
3.04	60114	100.0	7.5	0.089	0.083	0.032
2.81	65199	100.0	7.5	0.115	0.107	0.042
2.63	69953	100.0	7.5	0.164	0.153	0.060
2.48	74609	100.0	7.5	0.210	0.196	0.076
2.35	78477	100.0	7.5	0.281	0.261	0.102
2.24	82734	100.0	7.5	0.362	0.337	0.132
2.15	73909	100.0	6.4	0.448	0.411	0.175
2.06	54496	100.0	4.6	0.573	0.506	0.264
1.99	43694	96.5	3.6	0.766	0.651	0.392
1.92	34681	87.1	3.1	0.951	0.782	0.527
1.86	26800	74.0	2.7	1.126	0.898	0.661
1.80	19891	59.2	2.5	1.292	0.998	0.803
1.75	13579	43.4	2.2	1.444	1.081	0.942
1.71	7730	26.5	2.0	1.671	1.218	1.135
1.66	2358	8.9	1.8	2.554	1.831	1.773
Overall	897723	78.4	5.6	0.119	0.108	0.048

Table 12. Reflection statistics for the data used in the refinement. D_{min} is the nominal resolution.

7.5.3.2.Model and refinement

Structure solution search used Molecular Replacement in space group I_{222} (no solutions were found in space group I_{212121}) with the program Phaser [237]. The model for the refinement was based on the structure of *Bs* DHFR deposited in the protein data bank (accession code 1ZDR). The search was set with 1 monomer of the DHFR from *Bacillus stearothermophilus* with protein atoms only (no solvent molecules) and a resolution range of 40-4 Å. Phaser was run in the "automated molecular replacement mode" by searching for 10 solutions (10 monomers) in the asymmetric unit (according to most probable number as given by the Matthew's coefficient).

Solution number	RFZ	TFZ	PAK	LLG
1	4.0	10.3	0	74
2	4.0	19.5	0	262
3	5.2	18.5	0	534
4	4.4	21.7	0	845
5	3.4	24.6	0	1236
6	3.6	26.4	0	1694
7	4.0	23.8	0	2062
8	3.0	23.2	0	2375

Table 13. Summary of the Phaser runs.

PAK= Number of C-alpha atoms that can clash within 2 Å.
TFZ= Translation Function Z-score
RFZ= Rotation Function Z-score
LLG= Log Likelihood Gain

8 solutions were found as indicated by table 13. With MR, the problem is to define which solution is correct. The ranking of the various solutions is assessed by various coefficients such as TFZ, LLG or packing arrangement. The first possibility to find the solution is to look at the packing of the top 8 solutions. This is a very discriminating criterion since it discards physically clashing models. Indeed, if one solution is well detached (see TFZ score or LLG values), it is essential to examine carefully the packing arrangement either graphically or through the summary of the logfile. Here no clashes were found as given by the zero values presented in table 13. This is therefore very encouraging for generating the correct solution. Using Phaser, other commonly used criteria are the Z-score and LLG. The log likelihood gain is the probability that the data would have been measured, given the model. It allows us to measure how much better the data can be predicted with your model than with a random distribution of the same atoms. The LLG should always be positive, and it should increase as the solution progresses. Our data clearly demonstrate an increasing LLG when new components are added to the solution. It peaks at 2375 for solution 8. Additionally, the TFZ (the translation function Z-score) for the correct solution is generally (number of standard deviations above the mean value) over 5 and well separated from the rest of the solutions. The *Z* score for the final translation function is the highest: 23.2 and therefore it reinforces the probability that the correct solution

has been found. For, the rotation function, the correct solution is within a Z-score under 4. As these criteria are fulfilled, it indicates that the solution number 8 is the correct solution.

7.5.3.3.Model building

Since the phases are known, the next step is to build the model from the current available electron density. A crucial point here is that the initial phase estimates are not of the best quality which results in a relatively incorrect electron density map. The model which can be built from such maps may be in parts incorrect or incomplete. It requires extra rounds of refinement combined with model rebuilding. Model building and refinement are therefore a phase improvement procedure. The protocol and softwares are chosen in order to optimize the parameters for the building of the macromolecular model. For a detailed comparison of some popular model building package see reference [260]. In our case, we have used the software ARP/wARP from the CCP4 package. ARP/wARP links model building and refinement in a general view [253]. Basically, it processes iteratively towards the model by optimizing the phases through the manipulation of atoms coordinates. The procedure is based on the concept that the electron density is made of a collection of free atoms. Free atoms account for the electrons without the need of applying chemical knowledge or distance restraints between them. The originality and power of ARP/wARP comes from the way it establishes the atoms connectivity. It proceeds iteratively through snapshots, positioning atoms according to the electron density and then building blocks when stereochemical informations (inter-atomic distances) and restraints are satisfied with confidence. The parts that are less certain are still allowed and can be filled at any point of the procedure. The generated ‘hybrid model’ which is a mixture of free atoms and fragment of the protein structure is refined using the maximum likelihood program REFMAC5 [235].

To monitor the refinement progress over the various refinement cycles, the crystallographic R factor defined previously in equation 52 is very useful. It measures the fit between the model and the observed data. In addition to the standard R factor, the so-called free R value is used as a cross-validating indicator to monitor the overall

progress and avoid ‘over’ fitting [244]. In table 14, the data used for the refinement are listed.

Resolution limits	42.000-2.000 Å
Number of used reflections	110893
Percentage observed	99.7360
Percentage of free reflections	5.0175
Overall R factor	0.4130
Free R factor	0.4172
Overall correlation coefficient	0.7515
Free correlation coefficient	0.7442
Table 14. Data statistics using REFMAC 5.	

20 refinement cycles were run in total. Further refinement using REFMAC in each ARP/wARP cycle. The starting model after REFMAC gives a R_{fact} = 0.413 and R_{free} = 0.417. The first cycle shows an improvement of both parameters, R_{fact} = 0.362 and R_{free} = 0.413. After 20 cycles a final R_{fact} of 0.254 and R_{free} of 0.318 were obtained. The refinement in modern crystallography aim at minimizing both the standard R factor and the R_{free} value however these factors are often biased according to the chosen protocol and sometimes do not give the model that best describes the data [226]. The R factor progression and its cross validation are not sufficiently sensitive to indicate whether or not a lateral chain is well positioned. The correct conformation need a graphical checking and therefore depends on user creativity. Preliminary pictures of the model are shown in figure 46. They were generated using the graphical interface Coot [239] and Pymol [54]. Coot, the crystallographic object oriented toolkit is a CCP4 module for model building, completion and validation.

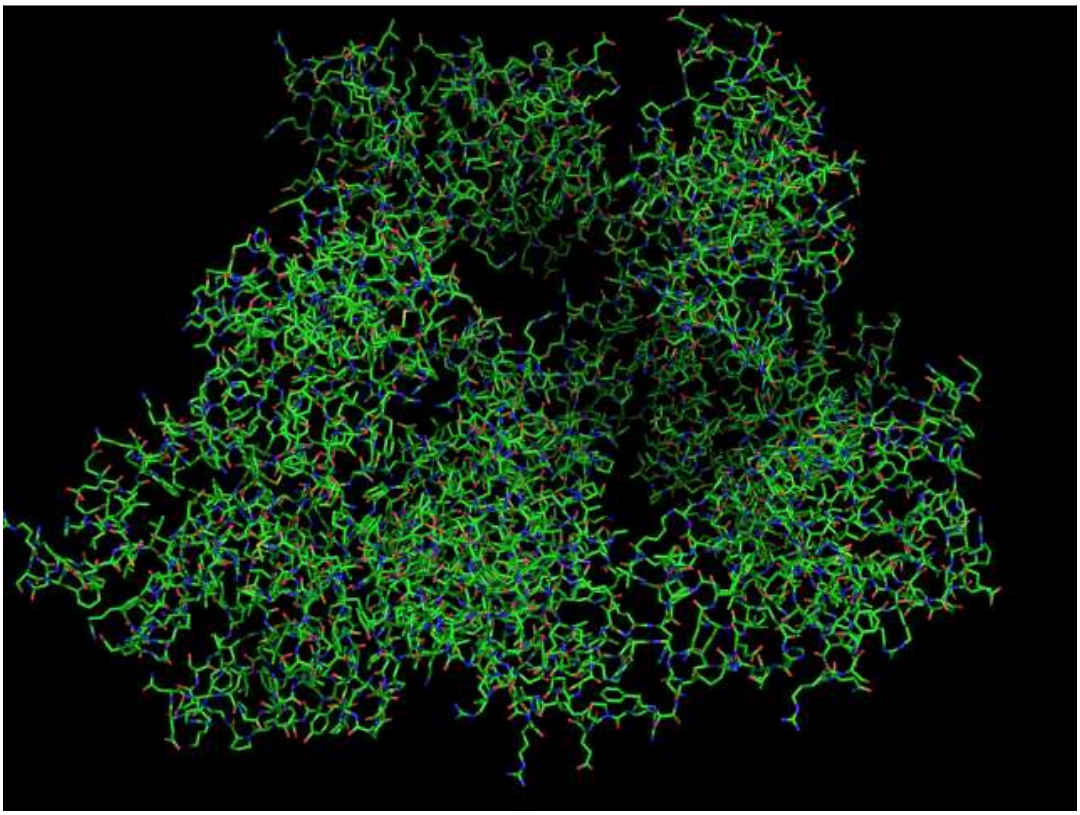


Figure 46. Content of the asymmetric unit. The eight molecules are represented in stick. The picture was made using Pymol . The asymmetric unit contains 8 molecules corresponding to eight monomers. Solvent molecules are omitted to provide a better overview.

Overall, the electronic density map of the MTX-*Bs* DHFR complexe is well defined and therefore gives a solid base for model building. Upon initial inspection, the electronic density shows in all monomers a significant nuclear density in the vicinity of the active site (positive peak in green color). Using the CCP4 library for small molecules, this density was fitted with the atomic coordinate of MTX. The result for chain C is presented in figure 47. It clearly outlines without ambiguity the presence of

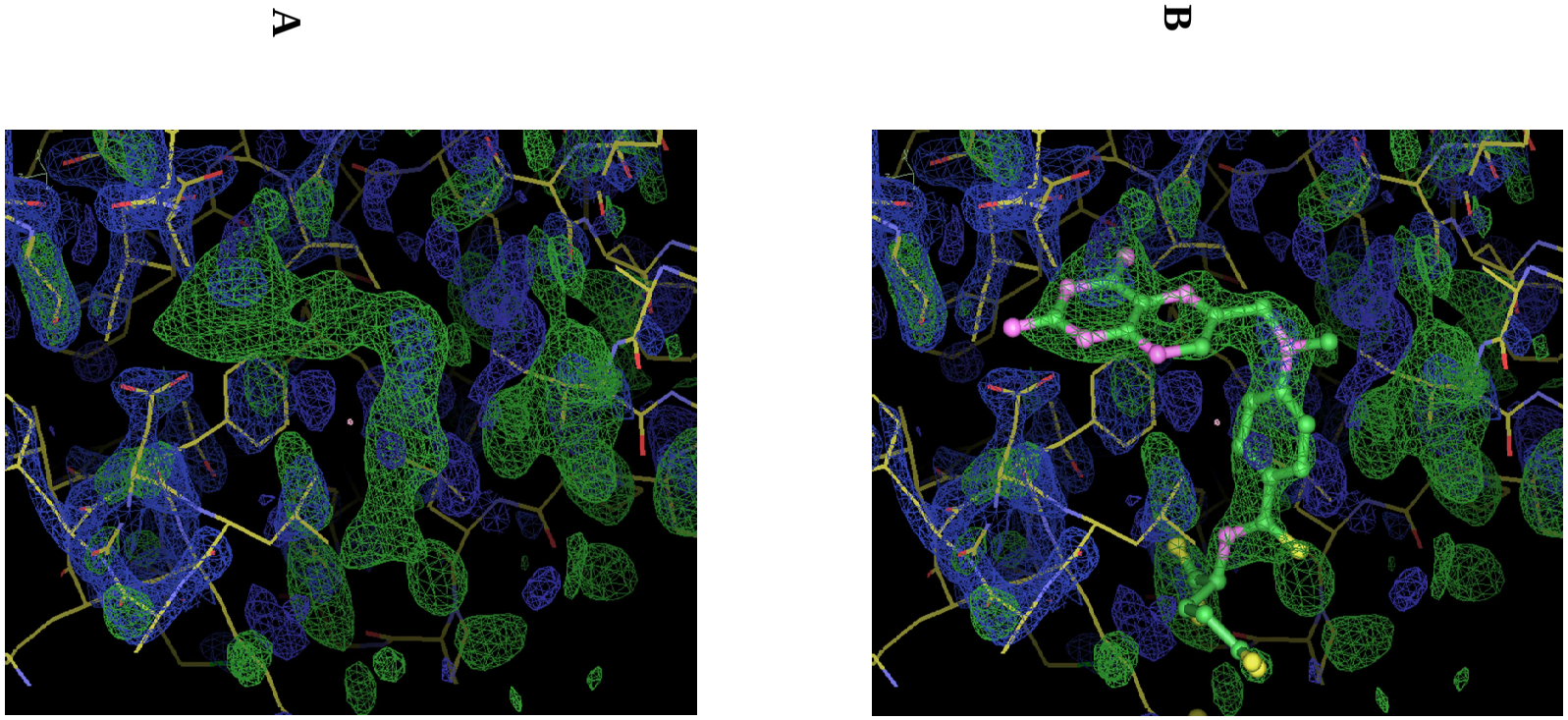


Figure 47. Positive difference maps reveal putative MTX position in the *Bs* DHFR active site. Shown in green is the F_{obs}-F_{calc} electron density map contoured at 3 σ and in blue is the $2F_{obs}$-F_{calc} electron density map contoured at 2.0 σ. In A, the positive density in green is not occupied. In B, this gap is filled with a MTX molecule using Coot [239].

MTX. Preliminary examination of the interaction of Asp27 and MTX at the *Bs* DFR active site confirms the common pattern observed (figure 48). The N1 atom of MTX is well positioned to interact with Oδ2 of Asp27. Although the precise features of the interaction can not be accurately highlighted, these results are very encouraging.

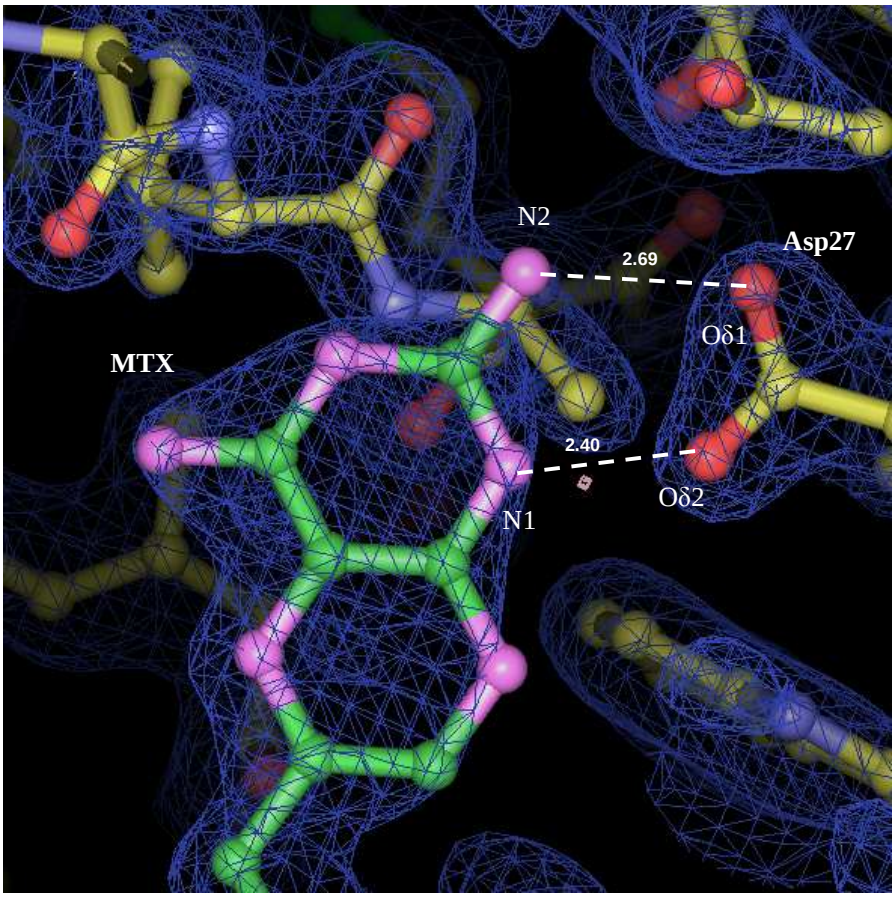


Figure 48. Electron density maps for the MTX pteridine ring and the DHFR active. Shown in blue is the $2F_{obs}$-F_{calc} electron density map contoured at 2.0 σ. The distance are given in Å and were computed using Coot [239].

The reconstruction of the model was therefore pushed further and figure 49 was generated. As commonly observed with other bacterial DHFR, these models possess the original structure consisting of a pseudo Rossman fold [51, 261] of alternating β-strands (8, named βA- βH) and α-helices (4, named αB,C, E, F) with three main loop regions.

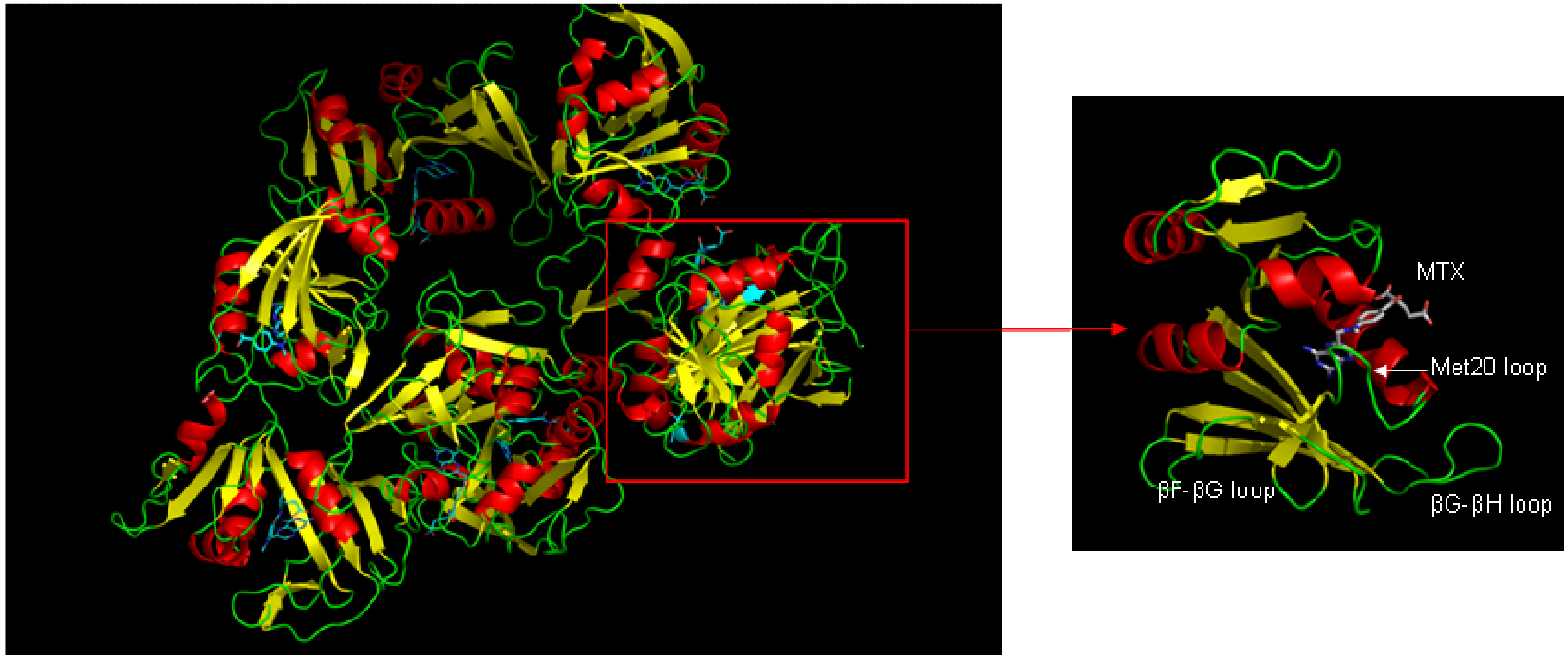


Figure 49. Secondary structure representation of the asymmetric unit content. Secondary structures in red correspond to α-helix, β-trands and loops are colored in yellow and green respectively. MTX molecules are colored in blue. In the inset, Chain C is magnified with a MTX molecule.

The next step in the analysis would certainly assess any possible difference between the 8 monomers and the description of the MTX binding interaction with *Bs* DHFR. At this stage it is possible to look at these hpotheseses but it is a preliminary analysis. The main difference between the monomers as discussed in [73, 262], often comes from the loop conformations adopting the closed and partially occluded conformations. A preliminary inspection of the Met20 loop region show potential differences in the loop conformation (see figure 50). However the differences are weak and need to be compared with the other available data from various DHFR sources.

Figure 50. Comparison of the eight monomers in the asymmetric unit (ribbon representation). On the left, the eight monomers are superposed. On the right, the MTX binding site in the vicinity of the Met20 loop region is magnified. This picture was made using Pymol [55].

7.6. Discussion

The primary goal of this work was to determine the X-ray structure the MTX-*Bs* DHFR complexe. The work presented here is incomplete, but we expect these preliminary results to lead to a final model, which after careful analysis will be deposited in the Protein Data Bank (PDB) [263]. The structure validation is crucial.

Grossly incorrect structures have been published according to the R factors and should be avoided with the progress of validation tools [226, 264]. The validation of macromolecular structures are quite standardized now [265]. The PDB's automated deposition system used PROCHECK [247] for checking bonds, angles, dihedrals and possible atom clashes. The SFCHECK [266] validates the structure factors. The number of parameters generates an important number of possibilities which must be carefully checked computationally and visually. The handling of this information is a problem in terms of time and computing power. Human intuition and experience is indispensable in validating atomic models. This work described here has not been completed in a satisfactory fashion since the lack of resources (financial, computing, experience) have drastically limited the author progress. The primary perspectives of this research are fairly simple as stated previously: further refinement and validation of the model before structure deposition. The restrained refinement is ongoing with REFMAC 5. The refinement is primarily graphical using the difference maps and aim at correcting rotamer conformations and water molecules positioning.

The model is at medium resolution between 2.0 and 2.4 Å and would not be suitable for determining hydrogen positions. However as presented in chapter 7.1 the ambiguity of hydrogen and water positions could be achieved by neutron crystallography. The X–ray model would serve as a model for the neutron structure building. Furthermore, the long term perspectives of this project is to provide a precise description of the energetics involved in the MTX binding. The cross-correlation of the ITC data from chapter 3.3 within the structural information from X-ray and neutron crystallography would definitively shed the light on the binding mechanism and its molecular basis.

8.

8. PEGylation of *Bs* DHFR

8.1. Introduction

The early goals of my project aimed at investigating the effect of small ligand molecules on fast protein dynamics. As described previously we have initiated this research project with the inhibitor MTX and the model protein DFHR. However, it was clear that projects based on a similar approach could be of interest too. *In vivo*, proteins are often modified and conjugated with non peptidic substance through process such as acylation [267]; methylation [268]; phosphorylation [269]. The properties of the conjugate are often drastically modified. In this respect we have tested the possibility of using a non specific, covalently bound molecule to perturb simultaneously; activity, stability and dynamics. However various classes of molecules can fulfilled these criteria since any hydrophilic polymer such as dextran, polysaccharides [270], or even insoluble matrices can be linked to a protein [271]. However one molecule, namely PEG: Poly-ethylene glycol receives our attention. It medical use over the last twenty years has made PEGylation a standard technique in protein modification [272]. The following chapter will deal with the production and purification of a PEGylated form of our protein model *Bs* DHFR. Two experimental objectives will be presented, first the investigation of PEG dynamics when coupled to *Bs* DHFR by neutron scattering and the solution structure of the conjugate by Small Angle Neutron Scattering (SANS).

8.2.The pegnology science

The technique aiming at coupling polymers to proteins originated in the 1950s and 1960s through 'gentle chemistry' but the 'origin of pegnology' came later, in the late 1970s in the laboratory of Professor Frank Davis of Rutgers University and also with the work of Professor Abuchowsky [272]. F. Davis qualified PEGylation in its early stage as a 'library shelves quest'. The initial problem was to provide the benefits of non-human bioactive proteins to humans. The first two key papers dealt with

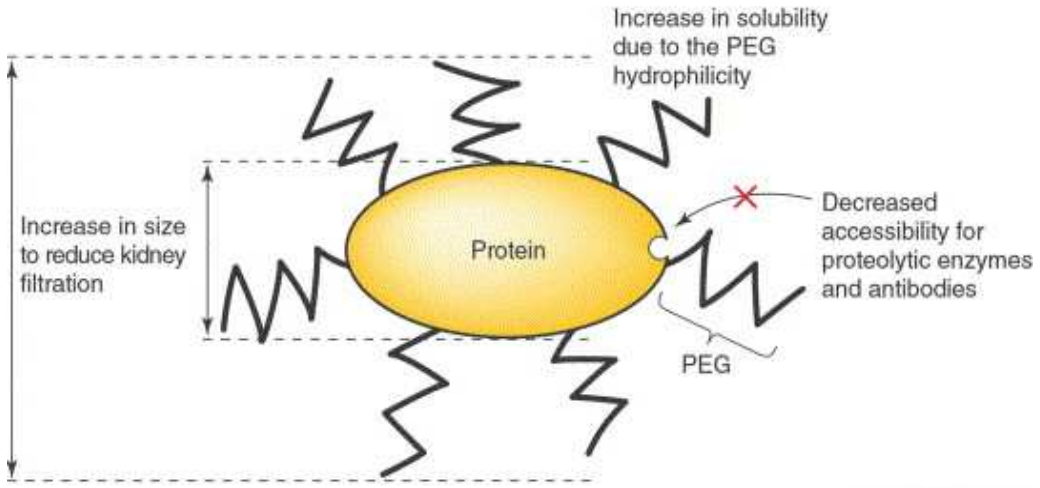


Figure 51. Schematical representation of a PEGylated protein. The main advantages of the conjugate are outlined. The figure was reproduced from [273].

catalase [274] and albumin [275]. These pioneering steps have made PEGylation a choice technique in medicine with a tremendous commercial power. It is well known that proteins can be used as therapeutic agents such as insulin or interferon [276]. However, the direct use of protein in the body presents a number of disadvantages. Protein are stable under certain conditions, a slight change in pH or ionic strength may have a dramatic effect on the protein bioactivity. The circulating body life is also a major drawback in protein pharmaceutical. As a general rule, the protein life span in blood circulation is very limited. Serum concentrations of interferons decay quickly during the initial 24 hours. The half life ($t_{1/2}$) of interferon-α-2a is 2.3 hours and 2.1 min for interleukin-6 [277]. These short-lived actions are the results of a number of effective clearance mechanisms such as glomerular filtration, proteolysis or

phagocytose (see figure 51). The action of protein therapeutic implies therefore repetitive doses, which increases, not only the amount of drug needed and treatment costs, but may also cause severe secondary effects for the patient [273, 277, 278]. PEGylation is a challenging strategy in drug delivery, which has the potential to overcome these problems. Among all the advantages provided by the PEGylation, there are: a decrease in toxicity, immunogenicity and increase in body life time circulation. For instance the half time circulating life of a PEGylated form of interferon-α-2a is 50 hours instead of 2.3 hours originally [277]. The pharmacokinetic profile of drugs can be profoundly reformulated. The direct outcomes clearly lead to the optimization of blood circulating life; drug elimination and time distribution of therapeutics [279].

Brand Name	Product	Company	Indication
PEGasys	PEG-IFN α-2a (interferon)	Hoffmann – La Roche	Hepatitis
PEG-Intron	PEG-IFN α-2b (interferon)	Enzon	Hepatitis
Neulasta	PEG-filgrastim (granulocyte colony-stimulating factor)	Amgen	Neutropenia
Adagen	PEG-adenosine deaminase	Enzon	Immuno-deficiency
Oncaspar	PEG-aspargase (asparaginase)	Enzon	Cancer
Somavert	PEG-visomant (growth hormone)	Pfizer	Acromegaly
PEG-Hirudin	PEG-recombinant hirudin	Abbot	Thrombosis (phase III)
PEG-monoclonal antibody	PEG-CDP 870	Pfizer	Rheumatoid Arthritis (phase III)
PEG-Axokine	PEG-ciliary neurotrophic Factor	Regeneron	Obesity (phase III)
Pre-clinical testing	PEG-erythropoietin	Prolong Phamaceuticals	Anemia

Table 15. Example of clinical applications of PEGylated protein therapeutics. The table was taken from [280] based on www.biospace.gov and www.fda.gov.

Direct applications in medicine (table 15) are illustrated with a number of commercially and FDA (Food and Drug Administration) approved drugs. Interferons, which are routinely used in the treatment of hepatitis C, are an example of the clinical challenges involved in protein PEGylation. One of the direct clinical application of interferons is PEGASYS® (Roche), which is a PEGylated form of interferon α-2a with a 40 kDa branched PEG [281]. Other drugs have been designed such as Exubera® from Pfizer, which consists of an inhaled form of insulin to cure patients suffering of type 1 and 2 diabetes [282, 283]. These examples illustrate the value of PEGylation as a methodology for drug delivery.

8.3. Polyethylene glycol's properties

PEG molecules are made of the repetition of ethylene oxide units. This polyether diol, highly soluble, amphiphilic is relatively chemically inert. PEGs are prepared by anionic polymerization of ethylene oxide, providing a variety of molecular sizes. PEGs are named based on the number of ethylene oxide units in the polymer chain. PEG is a viscous liquid at molecular weight below 1000 Da and solid at higher molecular weights [278]. PEGs can be categorized in two main classes: linear and branched polymers (figure 52). In addition PEGs can be divided into two groups: PEGs with free hydroxyl (-OH) groups at both ends and PEGs with one or two methoxylated end group(s) (-OH replaced by $-OCH_3$).

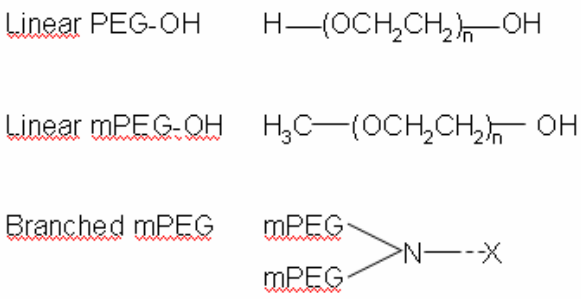


Figure 52. Typical molecular structures of poly-ethylene glycols (PEGs). The structure of "Y-shaped" branched PEG is shown at the bottom where X is the functional or active group [278].

If figure 52 presents the ‘raw’ materials used in PEGylation, PEG molecules can not react directly with proteins, the molecule needs to be ‘activated’. The basis of the activation is to modify at least one of the free hydroxyl groups into a protein reactive end.

8.4. The PEGylation reaction

Basically, the PEGylation reaction is to conjugate the protein to PEG molecule(s) but this task is not straightforward [284]. For instance, the first activated PEG molecules suffer from a number of disadvantages. Since PEG is obtained by chemical synthesis, the PEG batch yielded was polydispersed (polydispersity is the ratio between the observed molecular weight (Mw) of the PEG molecules and their calculated value according to the number-average molecular weight (Mn)) in terms of molecular weights and the choice of PEG chain lengths was restricted. Nowadays with the development of synthetic and purification procedures polydipersity has decreased and activated PEG molecules over a large range of molecular weights (2 kDa up to 60 kDa) with original architecture (branched PEG, comb-shaped copolymer PEG) [285] are available. Another crucial drawback was the chemistry behind the PEGylation reaction. Protein amino groups were the first entities to be PEGylated by acylation or alkylation but these procedures presents a lack of specificity, stability and reproducibility [286]. A number of PEGylation procedures are now at hand. Various protocol using enzymatic [287] or chemical procedures offers a wide variety of ways to conjugate a specific target protein [284, 286]. The choice of the appropriate functional group for the PEG derivative is based on the type of available reactive group on the molecule that will be coupled to. For proteins, typical reactive amino acids include lysine, cysteine, histidine, arginine, aspartic acid, glutamic acid, serine, threonine, tyrosine. The N-terminal amino group and the C-terminal carboxylic acid can be also used. With the help of chemistry this large variety of reactive groups has increased the spectrum of PEGylation [284]. Site-specific PEGylation is therefore possible, when for instance free cysteines are available naturally or via protein engineering. Site specific binding avoids the number of isomers obtained, which made PEGylated product purification easier and also their

FDA approval. The choice and the guidance toward a specific reactive PEG are therefore more flexible and have increased the potential of this technology. PEGylation chemistry is well documented now, details of the conjugation methods can be found in [284, 286]. Additionally, a number of companies offer activated PEG for research and commercial use. To review the important catalogue of available PEG, consult the NOF website (www.peg-drug.com) which is one of the leading companies in this growing and important market. Continuous research efforts from the manufacturers have led to the so-called second generation PEG chemistry. It utilizes branched PEG with a central protein conjugation functional group surrounded by two or more identical methoxy-PEG chains. However it seems that multi-arm activated PEG are challenged by bi or heterofunctional PEG or even comb-shaped copolymers leading ultimately to a third generation of PEG (www.peg-drug.com).

8.4.1. PEGylation at cysteine residue

For proteins that possess thiols suitable for PEG binding, PEGylation can be very specific and reversible [286]. In fact, cysteines are rarely present in proteins and the limited cross reactivity with other reactive groups present in proteins made PEGylation at thiol group very relevant. A number of thiol PEGylating agents are presented in table 16. Among these, PEG-orthopyridyl-disulfide reacts with thiol groups forming a stable symmetric disulfide [288]. Another agent, PEG-maleimide based on the well known Michael reaction takes advantage of thiol addition to the activated double bond [289]. PEG-vinylsulfone reacts slowly with thiols to form a stable thioether linkage to the protein. PEG-iodoacetamide reacts with free thiols by nucleophilic substitution, creating a stable thioether linkage. All these derivatives targeting specifically thiol groups have their own advantages and disadvantages [288, 290, 291]. For instance, PEG-vinylsulfone is stable in an aqueous environment but may react with lysine at elevated pH. PEG-orthopyridyl-disulfide reacts specifically with sulfhydryl groups under both acidic and basic conditions (pH 3-10) to form a stable bond but in reducing environment the linkage is readily converted to thiols. The thioether bond between PEG-maleimide and the protein is stable but slow cleavage of the amide linkage can occur in aqueous solution. Our choice led us to use

PEG maleimide because of its price and availability in a number of molecular weights. It has been reported that a relatively stable linkage with a good specificity toward sulfur groups.

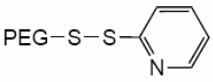
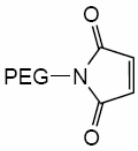
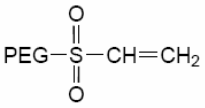
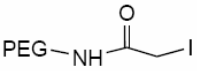
Structure	Thioreactive PEGs	Properties
	PEG–pyridildisulphide	The most specific towards thiol but yields a cleavable linkage by a reducing agent also *in vivo*.
	PEG–maleimide	Gives stable linkage by double bond addition but can also react with amines at pH >8.
	PEG–vinylsulfone	
	PEG–iodo acetamide	Less reactive, not much used

Table 16. PEGs reactives towards thiol group from [273].

8.5. STRUCTURAL CONFORMATION OF PEGYLATED PROTEIN SPECIES

The tremendous value of PEGylated therapeutic proteins is unquestionable however, it is inversely related to the amount of data available on the molecular basis of its biological effectiveness. The effects of PEGylation are unpredictable and their molecular basis and dynamic properties are unclear. It is believed that the improved properties of the PEG-protein conjugates derive mainly from the PEG molecule [280, 292, 293]. Significantly, PEG molecules are hypothesised to act *in vivo* as a 'molecular shield' [273, 294], therefore the spatial conformation and dynamical behaviour of PEG molecules seem crucial (figure 51). *In vivo* branched PEG-protein conjugates have been found to be more effective than their linear counterparts with an equivalent total molecular weight [295]. The longer circulation half-life of the branched PEG was attributed to a more efficient masking mechanism of the protein surface but also to a larger effective size [296] or hydrodynamic volume of the

branched PEGylated protein [295]. The structure details of the PEG molecules when conjugated to a protein are therefore the key stones in the elucidation of PEG bioeffectiveness. X-ray structural analysis shows that free PEG molecules can assume at least two distinct conformations according to the chain length: first the zigzag or random coil structure for short chains and secondly a helical structure for longer chains [297]. These conformations are interchangeable in water and depend on the solution conditions. Another intriguing property of PEG is its solubility. It is either soluble in aqueous or organic solvent. The PEG solubility properties are extensively used in protein crystallography. PEG is a choice precipitating agent. In fact, the presence of the ether oxygen atom allows the hydration of the carbon chains through the formation of the oxonium ion (The oxonium ion in chemistry is any positive oxygen cation, which has three bonds). This specific chemistry makes PEG chains heavily hydrated. In aqueous solution PEG monomers can established up to three water bridges [298]. The large exclusion volume is believed to drive the overall hydrodynamic properties of the PEG conjugate and reinforce the concept of 'molecular shield'. These structural information are useful but until now the atomic structure of a PEG protein conjugate is still not known: does the (extremely mobile) PEG chain shroud the protein, or does it sit adjacent to the protein as an unperturbed random coil? In what way, if any, does the conformation change when the conjugate approaches other molecules, substrates, binding sequences or surfaces?

As stated previously, the effect of PEG polymer on protein structure and the three-dimensional structure of the PEG-protein conjugate are unknown. The relatively large, highly flexible PEG polymer conjugate impedes crystallization and structural determination by protein crystallography. However two models have been suggested (figure 53).

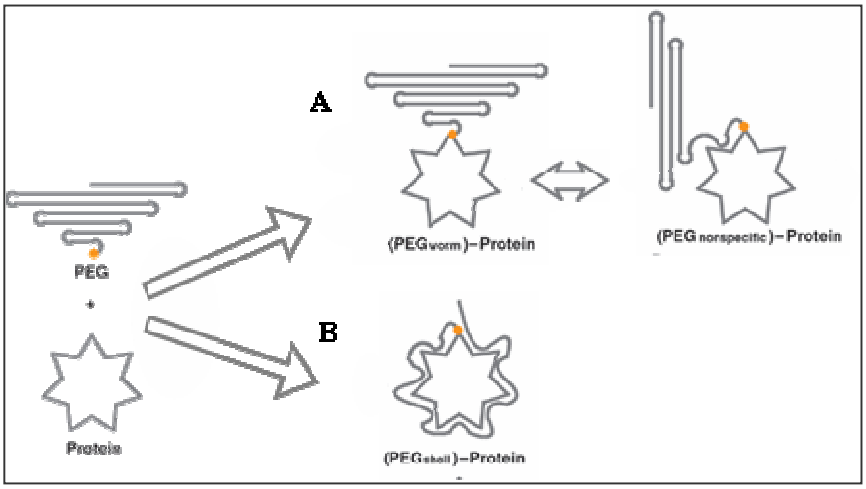


Figure 53. Schematical representation of potential structures of a PEG-protein conjugate. In A, the worm-like model is in equilibrium with an intermediate model where PEG interacts locally with the protein. The B model illustrates the shell-like structure. The PEG molecule wraps and interacts more thoroughly with the protein. This picture was modified and reproduced from [280].

The first model supports a shell like structure where water molecules solvate hydrophilic regions around the protein whereas hydrophobic patches interacts with hydrophobic PEG clusters. According to this view, PEG molecules are therefore wrapping the protein. There is also an alternative model in which there is no interaction between the PEG and the protein. The conjugate forms a worm like helical structure that fluctuates freely in solution. A few studies have approached this issue. Using Size Exclusion Chromatography (SEC) [299] Fee *et al* have suggested that the viscosity radius of a PEGylated protein depends only on the molecular weight of the native protein and the total weight of grafted PEG. They also suggested that PEG molecules form a dynamic layer over the surface of the protein. In conclusion, the PEG-protein hybrid can adopt various conformations but the stabilising interactions are unknown. On the other hand SAXS (Small Angle X-ray Scattering) and analytical centrifugation have given the first lines of evidences of PEGylated protein solution structure [292, 293]. A model of the PEGylated structure of hemoglobin was proposed by Svergun *et al* using an *ab initio* method [292]. The proposed model clearly outlines no change in the tertiary structure of the protein and a compaction of the quaternary structure. Additionally, PEG molecules were found to interact with the

core of the protein, cavities between subunits and surface to form either a dual-mushroom or a dual-mushroom to brush. The transition between the two PEG conformations was proposed to depend on the distance between the grafting sites (D_G) and the Flory dimension (R_F) of PEG chain length.

$$R_F = aN^{3/5} \qquad \textbf{Equation 63}$$

Where N is the number of units per polymer, *a* the effective length of an oxyethylene unit (*a*=3.5 Å) [300]. Basically, the brush conformation is the predominant form when $D_G < R_F$ whereas the mushroom like structure occurs at $D_G > R_F$. The PEG conformation depends primarily of the chain length but also on the available space between them. When sterical hindrance is high, the PEG chain adopts an elongated form, brush like, whereas a non crowded area favours the folding of the PEG chain on itself. This study has unambiguously brought some evidence of the three-dimensional structure of PEGylated conjugate, however SAXS is a low resolution technique; model building is not accurate to determine the architecture of the complex at atomic resolution. Moreover, the PEG chains used in this study were relatively small, 5 kDa, in comparison with PEG molecules used in medicine. Generally PEGs with relatively high molecular weight (≥ 30 kDa) cause a sharp increase in the *in vivo* circulation half-life [301]. Lower molecular PEG may not have a relevance in medicine since their size would not be high enough to avoid kidney elimination [302].

Our first goal is to address this problem by means of Small Angle Neutron Scattering (SANS) which allows a direct measurement radius of gyration (R_g). With the help of contrast variation, we could extract specifically the PEG and or protein structure, which is not possible by SAXS. It will provide a detailed map of the three dimensional conformations of PEG molecules when bonded to a protein. The analysis will highlight the basis and the nature of the interactions that maintain and govern the PEG/protein complex. Additionally, these results will be implemented with an *ab initio* modelling in the hope of building a more thorough understanding of the molecular basis of PEG biological effectiveness.

Another goal is the dynamical behaviour of PEG molecules when grafted to a protein. A recent study has pointed out that a difference in molecular flexibility between linear and branched PEGylated protein could explain differences in

clearance rates by glomerular filtration [303]. The restricted dynamical behaviour of branched PEG would reduce the ability of PEG chain to deform and cross glomerular pores. If this hypothesis turns out to be correct, molecular flexibility is a milestone in PEG bio-effectiveness. Our second goal in this study is to investigate the effect of PEGylation on PEG dynamics. We will use the technique of neutron scattering to follow nano and picosecond dynamics at a sub-Ångstrom scale. As a subsidiary benefit, we hope to test the effect of PEG molecules on protein dynamics at a later stage.

The following chapters will describe the production and characterization of a PEGylated form of *Bs* DHFR.

8.6. Materials and Methods

8.6.1. PEGylation reaction

Hydrogenated *Bs* DHFR was purified as described previously in section 2.4.2 and fully deuterated *Bs* DHFR as described in section 6.3.1. Purified *Bs* DHFR protein was diluted to a final concentration of 6 µM in 100 mM sodium phosphate buffer at pH 7.5, 0.5 mM EDTA, 0.2 M NaCl and reduced by incubation with a 25-fold molar excess of dithiothreitol (DTT) for 1 hour at room temperature. The reduced mixture was then desalted by chromatography using a Hi-Prep G-25 26/10 desalting column (GE healthcare, Uppsala, Sweden) in 100 mM sodium phosphate buffer at pH 8, 0.5 mM EDTA, 0.2 M NaCl to remove the DTT. The PEGylation reaction was initiated by the addition of 6 µM *Bs* DHFR to a 5-fold molar excess of PEG maleimide (NOF corporation, Tokyo, Japan) reagent in 100 mM sodium phosphate at pH 8, and 0.5 mM EDTA, 0.2 M NaCl for 2 hours at room temperature under gentle stirring. PEGylated *Bs* DHFR was separated from unPEGylated protein and unreacted PEG by ion exchange chromatography using a customed MacroCap SP column (GE, Healthcare, Uppsala, Sweden) equilibrated in 50 mM MES pH 6. The bound protein was eluted using a linear gradient of 1 molar NaCl. Column fractions containing the PEGylated *Bs* DHFR protein were identified by non-reducing SDS−PAGE and freeze-dried for 24 hours in the dark. This method was used to

prepare *Bs* DHFR modified with a 5 and 12.5 kDa maleimide PEG (NOF corporation, Tokyo, Japan).

8.6.2. Protein determination

Protein concentration was estimated with a modified Bradford method [304] using the Bio-Rad protein assay kit (Bio-Rad, Hercules, United States). The micro assay procedure was followed to determine protein concentration. The method consists of mixing 0.8 mL of the the unknown protein solution concentration (ranging from 1 µg/mL to 10 µg/mL) with 0.2 mL of Bio-Rad reagent. The mix is let to stand for at least 15 min and no more than an hour. Absorbance is read at 595 nm and the protein concentration is determined against the protein standard curve. Samples are made in duplicate. The protein standard curve is made by varying concentrations of stock standard bovine serum albumin solution from 1 to 10 µg/mL. The Bio-Rad reagent was purchased from Bio-Rad (Hercules, United States).

8.6.3. Enzyme assays

The same procedure as described in chapter 2.4.2 was applied to obtain hydrogenated *Bs* DHFR and 6.3.1 for fully perdeuterated *Bs* DHFR. Specific activities of *Bacillus stearothermophilus* dihydrofolate reductase were measured continuously by following the decrease in absorbance at 340 nm using a molar extinction coefficient of ε_{340}= 12300 M^{-1} L cm^{-1} [116]. The standard assay mixtures contains sodium phosphate buffer 100 mM at pH 8, at a fixed concentration of NADPH (100 µM) and H_2F (50 µM). The final enzyme concentration was typically ~ 0.3 nM. The reaction was initiated by the addition of 1 µL of enzyme to the reaction mixture (469 µL), pre-incubated 2 min at a temperature of 30 °C. Assays were run for 60 seconds using a thermospectronic helios γ-spectrophotometer interfaced with the Vision™ 32 software (version 1.25, Unicam Ltd.). Reaction progress curves were recorded in triplicate and repeated if the data deviate by more than 10 %. One unit of enzyme is defined as that amount of the enzyme that catalyzes the conversion of 1 micro mole of substrate per minute under the previous assay conditions.

8.6.4. SDS-PAGE electrophoresis

The same procedure was applied as described in section 2.4.3 except that no reductive agent was present

8.6.5. Neutron scattering sample preparation

PEGylated conjugates of fully perdeuterated *Bs* DHFR with PEG-maleimide of 5 kDa and 12.5 kDa were synthesized, purified and characterised using the same procedure described in section 8.6.3. To minimize the contribution of hydrogen scattering, the labile hydrogen atoms from the PEGylated *Bs* DHFR were exchanged with deuterium oxide (D_2O). PEGylated perdeuterated *Bs* DHFR was dissolved in D_2O (10 mg/mL), gently stirred 12 hours at 4 °C and then freeze-dried. The operation was repeated two more times with higher grade D_2O (purity 99.9 %), freeze-dried, and then stored at 4 °C until use. Activity was checked and less than 10 % of the original activity was lost during the process. After lyophilisation, the PEGylated perdeuterated *Bs* DHFR preparations (100 mg) were dried over phosphorus pentoxide. To reach a higher hydration level, the proteins were equilibrated in a desiccator over a saturated solution of NaBr (which gives a relative humidity of 50 % at 20 °C). After exposure at this relative humidity (for two days), the D_2O content was shown to be 20 %, (i.e.25 mg of D_2O per 100 mg of dry weight protein). At this hydration level the sample holder was quickly sealed and stored at 4° C until use. The NaBr salt was rotary evaporated before use to near dryness in order to remove as much as possible any remaining water molecules and then dissolved in high grade D_2O (99.9 %).

8.6.6. Small angle neutron scattering preparation

PEGylated conjugates of fully deuterated *Bs* DHFR as described previously were dissolved at 2.5 mg/mL in a series of solvent contrast 0, 25, 50, 75 and 100 % ^2H_2O:H_2O ratio. Quartz cuvettes of 1mm were used for measurements at room temperature. Data were collected using the D22 diffractometer at the ILL (with two detector distances covering a Q range of 0.01–0.25 $Å^{-1}$).

8.6.7. Size Exclusion chromatography

Chromatographic experiments were carried out on a AKTA Explorer 100™ liquid chromatography system with protein and PEG-protein conjugate detected by UV absorbance at 280 nm (GE Healthcare Technologies, Uppsala, Sweden). The chromatography system was controlled and data analysed using the Unicorn software, version 4.1 (GE Healthcare Technology, Uppsala, Sweden). Tricorn™ 10/30 GL Superdex 200 pre-packed SEC columns were obtained from GE Healthcare Technologies (Uppsala, Sweden). 40 μL samples were injected onto each size exclusion column, using a 100 mM sodium phosphate buffer pH 7 with a flow rate of 1 mL/min. Size-exclusion column was calibrated against a High Molecular Weight (HMW) gel filtration calibration kits containing ovalbumin (43 kDa), conalbumin (75 kDa), aldolase (158 kDa), ferritin (440 kDa) and thyroglobulin (669 kDa), obtained from GE Healthcare Technologies (Uppsala, Sweden). Blue Dextran (2000 kDa) was purchased from Sigma (Saint Louis, MO, USA). The extraparticle void volume, V_0, was measured by the elution volume of Blue Dextran marker. K_{av} values were calculated for each sample using equation 64.

$$K_{av} = \frac{(V_e - V_0)}{(V_c - V_0)} \qquad \textbf{Equation 64}$$

Where V_e is the elution volume of the sample and V_c is the total column liquid volume, V_0 the column void volume.

8.6.8. Mass spectrometry

Molecular weights of PEGylated conjugates were determined using an Autoflex II MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). Sample were dissolved in deionised water and mixed 1:1 with 5 mg/mL dihydroxybenzoic acid matrix (Bruker Daltonics, Bremen, Germany) in 2:1 0.1 % of trifluoroacetic acid/acetonitril. 1 μL of sample was loaded onto an AnchorChip™ target plate and allowed to dry at room temperature. A protein Calibration standard I (Brucker Daltonics, Bremen, Germany) kit was mixed with matrix in the same manner and 0.5 μL was loaded on the calibration spots on the target. Spectra were collected in linear detection mode as described previously in section 6.3.3.

8.7. Results

8.7.1. Purification and analysis of PEGylated *Bs* DHFR

The primary structure of *Bs* DHFR presents only one cysteine in position 73. After inspection of the three-dimensional structure, it was hypothesised this single cysteine could potentially react with the PEG-maleimide following the reaction described in figure 54.

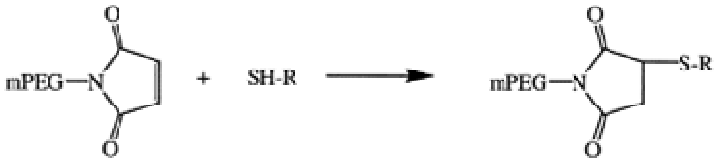


Figure 54. Schematical representation of the reaction between PEG-maleimide (on the left) and a protein thiol group (SH-R) from [286].

The purification of mono-PEGylated *Bs* DHFR was performed using ion exchange chromatography with a specifically designed resin: MacroCap SP (GE healthcare, Uppsala, Sweden) in a single run. The chromatogram shows three main peaks. The unbound material (figure 55, peak A) was identified as unreacted PEG. The SDS-PAGE analysis of this fraction did not reveal any band (figure 56, lane 6) when stained by Coomassie blue. Indeed PEG molecules do not absorb at 280 nm but maleimide groups do, therefore it is not surprising to have a UV signal at 280 nm. As expected the first peak (figure 55, peak B) obtained during elution corresponded to a mono-PEGylated form of *Bs* DHFR (lanes 4 and 5, figure 56) whereas the second peak (figure 55, peak C) was unambiguously free *Bs* DHFR as assesed by the SDS-PAGE (lanes 1, 2, 3, figure 56). This selective elution was reported in a number of studies [305, 306]. This specificity of PEG molecules to modify the physical and chemical properties of the protein conjugate such as the iso-electeric point, pKa or hydrophobicity is clear but the mechanism is not completely understood. Indeed, it is believed that PEG molecule when grafted to a protein modify these properties [280] but the interactions remain complex to assess [307]. The MacroCap SP resin, which is a cation exchanger, is designed to purify PEGylated proteins with high recovery yields (GE, healthcare information).

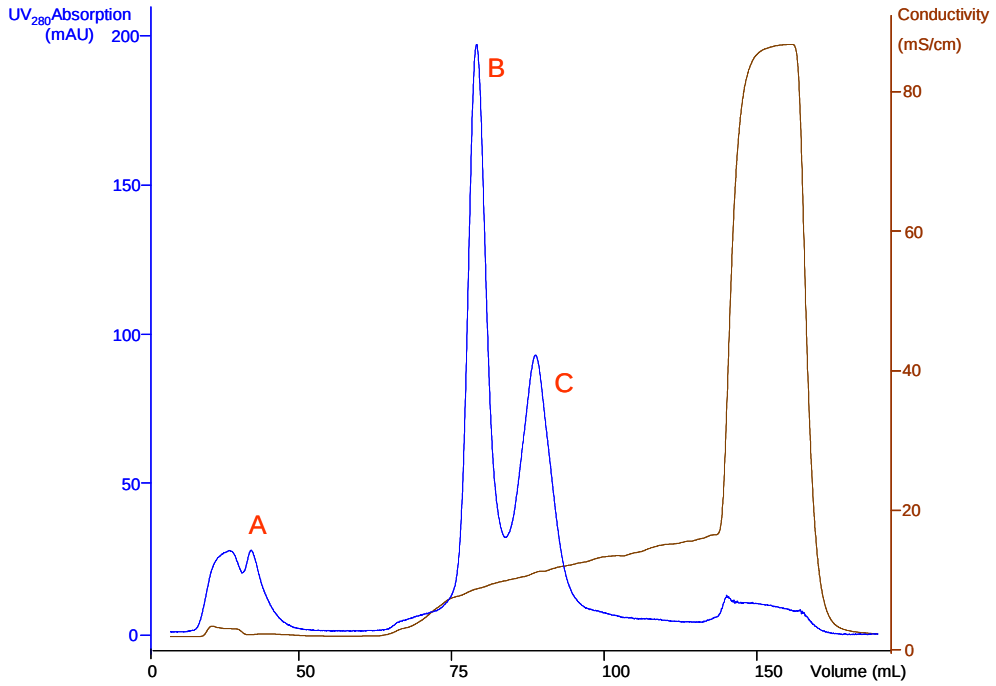


Figure 55. HPLC chromatogram of the purification of *Bs* DHFR conjugated with a 5 kDa PEG-maleimide on MacroCap SP. Peak (A) corresponds to the unbound material whereas peak B and C are the eluted fractions.

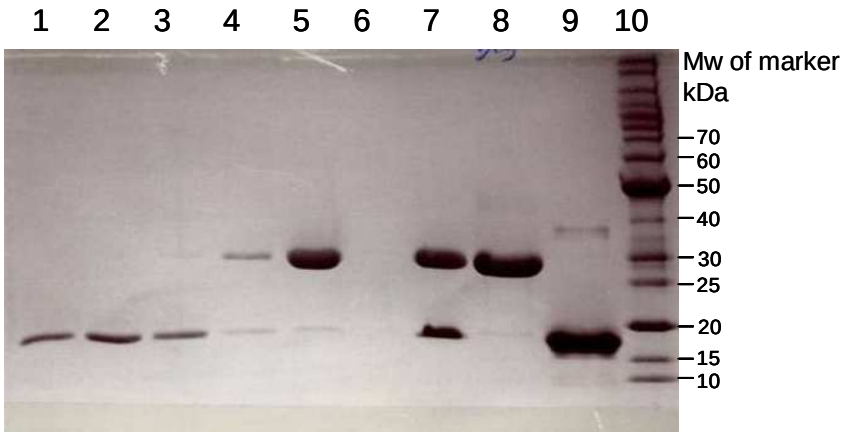


Figure 56. SDS-PAGE analysis of the purification PEGylated *Bs* DHFR with 5 kDa PEG-maleimide in non reducing conditions. Lane 10 corresponds to the molecular weight marker (BenchMark Protein Ladder, Invitrogen), lane 9 to native *Bs* DHFR, lane 8 to the reaction mixture, lane 7 desalted reaction mixture, lane 6 unbound material, lane 5 and 4 first eluted peak, lane 1, 2, 3 correspond to fraction collected in the last eluted peak.

The SDS-PAGE analysis is rich in information. As implied by lane 8 in figure 56, almost all *Bs* DHFR is PEGylated in a single and homogenous conjugate. The desalting step has a slight effect on the stability of the linkage established between the protein and the PEG molecules. It seems that approximately 20 to 30 % of PEG molecules leach the protein as shown by the extra band of lower molecular weight. This band corresponds to free *Bs* DHFR as it migrates as the same position of the *Bs* DHFR control. Lane 6, which is empty, corresponds to the non reacted PEG. Coomassie blue staining can not stain free PEG molecules, a specific staining designed by Kurfurst can probe specifically PEG moieties [308]. Lanes 4 and 5 correspond to pure mono-PEGylated *Bs* DHFR albeit a very slight amount of free enzyme is present. The estimated molecular weight of the conjugate is according to

	Protein			**Enzyme**	
Step	PEG-maleimide 5 kDa (mg)	Protein quantity (mg)	Yield (%)	Specific activity (units/mg of protein)	% activity
PEGylation	150	100	100	3.9×10^{6} $\pm 0.2 \times 10^{6}$	100
Desalting step		90	90	3.3×10^{6} $\pm 0.3 \times 10^{6}$	84
Anion exchange		73	73	3.2×10^{6} $\pm 0.2 \times 10^{6}$	81

Table 17. PEGylated *Bs* DHFR purification table summary. The PEGylation reaction was done with a PEG-maleimide of 5 kDa. Protein quantities were determined using the Bradford method. One unit of enzyme is defined as the amount of enzyme that catalyzes the conversion of 1 micro mole of substrate per minute under our assay conditions.

the protein ladder approximately 30 kDa which does not strictly corresponds to one monomer of *Bs* DHFR with a 5 kDa PEG molecule. The last lanes (1, 2, 3) show a single band corresponding of the free *Bs* DHFR which has leached the PEG-*Bs* DHFR conjugate. The PEGylation yields for both *Bs* DHFR, deuterium labelled or

hydrogenated were very similar. For us, it was critical to recover as much enzyme as we can. The cost of the PEG reagent (~250 US $ per gram) was relatively low in comparison to fully deuterated *Bs* DHFR which approximately costs 7000 US $ per gram of pure protein. From 100 mg of pure protein, the final yield of PEGylated protein in terms of mass was 73 mg (see table 17). The procedure was therefore very efficient. Almost all the protein was PEGylated under our conditions and only 10 % was lost during the desalting step. The cation exchange procedure was experimentally the most 'expensive'step but the free protein (as seen in peak C, figure 55) was almost entirely recovered and kept for further use if additional quantities of PEGylated material were needed. Looking at the specific activities of the PEGylated conjugate, it was clear under our assay conditions that PEGylation has a very limited impact on the enzyme catalysis (see table 18). The control, purified *Bs* DHFR has an activity of 3.8 x 10^6 units per mg. The activity of the PEGylated samples with either a 5 kDa or 12.5 kDa PEG molecule is decreased (3.2 x 10^6 and 3.1 x 10^6 units per mg respectively). However, the lost of specific activity does not exceed 20 % in all tested samples. These results also indicate that the PEG chain length does not play a significant role in the lost of activity. The differences between both samples, hydrogenated and perdeuterated *Bs* DFHR does not exceed 5 %. These limited alteration of *in vitro* activities are not surprising since PEGylation has been found to be a none invasive method in respect of protein bio-activity [287, 309, 310].

Protein	Units/mg of protein	% activity
Bs DFHR	3.8 x 10^6±0.4 x 10^6	100
Bs DFHR-PEG_{mal} 5 kDa	3.2 x 10^6±0.2 x 10^6	85
Bs DFHR- PEG_{mal} 12.5 kDa	3.1 x 10^6±0.3 x 10^6	83
dBs DFHR	2.6 x 10^6±0.2 x 10^6	100
dBs DFHR-PEG_{mal} 5 kDa	2.1 x 10^6±0.1 x 10^6	81
dBs DFHR- PEG_{mal} 12.5 kDa	2.3 x 10^6±0.2 x 10^6	87

Table 18. Bioactivity of PEG-modified *Bs* DHFR and d*Bs* DHFR (perdeuterated *Bs* DHFR). PEG*mal* is PEG-maleimide.

In conclusion, the PEGylation of *Bs* DHFR with PEG-maleimide was very efficient and enable us to obtain at a single attachment point a PEGylated form of the enzyme *Bs* DHFR with both PEG-maleimide reagents, 5 kDa and 12.5 kDa respectively. In order to have a deeper characterization of the produced samples, we have performed a mass spectrometry analysis using the MALDI-TOF technique

8.7.2. Mass spectrometry: MALDI TOF

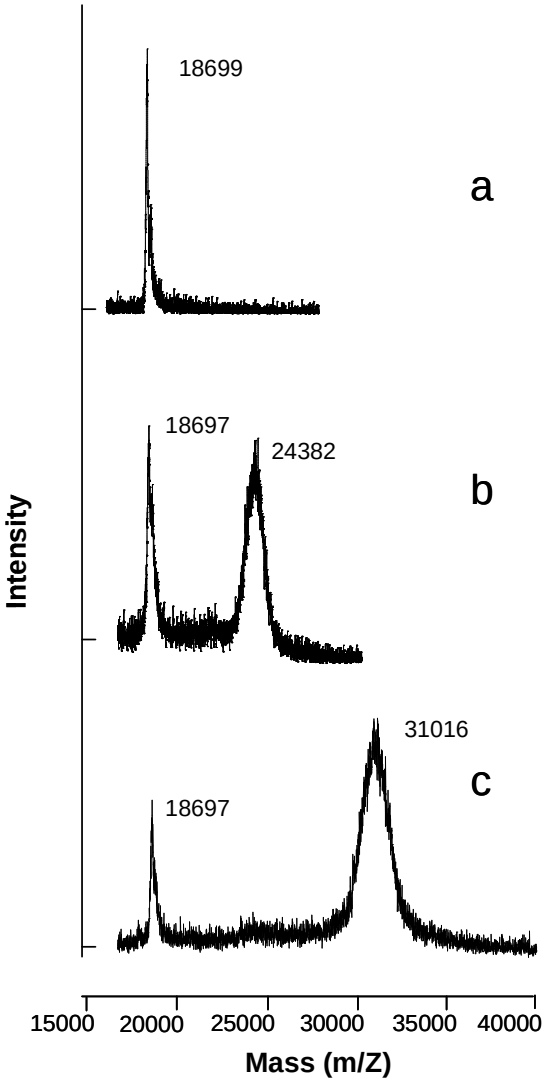


Figure 57. MALDI-TOF mass spectra of (a) native *Bs* DHFR, (b) *Bs* DHFR PEGylated with a 5kDa maleimide-PEG and (c) *Bs* DHFR PEGylated with a 12.5 kDa PEG-maleimide.

Figure 57 shows the MALDI-TOF mass spectra of *Bs* DHFR before and after PEGylation with PEG-maleimide. The MALDI-TOF mass spectrum of the native *Bs* DHFR show a single peak at m/Z 18697 which is relatively consistent with the calculated mass of a monomer (18695 Da) [56]. As expected the MALDI-TOF mass spectrum of PEGylated *Bs* DHFR show two peaks, one corresponding to an unPEGylated form of *Bs* DHFR and a second peak with a higher m/Z corresponding to a homogeneously mono-PEGylated form of *Bs* DHFR with PEG-maleimide. The mono-PEGylated *Bs* DHFR with 5 kDa PEG-maleimide produces a peak at m/Z 24382 which corresponds to a conjugate of one PEG maleimide to one *Bs* DHFR monomer. As shown in (c) a very similar spectrum is observed. Two peaks, one corresponding to the native *Bs* DHFR at m/Z 18699 and a second peak at m/Z of 31016 corresponding to a mono-PEGylated conjugate of *Bs* DHFR with a PEG-maleimide of 12.5 kDa. In all inspected spectra, the peaks corresponding to the PEGylated conjugates have a bell shape which is due to the mass distribution of the raw PEG [311]. The mass distribution is usually quantified with the term polydispersity which is the ratio between the observed molecular weight (Mw) of the PEG molecules and their calculated value according to the number-average molecular weight (Mn). For the PEG-maleimide used in this present studies, the polydispersity given by the supplier company was 1,008 and their M_w were 5513 and 12645 Da for the 5 and 12.5 kDa PEG respectively. Our experimental data yield approximate values of 5.685 kDa and 12.317 kDa for the 5 and 12.5 kDa PEG respectively. These values are therefore consistent with the supplier data. The relative importance of peak heights for the native and PEGylated *Bs* DHFR depends primarily of the sample to matrix ratio and is not linked to the abundance of the molecule of interest. The best results were obtained with a sample to matrix ratio of 1:8. A similar analysis was performed for the fully deuterated *Bs* DHFR. As shown in figure 58, very similar spectra were observed for the native and PEGylated enzyme with both PEG-maleimide reagent. The native enzyme produced a single peak at m/Z of 19740. The mono-PEGylated conjugate of fully deuterated *Bs* DHFR with a PEG-maleimide of 5 kDa produced two peaks: one at *m/Z* of 19741 corresponding to the free enzyme and a second one centered at m/Z 25400. The mass difference is 5660, which again signs

a mono-PEGylation of the enzyme by a 5 kDa PEG. The PEGylation of *Bs* DHFR with the 12.5 kDa PEG-maleimide shows similarly two peaks, one at m/Z of 19738 and a second one at *m/Z* of 32153. Again, the enzyme was successfully PEGylated at a single point attachment.

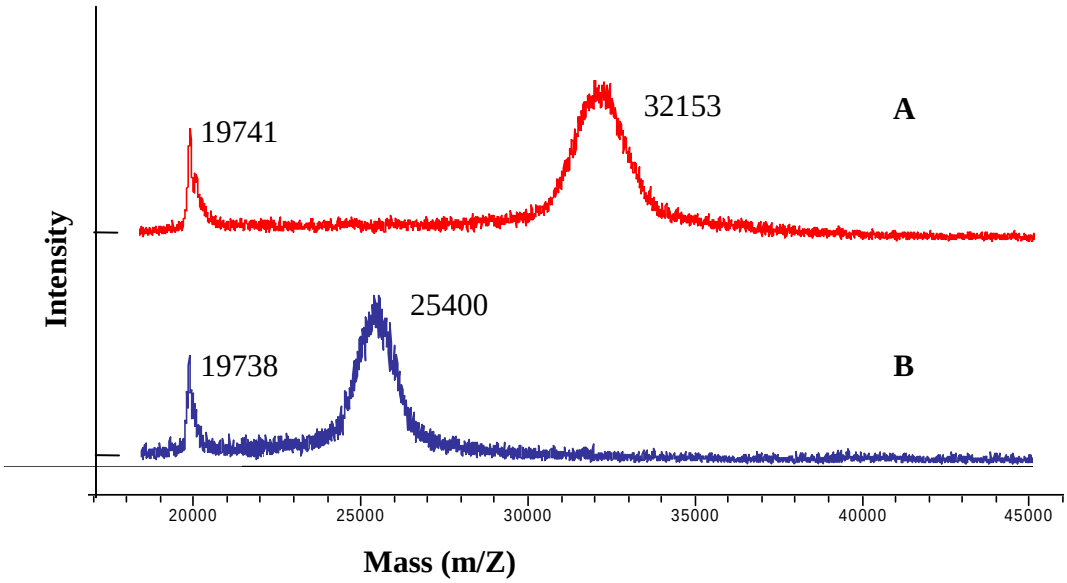


Figure 58. MALDI-TOF mass spectra of (A) perdeuterated *Bs* DHFR PEGylated with a 12.5 kDa PEG-maleimide, (B) perdeuterated *Bs* DHFR PEGylated with a 5 kDa maleimide-PEG.

In conclusion, the MALDI-TOF analysis was very valuable to characterize our sample and follow the PEGylation reaction. It is a method of choice for PEGylated therapeuthic since it does not require large amounts of sample and time.

8.7.3. Characterisation of PEGylated *Bs* DHFR by SDS-PAGE and chromatography

Both samples were analysed by SDS-PAGE in non-reducing conditions. Under these conditions, only the size of protein affects the migration. As expected the native enzyme migrates at a position close to 20 kDa as indicated by the protein ladder in lane 9, figure 59. A faint band was also observed around 40 kDa, which is probably the result of the dimerization of the native enzyme as no reductive agent was present in the medium. Looking at lanes 7, 4 and 3, we can give an estimate of M_w of the PEGylated conjugate of *Bs* DHFR with a 5 kDa PEG-maleimide. It is

approximately 30 kDa as reported previously. At high concentration (5 mg/mL) a smeared band (around 40 kDa) is observed in lane 7 with a small quantity of free enzyme as revealed by a band around 20 kDa. The smear is probably a contamination from the adjacent well (lane 6) containing the PEGylated *Bs* DHFR with a 12.5 kDa PEG-maleimide. In addition, a similar trend is observed when *Bs* DHFR is PEGylated with a 12.5 kDa PEG-maleimide (lanes 1, 2 and 6). The M_w of the PEGylated conjugate is approximately 50 kDa. The M_w estimation of the hybrid molecule between PEG and *Bs* DHFR can not be determined using protein molecular standards in SDS-PAGE. Again, the direct sum of the PEG and the protein molecular weights does not match the M_w given by SDS-PAGE analysis.

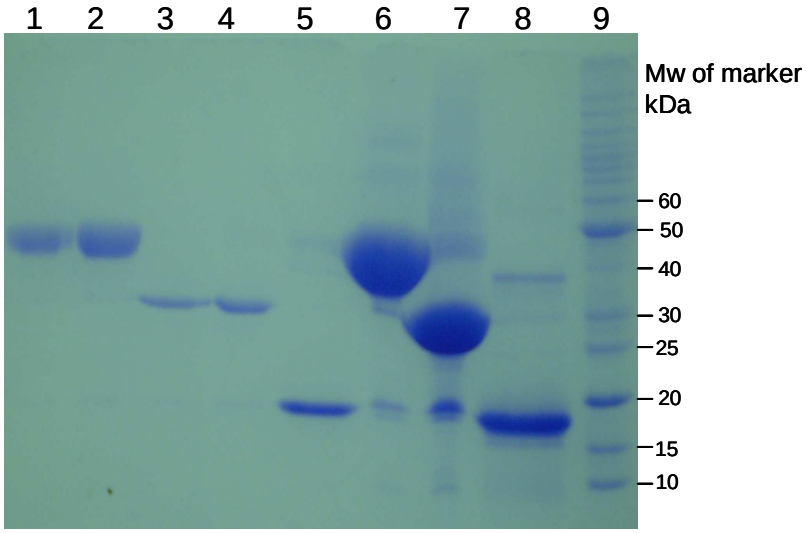


Figure 59. SDS-PAGE analysis of PEGylated *Bs* DHFR with 5 and 12.5 kDa PEG-maleimide in non-reducing conditions. Lane 9 corresponds to the molecular weight marker (BenchMark Protein Ladder, Invitrogen), lane 8 to native *Bs* DHFR at 5 mg/mL, lane 7 to PEGylated *Bs* DHFR with PEG maleimide 5 kDa at 10 mg/mL, lane 6 to PEGylated *Bs* DHFR with PEG maleimide 12.5 kDa at 10 mg/mL. Lane 5 is the native *Bs* DHFR at 2.5 mg/mL, lanes 3 and 4 contain PEGylated *Bs* DHFR with PEG maleimide 5 kDa at 2.5 and 5 mg/mL respectively. Lanes 1, 2 correspond to PEGylated *Bs* DHFR with PEG maleimide 12.5 kDa at 2.5 and 5 mg/mL respectively.

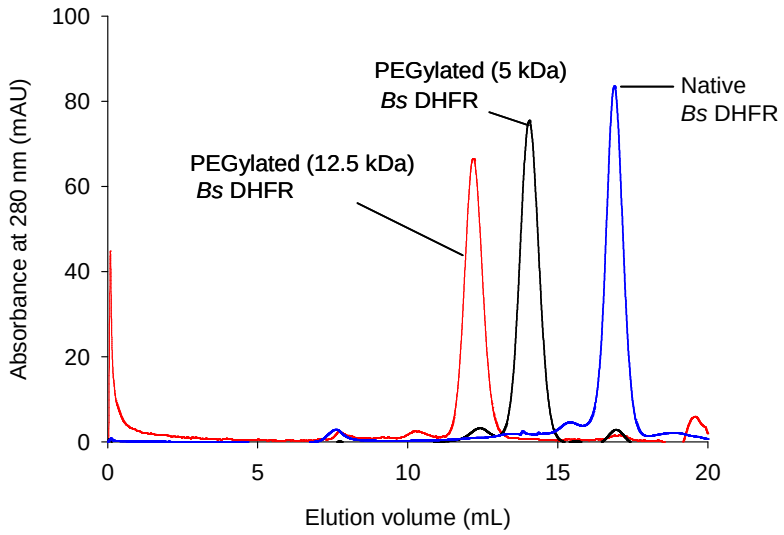


Figure 60. Elution profile of native *Bs* DHFR (blue curve), PEGylated *Bs* DHFR with PEG-maleimide 5 kDa (black curve) and PEGylated *Bs* DHFR with PEG-maleimide 12.5 kDa (red curve) on a Superdex 200 10/30.

A second way to characterize our PEGylated samples was carried out using Size Exclusion Chromatography (SEC) also called gel filtration (figure 60). Gel filtration is a useful technique to assess the purifity and homogeneity of a protein sample. We have therefore analysed the elution profile of our PEGylated conjugates of *Bs* DHFR against its native form. As shown in figure 60, the SEC curves of the native *Bs* DHFR present a single and well-defined peak with a K_{av} of 0.58. $K_{av} = \frac{(V_e - V_0)}{(V_c - V_0)}$ is the distribution coefficient where V_e, V_0 and V_c represent solute elution volume, void volume, and the total bed volume of fluid and SEC combined, respectively. The PEGylated forms of *Bs* DHFR gave a similar pattern with a single and well behaving UV peak. As expected, the modifications of the protein with PEG moieties have shifted the peak position. The K_{av} were respectively 0.40 and 0.28 for the 5 kDa and

12.5 kDa modified *Bs* DHFR. These shifts are therefore proportional to the size of grafted PEG molecule. The distribution coefficient decreases when PEG molecules size increase. These profiles demonstrate that first; the preparation is pure and homogenous as reported previously by SDS-PAGE analysis. Secondly, PEGylation modifies drastically the conjugates size or shape since their K_{av} coefficients are lower than the *Bs* DHFR. Such behaviour has been extensively reported in the literature [294, 299]. It beyond the scope of this chapter to model the PEGylated conjugate however gel filtration chromatography is also often uses in the determination of molecular weight and size of proteins [312]. It is based on the comparison of the elution volume parameter, such as the gel phase distribution coefficient K_{av} of the protein of interest against a series of known standards.

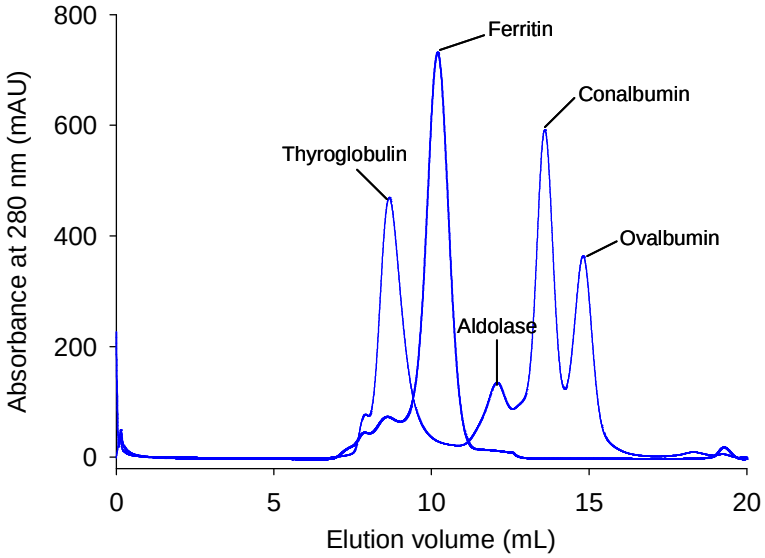


Figure 61. Chromatographic separation and calibration curve for the gel filtration calibration kit HMW on the Tricorn™ Superdex 200 10/30 GL column.

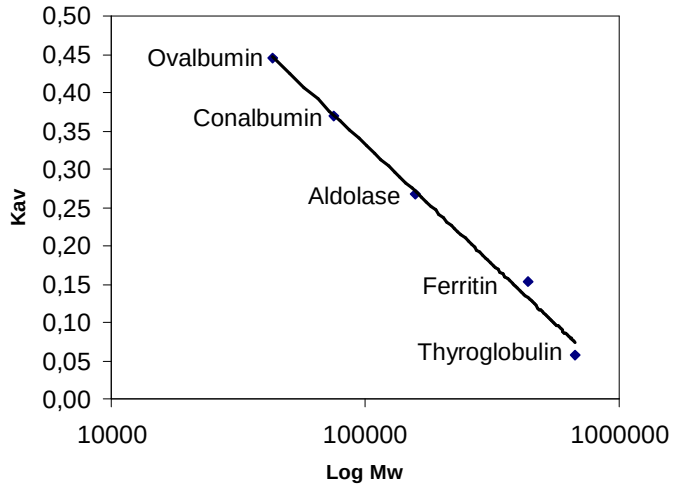


Figure 62. Molecular size plot of the protein calibration kit HMW (GE Healthcare, Uppsala, Sweden).

Protein	Molecular weight (M_w)	K_{av}
ovalbumin	43000	0.45
Conalbumin	75000	0.37
Aldolase	158000	0.27
Ferritin	440000	0.15
Thyroglobulin	669000	0.06

Table 19. Gel calibration kit HMW from GE Healthcare, Uppsala, Sweden. M_w were obtained from [172].

As shown in figure 62 and table 19, a series of standards have been established. It is common to produce a chart in terms of K_{av} versus log (M_w). The relationship is approximately linear over a wide range of molecular weights [313]. From our experimentally determined M_W, it seems that the PEGylated conjugates of *Bs* DHFR with a 5 kDa PEG or 12.5 kDa would have behaved as proteins with a M_W of 59109 and 151367 Da respectively. These extrapolated findings do not support the results

obtained by SDS-PAGE analysis, which have estimated much lower M_w: approximately 30 kDa for *Bs* DHFR with a 5 kDa PEG and 50 kDa for *Bs* DHFR with a 12.5 kDa PEG. A number of studies have reported such deviations when using SDS-PAGE analysis [314]. It has been suggested that PEG and SDS form a kind of complex [315]. The direct evidence of such interactions is given by the presence of broad or even smeared bands in some SDS-PAGE of PEGylated proteins. The migration property of PEGylated protein could be very different from that of the same molecular weight. Therefore, a simple comparison with the protein ladder will not give the correct information. On the other hand, if 'ideal' SEC separates solutes only on the basis of size, nonideal cases occur too. For instance, hydrophobic interactions between the support and the solute molecules can occur and therefore modify the elution profile of a given molecule. Accordingly, solute molecules with the same charge as that of the support will experience electrostatic repulsion from the pores, referred to as "ion exclusion," and emerge earlier than expected on the basis of size [316]. To complicate the equation, PEG-proteins are hybrid molecules and their properties result of a mixed effect resulting from the direct association of a protein and a polymer: PEG. It has been noted by a number of SEC studies that determination of PEGylated conjugates M_w with protein M_w standards was unreliable [294]. The routinely used methods in SEC can not be strictly applied and need to be specifically designed for PEG-protein conjugates. Overall, if SEC is relatively cheap and easy to set up SEC chromatography reflects an average of many molecular level interactions depending on the solvent, ionic force, type of column, temperature. For instance the viscosity radius (R_h) which is often used to define the size (in Å) of a protein by SEC chromatography makes the assumption that all proteins of polymers in solution behave as spheres, whatever is the true shape of the molecule [317]. Therefore, we must be cautious in interpreting the gel filtration data. In other words, SEC can not be applied universally to obtain good estimates of molecular weights and the need of alternative techniques is crucial. Actually a number of techniques are available to investigate the possible solution behaviour of PEG when grafted to globular protein surface. For instance, analytical centrifugation, NMR, SAXS or SANS would

obtained by SDS-PAGE analysis, which have estimated much lower M_w: approximately 30 kDa for *Bs* DHFR with a 5 kDa PEG and 50 kDa for *Bs* DHFR with a 12.5 kDa PEG. A number of studies have reported such deviations when using SDS-PAGE analysis [314]. It has been suggested that PEG and SDS form a kind of complex [315]. The direct evidence of such interactions is given by the presence of broad or even smeared bands in some SDS-PAGE of PEGylated proteins. The migration property of PEGylated protein could be very different from that of the same molecular weight. Therefore, a simple comparison with the protein ladder will not give the correct information. On the other hand, if 'ideal' SEC separates solutes only on the basis of size, nonideal cases occur too. For instance, hydrophobic interactions between the support and the solute molecules can occur and therefore modify the elution profile of a given molecule. Accordingly, solute molecules with the same charge as that of the support will experience electrostatic repulsion from the pores, referred to as "ion exclusion," and emerge earlier than expected on the basis of size [316]. To complicate the equation, PEG-proteins are hybrid molecules and their properties result of a mixed effect resulting from the direct association of a protein and a polymer: PEG. It has been noted by a number of SEC studies that determination of PEGylated conjugates M_w with protein M_w standards was unreliable [294]. The routinely used methods in SEC can not be strictly applied and need to be specifically designed for PEG-protein conjugates. Overall, if SEC is relatively cheap and easy to set up SEC chromatography reflects an average of many molecular level interactions depending on the solvent, ionic force, type of column, temperature. For instance the viscosity radius (R_h) which is often used to define the size (in Å) of a protein by SEC chromatography makes the assumption that all proteins of polymers in solution behave as spheres, whatever is the true shape of the molecule [317]. Therefore, we must be cautious in interpreting the gel filtration data. In other words, SEC can not be applied universally to obtain good estimates of molecular weights and the need of alternative techniques is crucial. Actually a number of techniques are available to investigate the possible solution behaviour of PEG when grafted to globular protein surface. For instance, analytical centrifugation, NMR, SAXS or SANS would

certainly help to draw a comprehensive picture of the solution conformation of PEGylated conjugates.

8.8. Conclusions and perspectives

The aim is of this chapter was to describe the different steps to provide a pure and homogenous preparation of a mono-PEGylated conjugate of the enzyme *Bs* DHFR with either a 5 kDa PEG-maleimide or 12.5 kDa PEG-maleimide. We have been successful in this work and we have produced in mg quantities both conjugates. Unfortunately the second step of this research effort was initiated and not fully achieved. Two projects were planned. As previously stated, PEG bio-effectiveness is believed to come from primarily its solution structure as it acts as a molecular shield [273, 318]. On the other hand, PEG molecules must be flexible so they will not impair the bioactivity of the protein, as for instance: in ligand-receptor recognition or enzyme catalysis. There is therefore an intricate balance between both properties. Because of our involvement with neutron scattering techniques, we thought that Small Angle Neutron Scattering was a potential technique to probe the effect of PEG on protein surface properties. Briefly SANS is a solution scattering technique probing the shape of nano-particle at a moderate resolution (1 to 1000 nm) [81]. The Guinier approximation of the scattering profile yields the radius of gyration (R_g) of the particle in solution and indicate its elongation or compactness [101]. The power of SANS resides also in the application of the contrast variation technique which is a powerful auxiliary that could highlight specifically element of a structure when the molecule is labelled. Contrast variation is based on the substitution of hydrogen atoms for deuterium atoms either in the solvent or in the molecule itself. For a detailed description of the technique see [319]. We have already began this work in collaboration with Pr. John Finney; Pr. Jeremy Smith; Dr. Moeava Tehei; Dr. Michael Hartlein and Pr. Fee Conan. We have produced the deuterated form of the enzyme *Bs* DHFR (since SANS require ^{1}H labelled protein) in collaboration with the deuteration laboratory facility in Grenoble. We have also been successfully allocated 8 days on the Diffractometer D22 at the Institut Laue Langevin which is the world leading neutron source in 2007. My role in this experiment was to produce and

prepare the PEGylated samples and assist the SANS experiment when possible. However, the data processing has not been performed entirely since the presence of aggregates has impaired the interpretation of the scattering curves. Since my role in this established collaboration with Dr. Hartlein (ILL) and Dr. Timmins (ILL) was not to process the data, I do not have the material from this specific experience and therefore I'm not able to discuss this issue here. However, it is critical for a successful SANS experience to test the stability and solubility of the sample in a deuterated media. Indeed the technique of contrast variation implies the substitution of aqueous buffer by a fully deuterated buffer from 0 % up to 100 % D_2O content. It is known that proteins tend to aggregate more easily in deuterated media [320]. A simple experience using native gel electrophoresis or light scattering would be of great benefit to find the optimal conditions for protein solubility and for success. If a few studies have approached and shed the light on the aspects of PEGylated conjugates [292, 293], the structure of such hybrid polymers is still a matter of debate. Efforts are needed. With the development of bifunctional molecules and complexes PEG architectures the structural milestone underlying the power of Pegnology is crucial.

As a subsidiary benefit, we have also investigated the dynamic properties of PEG molecules when conjugated to a protein by neutron scattering. We have used three different spectrophotometers, the backscattering spectrophotometers IN16 and IN13 and the time-of -flight IN6 for their complementary time and space windows. This experiment has taken place in May 2007 in collaboration with Dr Moeava Tehei. The data have not been processed and it is not possible to discuss them here until progresses are made. Depending on the previous results, the long term goal of PEGylation has definitively for us a role to play in the elucidation of protein dynamic toward activity and protein stability. We would like to set a similar experience on a hydrogenated form of *Bs* DHFR conjugated wih fully deuterated PEG molecules. In this respect, we would have access to the changes of fast protein dynamics when a non specific and covalently ligand is bound.

In conclusion, the field of PEGylation which is increasingly popular in protein therapeutics has also a strong potential in fundamental research. The cross linking of applied and more fundamental research activities may reveal here a general consensus

which unambiguously would stimulate a constant development of new technologies in protein therapeutics.

9. General Conclusion & Perspectives

The primary goal of this work was to assess the role of fast protein dynamics and their inter correlation with protein activity and stability. We have therefore focused on the protein model DHFR and its dynamics dependency on the ps-ns timescales when bound to the anticancer drug MTX using the technique of neutron scattering and correlated these findings with activity and stability measurements. In addition, we have investigated the role of fast fluctuations in view of protein thermo-adaptation by comparing two thermal variants of the enzyme DHFR, the moderate thermophilic: *Bs* DHFR and its mesophilic counterpart: *E. coli* DHFR. Collectively, the data have identified several features of the linkage established between the properties enumerated earlier.

The first chapters of the thesis have focused on the effect of the anti-cancer drug methotrexate. We have first explored the binding energetics in the interaction between the drug and the enzyme by isothermal titration calorimetry and steady state kinetics. ITC has revealed that the protein ligand recognition was largely favorable with an affinity constant in the nM range, enthalpy driven with unfavorable entropy. In a second step, we have assessed the type of inhibition mediated by MTX using steady state kinetics. The inhibition was found to be competitive with respect to the natural substrate H_2F and MTX was characterized as a tight binder. This study has shown that even a small change in a ligand structure can have profound effects on the interaction and binding energetics. It also highlighted the prodigious efficiency of MTX to block DHFR activity at levels flirting with the nano-molar range.

Chapter 4 was focusing on the effect H_2F or the inhibitor MTX on DHFR thermal stability. From circular dichroism, we found that both ligands can significantly increase the thermal stability of the enzyme with a more pronounced effect with MTX. There are numerous ways to stabilize the enzyme structure and it is very difficult to accurately rationalize which contributions have led to the change in the thermal stability. Explanation for these shortcomings include the effect of dynamics, buried water accessibility, folding coupled with binding and other conformational change or structural rearrangements [82, 321]. As judged by the

crystal structure for the two complexes, the difference in thermal stability is apparently due to the unique binding geometry of MTX and the protein. The inverted position of the pteridin ring and the specific interaction with Asp 27 [51] has necessarily involved a structural rearrangement of the protein architecture but it might have also redistributed other discrete forces concurring in the stabilization of the chain upon the thermal effect. We have hypothesized that dynamics could be one of these underlying forces. The work presented in this chapter is therefore an important step toward the understanding of structural energetics toward ligand protein interaction and protein structure.

The data presented in this chapter 3 assess the effect of the binding of MTX on protein dynamics on the ps-ns timescales, and specifically the significance of this interaction to the overall macromolecular dynamics parameters: the flexibility $<u^2>$ and resilience $<k>$. Incoherent neutron scattering data obtained on the backscattering spectrophotometer IN13 have provided a global picture of the side chain dynamics on the ps timescale. Overall, a global decrease in flexibility and an increase in protein resilience (only below the dynamical transition) were observed in the presence of MTX. These changes provide support for the hypothesis that ligand binding induces changes in dynamics and these changes can be propagated through the structure. From a combination of sequence conservation, mutational analysis and molecular dynamics simulation Agarwal *et al* have suggested a concept of a correlated network of residues in DHFR critical for the hydride transfer step of the reaction [18]. Interestingly many of the strictly conserved residues are responsible for maintaining the hydrogen bonding contacts with ligands such as Thr 113 (threonine) and Asp 27 in *E. coli*. or Glu 30 in *H. sapiens* [47]. More surprisingly, several of these key residues were quite distant from the active site, like Tyr 100, Phe 31 and Met 42 in *E. coli* DHFR. It implies that those residues distal from the active site would have been positioned in a particular way to maintain the synergy between the structure and catalytic efficiency of the enzyme. It suggests that there is a coupled 'promoting motions' exploit by DHFR to enhance its catalytic rate. Recent NMR relaxation studies have elegantly depicted the DHFR catalytic cycle as dynamic energy landscape [16, 19, 37]. It suggests that the progression of the enzyme along its

reaction coordinate is promoted by a succession of excited and grounding energy states induced by the binding or release of the products. It clearly outlines a direct correlation between the fluctuations and the catalytic efficiency of the enzyme. Thus optimization of the free energy landscape by selection of the substates and pathways needed for function is an element of efficient catalysis. If this hypothesis turns out to be correct, the effect of a specific inhibitor in the DHFR cycle might also trigger a specific effect. It is therefore very tempting to suggest that the preferred fluctuations observed along the reaction coordinate discussed earlier could have their respective counterpart in presence of inhibitor. The energy landscape of a protein is usually depicted as a 3D surface with peak and valleys [322]. We suggest that some of these peaks and valleys could be seen as 'dead end pathways'. In other words, if the hierarchy in time and space indicates that slow conformational transitions occurring on the timescale of catalytic turnover are facilitated by collective high frequency fluctuations, the reduction or modification of these fast motions could also have impaired the biological function of proteins. It may explain why transition state analogs exert such an unprecedented power on protein function and are used successfully in cancer therapy for instance [36]. We could also extend this view as a plausible determinant for allosteric regulation of enzymes and signal transduction but the correlation is not easy to draw.

A number of complementary projects have been also initiated during this PhD. The determination of the X-ray structure of the MTX/DHFR complex is one of them. This particular work has not been completed, so the discussion is not yet possible even if the results are quite promising. To exploit the findings from the dynamic study, we though that an unspecific and covalent ligand such as PEG could yield original and unexpected results. The basic idea was to replace MTX by PEG molecules but the cost of the raw materials (fully deuterated PEG) was out of reach for our laboratory so we decide to examine PEG dynamics instead. Chapter 8 gives an overview of the PEGylation reaction and sample preparation for neutron scattering and SANS. The neutron scattring was carried out in 2007 at the ILL, Grenoble but the results have not been processed. As a subsidiary benfit, we also try to determine the solution of the PEG-DHFR conjugate using SANS but unfortunately the experiment

fails to give the expected results. Last but not least, since neutron scattering is an isotope sensitive technique, the need of deuterium labelled proteins is often crucial. We have therefore examined the effect of perdeuteration on the temperature dependence of the enzyme *Bs* DHFR in chapter 6. If similar studies have been done on the effect of perdeuteration on protein function [180, 207, 214], stability [211], and structure [180]. It was original to use the equilibirum model [323] to test the interplay between enzyme temperature dependency and perdeuteration. This experiment has clearly outlined a decrease in activity, stability when the enzyme was perdeuterated. However, this piece of work needs the input of a biophysical method to rationalize in terms of structure or energetics these kinetics findings.

10. Future directions

One of the long term goals of this line of research was to highlight the strength of neutron scattering in the field of enzymology and catalysis. We show that, DHFR as a model system provides a useful tool for the study of fast protein dynamics and the effect of ligand binding. However, an obvious missing piece in the present work is a description of the other substrate/cofactor combination. It would be clearly advantageous to complement our set, with for instance, the natural substrate H_2F or with other potent inhibitors of DHFR. It well known that bacterial and eucaryotes DHFR variants do not respond similarly to anti-folate drugs but the key stones of this powerful inhibition are not entirely known [324]. Protein flexibility might be one of the milestones of this intricate puzzle.

Additionally, the binding signature given by the ITC measurement has provided some interesting line of evidence about the underlying binding energetics that drive the interaction between the enzyme and its stoichoimetric inhibitor. However to fully exploit these data, we need a structural input for the reliable interpretations of the binding determinants. We have already started the structural determination of the complex between the *Bs* DHFR and the drug MTX using X-ray crystallography. Our model is under refinement (R_{fact}: 22.1%, R_{free} : 31% at a resolution of 2.1 Å). In light of this evidence, it is interesting to raise the question of drug design. The discovery and marketing of drugs is a time and money consuming process for pharmaceutical groups. In proteomics, the common strategies are based on the X-ray structures which allow the docking of active molecules in the active site of enzymes or binding domain of receptors [325, 326]. Nevertheless, through this study and a number or others [327, 328], it is becoming clear that protein flexibility is a relevant factor in ligand recognition, binding or chemical turnover. The implementation of drug design strategies must exploit protein flexibility. For instance, Lerner *et al* employ multiple protein structures to incorporate protein flexibility in structure-based drug discovery [329]. Their pharmacophore model improves its performance with increased dynamic sampling, indicating that including a greater degree of flexibility can enhance the quest for potent inhibitors. It emphasizes the growing importance of protein

flexibility in the quest of identifying novel inhibitors. It is also important to note that distal mutations can impact active site flexibility and protein function [330, 331]. It is very tempting to test this hypothesis with PEG molecules and find out whether or not protein dynamics, protein function and or protein stability are enhanced or modulated by such unspecific and distal ligands. This new field will undoubtedly open new strategies in drug design.

11. Appendix

Enzyme	*Bs* DHFR	*dBs* DHFR	
Temperature (K)	k_{cat} (s^{-1})	k_{cat} (s^{-1})	k_{cat}^{D}/k_{cat}^{H}
298	13	8	0.57
303	21	12	0.57
308	33	19	0.58
313	51	30	0.58
318	78	46	0.59
323	117	69	0.59
328	173	104	0.60
333	254	153	0.60
338	367	223	0.61
343	526	322	0.61
348	747	460	0.62
353	1049	651	0.62
358	1460	912	0.62

Table A. Summary of experimentally determined k_{cat} and k_{cat}^{D}/k_{cat}^{H} in sodium phosphate buffer 100 mM pH 8 in H_2O

Enzyme	*Bs* DHFR	*dBs* DHFR	
Temperature (K)	k_{cat} (s^{-1})	k_{cat} (s^{-1})	k_{cat}^{D}/k_{cat}^{H}
298	12	6	0.52
303	19	10	0.53
308	30	16	0.54
313	46	25	0.54
318	70	38	0.55
323	105	58	0.55
328	155	86	0.56
333	228	128	0.56
338	330	187	0.57
343	474	270	0.57
348	673	387	0.58
353	947	549	0.58
358	1320	771	0.58

Table B. Summary of experimentally determined k_{cat} and k_{cat}^{D}/k_{cat}^{H} in sodium phosphate buffer 100 mM pH 8 in D_2O

12. References

1. Kraut, D.A., K.S. Carroll, and D. Herschlag. *Challenges in enzyme mechanism and energetics.* Annual Review of Biochemistry, 2003. **72**(1): p. 517-571.

2. Laidler, K.J. *Chemical-Kinetics and the Origins of Physical-Chemistry.* Archive for History of Exact Sciences, 1985. **32**(1): p. 43-75.

3. Laidler, K.J.S., J. M.Laurence, D. J. R,Gutfreund, H.Neurath, H.Lumry, R.Rabin, B. R.Smith, E. L.Williams, R. J. P.Vallee, B. L.Robert, L.Blum, J. J.Bernhard, S. A.Bergmann, P. D.Davies, D. R.Green, A. L.Morton, R. K.Koshland, D. E.Swoboda, P. A. T.Hoffmannostenhof, O.Eley, D. D.Bunn, D.Butler, J. A. V.Sangster, M.Klug, A.Roughton, F. J. W., J. G.Couper, A.Rydon, H. N.Kacser, H.Haurowitz, F. *The Physical Chemistry of Enzymes - General Discussion.* Discussions of the Faraday Society, 1955(20): p. 254-316.

4. Pauling, L.V., *The nature of the chemical bond and the structure of molecules and crystals*. 1939, New York Cornwall University Press.

5. James, L.C. and D.S. Tawfik. *Conformational diversity and protein evolution - a 60-year-old hypothesis revisited.* Trends in Biochemical Sciences, 2003. **28**(7): p. 361-368.

6. Blake, C.C.F., D.F. Koenig, G.A. Mair, A.C.T. North, D.C. Phillips, and V.R. Sarma. *Structure of Hen Egg-White Lysozyme - a 3-Dimensional Fourier Synthesis at 2a Resolution.* Nature, 1965. **206**(4986): p. 757-759.

7. Koshland, D.E. *Conformational changes: How small is big enough?* Nature Medicine, 1998. **4**(10): p. 1112-1114.

8. Wyckoff, H.W., K.D. Hardman, N.M. Allewell, T. Inagami, Tsernog l.D, L.N. Johnson, and F.M. Richards. *Structure of Ribonuclease-S at 6 a Resolution.* Journal of Biological Chemistry, 1967. **242**(16): p. 3749-3751.

9. Koshland, D.E. *Application of a Theory of Enzyme Specificity to Protein Synthesis.* Proceedings of the National Academy of Sciences of the United States of America, 1958. **44**(2): p. 98-104.

10. Goh, C.S., D. Milburn, and M. Gerstein. *Conformational changes associated with protein-protein interactions.* Current Opinion in Structural Biology, 2004. **14**(1): p. 104-109.

11. Flores, S., N. Echols, D. Milburn, B. Hespenheide, K. Keating, J. Lu, S. Wells, E.Z. Yu, M. Thorpe, and M. Gerstein. *The database of macromolecular motions: new features added at the decade mark.* Nucleic Acids Research, 2006. **34**: p. D296-D301.

12. Gerstein, M. and W. Krebs. *A database of macromolecular motions.* Nucleic Acids Research, 1998. **26**(18): p. 4280-4290.

13. Cannon, W.R. and S.J. Benkovic. *Solvation, reorganization energy, and biological catalysis.* Journal of Biological Chemistry, 1998. **273**(41): p. 26257-26260.

14. Feeney, J. *NMR studies of ligand binding to dihydrofolate reductase.* Angewandte Chemie-International Edition, 2000. **39**(2): p. 290-312.

15. Eisenmesser, E.Z., D.A. Bosco, M. Akke, and D. Kern. *Enzyme dynamics during catalysis.* Science, 2002. **295**(5559): p. 1520-1523.

16. Boehr, D.D., H.J. Dyson, and P.E. Wright. *An NMR perspective on enzyme dynamics.* Chemical Reviews, 2006. **106**(8): p. 3055-3079.

17. Daniel, R.M., R.V. Dunn, J.L. Finney, and J.C. Smith. *The role of dynamics in enzyme activity.* Annu Rev Biophys Biomol Struct, 2003. **32**: p. 69-92.

18. Agarwal, P.K., S.R. Billeter, P.T. Rajagopalan, S.J. Benkovic, and S. Hammes-Schiffer. *Network of coupled promoting motions in enzyme catalysis.* Proc Natl Acad Sci U S A, 2002. **99**(5): p. 2794-9.

19. Vendruscolo, M. and C.M. Dobson. *Dynamic visions of enzymatic reactions.* Science, 2006. **313**(5793): p. 1586-1587.

20. Wolf-Watz, M.T., V.Henzler-Wildman, K.Hadjipavlou, G.Eisenmesser, E. Z.Kern, D. *Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair.* Nature Structural & Molecular Biology, 2004. **11**(10): p. 945-949.

21. Jardetzky, O. and J.-F. Lefèvre, *Protein Dynamics, Function and Design*. Proceedings of a NATO ASI and International School of Structural Biology and magnetic resonance. Vol. **301**. 1999.

22. Daniel, R.M., J.L. Finney, and J.C. Smith. *The dynamic transition in proteins may have a simple explanation.* Faraday Discussions, 2003. **122**: p. 163-169.

23. Daniel, R.M., J.C. Smith, M. Ferrand, S. Hery, R. Dunn, and J.L. Finney. *Enzyme activity below the dynamical transition at 220 K.* Biophys J, 1998. **75**(5): p. 2504-7.

24. Dunn, R.V., V. Reat, J. Finney, M. Ferrand, J.C. Smith, and R.M. Daniel. *Enzyme activity and dynamics: xylanase activity in the absence of fast anharmonic dynamics.* Biochem J, 2000. **346 Pt 2**: p. 355-8.

25. Tehei, M., B. Franzetti, D. Madern, M. Ginzburg, B.Z. Ginzburg, M.T. Giudici-Orticoni, M. Bruschi, and G. Zaccai. *Adaptation to extreme environments: macromolecular dynamics in bacteria compared in vivo by neutron scattering.* Embo Reports, 2004. **5**(1): p. 66-70.

26. Tehei, M., B. Franzetti, K. Wood, F. Gabel, E. Fabiani, M. Jasnin, M. Zamponi, D. Oesterhelt, G. Zaccai, M. Ginzburg, and B.Z. Ginzburg. *Neutron scattering reveals extremely slow cell water in a Dead Sea organism.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(3): p. 766-771.

27. Tehei, M., D. Madern, C. Pfister, and G. Zaccai. *Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability.* Proc Natl Acad Sci U S A, 2001. **98**(25): p. 14356-61.

28. Tehei, M., J.C. Smith, C. Monk, J. Ollivier, M. Oettl, V. Kurkal, J.L. Finney, and R.M. Daniel. *Dynamics of immobilized and native Escherichia coli dihydrofolate reductase by quasielastic neutron scattering.* Biophysical Journal, 2006. **90**(3): p. 1090-1097.

29. Frauenfelder, H., B.H. McMahon, R.H. Austin, K. Chu, and J.T. Groves. *The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(5): p. 2370-2374.

30. Choy, W.Y. and J.D. Forman-Kay. *Calculation of ensembles of structures representing the unfolded state of an SH3 domain.* Journal of Molecular Biology, 2001. **308**(5): p. 1011-1032.

31. Monod, J., J. Wyman, and J.P. Changeux. *On Nature of Allosteric Transitions - a Plausible Model.* Journal of Molecular Biology, 1965. **12**(1): p. 88-95.

32. Dill, K.A. and H.S. Chan. *From Levinthal to pathways to funnels.* Nature Structural Biology, 1997. **4**(1): p. 10-19.

33. Frauenfelder, H., F. Parak, and R.D. Young. *Conformational Substates in Proteins.* Annual Review of Biophysics and Biophysical Chemistry, 1988. **17**: p. 451-479.

34. Frauenfelder, H. *Proteins: Paradigms of complexity.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**: p. 2479-2480.

35. Frauenfelder, H. and B.H. McMahon. *Energy landscape and fluctuations in proteins.* Annalen Der Physik, 2000. **9**(9-10): p. 655-667.

36. Benkovic, S.J., G.G. Hammes, and S. Hammes-Schiffer. *Free-Energy Landscape of Enzyme Catalysis.* Biochemistry 2008. **47**(11): p. 3317-3321.

37. Boehr, D.D., McElheny, D.,Dyson, H. J. and Wright, P. E. *The dynamic energy landscape of dihydrofolate reductase catalysis.* Science, 2006. **313**(5793): p. 1638-1642.

38. Rajagopalan, P.T.R. and S.J. Benkovic. *Preorganization and protein dynamics in enzyme catalysis.* Chemical Record, 2002. **2**(1): p. 24-36.

39. Smith, J. *Protein dynamics: comparison of simulations with inelastic neutron scattering experiments.* Q Rev Biophys., 1991. **24**(3): p. 227-91.

40. Klinman, J.P. *Tunneling in enzyme catalyzed hydrogen atom transfer reactions.* Abstracts of Papers of the American Chemical Society, 2000. **220**: p.

41. Klinman, J.P. *Dynamic barriers and tunneling. New views of hydrogen transfer in enzyme reactions.* Pure and Applied Chemistry, 2003. **75**(5): p. 601-608.

42. Bruice, T.C. and S.J. Benkovic. *Chemical basis for enzyme catalysis.* Biochemistry, 2000. **39**(21): p. 6267-6274.

43. Radkiewicz, J.L. and C.L. Brooks. *Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase.* Journal of the American Chemical Society, 2000. **122**(2): p. 225-231.

44. Liang, Z.X. and J.P. Klinman. *Structural bases of hydrogen tunneling in enzymes: progress and puzzles.* Current Opinion in Structural Biology, 2004. **14**(6): p. 648-655.

45. Benkovic, S.J. and S. Hammes-Schiffer. *A perspective on enzyme catalysis.* Science, 2003. **301**(5637): p. 1196-1202.

46. Futterman, S. *Enzymatic reduction of folic acid and dihydrofolic acid to tetrahydrofolic acid.* Jounal of Biological Chemistry, 1957. **228**(2): p. 1031-1038.

47. Schnell, J.R., H.J. Dyson, and P.E. Wright. *Structure, dynamics, and catalytic function of dihydrofolate reductase.* Annual Review of Biophysics and Biomolecular Structure, 2004. **33**: p. 119-140.

48. Baccanari, D.P. and R.L. Tansik. *Kinetics of methotrexate binding to dihydrofolate reductase from Neisseria gonorrhoeae.* Biochemical Pharmacology, 1984. **33**(13): p. 2101-2107.

49. McGuire, J.J. *Anticancer Antifolates: Current Status and Future Directions.* Current Pharmaceutical Design, 2003. **9**(31): p. 2593-2613.

50. Taylor, E.C., D. Kuhnt, C. Shih, S.M. Rinzel, G.B. Grindey, J. Barredo, M. Jannatipour, and R.G. Moran. *A dideazatetrahydrofolate analog lacking a chiral center at C-6: N-[4-[2-(2-amino-3,4-dihydro-4-oxo-7H-pyrrolo[2,3-d]pyrimidin-5yl)ethyl[benzoyl]-L-glutamic acid is an inhibitor of thymidylate synthase.* Journal of Medicinal Chemistry, 1992. **35**(23): p. 4450-4454.

51. Sawaya, M.R. and J. Kraut. *Loop and subdomain movements in the mechanism of Escherichia coli dihydrofolate reductase: Crystallographic evidence.* Biochemistry, 1997. **36**(3): p. 586-603.

52. Swanwick, R.S., G. Maglia, L.-h. Tey, and R.K. Allemann. *Coupling of protein motions and hydrogen transfer during catalysis by Escherichia coli dihydrofolate reductase.* Biochemical Journal, 2006. **394**(1): p. 259-265.

53. Shrimpton, P., A. Mullaney, and R.K. Allemann. *Functional role for Tyr 31 in the catalytic cycle of chicken dihydrofolate reductase.* Proteins-Structure Function and Genetics, 2003. **51**(2): p. 216-223.

54. DeLano, W.L. *Use of PYMOL as a communications tool for molecular science*. in *228th National Meeting of the American-Chemical-Society*. 2004. Philadelphia, PA.

55. DeLano, W.L. and J.W. Lam. *PyMOL: A communications tool for computational models*. in *230th National Meeting of the American-Chemical-Society*. 2005. Washington, DC.

56. Kim HS, D.S., Lee SY, Wemmer D, Klinman JP. *Structure and hydride transfer mechanism of a moderate thermophilic dihydrofolate reductase from Bacillus stearothermophilus and comparison to its mesophilic and hyperthermophilic homologues.* Biochemistry, 2005. **44**(34): p. 11428-39.

57. Gouet, P., X. Robert, and E. Courcelle. *ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins.* Nucleic Acids Research, 2003. **31**(13): p. 3320-3323.

58. Corpet, F. *Multiple Sequence Alignment with Hierarchical-Clustering.* Nucleic Acids Research, 1988. **16**(22): p. 10881-10890.

59. Krissinel, E. and K. Henrick. *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions.* Acta

Crystallographica Section D-Biological Crystallography, 2004. **60**: p. 2256-2268.

60. Bailey, S. *The Ccp4 Suite - Programs for Protein Crystallography.* Acta Crystallographica Section D-Biological Crystallography, 1994. **50**: p. 760-763.

61. Osborne MJ, S.J., Benkovic SJ, Dyson HJ, Wright PE. *Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism.* Biochemistry, 2001. **40**(33): p. 9846-59.

62. Schnell, J.R., H.J. Dyson, and P.E. Wright. *Effect of cofactor binding and loop conformation on side chain methyl dynamics in dihydrofolate reductase.* Biochemistry, 2004. **43**(2): p. 374-383.

63. Polshakov, V.I. *Dihydrofolate reductase: structural aspects of mechanisms of enzyme catalysis and inhibition.* Russian Chemical Bulletin, 2001. **50**(10): p. 1733-1751.

64. Huennekens, F.M. *The methotrexate story: A paradigm for development of cancer chemotherapeutic agents.* Advances in Enzyme Regulation, 1994. **34**: p. 397-419.

65. Williams, J.W., J.F. Morrison, and R.G. Duggleby. *Methotrexate, a high-affinity pseudosubstrate of dihydrofolate reductase.* Biochemistry, 1979. **18**(12): p. 2567-2573.

66. Polshakov, V.I., B. Birdsall, T.A. Frenkiel, A.R. Gargaro, and J. Feeney. *Structure and dynamics in solution of the complex of Lactobacillus casei dihydrofolate reductase with the new lipophilic antifolate drug trimetrexate.* Protein Science, 1999. **8**(3): p. 467-481.

67. Bolin, J.T., D.J. Filman, D.A. Matthews, R.C. Hamlin, and J. Kraut. *Crystal-Structures of Escherichia-Coli and Lactobacillus-Casei Dihydrofolate-Reductase Refined at 1.7 a Resolution .1. General Features and Binding of Methotrexate.* Journal of Biological Chemistry, 1982. **257**(22): p. 3650-3662.

68. Reyes, V.M., M.R. Sawaya, K.A. Brown, and J. Kraut. *Isomorphous Crystal-Structures of Escherichia-Coli Dihydrofolate-Reductase Complexed with Folate, 5-Deazafolate, and 5,10-Dideazatetrahydrofolate - Mechanistic Implications.* Biochemistry, 1995. **34**(8): p. 2710-2723.

69. Gargaro, A.R., A. Soteriou, T.A. Frenkiel, C.J. Bauer, B. Birdsall, V.I. Polshakov, I.L. Barsukov, G.C.K. Roberts, and J. Feeney. *The solution structure of the complex of Lactobacillus casei dihydrofolate reductase with methotrexate.* Journal of Molecular Biology, 1998. **277**(1): p. 119-134.

70. Cocco, L., B. Roth, C. Temple, J.A. Montgomery, R.E. London, and R.L. Blakley. *Protonated state of methotrexate, trimethoprim, and pyrimethamine bound to dihydrofolate reductase.* Archives of Biochemistry and Biophysics, 1983. **226**(2): p. 567-577.

71. Chen, J.T., K. Taira, C.P.D. Tu, and S.J. Benkovic. *Probing the Functional-Role of Phenylalanine-31 of Escherichia-Coli Dihydrofolate-Reductase by Site-Directed Mutagenesis.* Biochemistry, 1987. **26**(13): p. 4093-4100.

72. Rajagopalan, P.T.R., Z.Q. Zhang, L. McCourt, M. Dwyer, S.J. Benkovic, and G.G. Hammes. *Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(21): p. 13481-13486.

73. Bennett, B., P. Langan, L. Coates, M. Mustyakimov, B. Schoenborn, E.E. Howell, and C. Dealwis. *Neutron diffraction studies of Escherichia coli dihydrofolate reductase complexed with methotrexate.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(49): p. 18493-18498.

74. Birdsall, B., A.S.V. Burgen, E.I. Hyde, G.C.K. Roberts, and J. Feeney. *Negative cooperativity between folinic acid and coenzyme in their binding to L. casei dihydrofolate reductase.* Biochemistry, 1981. **20**(25): p. 7186-7195.

75. Bystroff, C. and J. Kraut. *Crystal structure of unliganded Escherichia coli dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding.* Biochemistry, 1991. **30**(8): p. 2227-2239

.

76. Parak, F., E.W. Knapp, and D. Kucheida. *Protein dynamics : Mössbauer spectroscopy on deoxymyoglobin crystals.* Journal of Molecular Biology, 1982. **161**(1): p. 177-194.

77. Pynn, R., *Neutron scattering: A Primer*. 1990.

78. Helliwell, J.R. *Neutrons in biology.* Nature Structural Biology, 1998. **5**(7): p. 538.

79. Zaccai, G. and B. Jacrot. *Small Angle Neutron Scattering.* Annual Review of Biophysics and Bioengineering, 1983. **12**(1): p. 139-157.

80. Teixeira, S.C.M., G. Zaccai, J. Ankner, M.C. Bellissent-Funel, R. Bewley, M.P. Blakeley, P. Callow, L. Coates, R. Dahint, R. Dalgliesh, N.A. Dencher, V.T. Forsyth, G. Fragneto, B. Frick, R. Gilles, T. Gutberlet, M. Haertlein, T. Hauß, W. Häußler, W.T. Heller, K. Herwig, O. Holderer, F. Juranyi, R. Kampmann, R. Knott, S. Krueger, P. Langan, R.E. Lechner, G. Lynn, C. Majkrzak, R.P. May, F. Meilleur, Y. Mo, K. Mortensen, D.A.A. Myles, F. Natali, C. Neylon, N. Niimura, J. Ollivier, A. Ostermann, J. Peters, J. Pieper,

A. Rühm, D. Schwahn, K. Shibata, A.K. Soper, T. Strässle, J. Suzuki, I. Tanaka, M. Tehei, P. Timmins, N. Torikai, T. Unruh, V. Urban, R. Vavrin, and K. Weiss. *New sources and instrumentation for neutrons in biology.* Chemical Physics, 2008. **345**(2-3): p. 133-151.

81. Pynn, R., *Neutron Scattering-A Primer*, L.A.N.S. Center, Editor. 1990.

82. Creighton, T.E. *Stability of folded conformations : Current opinion in structural biology 1991, 1: 5-16.* Current Opinion in Structural Biology, 1991. **1**(1): p. 5-16.

83. Tehei, M., R. Daniel, and G. Zaccai. *Fundamental and biotechnological applications of neutron scattering measurements for macromolecular dynamics.* European Biophysics Journal with Biophysics Letters, 2006. **35**(7): p. 551-558.

84. Petoukhov, M. and D. Svergun. *Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution.* European Biophysics Journal, 2006. **35**(7): p. 567-576.

85. Bée, M., *Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science*. 1988: Institute of Physics Publishing.

86. Cotton, J.P. *Diffraction et spectroscopie des neutrons.* Sciences et techniques de l'ingenieur. **PE 1 095** p. 1-12.

87. Schoenborn, B.P. and R.B. Knott, *Neutrons in biology*. 1996: New York : Plenum Press, c1996.

88. Holderer, O., H. Frielinghaus, D. Byelov, M. Monkenbusch, J. Allgaier, and D. Richter. *Dynamic properties of microemulsions modified with homopolymers and diblock copolymers: The determination of bending moduli and renormalization effects.* The Journal of Chemical Physics, 2005. **122**(9): p. 094908-8.

89. Biehl, R., B. Hoffmann, M. Monkenbusch, P. Falus, S. Preost, R. Merkel, and D. Richter. *Direct Observation of Correlated Interdomain Motion in Alcohol Dehydrogenase.* Physical Review Letters, 2008. **101**(13): p. 138102-4.

90. DiCapua, E., M. Schnarr, R.W.H. Ruigrok, P. Lindner, and P.A. Timmins. *Complexes of RecA protein in solution: A study by small angle neutron scattering.* Journal of Molecular Biology, 1990. **214**(2): p. 557-570.

91. Doster, C.a. *'Expanding Frontiers in Polypeptide and Protein structural research'*. 1990.

92. Doster, W., S. Cusack, and W. Petry. *Dynamical Transition of Myoglobin Revealed by Inelastic Neutron-Scattering.* Nature, 1989. **337**(6209): p. 754-756.

93. Agarwal, P. *Enzymes: An integrated view of structure, dynamics and function.* Microbial cell factories, 2006. **5**(1): p. 2.

94. Gabel, F., D. Bicout, U. Lehnert, M. Tehei, M. Weik, and G. Zaccai. *Protein dynamics studied by neutron scattering.* Q Rev Biophys, 2002. **35**(4): p. 327-67.

95. Zaccai, G. *How soft is a protein? A protein dynamics force constant measured bby neutron scattering.* science, 2000. **2**(288): p. 1604-7.

96. Natali, F., M. Bee, A. Deriu, C. Mondelli, L. Bove, C. Castellano, and S. Labbe-Lavigne. *IN13 backscattering spectrometer: an instrument in evolution.* Physica B: Condensed Matter, 2004. **350**(1-3, Supplement 1): p. E819-E822.

97. Dilip, M., *Principles and Chemical applications of Neutron scattering*, T.U.o. Alabama, Editor. 2004.

98. Laemmli, U. *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-685.

99. Réat, V.Z., G.;Ferrand, M.; Pfister, C. , *Biological Macromolecular Dynamics*. Adenine, ed. N. Guilderland. 1997. 117-22.

100. Gabel, F. *Protein dynamics in solution and powder measured by incoherent elastic neutron scattering: the influence of Q-range and energy resolution.* European Biophysics Journal, 2005. **34**(1): p. 1-12.

101. Guinier, A., *Small Angle Scattering of X-rays*. Vol. **268**. 1955: John Wiley & Sons, Inc., New York.

102. Bicout, D.J. and G. Zaccai. *Protein flexibility from the dynamical transition: A force constant analysis.* Biophysical Journal, 2001. **80**(3): p. 1115-1123.

103. Daniel, R.M., J.L. Finney, and J.C. Smith. *The dynamic transition in proteins may have a simple explanation.* Faraday Discuss, 2003. **122**: p. 163-9; discussion 171-90.

104. Nienhaus K, O.A., Nienhaus GU, Parak FG, schmidt M. *Ligand migration and protein fluctuations in myoglobin mutant L29W.* Biochemistry, 2005. **44**: p. 5095-5105.

105. Tehei, M., D. Madern, C. Pfister, and G. Zaccai. *Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(25): p. 14356-14361.

106. Chunduru, S.K., V. Cody, J.R. Luft, W. Pangborn, J.R. Appleman, and R.L. Blakley. *Methotrexate-resistant variants of human dihydrofolate reductase. Effects of Phe31 substitutions.* Journal of Biological Chemistry, 1994. **269**(13): p. 9547-9555.

107. Junker, J.P., K. Hell, M. Schlierf, W. Neupert, and M. Rief. *Influence of Substrate Binding on the Mechanical Stability of Mouse Dihydrofolate Reductase.* Biophysical Journal, 2005. **89**(5): p. L46-48.

108. Zaccai, G., *Biochemistry - How soft is a protein? A protein dynamics force constant measured by neutron scattering*. 2000. p. 1604 - 1607.

109. Gabel, F., D. Bicout, U. Lehnert, M. Tehei, M. Weik, and G. Zaccai. *Protein dynamics studied by neutron scattering.* Quarterly Reviews of Biophysics, 2002. **35**(4): p. 327-367.

110. Balog, E., T. Becker, M. Oettl, R. Lechner, R. Daniel, J. Finney, and J.C. Smith. *Direct determination of vibrational density of states change on ligand binding to a protein.* Phys Rev Lett, 2004. **93**(2): p. 028103.

111. M.J.Danson, R.E.a., *Enzymes assays: a practical approach*. The practical approach series, ed. D.R.a.B.D. Hames. 1992.

112. Lineweaver, H. and D. Burk. *The Determination of Enzyme Dissociation Constants.* Journal of the American Chemical Society, 1934. **56**(3): p. 658-666.

113. Clausen, T., R. Huber, A. Messerschmidt, H.D. Pohlenz, and B. Laber. *Slow-Binding Inhibition of Escherichia coli Cystathionine beta-Lyase by L-Aminoethoxyvinylglycine: A Kinetic and X-ray Study.* Biochemistry 1997. **36**(41): p. 12633-12643.

114. Marko, G. and S. Jure. *Slow-binding inhibition: A theoretical and practical course for students.* Biochemistry and Molecular Biology Education, 2004. **32**(4): p. 228-235.

115. Cha, S., S.Y.R. Kim, S.G. Kornstein, P.W. Kantoff, K.H. Kim, and F.N.M. Naguib. *Tight binding inhibitors--IX: Kinetic parameters of dihydrofolate reductase inhibited by methotrexate, an example of equilibrium study.* Biochemical Pharmacology, 1981. **30**(12): p. 1507-1515.

116. Hillcoat, B.l., Nixon, P.F, Blakeley, R.L. *Effect of substrate decomposition on the spectrophotometric assay of dihydrofolate reductase.* Analytical Biochemistry, 1967. **21**(2): p. 178-89.

117. Dixon, M. *The determination of enzyme inhibitor constants.* Biochemical Journal, 1953. **55**(1): p. 170-180.

118. Baccanari, D.P., R.L. Tansik, S.S. Joyner, M.E. Fling, P.L. Smith, and J.H. Freisheim. *Characterization of Candida albicans dihydrofolate reductase.* Journal of Biological Chemistry, 1989. **264**(2): p. 1100-1107.

119. Jarabak, J. and N.R. Bachur. *A soluble dihydrofolate reductase from human placenta: Purification and properties.* Archives of Biochemistry and Biophysics, 1971. **142**(2): p. 417-425.

120. Rao, K.N. *Purification and characterization of dihydrofolate reductase from Lactobacillus leichmannii*. Indian Journal of Biochemistry & Biophysics, 2000. **37**(2): p. 121-129.

121. Cha, S. *Tight-binding inhibitors: Kinetic behavior*. Biochemical Pharmacology, 1975. **24**(23): p. 2177-2185.

122. Perdicakis, B., H.J. Montgomery, J.G. Guillemette, and E. Jervis. *Analysis of slow-binding enzyme inhibitors at elevated enzyme concentrations.* Analytical Biochemistry, 2005. **337**(2): p. 211-223.

123. Garrido-del Solo, C., F. García-Cánovas, B.H. Havsteen, and R. Varón Castellanos. *Kinetic analysis of enzyme reactions with slow-binding inhibition.* Biosystems, 1999. **51**(3): p. 169-180.

124. Sculley, M.J., J.F. Morrison, and W.W. Cleland. *Slow-binding inhibition: the general case.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1996. **1298**(1): p. 78-86.

125. Kuzmic, P. *A steady state mathematical model for stepwise "slow-binding" reversible enzyme inhibition.* Analytical Biochemistry, 2008. **380**(1): p. 5-12.

126. Senkovitch O., V.B., Nisha G.,Debasish C. *Lipophilic Antifolate Trimetrexate is a potent inhibitor of trypanosoma cruzi: Prospect for chemotherapy of Chagas' disease.* antimicrobial agents and chemotherapy, 2005. **49**(8): p. 3234-38.

127. Baker, B.M. and K.P. Murphy. *Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry.* Biophysical Journal, 1996. **71**(4): p. 2049-2055.

128. Horn, J.R., S. Ramaswamy, and K.P. Murphy. *Structure and Energetics of Protein-Protein Interactions: The Role of Conformational Heterogeneity in OMTKY3 Binding to Serine Proteases.* Journal of Molecular Biology, 2003. **331**(2): p. 497-508.

129. O'Brien, R., J.E. Ladbury, and C. B.Z, *Isothermal titration calorimetry of biomolecules, Chapter 10 in Protein-Ligand interactions: hydrodynamics and calorimetry*, ed. S. Harding and B. Chowdry. 2000: Oxford University Press.

130. Leavitt, S. and E. Freire. *Direct measurement of protein binding energetics by isothermal titration calorimetry.* Current Opinion in Structural Biology, 2001. **11**(5): p. 560-566.

131. Velazquez-Campoy, A. and E. Freire. *Isothermal titration calorimetry to determine association constants for high-affinity ligands.* Nat. Protocols, 2006. **1**(1): p. 186-191.

132. Pierce, M.M., C.S. Raman, and B.T. Nall. *Isothermal titration calorimetry of protein-protein interactions.* Methods-a Companion to Methods in Enzymology, 1999. **19**(2): p. 213-221.

133. Velazquez-Campoy, A., S. Muzammil, H. Ohtaka, A. Schon, S. Vega, and E. Freire. *Structural and Thermodynamic Basis of Resistance to HIV-1 Protease Inhibition: Implications for Inhibitor Design.* Current Drug Targets - Infectious Disorders, 2003. **3**: p. 311-328.

134. Turnbull, W.B., B.L. Precious, and S.W. Homans. *Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry.* Journal of the American Chemical Society, 2004. **126**(4): p. 1047-1054.

135. Delves, C.J., S.P. Ballantine, R.L. Tansik, D.P. Baccanari, and D.K. Stammers. *Refolding of Recombinant Pneumocystis carinii Dihydrofolate Reductase and Characterization of the Enzyme.* Protein Expression and Purification, 1993. **4**(1): p. 16-23.

136. Giladi, M., N. Altman-Price, I. Levin, L. Levy, and M. Mevarech. *FolM, A New Chromosomally Encoded Dihydrofolate Reductase in Escherichia coli.* Journal of Bacteriology, 2003. **185**(23): p. 7015-7018.

137. Sigurskjold, B.W. *Exact Analysis of Competition Ligand Binding by Displacement Isothermal Titration Calorimetry.* Analytical Biochemistry, 2000. **277**(2): p. 260-266.

138. Wang, Z.-X. *An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule.* FEBS Letters, 1995. **360**(2): p. 111-114.

139. Irene Luque, E.F. *Structural parameterization of the binding enthalpy of small ligands.* Proteins, 2002. **49**(2): p. 181-190.

140. Jackson, M., S. Chopra, R.D. Smiley, P.O. Maynord, A. Rosowsky, R.E. London, L. Levy, T.I. Kalman, and E.E. Howell. *Calorimetric Studies of Ligand Binding in R67 Dihydrofolate Reductase.* Biochemistry, 2005. **44**(37): p. 12420-12433.

141. Stone, S.R. and J.F. Morrison. *Kinetic mechanism of the reaction catalyzed by dihydrofolate reductase from Escherichia coli.* Biochemistry, 1982. **21**(16): p. 3757-3765.

142. Birdsall, B., V.I. Polshakov, and J. Feeney. *NMR Studies of Ligand Carboxylate Group Interactions with Arginine Residues in Complexes of Lactobacillus casei Dihydrofolate Reductase with Substrates and Substrate Analogues.* Biochemistry, 2000. **39**(32): p. 9819-9825.

143. Li, D., L.A. Levy, S.A. Gabel, M.S. Lebetkin, E.F. DeRose, M.J. Wall, E.E. Howell, and R.E. London. *Interligand Overhauser Effects in Type II Dihydrofolate Reductase.* Biochemistry, 2001. **40**(14): p. 4242-4252.

144. Narayana N, M.D., Howell EE, Nguyen-huu X. *A plasmid-encoded dihydrofolate reductase from trimethoprim-resistant bacteria has a novel D2-symmetric active site.* Nat Struct Biol. , 1995. **2**(11): p. 1018-25.

145. Velazquez-Campoy, A., Y. Kiso, and E. Freire. *The Binding Energetics of First- and Second-Generation HIV-1 Protease Inhibitors: Implications for Drug Design.* Archives of Biochemistry and Biophysics, 2001. **390**(2): p. 169-175.

146. Stammers, D.K., J.N. Champness, C.R. Beddell, J.G. Dann, E. Eliopoulos, A.J. Geddes, D. Ogg, and A.C.T. North. *The structure of mouse L1210 dihydrofolate reductase-drug complexes and the construction of a model of human enzyme.* FEBS Letters, 1987. **218**(1): p. 178-184.

147. Yamamoto, T., S. Izumi, and K. Gekko. *Mass Spectrometry on Hydrogen/Deuterium Exchange of Dihydrofolate Reductase: Effects of Ligand Binding.* Journal of Biochemistry, 2004. **135**(6): p. 663-671.

148. Bushmarina, N.A., C.E. Blanchet, G. Vernier, and V. Forge. *Cofactor effects on the protein folding reaction: Acceleration of {alpha}-lactalbumin refolding by metal ions.* Protein Science, 2006. **15**(4): p. 659-671.

149. Greenfield, N.J. *Using circular dichroism spectra to estimate protein secondary structure.* Nat. Protocols, 2007. **1**(6): p. 2876-2890.

150. Greenfield, N.J. *Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism.* Nat. Protocols, 2007. **1**(6): p. 2733-2741.

151. Per Unneberg, J.J.M.P.C.F.M., *SOMCD: Method for evaluating protein secondary structure from UV circular dichroism spectra*. 2001. p. 460-470.

152. Yang JT, Wu CS, and H. Martinez, *Calculation of protein conformation from circular dichroism.* Methods Enzymol. Vol. **130**. 1986. 208-69.

153. Wolfgang Kabsch, C.S. *Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features.* Biopolymers, 1983. **22**(12): p. 2577-2637.

154. Benjwal, S., S. Verma, K.-H. Rohm, and O. Gursky. *Monitoring protein aggregation during thermal unfolding in circular dichroism experiments.* Protein Science, 2006: p. ps.051917406.

155. Gekko, K., K. Yamagami, Y. Kunori, S. Ichihara, M. Kodama, and M. Iwakura. *Effects of Point Mutation in a Flexible Loop on the Stability and Enzymatic Function of Escherichia coli Dihydrofolate Reductase.* Biochemical Journal, 1993. **113**(1): p. 74-80.

156. Hatsuho Uedaira, Shun Ichi Kidokoro, Masahiro Iwakura, Shinya Honda, and S. Ohashi. *Thermal Stability of Dihydrofolate Reductase and Its Fused Proteins with Oligopeptides.* Annals of the New York Academy of Sciences, 1990. **613**(10): p. 352-357.

157. Brandts, J.F. and L.N. Lin. *Study of strong to ultratight protein interactions using differential scanning calorimetry.* Biochemistry, 1990. **29**(29): p. 6927-6940.

158. Sturtevant, J.M. *Biochemical Applications of Differential Scanning Calorimetry.* Annual Review of Physical Chemistry, 1987. **38**: p. 463-488.

159. Hood K, B.P., Roberts GC. *Circular-dichroism studies of ligand binding to dihydrofolate reductase from Lactobacillus casei MTX/R.* Biochem J., 1979. **2**(177): p. 425-32.

160. Deléage G and Geourjon C. *An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum.* Computing Applied Biosciences, 1993. **2**: p. 197-199.

161. Merelo, J.J., M.A. Andrade, A. Prieto and F. Morán. *Proteinotopic Feature Maps.* Neurocomputing, 1994. **6**: p. 443-454.

162. Yuk Yin Sham, B.M.C.-J., Tsai Ruth Nussinov. *Thermal unfolding molecular dynamics simulation of Escherichia coli dihydrofolate reductase: Thermal stability of protein domains and unfolding pathway.* Proteins, 2002. **46**(3): p. 308-320.

163. Georlette, D., B. Damien, V. Blaise, E. Depiereux, V.N. Uversky, C. Gerday, and G. Feller. *Structural and Functional Adaptations to Extreme Temperatures in Psychrophilic, Mesophilic, and Thermophilic DNA Ligases.* Journal of Biological Chemistry, 2003. **278**(39): p. 37015-37023.

164. Siddiqui, K.S., G. Feller, S. D'Amico, C. Gerday, L. Giaquinto, and R. Cavicchioli. *The Active Site Is the Least Stable Structure in the Unfolding Pathway of a Multidomain Cold-Adapted alpha-Amylase.* Journal of Bacteriology, 2005. **187**(17): p. 6197-6205.

165. Sasso, S.P., R.M. Gilli, J.C. Sari, O.S. Rimet, and C.M. Briand. *Thermodynamic study of dihydrofolate reductase inhibitor selectivity.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1994. **1207**(1): p. 74-79.

166. Ohmae, E., T. Kurumiya, S. Makino, and K. Gekko. *Acid and Thermal Unfolding of Escherichia coli Dihydrofolate Reductase.* Journal of Biochemistry, 1996. **120**(5): p. 946-953.

167. Ainavarapu, S.R.K., L. Li, C.L. Badilla, and J.M. Fernandez. *Ligand Binding Modulates the Mechanical Stability of Dihydrofolate Reductase.* Biophysical Journal, 2005. **89**(5): p. 3337-3344.

168. Ragone, R. *Hydrogen-bonding classes in proteins and their contribution to the unfolding reaction.* Protein Science, 2001. **10**: p. 2075-2082.

169. Stuhrmann, H. *Small-angle scattering and its interplay with crystallography, contrast variation in SAXS and SANS.* Acta Crystallographica Section A, 2008. **64**(1): p. 181-191.

170. Wobbe, C.R., S. Mitra, and V. Ramakrishnan. *Structure of the capsid of Kilham rat virus from small-angle neutron scattering.* Biochemistry, 1984. **23**(26): p. 6565-6569.

171. Callow, P., A. Sukhodub, J.E. Taylor, and G.G. Kneale. *Shape and Subunit Organisation of the DNA Methyltransferase M.AhdI by Small-angle Neutron Scattering.* Journal of Molecular Biology, 2007. **369**(1): p. 177-185.

172. Lide, D.R. *Physical Constants of Organic Compounds.* CRC Handbook of Chemistry and Physics, 88th Edition (Internet Version 2008) CRC Press/Taylor and Francis, Boca Raton, FL.

173. Kushner, D., A. Baker, and T. Dunstall. *Pharmacological uses and perspectives of heavy water and deuterated compounds.* Canadian journal of physiology and pharmacology, 1999. **77**(2): p. 79-88.

174. Guild, W.R. and R.P. Van Tubergen. *Heat Inactivation of Catalase in Deuterium Oxide.* Science, 1957. **125**(3254): p. 939

175. Harrington, W.F. and P.H. von Hippel. *Formation and stabilization of the collagen-fold.* Archives of Biochemistry and Biophysics, 1961. **92**(1): p. 100-113.

176. Hermans, J.J. and H.A. Scheraga. *The thermally induced configurational change of ribonuclease in water and deuterium.* Biochimica Biophysica Acta, 1959. **36**: p. 534-5.

177. Brockwell, D., L. Yu, S. Cooper, S. McCleland, A. Cooper, D. Attwood, S.J. Gaskell, and J. Barber. *Physicochemical consequences of the perdeuteriation of glutathione S-transferase from S. japonicum.* Protein Science, 2001. **10**(3): p. 572-580.

178. Kresheck, G.C., H. Schneider, and H.A. Scheraga. *The Effect of D2O on the Thermal Stability of Proteins. Thermodynamic Parameters for the Transfer of Model Compounds from H2O to D2O.* The Journal of Physical Chemistry, 1965. **69**(9): p. 3132-3144.

179. Ahmed, A.I., D.T. Osuga, and R.E. Feeney. *Antifreeze proteins from fishes: freezing behavior in H2O and D2O.* Biochem int, 1980. **1**: p. 41-46.

180. Di Costanzo, L., M. Moulin, M. Haertlein, F. Meilleur, and D.W. Christianson. *Expression, purification, assay, and crystal structure of perdeuterated human arginase I.* Archives of Biochemistry and Biophysics, 2007. **465**(1): p. 82-89.

181. Meilleur, F., J. Contzen, D.A.A. Myles, and C. Jung. *Structural Stability and Dynamics of Hydrogenated and Perdeuterated Cytochrome P450cam (CYP101).* Biochemistry, 2004. **43**(27): p. 8744-8753.

182. Fisher, S.J. and J.R. Helliwell. *An investigation into structural changes due to deuteration.* Acta Crystallographica Section A, 2008. **64**(3): p. 359-367.

183. Wishnia, A. and M. Saunders. *The Nature of the Slowly Exchanging Protons of Ribonuclease.* Journal of the American Chemical Society, 1962. **84**(22): p. 4235-4239.

184. Halevi, E. *The secondary hydrogen isotope effect.* Int J Appl Radiat Isot., 1960. **7**: p. 192-7.

185. Daniel, R.M., M.J. Danson, and R. Eisenthal. *The temperature optima of enzymes: a new perspective on an old phenomenon.* Trends in Biochemical Sciences, 2001. **26**(4): p. 223-225.

186. Peterson, M.E., R. Eisenthal, M.J. Danson, A. Spence, and R.M. Daniel. *A new intrinsic thermal parameter for enzymes reveals true temperature optima. (vol 279, pg 20717, 2004).* Journal of Biological Chemistry, 2005. **280**(50): p. 41784-41784.

187. Arrhenius, S. *On the Reaction Velocity of the Inversion of Cane Sugar by Acids.* Zeitschrift für Physikalische Chemie, 1889. **4**: p. 226.

188. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*. Portland Press, London,UK,(1995), pp193-197. 1995.

189. Thomas, T.M. and R.K. Scopes. *The effects of temperature on the kinetics and stability of mesophilic and thermophilic 3-phosphoglycerate kinases.* Biochemical Journal, 1998. **330**(3): p. 1087-1095.

190. Eisenthal, R., M.E. Peterson, R.M. Daniel, and M.J. Danson. *The thermal behaviour of enzyme activity: implications for biotechnology.* Trends in Biotechnology, 2006. **24**(7): p. 289-292.

191. Michelle E. Peterson, R.E., Michael J. Danson, Alastair Spence, and Roy M. Daniel. *A New Intrinsic Thermal Parameter for Enzymes Reveals True Temperature Optima.* Journal of biological chemistry, 2004. **279**(20): p. 20717-20722.

192. Peterson, M.E., R. Eisenthal, M.J. Danson, A. Spence, and R.M. Daniel. *A new intrinsic thermal parameter for enzymes reveals true temperature optima.* Journal of Biological Chemistry, 2004. **279**(20): p. 20717-20722.

193. Lee, C.K., R.M. Daniel, C. Shepherd, D. Saul, S.C. Cary, M.J. Danson, R. Eisenthal, and M.E. Peterson, *Eurythermalism and the temperature dependence of enzyme activity*. 2007. p. 1934-1941.

194. Daniel, R., M. Danson, R. Eisenthal, C. Lee, and M. Peterson. *The effect of temperature on enzyme activity: new insights and their implications.* Extremophiles, 2008. **12**(1): p. 51-59.

195. Daniel, R.M., Danson, M. J., Hough, D. W., Lee, C. K., Peterson, M. E. and Cowan, D. A., *Protein Adaptation in Extremophiles*. 2008: Nova Science Publishers Inc. pp 1-34.

196. Artero, J.-B., M. Hartlein, S. McSweeney, and P. Timmins. *A comparison of refined X-ray structures of hydrogenated and perdeuterated rat [gamma]E-*

crystallin in H2O and D2O. Acta Crystallographica Section D, 2005. **61**(11): p. 1541-1549.

197. Gornall, A.G., C.J. Bardawill, and M.M. David. *Determination of serum proteins by means of the Buiret reaction.* Journal of Biological Chemistry, 1949. **177**(2): p. 751-766.

198. Peterson, M.E., R.M. Daniel, M.J. Danson, and R. Eisenthal. *The dependence of enzyme activity on temperature: determination and validation of parameters.* Biochemical Journal, 2007. **402**: p. 331-337.

199. Lee, C.K., S.C. Cary, A.E. Murray, and R.M. Daniel. *Enzymic Approach to Eurythermalism of Alvinella pompejana and Its Episymbionts.* Applied and Environmental Microbiology, 2008. **74**(3): p. 774-782.

200. Lee, C.K., R.M. Daniel, C. Shepherd, D. Saul, S.C. Cary, M.J. Danson, R. Eisenthal, and M.E. Peterson. *Eurythermalism and the temperature dependence of enzyme activity.* Faseb Journal, 2007. **21**(8): p. 1934-1941.

201. Brockwell D, Y.L., Cooper S, McCleland S, Cooper A, Attwood D, Gaskell SJ, Barber J. . *Physicochemical consequences of the perdeuteriation of glutathione S-transferase from S. japonicum.* Protein Sci. , 2001. **10**(3): p. 572-80.

202. Meilleur F, C.J., Myles DA, Jung C *Structural stability and dynamics of hydrogenated and perdeuterated cytochrome P450cam (CYP101).* Biochemistry, 2004. **43**: p. 8744-8753.

203. Rokop S, G.L., Parmerter S, Crespi HL, Katz JJ. . *Purification and characterization of fully deuterated enzymes.* Biochim Biophys Acta. , 1969. **191**(3): p. 707-715.

204. Northrop, D.B. *Uses of Isotope Effects in the Study of Enzymes.* Methods, 2001. **24**(2): p. 117-124.

205. Northrop, D.B. and L.P. Daniel, *Deuterium and tritium kinetic isotope effects on initial rates,* in *Methods in Enzymology*. 1982, Academic Press. p. 607-625.

206. Rokop, S., L. Gajda, S. Parmerter, H.L. Crespi, and J.J. Katz. *Purification and characterization of fully deuterated enzymes.* Biochimica et Biophysica Acta (BBA) - Enzymology, 1969. **191**(3): p. 707-715.

207. Xuying, L., P. Langan, B. Hanson, and R. Viola. *The effect of deuteration o protein structure: a high-resolution comparison of hydrogenous and perdeuterated haloalkane dehalogenase.* Acta Crystallographica Section D, 2007. **63**: p. 1000-1008.

208. Meng, F.-G., Y.-K. Hong, H.-W. He, A.E. Lyubarev, B.I. Kurganov, Y.-B. Yan, and H.-M. Zhou. *Osmophobic Effect of Glycerol on Irreversible Thermal Denaturation of Rabbit Creatine Kinase.* Biophysical Journal, 2004. **87**(4): p. 2247-2254.

209. Lopez, M.M. and G.I. Makhatadze. *Solvent isotope effect on thermodynamics of hydration.* Biophysical Chemistry, 1998. **74**(2): p. 117-125.

210. Hattori, A., H.L. Crespi, and J.J. Katz. *Association and Dissociation of Phycocyanin and the Effects of Deuterium Substitution on the Processes.* Biochemistry, 1965. **4**(7): p. 1225-1238.

211. Hattori, A., H.L. Crespi, and J.J. Katz. *Effect of Side-Chain Deuteration on Protein Stability.* Biochemistry, 1965. **4**(7): p. 1213-1225.

212. Yokogaki, S., K. Unno, N. Oku, and S. Okada. *Chaperonin-Repairable Subtle Incompleteness of Protein Assembly Induced by a Substitution of Hydrogen with Deuterium: Effect of GroE on Deuterated Ribulose 1,5-Bisphosphate Carboxylase.* Plant and cell physiology, 1995. **36**(3): p. 419-423.

213. Makhatadze GI, C.G., Gronenborn AM. *Solvent isotope effect and protein stability.* Nature Structural Biology, 1995. **2**(10): p. 852-55.

214. Scott, E. and D.S. Berns. *Completely Deuterated Proteins. III. Deuteration Effects on Protein-Protein Interaction in Phycocyanin.* Biochemistry, 1967. **6**(5): p. 1327-1334.

215. Banerjee, D. and S.K. Pal. *Conformational Dynamics at the Active Site of alpha-Chymotrypsin and Enzymatic Activity.* 2008. **24**(15): p. 8163-8168.

216. Hoi Sung Chung, A.T., *Temperature-dependent downhill unfolding of ubiquitin. I. Nanosecond-to-millisecond resolved nonlinear infrared spectroscopy*. 2008. p. 474-487.

217. Cleland, W.W. *The Use of Isotope Effects to Determine Enzyme Mechanisms.* Journal of Biological Chemistry, 2003. **278**(52): p. 51975-51984.

218. Appleman, J.R., E.E. Howell, J. Kraut, M. Kuhl, and R.L. Blakley. *Role of aspartate 27 in the binding of methotrexate to dihydrofolate reductase from Escherichia coli.* Journal of Biological Chemistry, 1988. **263**(19): p. 9187-9198.

219. Ozaki, Y., R.W. King, and P.R. Carey. *Methotrexate and folate binding to dihydrofolate reductase. Separate characterization of the pteridine and p-aminobenzoyl binding sites by resonance Raman spectroscopy*. Biochemistry, 1981. **20**(11): p. 3219-3225.

220. Poe, M., N.J. Greenfield, J.M. Hirshfield, M.N. Williams, and K. Hoogsteen. *Dihydrofolate reductase. Purification and characterization of the enzyme from an amethopterin-resistant mutant of Escherichia coli.* Biochemistry, 1972. **11**(6): p. 1023-1030.

221. Cannon, W.R., B.J. Garrison, and S.J. Benkovic. *Consideration of the pH-dependent inhibition of dihydrofolate reductase by methotrexate.* Journal of Molecular Biology, 1997. **271**(4): p. 656-668.

222. Blakeley, M.P., P. Langan, N. Niimura, and A. Podjarny. *Neutron crystallography: opportunities, challenges, and limitations.* Current Opinion in Structural Biology, 2008. **18**(5): p. 593-600.

223. Glusker, J.P., M. Lewis, and M. Rossi, *Crystal Structure Analysis for Chemists and Biologists*. Methods in Stereochemical Analysis, ed. A.P. Marchand. 1994: VCH publishers.

224. Smyth, M.S. and J.H.J. Martin. *X Ray crystallography.* Journal of Clinical Pathology 2000. **53**(1): p. 8-14.

225. Blow, D., *Outline of crystallography for biologists*. 2002, New york: Oxford University Press.

226. Mark R Sanderson and J.V. Skelly, *Macromolecular crystallography: conventional and high-throuhput methods*. 2007: Oxford biosciences.

227. Lorber, B. *To save crystallization data.* Acta Crystallographica Section D, 2001. **57**(4): p. 479.

228. Gibbs, J.W. *Vector analysis.* Reprinted in his collected works, 2nd edition, 1928.

229. Ewald P, P. *The reciprocal lattice in the theory of structure* Zeit Krist, 1921. **56**: p. 129-156.

230. Hauptman, H. *Phasing methods for protein crystallography.* Current Opinion in Structural Biology, 1997. **7**(5): p. 672-680.

231. Woolfon, V.G. *Direct methods in crystallography by Woolfson.* Acta Crystallographica, 1962. **15**(6): p. 626.

232. Rossmann, M.G. and D.M. Blow. *The detection of sub-units within the crystallographic asymmetric unit.* Acta Crystallographica, 1962. **15**(1): p. 24-31.

233. Patterson, A. *An alternative interpretation for vector maps.* Acta Crystallographica, 1949. **2**(5): p. 339-340.

234. Navaza, J., *Implementation of molecular replacement in AMoRe*, in *Acta Crystallographica Section D*. 2001. p. 1367-1372.

235. Murshudov, G.N., A.A. Vagin, and E.J. Dodson. *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallographica Section D-Biological Crystallography, 1997. **53**: p. 240-255.

236. Vagin, A. and A. Teplyakov. *MOLREP: an Automated Program for Molecular Replacement.* Journal of Applied Crystallography, 1997. **30**(6): p. 1022-1025.

237. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read. *Phaser crystallographic software.* Journal of Applied Crystallography, 2007. **40**(4): p. 658-674.

238. Fisher, R.A. *On the Mathematical Foundations of Theoretical Statistics.* Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 1922. **222**: p. 309-368.

239. Emsley, P. and K. Cowtan. *Coot: model-building tools for molecular graphics.* Acta Crystallographica Section D, 2004. **60**(12 Part 1): p. 2126-2132.

240. Jones, T. *Interactive electron-density map interpretation: from INTER to O.* Acta Crystallographica Section D, 2004. **60**(12 Part 1): p. 2115-2125.

241. Sheldrick, G. *A short history of SHELX.* Acta Crystallographica Section A, 2008. **64**(1): p. 112-122.

242. Branden, C.-I. and T. Alwyn Jones. *Between objectivity and subjectivity.* Nature, 1990. **343**(6260): p. 687-689.

243. Kleywegt, G.J. and T.A. Jones. *Where freedom is given, liberties are taken.* Structure, 1995. **3**(6): p. 535-540.

244. Brunger, A. *Assessment of phase accuracy by cross validation: the free R value. Methods and applications.* Acta Crystallographica Section D, 1993. **49**(1): p. 24-36.

245. Ramachandran, G.N. and G. Kartha. *Structure of Collagen.* Nature, 1954. **174**(4423): p. 269-270.

246. Subramanian, E. *G.N. Ramachandran.* Nat Struct Mol Biol, 2001. **8**(6): p. 489-491.

247. Laskowski, R.A., M.W. MacArthur, D.S. Moss, and J.M. Thornton. *PROCHECK: a program to check the stereochemical quality of protein structures.* Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.

248. Luthy, R., J.U. Bowie, and D. Eisenberg. *Assessment of protein models with three-dimensional profiles.* Nature, 1992. **356**(6364): p. 83-85.

249. Garman, E. *'Cool' crystals: macromolecular cryocrystallography and radiation damage.* Current Opinion in Structural Biology, 2003. **13**(5): p. 545-551.

250. Blewett, J.P. *Synchrotron Radiation; Early History.* Journal of Synchrotron Radiation, 1998. **5**(3): p. 135-139.

251. van der Veen, J.F. *Synchrotron light of the third and fourth generation --- how to fill the generation gap.* SYNCHROTRON RADIATION INSTRUMENTATION: Eighth International Conference on Synchrotron Radiation Instrumentation, 2004. **705**: p. 3-12.

252. Baruchel J, J.L. Hodeau, M.S. Lehman, J.-R. Regnard, and C. Schlenker, *HERCULES, neutron and synchrotron radiation for condensed matter studies: Theory, instruments and methods* Vol. **1**. 1993.

253. Perrakis, A., M. Harkiolaki, K.S. Wilson, and V.S. Lamzin. *ARP/wARP and molecular replacement.* Acta Crystallographica Section D, 2001. **57**(10): p. 1445-1450.

254. *The CCP4 suite: programs for protein crystallography.* Acta Crystallographica Section D, 1994. **50**(5): p. 760-763.

255. Evans, P. *Scaling and assessment of data quality.* Acta Crystallographica Section D, 2006. **62**(1): p. 72-82.

256. Kantardjieff, K.A. and B. Rupp. *Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals.* Protein Science, 2003. **12**(9): p. 1865-1871.

257. Weiss, M. *Global indicators of X-ray data quality.* Journal of Applied Crystallography, 2001. **34**(2): p. 130-135.

258. Diederichs, K. and P.A. Karplus. *Improved R-factors for diffraction data analysis in macromolecular crystallography.* Nature Structural Biology, 1997. **4**: p. 269-275.

259. Hirshfeld, F.L. and D. Rabinovich. *Treating weak reflexions in least-squares calculations.* Acta Crystallographica Section A, 1973. **29**(5): p. 510-513.

260. Badger, J. *An evaluation of automated model-building procedures for protein crystallography.* Acta Crystallographica Section D, 2003. **59**(5): p. 823-827.

261. Matthews, D.A., R.A. Alden, J.T. Bolin, S.T. Freer, R. Hamlin, N. Xuong, J. Kraut, M. Poe, M. Williams, and K. Hoogsteen. *Dihydrofolate reductase: x-ray structure of the binary complex with methotrexate.* Science, 1977. **197**(4302): p. 452-455.

262. Bennett, B.C., F. Meilleur, D.A.A. Myles, E.E. Howell, and C.G. Dealwis. *Preliminary neutron diffraction studies of Escherichia coli dihydrofolate reductase bound to the anticancer drug methotrexate.* Acta Crystallographica Section D-Biological Crystallography, 2005. **61**: p. 574-579.

263. Berman, H.M., J.D. Westbrook, M.J. Gabanyi, W. Tao, R. Shah, A. Kouranov, T. Schwede, K. Arnold, F. Kiefer, L. Bordoli, J. Kopp, M. Podvinec, P.D. Adams, L.G. Carter, W. Minor, R. Nair, and J.L. Baer. *The protein structure initiative structural genomics knowledgebase.* Nucleic Acids Research, 2009. **37**(suppl_1): p. D365-368.

264. Anne Louise Morris, Malcolm W MacArthur, E Gail Hutchinson, and J.M. Thornton. *Stereochemical quality of protein structure coordinates.* 1992. **12**(4): p. 345-364.

265. Kleywegt, G. *Validation of protein crystal structures.* Acta Crystallographica Section D, 2000. **56**(3): p. 249-265.

266. Vaguine, A.A., J. Richelle, and S.J. Wodak. *SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model.* Acta Crystallographica Section D, 1999. **55**(1): p. 191-205.

267. Farazi, T.A., G. Waksman, and J.I. Gordon. *The Biology and Enzymology of Protein N-Myristoylation.* 2001. **276**(43): p. 39501-39504.

268. Wood, A., A. Shilatifard, and C.C.a.J.W.C. Ronald, *Posttranslational Modifications of Histones by Methylation*, in *Advances in Protein Chemistry*. 2004, Academic Press. p. 201-222.

269. Hunter, T. *Tyrosine phosphorylation: thirty years and counting.* Current Opinion in Cell Biology. **In Press, Corrected Proof**.

270. Melton, R.G., C.N. Wiblin, R.L. Foster, and R.F. Sherwood. *Covalent linkage of carboxypeptidase G2 to soluble dextrans--I: Properties of conjugates and effects on plasma persistence in mice.* Biochemical Pharmacology, 1987. **36**(1): p. 105-112.

271. Bickerstaff, G.F., *Immobilization of Enzymes and Cells*. 1996, Humana Press. p. 1-11.

272. Davis, F.F. *The origin of pegnology.* Advanced Drug Delivery Reviews, 2002. **54**(4): p. 457-458.

273. Veronese, F.M. and G. Pasut. *PEGylation, successful approach to drug delivery.* Drug Discovery Today, 2005. **10**(21): p. 1451-1458.

274. Abuchowski, A., J.R. McCoy, N.C. Palczuk, T. van Es, and F.F. Davis. *Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase.* Journal of Biological Chemistry, 1977. **252**(11): p. 3582-3586.

275. Abuchowski, A., T. van Es, N.C. Palczuk, and F.F. Davis. *Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol.* Journal of Biological Chemistry, 1977. **252**(11): p. 3578-3581.

276. Leader, B., Q.J. Baca, and D.E. Golan. *Protein therapeutics: a summary and pharmacological classification.* Nat Rev Drug Discov, 2008. **7**(1): p. 21-39.

277. Harris JM, C.R. *Effect of pegylation on pharmaceuticals.* Nature Review Drug Discovery, 2003. **2**(3): p. 214-221.

278. Hamidi, M., A. Azadi, and P. Rafiei. *Pharmacokinetic Consequences of Pegylation.* Drug delivery, 2006. **13**(6): p. 399 - 409.

279. Veronese, F. and A. Mero. *The impact of PEGylation on biological therapies.* BioDrugs, 2008. **22**(5): p. 315-329.

280. Morar, A.S., J.L. Schrimsher, and M.D. Chavez. *PEGylation of Proteins: A Structural Approach.* BioPharm International, 2006.

281. Modi, M., N. Martin, J.E. Heathcote, K. Nieforth, and S. Zeuzem. *Peginterferon alfa-2a (40 kDa) monotherapy: a novel agent for chronic hepatitis C therapy.* Expert Opinion on Investigational Drugs, 2001. **10**(12): p. 2201-2213.

282. Selam, J.L. *Inhaled insulin for the treatment of diabetes: projects and devices.* Expert opinion in pharmacotherapy, 2003. **4**(8): p. 1373-1377.

283. Siekmeier, R. and G. Scheuch. *Inhaled insulin-does it become reality?* Jounal of Physiology and Pharmacology, 2008. **59**(6): p. 81-113.

284. Veronese, F.M. *Peptide and protein PEGylation: a review of problems and solutions.* Biomaterials, 2001. **22**(5): p. 405-417.

285. Bailon, P. and C.-Y. Won. *PEG-modified biopharmaceuticals.* Expert opinion on drug delivery, 2009. **6**(1): p. 1-16.

286. Roberts, M.J., M.D. Bentley, and J.M. Harris. *Chemistry for peptide and protein PEGylation.* Advanced Drug Delivery Reviews, 2002. **54**(4): p. 459-476.

287. Sato, H. *Enzymatic procedure for site-specific pegylation of proteins.* Advanced Drug Delivery Reviews, 2002. **54**(4): p. 487-504.

288. Woghiren C, B. Sharma, and S. Stein. *Protected thiol-polyethylene glycol: a new activated polymer for reversible protein modification.* Bioconjugate chemistry, 1993. **4**(5): p. 314-318.

289. Ishii, Y. and S.S. Lehrer. *Effects of the state of the succinimido-ring on the fluorescence and structural properties of pyrene maleimide-labeled alpha alpha-tropomyosin.* Biophysical Journal, 1986. **50**(1): p. 75-80.

290. Goodson, R. and N. Katre. *Site-directed pegylation of recombinant interleukin-2 at its glycosylation site.* Biotechnology 1990. **8**(4): p. 343-346.

291. Morpurgo, M., F.M. Veronese, D. Kachensky, and J.M. Harris. *Preparation and Characterization of Poly(ethylene glycol) Vinyl Sulfone.* Bioconjugate chemistry, 1996. **7**(3): p. 363-368.

292. Svergun, D.I., F. Ekström, K.D. Vandegriff, A. Malavalli, D.A. Baker, C. Nilsson, and R.M. Winslow. *Solution Structure of Poly(ethylene) Glycol-Conjugated Hemoglobin Revealed by Small-Angle X-Ray Scattering: Implications for a New Oxygen Therapeutic.* Biophysical Journal, 2008. **94**(1): p. 173-181.

293. Yanling Lu, S.E. Harding, A. Turner, B. Smith, D.S. Athwal, J.G. Grossmann, K.G. Davis, and A.J. Rowe. *Effect of PEGylation on the solution conformation of antibody fragments.* Journal of pharmaceutical sciences, 2008. **97**(6): p. 2062-2079.

294. Fee, C.J. and J.M. Van Alstine. *Prediction of the Viscosity Radius and the Size Exclusion Chromatography Behavior of PEGylated Proteins.* Bioconjugate chemistry, 2004. **15**(6): p. 1304-1313.

295. Veronese, F.M., P. Caliceti, and O. Schiavon. *Branched and Linear Poly(Ethylene Glycol): Influence of the Polymer Structure on Enzymological, Pharmacokinetic, and Immunological Properties of Protein Conjugates.* Journal of Bioactive and Compatible Polymers, 1997. **12**(3): p. 196-207.

296. Harris JM, Veronese FM, Caliceti P, and Schiavon O. *Multiarmed, monofunctional, polymer for coupling to molecules and surfaces.* USA patent US5932462, 1999.

297. Heymann, B. and H. Grubmüller. *Elastic properties of poly(ethylene-glycol) studied by molecular dynamics stretching simulations.* Chemical Physics Letters, 1999. **307**(5-6): p. 425-432.

298. Oesterhelt, F., M. Rief, and H.E. Gaub. *Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water.* New Journal of Physics, 1999. **1**(6): p. 6-11.

299. Fee, C.J. *Size comparison between proteins PEGylated with branched and linear poly(ethylene glycol) molecules.* Biotechnology and Bioengineering, 2007. **98**(4): p. 725-731.

300. Flory, P. *Principles of polymer chemistry.* Ithaca, New York: Cornell University Press, 1953.

301. Nakaoka, R., Y. Tabata, T. Yamaoka, and Y. Ikada. *Prolongation of the serum half-life period of superoxide dismutase by poly(ethylene glycol) modification.* Journal of Controlled Release, 1997. **46**(3): p. 253-261.

302. Venturoli, D. and B. Rippe. *Ficoll and dextran vs. globular proteins as probes for testing glomerular permselectivity: effects of molecular size, shape, charge, and deformability.* American journal of physiology, renal physiology, 2005. **288**(4): p. F605-613.

303. Molek, J.R. and A.L. Zydney. *Ultrafiltration characteristics of pegylated proteins.* Biotechnology and Bioengineering, 2006. **95**(3): p. 474-482.

304. Bradford, M.M. *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Analytical Biochemistry, 1976. **72**(1-2): p. 248-254.

305. Pabst, T.M., J.J. Buckley, and A.K. Hunter. *Comparison of strong anion-exchangers for the purification of a PEGylated protein.* Journal of Chromatography A, 2007. **1147**(2): p. 172-182.

306. Yamamoto S, Fujii S, Yoshimoto N, and Akbarzadehlaleh P. *Effects of protein conformational changes on separation performance in electrostatic interaction chromatography: unfolded proteins and PEGylated proteins.* Journal of Biotechnology, 2007. **132**(2): p. 196-201.

307. Digilio, G., L. Barbero, C. Bracco, D. Corpillo, P. Esposito, G. Piquet, S. Traversa, and S. Aime. *NMR Structure of Two Novel Polyethylene Glycol*

Conjugates of the Human Growth Hormone-Releasing Factor, hGRF(1-29)-NH2. Jounal of the American Chemical Society, 2003. **125**(12): p. 3458-3470.

308. Kurfürst, M. *Detection and molecular weight determination of polyethylene glycol-modified hirudin by staining after sodium dodecyl sulfate-polyacrylamide gel electrophoresis.* Analytical Biochemistry, 1992. **200**(2): p. 244-248.

309. Hinds, K.D. and S.W. Kim. *Effects of PEG conjugation on insulin properties.* Advanced Drug Delivery Reviews, 2002. **54**(4): p. 505-530.

310. Long, D.L., D.H. Doherty, S.P. Eisenberg, D.J. Smith, M.S. Rosendahl, K.R. Christensen, D.P. Edwards, E.A. Chlipala, and G.N. Cox. *Design of homogeneous, monopegylated erythropoietin analogs with preserved in vitro bioactivity.* Experimental Hematology, 2006. **34**(6): p. 697-704.

311. Dong Hee Na, Y.S.Y.K.C.L. *Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for monitoring and optimization of site-specific PEGylation of ricin A-chain.* Rapid communications in mass spectrometry, 2004. **18**(18): p. 2185-2189.

312. Hagel, L., *Gel filtration. In Protein Purification: Principles,High Resolution Methods and Applications* 1989, J.C. Janson and L. Ryden editions, VCH Publishers: New York. p. 63-106.

313. Brent Irvine, G., *Unit 5.5: Determination of Molecular Size by Size-Exclusion Chromatography (Gel Filtration)*, in *Current Protocols in Cell Biology*. 2000, John Wiley & Sons.

314. Chunyang Y Zheng, Guanghui Ma, and Z. Su. *Native PAGE eliminates the problem of PEG-SDS interaction in SDS-PAGE and provides an alternative to HPLC in characterization of protein PEGylation.* Electrophoresis, 2007. **28**(16): p. 2801-2807.

315. Odom, O.W., W. Kudlicki, G. Kramer, and B. Hardesty. *An Effect of Polyethylene Glycol 8000 on Protein Mobility in Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis and a Method for Eliminating This Effect.* Analytical Biochemistry, 1997. **245**(2): p. 249-252.

316. Irvine, G., *Unit 5.5: Determination of Molecular Size by Size-Exclusion Chromatography (Gel Filtration)*, in *Current Protocols in Cell Biology*. 2000, John Wiley & Sons.

317. Teraoka, I. *Calibration of Retention Volume in Size Exclusion Chromatography by Hydrodynamic Radius.* Macromolecules, 2004. **37**(17): p. 6632-6639.

318. Veronese FM and J. Harris. *Introduction and overview of peptide and protein pegylation.* Advanced Drug Delivery Reviews, 2002. **54**(4): p. 453-456.

319. L.A. Feigin and D.I. Svergun, *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*, ed. Plenum Press. 1987, New York.

320. Lee, J. and D. Berns. *Protein aggregation. The effect of deuterium oxide on large protein aggregates of C-phycocyanin.* Biochemical Journal, 1968. **110**(3): p. 465-470.

321. Pace, C.N., B.A. Shirley, M. McNutt, and K. Gajiwala. *Forces contributing to the conformational stability of proteins.* Faseb Journal, 1996. **10**(1): p. 75-83.

322. Frauenfelder, H. *Energy landscape and dynamics of biomolecules Extended abstract.* Journal of Biological Physics, 2005. **31**(3-4): p. 413-416.

323. Daniel, R.M., M.J. Danson, R. Eisenthal, C.K. Lee, and M.E. Peterson. *New parameters controlling the effect of temperature on enzyme activity.* Biochemical Society Transactions, 2007. **035**(6): p. 1543-1546.

324. Hitchings, G., H. *Selective Inhibitors of Dihydrofolate Reductase (Nobel Lecture).* Angewandte Chemie International Edition in English, 1989. **28**(7): p. 879-885.

325. Kroemer, R.T. *Structure-Based Drug Design: Docking and Scoring.* Current Protein & Peptide Science, 2007. **8**(4): p. 312-328.

326. Sousa, S.F., P.A. Fernandes, and M.J. Ramos. *Protein-ligand docking: Current status and future challenges.* Proteins: Structure, Function, and Bioinformatics, 2006. **65**(1): p. 15-26.

327. B-Rao, C., J. Subramanian, and S.D. Sharma. *Managing protein flexibility in docking and its applications.* Drug Discovery Today. **In Press, Corrected Proof**.

328. Totrov, M. and R. Abagyan. *Flexible ligand docking to multiple receptor conformations: a practical alternative.* Current Opinion in Structural Biology, 2008. **18**(2): p. 178-184.

329. Lerner, M.G., A.L. Bowman, and H.A. Carlson. *Incorporating Dynamics in E. coli Dihydrofolate Reductase Enhances Structure-Based Drug Discovery.* Journal of chemical information and modeling, 2007. **47**(6): p. 2358-2365.

330. Wang, L., S. Tharp, T. Selzer, S.J. Benkovic, and A. Kohen. *Effects of a Distal Mutation on Active Site Chemistry.* Biochemistry, 2006. **45**(5): p. 1383-1392.

331. Wong, K.F., T. Selzer, S.J. Benkovic, and S. Hammes-Schiffer. *Impact of distal mutations on the network of coupled motions correlated to hydride transfer in dihydrofolate reductase.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(19): p. 6807-6812.

Abstract

The aim of this thesis is to investigate the role of fast protein dynamics (pico-second timescale) in enzyme activity and stability, and specifically to test the hypothesis that enzyme activity and stability are inversely related by their internal dynamics.

Activity ∝ Dynamics (flexibility) ∝ 1/Stability

In order to test this hypothesis, the well known anti-cancer drug: methotrexate was used as an informative ligand in the network established between these properties. A multidisciplinary approach combining neutron scattering, circular dichroism, UV absorption, isothermal titration calorimetry and X-ray crystallography was undertaken to examine the current paradigm using the enzyme: dihydrofolate reductase as a model.

As inferred by neutron spectroscopy, the binding of MTX influences the dynamical behavior of DHFR. Macromolecular dynamics such as the resilience: $\langle k \rangle$ (i.e. structural rigidity) was found to be increased and, inversely, the flexibility decreased upon MTX binding. In addition, as revealed by circular dichroism, this dynamical dependency upon MTX binding was correlated with an enhanced thermal stability. Compared to the free enzyme, the melting temperature was found to be increased by 13.8 °C in the presence of MTX. The inhibitory power of MTX was also examined by steady state kinetics and isothermal titration calorimetry. The K_i for MTX was found to be in the nanomolar range K_i= 10.9 nM. Using isothermal titration calorimetry, the binding thermodynamic signature between MTX and DHFR was characterized. The binding event was found to be largely favourable (ΔG_b=-12.1 Kcal mol^{-1}), enthalpy driven (ΔH_b= -16.8 Kcal mol^{-1}) with an unfavourable entropy ΔS_b=-15.6 cal $K^{-1}mol^{-1}$.

In conclusion, the modulation of the macromolecular dynamics may reflect how specific conformations are favoured for subsequent protein function in response of the binding of specific ligand and how conformational substates approach to protein function. In this context the unprecedented power of transition state analogs

such as MTX on protein function might therefore be dependent on fast protein dynamics.

Keywords: Incoherent neutron scattering, circular dichroism, isothermal titration calorimetry, steady state kinetics, dihydrofolate reductase, methotrexate.