Computer Modelling and Simulations of Enzymes and their Mechanisms

Hernán Alonso

A thesis submitted for the degree of Doctor of Philosophy of the Australian National University

September 2006

Publications from this work

Alonso H, Bliznyuk AA and Gready JE. **Combining docking and molecular dynamic simulations in drug design**. Med Res Rev 26:531-568 (2006).

Alonso H and Gready JE. **Integron-sequestered dihydrofolate reductase, a recently redeployed enzyme**. Trends Microbiol 14:236-242 (2006).

Alonso H, Gillies MB, Cummins PL, Bliznyuk AA and Gready JE. **Multiple ligand-binding modes in bacterial R67 dihydrofolate reductase.** J Comput Aided Mol Des 19:165-187 (2005).

A pdf file of each publication can be found in a supplementary CD at the back of the thesis.

Statement of Originality

The contents of this thesis are the result of my original research, which has been conducted under the principal supervision of Prof. Jill E. Gready (The John Curtin School of Medical Research, ANU).

Prof. Elizabeth E. Howell (University of Tennessee) provided the coordinates of some published docked complexes of folate and NADPH within R67 DHFR, which were used to generate a complete R67•$DHFH^{+}$•NADPH structure, which I referred to as *howH* in the analysis of Chapter 3.

Dr Malcolm Gillies produced the topology files for the ligands dihydrofolate and NADPH used for the GROMACS MD simulations presented in Chapter 3.

The X-ray crystal structure of the enzyme MeTr complexed with methyl-tetrahydrofolate, used for the analysis in Chapter 4, was provided to Prof. Jill Gready in confidence by Prof. Stephen W. Ragsdale (University of Nebraska).

I declare that the work presented in this thesis is, to my belief, original, except as acknowledged in the text above, and that the material has not been submitted, either in whole or in part, for a degree at this or any other university.

Hernán Alonso

September 2006

Acknowledgements

This PhD thesis would have not been possible without the invaluable supervision of Prof Jill Gready. I am grateful for her honest and thorough guidance, constructive advice and constant support. Her enthusiasm for science and inquisitive nature have set a clear example for me to follow. I hope there will be many opportunities in the future to further work and learn from her.

My special thanks to Dr Peter Cummins and Dr Andrey Bliznyuk who have guided me during these three and a half years, and have provided me with helpful and essential comments on my work, while helping me to become a computational biologist. I am particularly thankful to Dr Malcolm Gillis, who acted as my nanny during the first months of my thesis, showing me how to deal with scientific problems and tackle them with computers.

I would like to express my gratitude to all those people who facilitated my work by sharing their results with me: to Prof Elizabeth Howell (University of Tennessee), for providing the coordinates of R67 DHFR complexes and disclosing unpublished experimental results; and to Prof Stephen Ragsdale and Dr Tzanko Doukov (University of Nebraska), for their unpublished X-ray structure of MeTr complexed with CH_3THF.

I would also like to acknowledge the benefit of the valuable discussions with Prof Ruth Hall (University of Sydney) regarding integrons and gene cassettes, as well as Prof Rowena Matthews (University of Michigan) for her constructive advice on experimental results for methionine synthase, and to Dr Ivan Rostov (ANU Supercomputer Facility), for his help and advice with calculations in general and the ONIOM method in particular.

From a personal point of view, all these years in Australia would not have been the same without the invaluable friendship of Alex Zelensky, and those hundreds of lunches shared with the coolest Indians ever, Babu Kannappan and Nagesh Chakka.

My gratitude extends to all the members of the Computational Proteomics and Therapy Design Group, for creating a great work environment where I have been able to grow both as a scientist and as a person over the past several years.

As a proud member of the Fenner Hall community, I am sincerely grateful to all those superb people that form this little world. Fenner Hall is a place where students from all over the world live together in a friendly and supportive environment, and a place that I have grown up to call my home.

Finally, I would like to thank my family, especially my Mum, my sister Lucía, my grandma Chola and my grandpa Coco, for their unconditional support and love.

This thesis is dedicated to the courageous reader, who dares to go through hundreds of pages in search of some (hopefully) valuable information…

Hernán Alonso,

September 2006.

Abstract

Although the tremendous catalytic power of enzymes is widely recognized, their exact mechanisms of action are still a source of debate. In order to elucidate the origin of their power, it is necessary to look at individual residues and atoms, and establish their contribution to ligand binding, activation, and reaction. Given the present limitations of experimental techniques, only computational tools allow for such detailed analysis. During my PhD studies I have applied a variety of computational methods, reviewed in Chapter 2, to the study of two enzymes: DfrB dihydrofolate reductase (DHFR) and methyltetrahydrofolate: corrinoid/iron-sulfur protein methyltransferase (MeTr).

The DfrB enzyme has intrigued microbiologists since it was discovered thirty years ago, because of its simple structure, enzymatic inefficiency, and its insensitivity to trimethoprim. This bacterial enzyme shows neither structural nor sequence similarity with its chromosomal counterpart, despite both catalysing the reduction of dihydrofolate (DHF) using NADPH as a cofactor. As numerous attempts to obtain experimental structures of an enzyme ternary complex have been unsuccessful, I combined docking studies and molecular dynamics simulations to produce a reliable model of the reactive DfrB•DHF•NADPH complex. These results, combined with published empirical data, showed that multiple binding modes of the ligands are possible within DfrB.

Comprehensive sequence and structural analysis provided further insight into the DfrB family. The presence of the *dfrB* genes within integrons and their level of sequence conservation suggest that they are old structures that had been diverging well before the introduction of trimethoprim. Each monomer of the tetrameric active enzyme presents an SH3-fold domain; this is a eukaryotic auxiliary domain never found before as the sole domain of a protein, let alone as the catalytic one. Overall, DfrB DHFR seems to be a poorly adapted catalyst, a 'minimalistic' enzyme that promotes the reaction by facilitating the approach of the ligands rather than by using specific catalytic residues.

MeTr initiates the Wood-Ljungdahl pathway of anaerobic CO_2 fixation. It catalyses the transfer of the N5-methyl group from N5-methyltetrahydrofolate (CH_3THF) to the cobalt centre of a corrinoid/iron-sulfur protein. For the reaction to occur, the N5 position of CH_3THF is expected to be activated by protonation. As experimental studies have led to conflicting suggestions, computational approaches were used to address the activation mechanism.

Initially, I tested the accuracy of quantum mechanical (QM) methods to predict protonation positions and pK_as of pterin, folate, and their analogues. Then, different protonation states of CH_3THF and active-site aspartic residues were analysed. Fragment QM calculations suggested that the pK_a of N5 in CH_3THF is likely to increase upon protein binding. Further, ONIOM calculations which accounted for the complete protein structure indicated that active-site aspartic residues are likely to be protonated before the ligand. Finally, solvation and binding free energies of several protonated forms of CH_3THF were compared using the thermodynamic integration approach. Taken together, these preliminary results suggest that further work with particular emphasis on the protonation state of active-site aspartic residues is needed in order to elucidate the protonation and activation mechanism of CH_3THF within MeTr.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

B3LYP	Becke 3-Parameter (Exchange), Lee, Yang and Parr (DFT)
C/Fe-SP	Corrinoid/iron-sulfur protein
CH_3THF	N5-methyl-5,6,7,8-tetrahydrofolate
DFT	Density functional theory
DHF	7,8-dihydrofolate
DHFR	Dihydrofolate reductase
DHPS	Dihydropteroate synthase
FEP	Free energy perturbation
GA	Genetic algorithm
HF	Hartree Fock theory
LBHB	Low-barrier hydrogen bond
LGA	Lamarckian genetic algorithm
LIE	Linear integration energy
MC	Monte Carlo
MD	Molecular dynamics
MetH	Cobalamin-dependent methionine synthase
$MetH^{(2-649)}$	Homocystein- and CH_3THF-binding domains of MetH
MeTr	Methyltetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase
MM	Molecular mechanics
MM-PBSA	Molecular Mechanics/Poisson-Boltzmann Surface Area
MP2	Second-order Moller-Plesset perturbation theory
NAC	Near attack conformers
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced form
$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidized form
NMR	Nuclear magnetic resonance
ONIOM	Our own N-Layered Integrated molecular Orbital and molecular Mechanics
ORF	Open reading frame
PBC	Periodic boundary conditions

PCM	Polarisable continuum model
PDB	Protein data bank
PME	Particle mesh Ewald
PMF	Potential of mean force
QM	Quantum mechanics
REMD	Replica exchange molecular dynamics
RESP	Restrained electrostatic potential
RMSD	Root mean square deviation
THF	5,6,7,8-tetrahydrofolate
TI	Thermodynamic integration
TMP	Trimethoprim
TS	Transition state
vdW	van der Waals