

Appendix A:

ANU VPP300, SGI Power Challenge and APAC National Facility Compaq AlphaServer SC

Individual Research Reports

This section presents individual reports on ANU projects supported by the Supercomputer Time Allocation Committee (STAC) on the VPP300 and SGI Power Challenge, as well as both STAC and Merit Allocation Scheme (MAS) supported projects on the Compaq SC led by ANU Principal Investigators.

The reports were prepared by the individual Principal Investigators in a standard format. In some cases several projects are combined into a single report. Reports were not received for a small number of projects.

The reports are listed alphabetically by surname of the Principal Investigator; as in the preceding Tables.

Estimating Trends in HIV Incidence for Australian Women

Principal Investigator

Niels Becker

Nat Centre for
Epidemiology and Pop
Health
ANU

Co-Investigator

Zhengfeng Li

NCEPH
ANU

The work is concerned with a project funded by an NHMRC grant for which Prof Niels Becker, NCEPH, is the principal investigator and Dr. Zhengfeng Li is a team member. The aim of the project is to apply a newly developed method, using linked data on HIV and AIDS diagnoses, to reconstruct the infection curve and thereby estimate trends in HIV incidence for Australian women who were infected heterosexually. In mathematical terms, we are estimating curves and computing the precision of the estimated curves. The computations use an algorithm adapted from the EM-algorithm, by adding a smoothing step and a Newton-Raphson step.

Project

x00

Facilities Used

SC

RFCD Codes

239901

Significant Achievements, Anticipated Outcomes and Future Work

We find that estimating the extra parameters is feasible, producing estimates with good precision. Including age as a covariate gives 90 % confidence intervals for the HIV incidence curve that are about 20 % narrower than those obtained when age data are not used. Our finding has been reported in a paper of “Age-specific backproject of HIV diagnosis data” which was submitted to the journal of Statistics In Medicine for publication.

Computational Techniques Used

The code was implemented in C++.

Publications, Awards and External Funding

The project was funded by NHMRC.

N.G. Becker, J.C. Lewis, Z.F. Li and A. McDonald, Age-specific backproject of HIV diagnosis data, submitted to Statistics in Medicine.

Applications of Integer Factorization

Principal Investigator

Richard Brent

Computing Lab
Oxford University, UK

Co-Investigators

John Cannon

Computer Sciences
Laboratory
University of Sydney

Brendan McKay

Computer Science
Faculty of Science
ANU

The integer factorization problem is an essential ingredient of public-key cryptosystems such as the RSA system, which is used to ensure the security of many transactions on the Internet. Thus, there is a great interest in discovering the best integer factorization algorithms, because the choice of key sizes in the RSA system depends on how quickly large integers can be split into their prime factors. This project involves the implementation and testing of integer factorization algorithms.

Project

k58

Facilities Used

PC, VPP, MDSS

RFCD Codes

230102, 280401

Significant Achievements, Anticipated Outcomes and Future Work

We have demonstrated improvements to the elliptic curve algorithm (ECM) which has been used to find factors up to 55 decimal digits.

Brian Murphy (former Ph.D. student), in collaboration with Peter Montgomery (Microsoft and CWI), has improved the polynomial selection phase of the number field sieve algorithm. Murphy's polynomial selection algorithm was used in the record-breaking factorisations of a 512-bit (155 decimal digit) number and (in January 2002) of a 158-decimal digit number by Bahr, Franke and Kleinjung.

We have extended the tables of factors of numbers of special forms and incorporated these tables in the Magma package of John Cannon. All relevant numbers with fewer than 108 decimal digits have now been factored.

Computational Techniques Used

The numbers of interest in cryptography and some other applications such as the analysis of very long period random number generators are very large, so a large amount of computation is required, even when the most efficient known algorithms are used. A quantum computer using Shor's algorithm would be helpful, but no such computer is available at the present time.

Our programs include a vectorized implementation of Lenstra's elliptic curve algorithm with a second phase, based on the "birthday paradox", which was developed at ANU. We also use the Magma package and an implementation in Magma of the PMPQS quadratic sieve algorithm which was developed by collaborators at the University of Sydney and elsewhere, and an implementation of the number field sieve (NFS) algorithm which was developed mainly by collaborators at the Centre for Mathematics and Computer Science (CWI), Amsterdam.

Publications, Awards and External Funding

R. P. Brent, Factorization of the tenth Fermat number, *Mathematics of Computation* 68 (1999), 429-451.

R. P. Brent, Large factors found by ECM, available from <ftp://ftp.comlab.ox.ac.uk/pub/Documents/techpapers/Richard.Brent/champs.txt>

R. P. Brent, R. E. Crandall, K. Dilcher and C. Van Halewyn, Three new factors of Fermat numbers, *Mathematics of Computation* 69 (2000), 1297-1304.

R. P. Brent, Some parallel algorithms for integer factorisation, *Lecture Notes in Computer Science* 1685 (1999), 1-22.

R. P. Brent, Recent progress and prospects for integer factorisation algorithms, *Lecture Notes in Computer Science* 1858 (2000), 3-22.

R. P. Brent, P. L. Montgomery and H. J. J. te Riele, Factorizations of Cunningham numbers with bases 13 to 99, Report PRG-TR-14-00, December 2000, 502pp.

Mechanism of the GSTT2 Enzyme

Principal Investigator

Gareth Chelvanayagam
Human Genetics
JCSMR
ANU

Co-Investigators

Phil Board
Jack Flanagan
Human Genetics
JCSMR
ANU
William King
Molecular Biology
JCSMR
ANU

Michael Parker
Ian Potter Foundation
Protein Crystallography
Laboratory
St Vincent's Institute of
Medical Research

Project
x06, x07

Facilities Used
PC, VPP, MDSS

RFCD Codes
270108

Glutathione transferases (GSTs) are enzymes that play a role in detoxication: removing harmful substances from a cell. They do this by linking a molecule, glutathione, to the harmful substance. GSTs allow this reaction to take place by facilitating the removal of a Hydrogen atom from a Sulfur atom in glutathione, forming a reactive thiolate anion. It is this anion that links to the substance to be removed. There are several classes of GSTs. The Theta and the Delta classes differ from the Alpha, Pi, Mu, and Sigma classes in that they lack the Tyr amino acid that stabilises the glutathione thiolate anion. Stabilisation is achieved through the Tyr hydroxyl group. Focusing on the Theta class enzyme GSTT2, the aim of this work is to investigate the role of a Ser amino acid in the active site that is also capable of donating a hydrogen bond through its hydroxyl group. To do this, high level ab initio calculations on model atomic systems of the active site were performed.

Significant Achievements, Anticipated Outcomes and Future Work

It has been previously established that the role of the Tyr residue in the Alpha, Pi, Mu, and Sigma class GSTs is one of a surrogate water molecule, donating a proton through a conventional hydrogen bond to stabilise the thiolate. We have established that the hydrogen bond strength of the interaction between the GSTT2 Ser 11 and the glutathione thiolate is much weaker than in classes with an active site Tyr. Further, we have discovered second sphere interactions that, depending on the amino acids involved, markedly influence the first sphere hydrogen bond. In GSTT2, the interaction of Ser 14 with Ser 11 significantly increases the energy of the bond.

Based on these studies it can be hypothesised that by altering the position of the first sphere interaction, the thiolate could be oriented differently, thereby facilitating catalysis or transition state stabilisation of particular reactions.

Computational Techniques Used

This study was performed using the Gaussian 98 Revision A.6 (Gaussian Inc.) package, running on the ANU Supercomputing Facility SGI Power Challenge. Atomic coordinates of the crystal structure of the GSTT2 enzyme were obtained from the authors. Based on this structure, model systems of the Ser:thiolate first sphere interaction were derived. The Ser residue was modeled as methanlo, whereas the thiol of the glutathione was modeled as methanethiol. Different orientations of the hydrogen bond donor systems were considered in both a neutral and anionic state. Similar second sphere interactions were considered involving Ser 14:Ser 11:thiolate and Cys 14:Ser 11:thiolate. Calculations were performed using RHF and MP2 levels of theory with different basis sets.

Publications, Awards and External Funding

External Funding:

This work was supported by the Ramaciotti Foundation, the ANU and the ARC.

Publications:

J.U. Flanagan, W. King, M.W. Parker, P.G. Board and G. Chelvanayagam, Ab Initio Calculations on Hidden Modulators of Theta Class Glutathione Transferase Activity, Proteins: Structure, Function, and Genetics 39, 2000, 235-243.

Permeation of Ions Through Membrane Channels - Molecular Dynamics Studies

Principal Investigator

Shin-Ho Chung

*Chemistry
Faculty of Science
ANU*

Co-Investigators

Turgut Bastug

Ben Corry

Matthew Hoyles

Taira Vora

*Chemistry
Faculty of Science
ANU*

Serdar Kuyucak

Megan O'Mara

*Theoretical Physics
RSPHysSE
ANU*

Roger Brown

*Supercomputer Facility
ANU*

David Saint

*Department of
Physiology
University of Adelaide*

Toby Allen

*Biochemistry /
Physiology and
Biophysics
Weill Medical College of
Cornell University
New York, USA*

Project

r06, x26, w16, w50, d32

Facilities Used

VPP, SC, PC, MDSS

RFCD Codes

270104

Because all electrical activities in the brain are regulated by opening and closing of ion channels, understanding their mechanisms at a microscopic level is a fundamental problem in biology. Making full use of some recent technological and conceptual advances, we are studying, using computer simulations, how ions navigate through a narrow pore formed by protein molecules. From the results of extensive computer simulations, we propose to deduce mathematically the inner working details of membrane channels, information that may lead us to find the causes of and, possibly, cures for many of the neurological and muscular disorders. The research projects we have embarked upon can only be carried out with fast supercomputers. Thus, the APAC National Facility is indispensable for our research.

Significant Achievements, Anticipated Outcomes and Future Work

The field of ion channels has entered into a rapid phase of development in the last few years, partly due to the breakthroughs in determination of the crystal structures of membrane proteins and advances in computer simulations of biomolecules. These advances have finally enabled the long-dreamed goal of relating function of a channel to its underlying atomic structure through the fundamental processes operating in electrolyte solutions. Using the newly unravelled structural information, we have succeeded in elucidating the mechanisms underlying the permeation of ions across the potassium channel, and explaining how the channel allows potassium ions to move across, while rejecting sodium ions. The experimentally-determined structure of the potassium channel is one of many different types found in nature, which differ widely in their conductances and gating characteristics. We employed molecular dynamics calculations and Brownian dynamics simulations to demonstrate that the widely differing properties of potassium channels found in nature can be understood by small modifications of the channel geometry.

On the basis of known experimental properties of other channels, whose crystal structures are not yet determined, and insights gathered from our studies on the potassium channel, we are planning to build their models and deduce some of the salient structural properties. Among the channels we wish to study are the L-type calcium channel, which is necessary for neurotransmission, muscle contraction and neurochemical modulation, the sodium channel, the activation of which generates nerve impulses, and the glycine channel, which by allowing chloride ions to move inside of the cell reduces neuronal excitability. We anticipate that each of the channel models we deduce using the computational methods will closely reproduce its experimental properties and capture some of the salient structural features of the channel protein.

Computational Techniques Used

The main computational tools that we make use of in our studies are molecular dynamics calculations and stochastic dynamics simulations. Of the several tools in statistical mechanics that treat the dynamics of nonequilibrium systems, Brownian dynamics (which we use) is the most widely known. The algorithm and code for Brownian dynamics were developed and gradually refined by our group over the past 5 years. In Brownian dynamics simulations, the Langevin equation is solved repeatedly to trace the trajectory of every ion in the assembly. Snapshots of the simulation system are taken at short time intervals for many millions of timesteps. At each time, the Langevin equation is integrated to obtain the velocity of each ion to determine to which position the ion will move in the next timestep. The new coordinates of all ions in the assembly are deduced, and the calculation is repeated. By repeating this process for a sufficiently long period of time, we can deduce how many ions move across the channel in a fixed period of simulation time. Without the computational power provided by the APAC National Facility, we would not have been able to carry out any of our studies.

For molecular dynamics calculations, we make use of two commercial packages - CHARMM and GROMAC. These packages enable us to follow the trajectories of N particles interacting via a many-body potential using Newton's equation of motion. The trajectory data so generated are stored at certain intervals, which are analyzed subsequently to determine the structural

and dynamical properties of a system. Quantities such as free energy, mean square displacement, radial distribution function and other correlations functions are calculated from an ensemble average of several simulations.

Publications, Awards and External Funding

External Funding

NHMRC (2000-2004) - Approx. \$1,000,000

ARC Discovery (2002-2004) - Approx. \$220,000

APAC (2002) - Approx. \$70,000

Publications

B. Corry, S. Kuyucak, S.H. Chung, Test of Poisson-Nernst-Planck Theory in Ion Channels, *Journal of General Physiology*, 114, 1999, 597-599.

T.W. Allen, M. Hoyles, S. Kuyucak, S.H. Chung, Molecular and Brownian Dynamics Study of Ion Selectivity and Conductivity in the potassium Channel, *Chemical Physics Letters*, 313, 1999, 358-365.

T.W. Allen, S. Kuyucak, S.H. Chung, Molecular Dynamics Study of the KcsA Potassium Channel, *Biophysical Journal* 77, 1999, 2502-2516.

S.H. Chung, T.W. Allen, M. Hoyles, S. Kuyucak, Permeation of Ions across the Potassium Channel: Brownian Dynamics Studies, *Biophysical Journal* 77, 1999, 2517-2533.

T.W. Allen, S. Kuyucak, S.H. Chung, The Effect of Hydrophobic and Hydrophilic Channel Walls on the Structure and Diffusion of Water and Ions, *Journal of Chemical Physics*, 111, 1999, 7985-7999.

D. Saint, S. H. Chung, Variability of Channel Subconductance States of the Cardiac Sodium Channel Induced by Protease, *Receptors and Channels*, 6, 1999, 283-294.

D. S. Poskitt, K. Dogancay, S. H. Chung, A New Analytical Method of Studying Post-Synaptic Currents in the Brain, *Mathematical Biosciences*, 161, 1999, 15-41, 1999.

D. Poskitt, K. Dogancay, S. H. Chung, Double Bblind Deconvolution: the Analysis of Post-Synaptic Currents in Nerve Cells, *Journal of the Royal Statistical Society*, B61, 1999, 191-212.

B. Corry, S. Kuyucak, S. H. Chung, Invalidity of Continuum Theories of Electrolytes in Nanopores. *Chemical Physics Letters*, 320, 2000, 35-41.

G. Moy, B. Corry, S. Kuyucak, S. H. Chung, Tests of Continuum Theories as Models of Ion Channels: I. Poisson-Boltzmann Theory Versus Brownian Dynamics. *Biophysical Journal*, 78, 2000, 2349-2363.

B. Corry, S. Kuyucak, S. H. Chung, Tests of Continuum Theories as Models of Ion Channels: II. Poisson-Nernst-Planck Theory Versus Brownian Dynamics, *Biophysical Journal*, 78, 2000, 2364-2381.

- T. Allen, S. Kuyucak, S. H. Chung, Molecular Dynamics Estimates of Ion Diffusion in Model Hydrophobic and the KcsA Potassium Channels. *Biophysical Chemistry*, 86, 2000, 1-14.
- T. Allen, A. Bliznyuk, A. P. Rendell, S. Kuyucak, S. H. Chung, The Potassium Channel: Structure, Selectivity and Diffusion, *Journal of Chemical Physics*, 112, 2000, 8191-8204.
- B. Corry, T. W. Allen, S. Kuyucak, S. H. Chung, A Model of Calcium Channels, *Biochimica et Biophysica Acta - Biomembranes*, 1509, 2000, 1-6.
- S. H. Chung, S. Kuyucak, Predicting Channel Function from Channel Structure Using Brownian Dynamics Simulations, *Clinical & Experimental Pharmacology & Physiology*, 28, 2001, 89-94.
- B. Corry, T. W. Allen, S. Kuyucak, S. H. Chung, Mechanisms of Permeation and Selectivity in Calcium Channels, *Biophysical Journal*, 80, 2001, 195-214.
- T. Allen, S. H. Chung, Brownian Dynamics Study of an Open-State KcsA Potassium Channel, *Biochimica et Biophysica Acta - Biomembranes* 1515, 2001, 83-91, 2001.
- S. H. Chung, T. W. Allen, S. Kuyucak. Conducting-State Properties of the KcsA Potassium Channel from Molecular and Brownian Dynamics Simulations, *Biophysical Journal*, 82, 2002, 628-645.
- S. Kuyucak, O. S. Andersen, S. H. Chung, Models of Permeation in Ion Channels, *Reports on Progress in Physics* 64, 2001, 1427-1472.
- A. A. Bliznyuk, A. P. Rendell, T. W. Allen, S. H. Chung, The Potassium Channel: Comparison of Linear Scaling Semiempirical and Molecular Mechanics Representations of the Electrostatic Potentiation, *Journal of Physical Chemistry*, 2001, in press.
- S. Kuyucak, S. H. Chung, Permeation Models and Structure-Function Relationships in Ion Channels, *Journal of Biological Physics*, 2001, in press.
- S. H. Chung, T. W. Allen, S. Kuyucak, Modeling Diverse Range of Potassium Channels with Brownian Dynamics, *Biophysical Journal*, 2002, in press.
- B. Corry, M. Hoyles, T. Allen, M. Walker, S. Kuyucak, S. H. Chung, Reservoir Boundaries in Brownian Dynamics Simulations of Ion Channels, *Biophysical Journal*, 2002, in press.
- S. H. Chung, S. Kuyucak, Ion Channels: Recent Progress and Prospects, *European Biophysics Journal*, 2002, in press.
- S. H. Chung, S. Kuyucak, Recent Advances in Ion Channel Research, *Biochimica et Biophysica Acta - Biomembranes*, 2002, in press.

Molecular potential energy surfaces

Principal Investigator

Michael Collins

*Research School of
Chemistry
ANU*

Co-Investigators

Ryan Bettens

Alexander Duncan
Christian Evenhuis
Michael Smith

*Research School of
Chemistry
ANU*

Keiran Lim

*Centre for Chiral and
Molecular Technologies
School of Biological and
Chemical Sciences
Deakin University*

Donghui Zhang

*Department of
Computational Science
National University of
Singapore*

This project develops and applies methods for constructing, from first principles, the molecular potential energy surface (PES) for molecules undergoing chemical reactions, exchanging energy in collisions, and clustering together. This allows us to calculate how the atoms move during the break-up and formation of new molecules. In this way we can understand the mechanisms and rates of important chemical reactions that take place in the atmosphere, in combustion processes and in inter-stellar space, for example. The study of energy exchange and clustering will ultimately allow us to understand chemical reactions at high pressure and in liquids.

Project

s01

Facilities Used

SC, VPP, PC, Vizlab

RFCD Codes

250603

Significant Achievements, Anticipated Outcomes and Future Work

Dr Collins, in collaboration with Dr Keiran Lim (Deakin University), developed a program to construct *ab initio* molecular potential energy surfaces (PES) to describe energy transfer between molecules in gas-phase collisions. The software can now describe the PES as a many body expansion. The code was extended to simulate gas phase collisions in the canonical ensemble for more direct comparison with experimental data. The accuracy of this approach

has been proved by comparison with known “model” potentials. *Ab initio* PESs have now been constructed at the Hartree-Fock level of theory with various basis sets. A series of PES at higher levels of *ab initio* theory will also be constructed to complete an investigation of the effects of electron correlation and basis set completeness on collisional energy transfer. Collins has written a new version of the program package which is, in principle, capable of constructing the many-body expansion potential energy surfaces for clusters of molecules. Future work will revolve around refining the methodology to provide accurate interactions of many water molecules, so that cooperative hydrogen bonding can be investigated from first principles.

Dr Collins, in collaboration with Assoc. Professor Donghui Zhang, completed and published the PES and quantum reaction dynamics for the $\text{H}_2 + \text{OH} \rightarrow \text{H} + \text{H}_2\text{O}$ reaction. The quantum reactive scattering calculations for $\text{H} + \text{CH}_4 \rightarrow \text{H}_2 + \text{CH}_3$ reaction are still in progress (the PES was completed last year). Meanwhile, Collins and Zhang have completed construction of the PES for OH_3^- , and are preparing publication of this surface and the quantum calculation of the photoelectron spectrum of this anion.

Collins and Mr Michael Smith (PhD student) have completed construction of a PES for the important combustion reaction, $\text{H} + \text{HCO} \rightarrow \text{H}_2 + \text{CO}$. Collins and Mr Christian Evenhuis (research assistant funded by APAC) have begun the development of a new program to evaluate the multiple potential energy surfaces and coupling surfaces for reactions involving multiple electronic states. Professor David Yarkony, an expert in this type of electronic structure, visited ANU (partly funded by APAC) to facilitate this new research direction.

Computational Techniques Used

The program packages developed in the Collins group use a number of methods. Molecular dynamics simulations and Bayesian statistics are employed along with *ab initio* quantum chemistry techniques.

Publications, Awards and External Funding

M. A. Collins, R. P. A. Bettens, Potential energy surface for the reactions

$\text{BeH}_2 + \text{H} \rightleftharpoons \text{BeH} + \text{H}_2$, Phys. Chem. Chemical Physics, 1, 1999, 939-945.

R. P. A. Bettens, M. A. Collins, Learning to interpolate molecular potential energy surfaces with confidence: A Bayesian approach, Journal of Chemical Physics, 111, 1999, 816-826.

R. P. A. Bettens, T. A. Hansen, M. A. Collins, Interpolated potential energy surface and reaction dynamics for $\text{O}(^3\text{P}) + \text{H}_3^+(^1\text{A}_1')$ and $\text{OH}^+(^3\text{S}^-) + \text{H}_2(^1\text{S}_g^+)$, Journal of Chemical Physics 111, 1999, 6322-6332.

A. H. Duncan, M. A. Collins, Construction of interpolated potential energy surfaces using constrained dynamics: Application to rotational inelastic scattering, Journal of Chemical Physics 111, 1999, 1346-1353.

M. A. Collins, D. H. Zhang, Application of interpolated potential energy surfaces to quantum reactive scattering, Journal of Chemical Physics, 111, 1999, 9924-9931.

- R. P. A. Bettens, M. J. T. Jordan, D. H. Zhang, M. A. Collins, Ab initio potential energy surface for the reactions between H_2O and H, *Journal of Chemical Physics* 112, 2000, 10162-10172.
- M. A. Collins, S. Petrie, A. J. Chalk,, L. Radom, Proton-transport catalysis and proton-abstraction reactions: An ab initio dynamical study of $\text{X} + \text{HOC}^+$ and $\text{XH}^+ + \text{CO}$ ($\text{X}=\text{Ne}, \text{Ar}, \text{Kr}$), *Journal of Chemical Physics* 112,2000, 6625-6634.
- D. H. Zhang, M. A. Collins, S.-Y. Lee, First-principles theory for the $\text{H} + \text{H}_2\text{O}$, D_2O reactions, *Science*, 290,2000, 961-963.
- K. Song, M. A. Collins, A classical trajectory study of sym-triazine photodissociation on an interpolated potential energy surface, *Chemical Physics Letters*, 335,2001, 481-488.
- M. Yang, D. H. Zhang, M. A. Collins, S.-Y. Lee, Quantum dynamics on new potential surfaces for the $\text{OH} + \text{H}_2 \rightarrow \text{H}_2\text{O} + \text{H}$ reaction, *Journal of Chemical Physics*, 114,2001, 4759-4762.
- M. Yang, D. H. Zhang, M. A. Collins, S.-Y. Lee, Ab initio potential energy surfaces for the reactions $\text{OH} + \text{H}_2 \rightleftharpoons \text{H}_2\text{O} + \text{H}$, *Journal of Chemical Physics*, 115,2001, 174-178.
- R. P. A. Bettens, M. A. Collins, Capture rates for collisions of C ($^3\text{P}_j$) and Ge ($^1\text{S}_0$) with unsaturated hydrocarbons, *Journal of Chemical Physics*, 114,2001, 10342-10354.
- R. O. Fuller, R. P. A. Bettens, M. A. Collins, Interpolated potential energy surface and reaction dynamics for $\text{BH}^+ + \text{H}_2$, *Journal of Chemical Physics* 114,2001, 10711-10716.

QM/MM Calculations on Solvated Molecules and Enzyme Reaction Mechanisms

Principal Investigator

Peter Cummins

Biochemistry &
Molecular Biology
JCSMR
ANU

An understanding of the mechanisms by which protein and solvent environments influence the stability of reactants, products and transition states in biologically significant reactions is important and has a number of applications, for example, in the rational design of enzyme inhibitors as possible drugs. Combined quantum mechanical and molecular mechanical (QM/MM) methods are required for the theoretical study of enzymic reaction mechanisms.

Co-Investigators

Jill Gready

Stephen Titmuss

Biochemistry &
Molecular Biology
JCSMR
ANU

Since enzymes are solvated systems, we adopt models for the interaction of the reaction centre (QM) with both the surrounding protein and solvent medium (MM). Given access to “state of the art” computers, we can use molecular dynamics (MD) simulation techniques to obtain local free energy minima (reactant and product states) and transition states with respect to a suitably chosen set of reaction coordinates. This project focuses on assessing the utility of such free energy calculations, and the development of computationally efficient strategies together with our computer code, Molecular Orbital Programs for Simulations (MOPS), for examining the free energy changes along reaction

Alistair Rendell

Computer Science
Faculty of Science
ANU

Ivan Rostov

Supercomputer Facility
ANU

Kritsana Sagarik

School of Chemistry
Institute of Science
Suranaree University of
Technology, Thailand

Project

u51, d52

Facilities Used

VPP, PC, SC

Significant Achievements, Anticipated Outcomes and Future Work

One reaction mechanism of interest is the hydride-ion transfer in dihydrofolate reductase (DHFR). Initial free-energy gradient calculations on the hydride-ion transfer in DHFR have proved very encouraging with the characterisation of both reactant and product free-energy minima and the transition state. We have already carried out quantum chemical studies of the relative stabilities of different protonation sites of the substrates and acidic side chains in the active site of DHFR. Considerable progress has now been made in identifying protonation states in the enzyme active site, and in simulating the hydride-ion transfer step using

semiempirical QM/MM methods. In addition, our recent results have shown the need to perform calculations using the more accurate ab initio or DFT methods on systems of up to 200 atoms, and possibly larger. This may also involve the use of ab initio or DFT QM/MM calculation such as Morokuma's ONIOM method implemented in the Gaussian 98 program. Future protocol development may focus on the use of multiple MD trajectories for speeding up the calculation of reaction free energies, allowing for a larger number of reaction pathways to be investigated using QM/MM methods.

Complementary to this work, we have been using ab initio quantum chemical and linear-scaling semiempirical (using the MOZYME program of J.J.P. Stewart) methods for the validation of certain aspects of the QM/MM methodology for biomolecular systems. This has led to the investigation of new approaches for improving the quality of QM/MM force-fields for simulating enzyme catalysed reactions within the semiempirical QM approximation. This has the potential to be a significant step forward in terms of both computational efficiency and reliability of the methodology, as it may allow us to perform MD simulations with ab initio QM quality force fields for the cost of semiempirical QM calculations.

Computational Techniques Used

The study of chemical reactions in solvated systems of biochemical interest requires large allocations of computer time. Our QM/MM computer (Fortran) code, Molecular Orbital Programs for Simulations (MOPS), is being continually developed and improved in both capability and performance. The computationally intensive terms arising from interaction between the QM and MM atoms, and between MM and MM atoms, have now been fully parallelized under the Message Passing Interface (MPI) standard. This will effectively render the QM-only part of the calculations the rate limiting step in the calculation of an MD trajectory. Work is in progress to test the effectiveness of running multiple trajectories under MPI in order to improve the efficiency of free energy calculations. These reductions should become greater for systems that are larger compared with the ones we have treated to date. The implications of such reductions in computation time are significant. We will be able to treat larger systems more accurately by both extending the cutoff for the neglect of electrostatic interactions in the MM region and increasing the numbers of water molecules that solvate the protein. The APAC National Facility represents a major advance for these types of computations. The parallel computations that we have started have reduced by at least one order of magnitude the time taken to complete the calculations of the QM/MM and MM/MM non-bonded interactions.

Publications, Awards and External Funding

P.L. Cummins, J.E. Gready, Combined quantum and molecular mechanics (QM/MM) study of the ionization state of 8-methylpterin substrate bound to dihydrofolate reductase. *J. Phys. Chem. B*, 104, 2000, 4503-4510 .

P.L. Cummins, J.E. Gready, QM/MM and SCRF Studies of the ionization state of 8-methylpterin substrate bound to dihydrofolate reductase: existence of a low-barrier hydrogen bond and implications for the catalytic mechanism. *J. Mol. Graph. Mod.*, 18, 2000, 42-49.

S.J. Titmuss, P.L. Cummins, A.A. Bliznyuk, A.P. Rendell, J.E. Gready, Comparison of linear scaling semiempirical methods and combined quantum mechanical/ molecular mechanical methods applied to enzyme reactions. *Chem. Phys. Lett.*, 320, 2000, 169-176.

P.L. Cummins, J.E. Gready, Energetically most likely substrate and active-site protonation sites and pathways in the catalytic mechanism of dihydrofolate reductase. *J. Am. Chem. Soc.*, 123, 2001, 3418-3428.

P.L. Cummins, S.J. Titmuss, D. Jayatilaka, A.A. Bliznyuk, A.P. Rendell, J.E. Gready, Comparison of semiempirical and ab initio QM decomposition analyses for the interaction energy between molecules. *Chem. Phys. Lett.*, 352, 2002, 245-251.

Titmuss, P.L. Cummins, A.P. Rendell, A.A. Bliznyuk, J.E. Gready, Comparison of linear scaling semiempirical methods and combined quantum mechanical/ molecular mechanical methods for enzymic enzyme reactions II: an energy decomposition analysis. *J. Comput. Chem.*, in press.

P.L. Cummins, S.P. Greatbanks, A.P. Rendell, J.E. Gready, Mechanism of the hydride transfer step in the catalytic mechanism of dihydrofolate reductase I role of conserved water and protein interactions. *J. Am. Chem. Soc.*, submitted.

Mantle Plumes

Principal Investigator

Geoff Davies

Geophysical Fluid

Dynamics

Research School of

Earth Sciences

ANU

Co-Investigator

Alison Leitch

Department of Earth

Sciences

Memorial University of

Newfoundland

Project

r01

Facilities Used

VPP

RFCD Codes

260299

The Earth's mantle is a solid rocky layer 3000km thick extending from the bottom of the crust (30km depth) to the core-mantle boundary (CMB) at the top of the hot molten iron core. The crust and the top 100km of the mantle form a stiff cold layer called the lithosphere. Underneath the lithosphere the mantle behaves like a viscous fluid. Over geological timescales of 10's to 100's of millions of years, the mantle convects. At the surface the lithosphere sinks into the mantle at subduction zones and spreads apart at mid-ocean ridges. At these ridges the mantle partially melts and basalt lava about 6km thick forms the new ocean crust. At the CMB, mantle material is heated by the core and hot mantle "plumes" rise through the mantle as large spherical "heads" (1000km in diameter) with thinner "tails" connecting the heads to the source region. When plume heads approach the surface they partially melt, leading to massive, geologically brief outpourings of lava. These "flood basalt" events occur about every 10-20 million years. The aim of this project is to simulate mantle plumes impinging on the bottom of the lithosphere and melting.

Significant Achievements, Anticipated Outcomes and Future Work

We studied the melting that occurs when a plume first approaches the bottom of the lithosphere, aiming to match the observed melt rates and volumes of flood basalts. Melt rates are very sensitive to temperature, depth of melting and composition. We found that to match the observation using reasonable plume temperatures and lithosphere thicknesses, the plume must contain a significant fraction (about 15%) of easily melted recycled ocean crust. This is consistent with the quantities of basalt that are expected to have been returned to the mantle during subduction throughout the 4.5 billion years of Earth's history. We also found that in order for the plume to reach shallow depths beneath the lithosphere it was important to include a realistic model of mantle viscosity. At a depth of 660km, the mantle viscosity increases (going down) by about an order of magnitude. As plumes encounter this viscosity step they neck down, and the small central, hottest part of the plume rises quickly to the bottom of the lithosphere. This behaviour is necessary to produce the very rapid voluminous melting that is geologically observed.

On-going work, which has involved significant code development, involves the impingement of plumes on lithosphere of varying thickness. Seismic investigations indicate that the bottom of the lithosphere is not smooth under continents, and ponding of plume material in the shallowest parts of the lithosphere could lead to localized melting. We are also studying the conduction of heat through the lithosphere from cooling, spreading mantle plume heads.

Computational Techniques Used

The simulations have been carried out using the finite-difference program CONMG, written by G.F. Davies and modified by A.M. Leitch. The code was optimized to run on the VPP and is very efficient. This is important because very high resolution is required to properly resolve the melting region in the plumes (10's of km) within a mantle-scale (3000km) simulation. The high core memory of the VPP allowed high resolution simulations to be carried out that could not be run on workstations. Lower resolution exploratory simulations run in tens of minutes rather than hours or days.

Publications, Awards and External Funding

A. M. Leitch, G. F. Davies, Mantle Plumes and Flood Basalts: Enhanced Melting from Plume Ascent and an Eclogite Component, *Journal of Geophysical Research*, 106, 2001, 2047-2060.

Work is partially funded through NSERC research operating grant of A.M. Leitch.

Plasma Turbulence in 3-Dimensional Magnetically Confined Plasmas

Principal Investigator

Robert Dewar

*Theoretical Physics
RSPHysSE
ANU*

Co-Investigator

Paul Cuthbert

*Theoretical Physics
RSPHysSE
ANU*

Project

s55

Facilities Used

VPP, MDSS, VizLab

RFCD Codes

240303, 240201

The goal of nuclear fusion power research is to use strong magnetic fields to insulate a plasma (hot gas) of hydrogen isotopes, at a temperature of hundreds of millions of degrees, from walls near room temperature. The enormous temperature gradient creates a highly non-equilibrium situation where various self-organisation phenomena tend to occur, destroying smooth density and temperature distributions and instead creating turbulent structures that allow heat and particles to escape faster than desired (“anomalous transport”).

The aim of advanced magnetic confinement designs, such as the H-1 Helic (now upgraded to a national facility, H-1NF, in which temperatures up to ten million degrees will be obtained), is to thwart this tendency towards formation of turbulence by designing twisted magnetic field configurations that make instability energetically unfavourable. The simulation of such a device forms a grand challenge to theory and computation because the full three-dimensional geometry, both of the equilibrium and the perturbation, must be accurately taken into account.

The project has concentrated on the first phase of a turbulence study — linear stability analysis using a simple magnetohydrodynamic (MHD) fluid model — to understand the unusual geometric effects that arise in low-magnetic-shear, strongly three-dimensional toroidal plasma containment devices such as the H-1NF Helic.

Significant Achievements, Anticipated Outcomes and Future Work

The analysis of a high beta (beta = plasma pressure/magnetic pressure) configuration of the H-1NF heliac to determine the nature of the local ballooning eigenvalue for ideal MHD ballooning modes has led to the successful completion of a PhD thesis and several papers in prestigious journals. Ballooning modes are known to limit the maximum pressure which can be obtained in a confinement device, and as such are an important consideration for the development of future fusion devices.

With the use of the VPP, the parametric structure of the local ballooning eigenvalue was revealed, and the role of the magnetic shear investigated. The major outcomes of the work were a) establishing the importance of Anderson localization in low-shear systems, b) showing that the complex structure of the branches of the ballooning mode eigenvalue could be understood using symmetry operations and a perturbation expansion in magnetic shear, and c) showing that the rays of a regularized version of 3-D ballooning WKB are chaotic, so that understanding the global spectrum will require a “quantum chaos” analysis.

Computational Techniques Used

The fundamental computation in the “WKB ballooning method” consists of solving a linear 2nd order, one-dimensional boundary-value eigenproblem on a magnetic field line; a simple exercise in principle, at least when a simple fluid model is used. However the ODE coefficients (derived from an equilibrium configuration also calculated on a supercomputer) must be calculated at up to 30,000 points on a field line. Each point requires summing between 600 and 1,100 Fourier components and calculation of coefficients on the full set of points takes tens of cpu-seconds.

The supercomputing needs come from computing sufficient of these localised one-dimensional eigensolutions to construct a three-dimensional array from which global eigenmodes and their growth rates can be constructed.

Publications, Awards and External Funding

R.L. Dewar and P. Cuthbert, Anderson Localization and Ballooning Eigenfunctions, Chinese Physics Letters ISSN 0256-307X, 2000, 33^35

P. Cuthbert, Ballooning Instabilities in Three-Dimensional Toroidal Plasmas, PhD Thesis, ANU, 2000

P. Cuthbert and R. L. Dewar, Anderson-localized ballooning modes in general toroidal plasmas, Phys. Plasmas 7, 2000, 2302-2305

R. L. Dewar, P. Cuthbert and R. Ball, Strong “quantum” chaos in the global ballooning mode spectrum of three-dimensional plasmas, Phys. Rev. Letters 86, 2001, 2321-2324

Calculation of the Stability of Phase Space Trajectories Using Molecular Dynamics Simulations

Principal Investigator

Denis Evans

*Physical and
Theoretical Chemistry
RSC
ANU*

The Lyapunov exponents of a liquid system are a measure of the dynamical stability of a system and can be related to transport properties of the liquid such as the viscosity and the thermal conductivity. If the spectrum of Lyapunov exponents is symmetric, calculation of these liquid state properties from the Lyapunov exponents is facilitated. Studies of the structure of the spectrum are therefore being carried out. A number of relationships between the Lyapunov exponents and properties of fluids have now been developed. In this project new relationships will be derived and checked using numerical simulation.

Co-Investigators

Jerome Delhommelle

Owen Jepps

Janka Petravac

*Physical and
Theoretical Chemistry
RSC
ANU*

Numerical simulations of fluids flowing through micropores have been carried out for various systems: atomic fluids, rigid molecules and molecular fluids. In order to achieve this the configurational temperature is used to measure and/or constrain the temperature. The configurational temperature is a new expression for the thermodynamic temperature that is obtained by performing a functional differentiation of the entropy of a microcanonical system with respect to the thermodynamic internal energy. The new temperature is entirely dependant on configurations and thereby circumvents difficulties in defining streaming velocities in complicated geometries.

Emil Mittag

*Computational
Quantum Chemistry
RSC
ANU*

Debra Bernhardt

*Department of
Chemistry
Griffith University*

Project

s02, r61

Facilities Used

VPP, PC, SC

RFCD Codes

250600

Significant Achievements, Anticipated Outcomes and Future Work

A new derivation of the fluctuation theorem that describes the probability of observing violations of the Second Law of Thermodynamics has been developed using the Lyapunov weight formula. This formula expresses the probability of trajectories in phase space being observed in terms of their escape from a phase space volume, which is related to the sum of the positive Lyapunov exponents of the system. Tests of this formula have been carried out and will be submitted for publication early in 2002.

The semi-empirical scaling law between transport coefficients and excess entropy has been systematically studied for 3 types of atomic potentials (soft, LJ and WCA) along an isochore and an isobar. The disagreement with the law is explained in terms of differences between 2-particle and total excess entropy, the relative contributions of interaction and thermal motion to transport at different state points and discrepancy between Enskog hard-sphere estimate of collision frequency and the Maxwell relaxation time.

In the second project simulations of confined fluids have been carried out for atomic and molecular fluids. The configurational temperature profile in these fluids has been obtained. This work has been published in high profile journals.

Computational Techniques Used

Equilibrium and nonequilibrium molecular dynamics simulation methods are being used and are developed in this project. Access to supercomputer facilities is required to obtain statistically valid data for small systems and to study large systems in investigations of the influence of system size.

Publications, Awards and External Funding

J. Delhommelle, D. J. Evans, Configurational temperature profile in confined fluids. I. Atomic fluid, *J. Chem. Phys.*, 114, 2001, 6229-6235.

J. Delhommelle, D. J. Evans, Configurational temperature profile in confined fluids. II. Molecular fluids, *J. Chem. Phys.*, 114, 2001, 6236-6241.

J. Delhommelle, D. J. Evans, Comparison of thermostatting mechanisms in NVT and NPT simulations of decane under shear, *J. Chem. Phys.*, 115, 2001, 43-49.

Spectra and Molecular Structures

Principal Investigator

Gad Fischer

Chemistry
Faculty of Science
ANU

Co-Investigator

Rebecca Jacob

Chemistry
Faculty of Science
ANU

Project

u01, u54

Facilities Used

PC, SC, VPP

RFCD Codes

250104, 250105,
250699

Spectroscopy and computation can be used to determine the structures of molecules. Recent work has been directed at two families of molecules, the azines, and the amino acids and their pseudo clusters, the dipeptides. For the former, ab initio calculations on the excited electronic states and their force fields have been conducted at high theoretical levels to assist in analyses of the spectra. The substituted azines are widespread in nature and knowledge of their stabilities and photochemical properties are of interest.

Three-dimensional models of protein structures depend largely on knowledge of accurate geometries of the component amino acids. A novel experimental approach has been successfully used to measure resolved infrared spectra of the zwitterions isolated in a matrix. Molecular orbital calculations to determine the vibrational spectra and that are appropriate to the amino acid in its solid solvent environment have allowed for confident analyses of the infrared spectra. This combined approach marks an important breakthrough in amino acid structure determination.

Significant Achievements, Anticipated Outcomes and Future Work

The first singlet¹ singlet and singlet¹ triplet band systems of the absorption spectrum of pyridazine vapour have been analysed using ab initio and vibronic coupling calculations. The lowest singlet¹ triplet absorption involves a comparatively unperturbed triplet B1 state, and contrasts with the highly perturbed singlet¹ singlet spectrum. The major source of vibronic perturbation in the singlet¹ singlet absorption is attributed to coupling between near-resonant singlet A2 and B1 states, with the former being slightly lower in energy. Many features of this complex and unusual spectrum, and its associated single vibronic level fluorescence spectrum, can be explained using a simple vibronic model. This provides experimental support for recent relaxed CASPT2 and EOM-CCSD calculations, but contrasts with earlier assignments of the spectrum. Similar studies have been carried out for pyrimidine. However, for pyrimidine the singlet and triplet spectra are not marked by major perturbations, and as a consequence the spectral analyses are relatively unambiguous. With the completion of the investigations of the diazines the triazines will be revisited using the more powerful computational techniques that are now available. The mid-IR spectra of a number of amino acids and their N-D₃ deuterated isotopomers, and dipeptides and their N-D deuterated isotopomers, as isolated zwitterions in a KBr solid solvent have been obtained by means of a newly developed infrared sampling technique, which involves the production of thin films. Ab initio calculations of the non-aqueous Self-Consistent Reaction Field type, using the Onsager dipole-sphere model have been carried out for the prediction of the vibrational spectra and molecular structures of the zwitterions isolated in KBr matrices. For all the amino acids and dipeptides studied more

than one conformer was determined by computation. Nevertheless, good agreement was established between the simulated and calculated spectra. This combined theoretical and experimental approach has allowed accurate determination of the amino acid structures. The work is to be extended to the polypeptides with a view to practical applications, such as the photosynthetic reaction centre in Photosystem II and synthetic photosynthesis.

Computational Techniques Used

Two computational packages were used in the two projects. In the one "Gaussian" ab initio calculations were conducted at the HF, HF/SCRF, DFT, DFT/SCRF, MP2 and CIS theoretical levels using basis sets ranging from 6-31G* to 6-311++G**. In the other, "Molpro", complete active space calculations were carried out using Multi and RS2. The high level calculations require much memory, large scratch disk space, and can take much CPU time, requirements successfully met by the available computers, PC, VPP and SC.

Publications, Awards and External Funding

G. Fischer, X. Cao, Vibrational analysis and ring-puckering inversion barrier of perfluorocyclopentene. *J. Phys. Chem. A* 103 (1999) 3726-3731.

X. Cao, G. Fischer, New infrared spectra and the tautomeric studies of purine and L-alanine with an innovative sampling technique. *Spectrochim. Acta.* 55A (1999) 2329-2342.

D.A. Kirkwood, M. Tulej, M.V. Pachkov, M. Schnaiter, F. Guthe, M. Grutter, M. Wyss, J.P. Maier, G. Fischer, Electronic Spectra of the Carbon Chain Anions- C₂nH⁻ (n=5-12). *J. Chem. Phys.* 111 (1999) 9280-9286.

M. Tulej, F. Guthe, M. Schnaiter, M.V. Pachkov, D.A. Kirkwood, J.P. Maier, G. Fischer, Electronic Spectra of the Carbon Chain Anions C₂n-1H⁻ (n=5-8) in the Gas Phase. *J. Phys. Chem. A* 103 (1999) 9712-9716.

X. Cao, G. Fischer, Infrared Spectral, Structural, and Conformational Studies of Zwitterionic L-Tryptophan. *J. Phys. Chem. A* 103 (1999) 9995-10003.

X. Cao, G. Fischer, The infrared spectra and molecular structure of zwitterionic L-phenylalanine. *J. Mol. Structure* 519 (2000) 153-163.

X. Cao, G. Fischer, Infrared spectra of monomeric L-alanine and L-alanine-N-D₃ zwitterions isolated in a KBr matrix. *Chem. Phys.* 255 (2000) 195-204.

G. Fischer, P. Wormell, Vibronic analyses of the lowest singlet-singlet and singlet-triplet band systems of pyridazine. *Chem. Phys.* 257 (2000) 1-20.

G. Fischer, R. Jacob and X. Cao, The infrared spectra and structures of the valyl-glycine zwitterion isolated in a KBr matrix. *Chem. Phys.* 263 (2001) 243-253.

X. Cao, G. Fischer, Reply to comment on new infrared spectra and the tautomeric studies of purine and L-alanine with an innovative sampling technique. *Spectrochim. Acta* A57 (2001) 2719-2720.

G. Fischer, T. Wydrzynski, Isotope effects in FTIR difference spectra of the photosynthetic oxygen-evolving catalytic site determined by ab initio calculations on model compounds. *J. Phys. Chem. B* In press.

X. Cao, G. Fischer, Conformational and infrared studies of L-methionine and its N-deuterated isotopomer as isolated zwitterions. *J. Phys. Chem. A* In press.

N-body Simulation of the Milky Way

Principal Investigator

Roger Fux

Astronomy

RSAA

ANU

Co-Investigator

Ken Freeman

Astronomy

RSAA

ANU

Project

b103

Facilities Used

MDSS

RFCD Codes

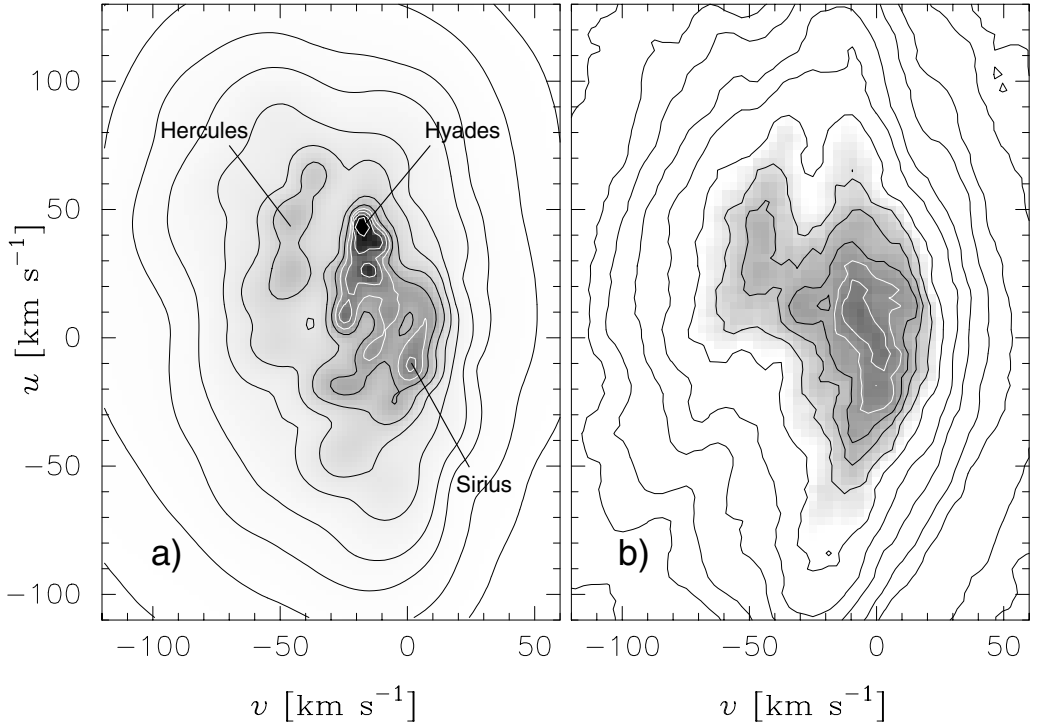
240101

The Milky Way, like most other spiral galaxies, is a barred galaxy. Its bar has a semi-major axis length of about half the galactocentric distance of the Sun and is rapidly rotating. In this project, the outputs of a very high-resolution N-body simulation (14 million particles) designed to mimic the Milky Way have been analysed to determine how the bar affects the stellar kinematics in the surrounding disc and whether the induced perturbations may explain some of the “streams” observed in the velocity distribution of stars in the Solar neighbourhood. The Mass Data Storage System was useful to store and have efficient access to the large files produced by the simulation. The present report will focus on the most outstanding results obtained.

Significant Achievements, Anticipated Outcomes and Future Work

Figure (a) shows the velocity distribution of predominantly old stars in the Solar neighbourhood, based on the Hipparcos input (radial velocities) and output (tangential velocities) catalogues. The horizontal (v) and vertical (u) axes represent the azimuthal and radial heliocentric velocities respectively, positive for motion towards galactic rotation and towards the Galactic anti-centre. These data for the first time clearly delineate the “Hercules” stream, characterised by a low angular momentum and a net u_0 average velocity. Such a net outward motion cannot be understood in terms of axisymmetric models of the Milky Way. Moreover, the involved stars are essentially older than 2 billion years, suggesting a dynamical origin of the stream.

The three-dimensional N-body simulation, started from bar-unstable axisymmetric initial conditions, as well as other test particle simulations within a rotating barred potential, have shown that a bar is able to induce Hercules-like streams at realistic space positions for the Sun, i.e. at roughly twice the galactocentric distance of the bar ends (close to the outer Lindblad resonance, hereafter OLR) and a trailing angle relative to the bar major axis of 10-50 degrees. Figure (b) gives an example taken from the N-body simulation, with the u -component of the velocity relative to the Galactic centre.



A detailed orbital study of the simulations has revealed two possible bar-related origins for the Hercules-like streams occurring in them. The first one, which corresponds to an older idea introduced by Agris Kalnajs (RSAA) to explain the Sirius and Hyades streams and recently revived by Walter Dehnen in the Hercules context, is that near the OLR of a barred potential, there exists two distinct families of elongated and perpendicularly-oriented regular orbits, resulting in a bimodality of the velocity distribution. The second origin, which is a completely new interpretation, is that at high orbital energies (in terms of Jacobi's integral), the velocity distribution is dominated by stars on chaotic orbits covering essentially the whole disc, including the Galactic bar region. According to this interpretation, the stars in the Hercules stream follow such orbits, and the $u < 0$ counterpart in the local velocity plane coincides with a regular orbit region that chaotic orbits are forced to avoid.

Most of the study was restricted to the two-dimensional motion of stars within the Galactic plane, i.e. to the u - v velocity components. A possible extension of this work would be to also take into account the vertical velocity component.

Publications, Awards and External Funding

External funding:

Visiting fellow grant from the Swiss National Foundation.

Publications:

R. Fux, Order and chaos in the local disc stellar kinematics induced by the Galactic bar, *Astronomy & Astrophysics*, 373, 2001, 511-535

R. Fux, Stellar streams in the Solar neighbourhood from high resolution N-body simulations, in “Galaxy Dynamics: from the Early Universe to the Present”, eds. F. Combes, G.A. Mamon, V. Charmandaris, ASP conference series vol. 197, 2000, 27-30

R. Fux, Order and chaos in the local disc stellar kinematics, in “Dynamics of Star Clusters and the Milky Way”, eds. S. Deiters, B. Fuchs, A. Just, R. Spurzem, R. Wielen, ASP Conference series vol. 228, 2001, 283-289

R. Fux, Order and chaos in the local disc stellar kinematics, in “Stellar Dynamics: from Classical to Modern”, eds. L.P. Ossipkov, I.I. Nikiporov, Saint Petersburg University, 2001, 209-212

Simulations of Spiral Galaxies

Principal Investigator

Roger Fux

Astronomy

RSAA

ANU

Co-Investigators

Ken Freeman

Isabel Perez

Astronomy

RSAA

ANU

Oak-Kyoung Park

IGPP

Lawrence Livermore

National Laboratory

USA

Project

x42

Facilities Used

SC

RFCD Codes

240101

The outer parts of the rotation curves of external galaxies provide the best evidence for the existence of dark matter. In almost all of them the rotation curve is flat or slowly rising out to the last measured point. Very few galaxies show falling rotation curves, and the ones that do either fall less than Keplerian or have nearby companions that perturb the velocity field. The simplest interpretation of these results is that spiral galaxies possess massive dark haloes that extend to larger radii than the optical disks.

The inner part of the rotation curve is crucial in determining the nature of the dark matter. The shape of the dark spherical halo is determined by the central density and the core radius. Haloes built up by hierarchical merging in dark matter cosmogonies are cusped at the very centre. Simulations by Navarro, Frenk & White (1997) found a nearly universal halo profile which rises as r^{-1} to the centre. These characteristics seem to disagree with a number of observations. Moreover, the number of subhaloes around typical galaxies, as identified by satellite galaxies, is an order of magnitude smaller than predicted by cold dark matter (CDM). The observed rotation curves of dwarf and low surface brightness galaxies seem to indicate that their dark matter haloes have a constant density core instead of steep cusps. For high surface brightness galaxies, the situation is not clear and there is not consensus as to whether bright galaxies are dark matter dominated at their very centres or not.

In order to clarify the problem of the mass distribution in the inner parts of bright galaxies and constrain the galaxy formation scenario, we are carrying out a project to test whether the luminous mass in the inner parts of spiral galaxies can account for their observed gas kinematics or whether an additional dark matter component is required. For this purpose we are planning to model the galaxy dynamics of a significant sample of barred spiral galaxies by running a 3-D composite N-body/hydrocode on the luminous underlying matter distribution of the sample galaxies and then compare it to the observe gas dynamics.

Significant Achievements, Anticipated Outcomes and Future Work

Due to the starting nature of the project, no final results have yet been obtained. However, we have explored and set the optimal initial conditions for the simulations. We have concluded that in order to reach gently a stationary state, without wobbling, we have to set the galaxy bar growth time to roughly 1.5 times the bar rotational period. This clashes with other people's

claim that a shorter time is enough to reach a stationary state. We are now ready to start all the simulations and soon after we will be able to publish the results.

Computational Techniques Used

We are running a serial 3-D composite N-body/hydrocode based on a code developed by the Geneva Observatory galactic dynamics group. At each time step, it computes gravitational forces on all particles using a particle mesh with fast Fourier transform method. Secondly, it evaluates pressure and artificial viscous forces on the gas particles, using the Lagrangian smooth particle hydrodynamics method to solve the equations of motions for a fluid. Finally, it moves all particles one time step further by integrating the equations of motion with an adaptive time-step algorithm. The maximum resolution used in our test simulations was a grid of 147456 cells and 500000 gas particles. Given the number of galaxies to simulate and the high resolution required it would be impossible to achieve this without the supercomputer capabilities.

3D MHD Equilibrium and Stability

Principal Investigator

Henry Gardner
Theoretical Physics
RSPHysSE
ANU

Co-Investigators

Bob Dewar
Sally Lloyd
Ben McMillan
Theoretical Physics
RSPHysSE
ANU

David Singleton
Supercomputer Facility
ANU

Project
k12

Facilities Used
VPP, SC, MDSS

RFCD Codes
240303

For some decades there has been an international scientific and engineering program to study the containment of plasmas by toroidal magnetic fields with the aim of developing a fusion power reactor. Australia presently makes a major contribution to this program through the H-1NF Helic which is a “stellarator” experiment located at ANU and is funded through the Major National Research Facility Program. Stellarators are an alternative to the better known “tokamak” types of experiments but have the great advantage of not needing large currents within the plasma in order to generate helical field lines. On the debit side, the lack of axial symmetry of stellarators makes their theory more complicated and simulation more computationally expensive than for tokamaks. In particular, the simply nested magnetic surfaces found in tokamaks can be broken and give rise to magnetic islands and regions of chaotic magnetic field lines. Activities under this project cover two distinct areas: (1) Magnetic Islands in the H-1NF Helic and (2) 3D Resistive MHD Resistive Stability (with R.G. Storer, Flinders University, S.A.).

Significant Achievements, Anticipated Outcomes and Future Work

Following extensive parameter studies with an accelerated version of HINT, we have a strong inference that magnetic island self-healing can be expected for Helicacs. Detailed comparisons with local theory based on expressions involving the resistive Mercier criterion have been more problematic and are still under investigation.

The development of a fully three dimensional resistive MHD stability code is underway. As part of this development the Jacobi-Davidson method has been used to compute a number of spectral eigenvalues at once. The code has been satisfactorily benchmarked in axisymmetric systems such as tokamaks. Additional new features of the resistive tokamak MHD spectrum seem to have been identified. Comparisons of our incompressible model with a compressible code used elsewhere have identified intriguing similarities and differences.

Computational Techniques Used

The HINT code, developed by T. Hayashi, solves time-dependent, resistive MHD equations on a specially-shaped coordinate grid. Relaxation along field lines is much slower than perpendicular to them and is treated using a special interpolation algorithm which has been

developed as part of this project (pressure is averaged along field-lines and interpolated back to the fixed grid).

The SPECTOR3D resistive stability code, which we are developing, computes matrix elements corresponding to a particular formulation of the incompressible, resistive, linearised MHD stability equations. Initially inverse iteration was used to compute eigenvalues of the system but this has been changed to implement the Jacobi Davidson method.

The VMEC equilibrium code, developed by SP Hirshman, is also extensively used to compute ideal MHD equilibrium. This code uses a form of steepest descent to find the minimum energy state of a plasma. The formulation is a combination of spectral, in the magnetic surfaces, and finite differences across the surfaces. SPECTOR3D has a similar, hybrid, representation of the plasma.

Local MHD stability is computed using the JMC code developed by J. Nuehrenberg. This code calculates surface averages of equilibrium quantities.

Field line tracing, and surface fitting, codes are used to set up an equilibrium modelling study from the real, H1NF magnetic field.

Publications, Awards and External Funding

R. Storer and H. Gardner, Resistive Magnetohydrodynamic Modelling for 3D Stellarator Geometry, Accepted for publication in Computer Physics Communications (December, 2001)

Tomographic Imaging of the Western Pacific

Principal Investigator

Alexei Gorbatov

Seismology

RSES

ANU

Project

d76

Tomographic imaging of the Western Pacific Region was performed based on the world-wide seismic data composed by travel-time and path of seismic rays. The purpose of this study was to estimate high-resolution structure of the Earth's mantle and to map subducting lithosphere at the plate boundaries. This information can improve our understanding about physical properties of the Earth's interior and tectonic fabric of the Western Pacific.

Facilities Used

VPP, SC

RFCD Codes

260299

Significant Achievements, Anticipated Outcomes and Future Work

A new inversion technique was developed to perform tomographic imaging of the Western Pacific. Two main innovations were introduced: a three-dimensional ray-tracing technique to determine the seismic ray path in the heterogeneous Earth mantle, and a combined inversion scheme that permits separation of bulk-sound and shear moduli in two images based on the P and S wave speeds. This newly developed tomographic scheme can not only detect seismic wave heterogeneities in the Earth mantle but can also give an idea about physical properties of those anomalies.

High-resolution tomographic image obtained from joint P and S tomographic inversion reveals penetration of the subducted slabs into the lower mantle below Philippine, Mariana, and Tonga-Kermadec subduction zones. These results are new evidence for deep subduction.

Relationships between bulk-sound and shear moduli point to the clear difference in the physical properties of subducted slabs near the age of 100 millions years. This peculiar feature suggests that physical properties of oceanic plates vary with time and age boundary near 100My is an important step in this evolution.

Further analysis of obtained results could reveal new tectonic features of the Western Pacific Region and, probably, the reason for the differences between bulk-sound and shear moduli anomalies with age and regions. Improvement of the inversion technique can refine our results. We plan to introduce new regularization approaches in order to increase resolution of tomographic images. Additionally, more data will be processed and incorporated into the inversion.

Computational Techniques Used

The pseudo-bending method (Koketsu & Sekine, GJI, 1998) was used for three-dimensional ray-tracing algorithm.

LSQR (Page and Saunders, ATMS, 1982) and Inversion for multiple parameter classes (B.L.N. Kennett & M. Sambridge, GJI, 1998) were used to solve large and sparse system of linear equations.

The computational power of the APAC National Facility allows us to perform all calculations in the relatively short duration of this project, which cannot be done on conventional workstations.

Graphical program GeoPlot, developed by Ajay Limaye (VIZLAB), was used to visualize all tomographic results.

Publications, Awards and External Funding

B.L.N. Kennett, A. Gorbatov, Why is there a Strong Shear Signature for Deep Slabs?, EOS Am Trans Act., S41-B, 2001.

A. Gorbatov, B. L. N. Kennett, Western Pacific Subduction - New Results from Regional Joint P and S Travel-Time Tomography., EOS Am Trans Act., T22D-08, 2001.

Definition of the Chemical Mechanism of the Photosynthetic Enzyme Rubisco

Principal Investigator

Jill E. Gready

Computational

Proteomics Group

John Curtin School of

Medical Research

ANU

Co-Investigators

Lars Andrees

Peter L. Cummins

Bill King

Harald Mauser

Computational

Proteomics Group

John Curtin School of

Medical Research

ANU

T. John Andrews

Research School of

Biological Science

ANU

Project

u53, d55

Facilities Used

SC, VPP, PC, MDSS

RFCD Codes

250106, 250601,

250699, 270108,

270402, 291102,

300203

*Virtually all of the carbon in the biosphere is the result of carbon dioxide (CO₂) fixation by the photosynthetic enzyme, D-ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase (Rubisco). Rubisco also has the dubious honour of being the most abundant protein on earth, a consequence of its woeful catalytic efficiency. Rubisco is both very slow and poorly selective for its substrate CO₂, binding also oxygen (O₂) and catalysing a competing wasteful oxygenation reaction. As one would expect Rubisco to have been subjected to extreme evolutionary pressures, its incompetency is a great puzzle. Understanding the reasons offers huge potential as a basis for reengineering Rubisco; even modest improvements in efficiency have major implications for improving light, water and nutrient utilization by plants, and, hence, applications for better agricultural crops, greening deserts and degraded land, and for soaking up green-house gases. As Rubisco catalyzes a multi-step reaction involving as many as four enzyme-bound intermediates, whose instabilities give rise to multiple side reactions which further compromise its efficiency, it is possible that the enzyme's complex structure and function is a compromise solution to effecting quite difficult chemistry. Not surprisingly, much of this chemistry is not approachable by experiment. This provides a (very challenging!) opportunity for computer simulation to try to define intractable issues, particularly the states and roles of the "invisibles" – protons and water molecules. Thus: protonation states of the forest of ionizable residue sidechains in the active site at different stages of the reactions (both carboxylation and oxygenation), as well as the protonation states for the reactant, intermediates and transition states; identities of the proton donors for the various steps; networks for channelling protons produced in the reaction away from the reaction centre; and origin of the water molecule consumed and produced (oxygenation only). We are addressing these issues using a range of computational methods, particularly *ab initio* QM studies of active-site fragment complexes and MD simulations with hybrid QM/MM potentials.*

Significant Achievements, Anticipated Outcomes and Future Work

Results. Because of the complexity of the active site and Rubisco chemistry, as well as difficulty in applying QM/MM methods to an enzyme as large as Rubisco, our initial computational approach has been to map the complete carboxylation reaction using an *ab initio* QM supermolecule fragment model, as a starting point for these other investigations

and simulation methods. In the first study [King *et al.* Biochemistry 37, 1998, 15414] we established a role as a base for the lysyl carbamate group in the first enolization step, a unique role for a carbamate not known in any other enzyme. This was the first theoretical study that took account of the likely roles of critical features within the active site. A prediction of this study has been tested and supported experimentally [[Harpel *et al.* Biochemistry 41, 2002, 1390]. We continued the computational dissection for the subsequent steps – the carboxylation, hydration and C–C bond cleavage reactions – using similar fragment complexes (29 atoms) at the DFT level [1]. The 29-atom model was defined by an Mg²⁺ ion coordinated to a 2-carbon atom fragment of RuBP (ethen-1,2-diol-2-ate), 3 water molecules (2 representing Asp203 and Glu204) and methylcarbamate representing the carbamylated Lys201. These calculations showed that CO₂ is added directly without assistance of a Michaelis complex, in agreement with experiment, and that hydration of the resultant β -keto acid occurs in a separate step with a discrete transition step. The carbamate appears very versatile, acting as a general base not only in the enolization reaction, but also in the hydration and cleavage reactions. While the step with the highest activation energy is the C–C bond cleavage, special arrangements of the metal coordination make it remarkably facile for such a bond-breaking reaction (28-37 kcal/mol depending on the *gem*-diol hydrate conformer).

QM/MM calculations to mirror the steps of the fragment calculations were then undertaken, to study the effect of the rest of the enzyme environment on the reaction and to calculate free energies [2]. A larger QM region of 165 atoms, rather than the 29-atom fragment-complex model, also allowed investigations of the roles of other active-site groups during the reaction, particularly in possible acid-base catalysis and proton channelling. The QM region included all residues of the first and second coordination sphere of the Mg²⁺ ion: Lys-175, Lys-177, carbamylated Lys-201, Asp-203, Glu-204, His-294, Lys-334, the Mg²⁺ ion and complete RuBP. Of interest was whether the β -ketoacid might be stabilized by protonation of the carboxylate group (from CO₂); experimentally there is no evidence for reversibility of CO₂ addition although the fragment-model calculations suggest it would be energetically accessible. The results showed that carboxylate protonation by Lys-334 was likely and did stabilize the β ketoacid, and that it also reduced the activation free energy for the next, hydration, step: the free energy difference of 16 kcal/mol suggests an almost irreversible CO₂ addition. The results also showed that the C–C cleavage reaction is stabilised by the enzyme environment. To fully understand the catalysis of Rubisco, it is essential to know how the two protons formed during each catalytic cycle are transported away from the active site. Despite many experimental efforts, the detailed mechanism is still unknown. We have found that the dianionic phosphate groups of RuBP are acting as strong proton acceptors in the active-site environment. In the first steps of the catalysis, we found strong H-bond coupling between the carbamylated Lys-201 and the P2-phosphate *via* the substrate water and RuBP-O4. This interaction could be responsible for the recycling of the intermediary protonated carbamate. In the same way, another path for proton transport from the carbamate to the P1-phosphate group *via* RuBP-O2 and Thr-173 is indicated. Thus, the QM/MM simulations suggest that these phosphates may be responsible for the recycling of the carbamate and the transport of the two resulting protons away from the active site.

In addition to these studies we have undertaken a pilot study to compare the energetics of

CO₂ or O₂ gas addition to the enediolate [King, Gready & Andrews, unpublished, 2000]. The oxygenation reaction, i.e. addition of triplet O₂ to a carbanion, is a puzzling reaction requiring a spin inversion, which is inherently improbable in the absence of spin-delocalising mechanisms. QM representations of the O₂ addition which would be adequate for treatment of the spin inversion require large correlated calculations, preferably MP2 or higher, which were not possible for our 29-atom fragment model until recently. Initial DFT calculations with a smaller fragment indicated a novel mechanism for the spin inversion for the oxygenation reaction, with much greater stabilization of the oxygenation product compared with the carboxylation product (-15.5 and -4.5 kcal/mol, respectively).

Future work. As we can now begin to attribute specific roles to active-site features and suggest the sources and fates of key protons in the acid/base chemistry that pervades the mechanism, we will expand the QM/MM studies to even larger representations of the active site. Such studies will also include comparisons of mechanistic details of Rubiscos from other species, and of mutants. We will also readdress the carboxylation *versus* oxygenation issues with DFT and MP2, and maybe higher methods, with a fragment model of comparable size (at least 30 atoms) to those used previously.

Computational Techniques Used

Transition-state, intermediate and product steps on the reaction pathway were performed with GAUSSIAN98 at DFT (B3LYP) level; stationary points were first optimised with the 6-31G* basis set, and single-point calculations with the 6-311G+** basis set were then used to calculate reliable energies. Starting coordinates for a minimal 29-atom active-site fragment model for the enediolate form of RuBP were generated from six sample sets of a 1 ns MD simulation of a Rubisco model. This model was based on a 1.6 \approx x-ray structure of activated spinach Rubisco L8S8, and contained two large (L) and two small (S) subunits plus \sim 500 water molecules; \sim 1200 residues and 13,500 atoms.

QM/MM calculations were performed with our locally-developed program MOPS, using the semiempirical PM3 method for the QM region and Amber94 forcefield for the MM region. Starting structures were constructed for each stationary point of the carboxylation reaction defined by the fragment calculations, by docking the DFT-optimized structures into the active-site of an L2 dimer model of Rubisco, the minimal functional unit of Rubisco. The L2 dimer was extracted from the 1.6 \approx crystal structure and solvated; 16,530 atoms, including 3040 water molecules. The co-crystallized CABP ligand and the Mg-coordination sphere of one active site were replaced by the analogous atoms of the DFT-optimised β ketoacid intermediate, and subjected to a 2 ns MD simulation with the DFT fragment coordinates frozen. For all other stationary points, the coordinates were constructed by replacing the β ketoacid active-site fragment with the analogous atoms of the respective DFT-optimised structures. QM/MM MD simulations of 200 ps were then performed for an expanded QM region of 165 atoms to equilibrate the system, using a 16 \approx sphere for the flexible region containing the QM active site, and free energies for each complex on the reaction coordinate calculated by averaging over 800 runs (corresponding to 160 ps) using the linear response approximation.

For status of optimization of GAUSSIAN98 and MOPS for the VPP300, PC and APAC National Facility Compaq SC see reports for u51/d52 and w05.

Publications, Awards and External Funding

1. H Mauser, WA King, JE Gready, TJ Andrews. CO₂ fixation by Rubisco: computational dissection of the key steps of carboxylation, hydration and C-C bond cleavage. *J Am Chem Soc* 123, 2001, 10821-10829.

2. H Mauser, L Andrees, JE Gready, TJ Andrews. QM/MM simulation of the steps of CO₂ fixation by Rubisco. M/S in preparation, 2002.

Funding:

DAAD (Deutscher Akademischer Austauschdienst) Postdoctoral Fellowship (2000-2001) to Dr Harald Mauser.

DAAD (Deutscher Akademischer Austauschdienst) Visiting Studentship (2001) to Mr Lars Andrees.

Simulation of the Structure of Sugar Chains of Glycoproteins

Principal Investigator

Jill E. Gready

Computational
Proteomics Group
John Curtin School of
Medical Research
ANU

Co-Investigators

Johannes Zuegg

Computational
Proteomics Group
John Curtin School of
Medical Research
ANU; and
Alchemia, Brisbane

Saburoh Sakai

RIKEN Genomics
Science Center
Yokohama
Japan

Project

x04, d55

Facilities Used

SC, VPP, PC, MDSS

RFCD Codes

250106, 250699,
270105, 270199

Deciphering the biological code for functions of oligosaccharides (glycans) represents the third frontier, after such characterization for proteins (oligopeptides) and DNA and RNA (oligonucleotides). At the level of glycan chemistry, i.e. composition, sequence, 3-D structure and biophysical properties, experimental “windows” are still relatively undeveloped. Although attachment of an oligosaccharide sidechain (glycosylation) to proteins is the normal state for most extracellular, and many intracellular, proteins, how such “decoration” affects the structure, dynamics, and functional properties of proteins is largely unknown. Indeed it is largely ignored, as nearly all high-resolution (NMR, x-ray crystallography) structural and other experimental molecular studies are performed on “undressed” protein, i.e. without attached glycans, for reasons of experimental tractability (avoid heterogeneous protein) or convenience (bacterial expression systems lack glycosylation machinery). For these reasons, computer simulation offers a particularly rich opportunity to complement the large amount of protein structural data, by studying the properties of proteins reclothed with their sugar chains. Simulation offers a second major advantage in being able to deal easily with glycan heterogeneity; glycans with variant chain length, sugar composition, branching pattern, and linkage type, as found experimentally by MS sequencing, can be simulated. Our interest started with prion protein (PrP), the unusual protein for which an aberrant conformational form (PrP^{Sc}) is associated with diseases such as mad-cow disease (BSE) and CJD in humans. Although the correlation of PrP glycosylation state with disease characteristics is well-known, NMR structures and other experimental molecular studies have all been done on unglycosylated protein, and, remarkably, interpreted without acknowledgement of their normal glycosylation. Such information is critical to understanding the potential of PrP to interact with other proteins and ligands, which underlies its normal, but as-yet-still-unknown, function, and its capacity to change into other conformational states with apparently neurotoxic or,

Significant Achievements, Anticipated Outcomes and Future Work

Molecular dynamics simulations have been undertaken to characterize the changes in the conformational flexibility and surface properties of PrP when sugar chains (*N*-glycans) are added to the two known glycosylation sites, and to define the properties of the sugar chains themselves. As PrP is also anchored, at its C-terminus, to the membrane by a GPI

(glycophosphatidylinositol)-anchor, a GPI anchor was also added. Using NMR structures to develop starting coordinates, our initial work established suitable simulation conditions for unglycosylated PrP [1]; the model contained the structured C-terminal region of ~110 residues only as, unusually, the complete N-terminal region is shown to be unstructured by NMR. Defining appropriate simulation conditions was not straightforward due to the unusually open structure of PrP, which has a relatively low composition of formal secondary structure (only ~50%) and a very large number of charged residues on the protein surface; at effective pH 7, a net charge of -1 e.u. from 15 positively- and 14 negatively-charged residues. Simulations showed high sensitivity to the correct treatment of the electrostatic interactions, with stable trajectories achieved only by including all the long-range electrostatic interactions using the (computationally expensive) Particle Mesh Ewald (PME) method. These initial results on the importance of electrostatics and, particularly, stability of salt bridges had instructive implications for PrP structure and stability: PrP encounters a range of pHs physiologically during normal or abnormal recycling from the neutral pH membrane surface into endosomes or lysosomes (acidic pHs). This study was the first reported MD simulation for PrP. No

Simulations were then undertaken on PrP models with attached sugars and GPI anchor, to probe their effects on the protein stability, electrostatics and dynamics and to clarify which regions of the protein might be occluded from solvent and approach of potential ligands or other proteins, either from the attached sugars themselves or by proximity to the membrane [2]. The *N*-glycans are highly branched, making up nearly 1/3 of the total mass. A large box of water containing >10,000 water molecules was required to solvate the total molecular system (~33,000 atoms in total), and trajectories were run for several nanoseconds (ns), i.e. these are computationally very demanding simulations. The results showed PrP is stabilised overall from addition of the glycans, but that the stabilization appears indirect, by reducing the mobility of the surrounding water molecules, and not from specific interactions such as H bonds or ion pairs. As the attached glycans have three negatively charged SiaLex groups each, the surface electrostatic properties of PrP are totally changed. A negative electrostatic field covers most of the surface, including the whole surface of two of the three helices, but, significantly, the unusual totally hydrophilic Helix-A is not affected. Modelling showed Helix-A could readily dimerise in anti-parallel fashion: this model was consistent with indirect evidence for how the experimentally observed dimer might form. Together with separate simulations of the GPI anchor in a membrane model, the results showed a highly flexible GPI anchor, which would maintain the protein at a distance between 9 and 13 Å from the membrane surface, with little influence on its structure or orientational freedom. The work was featured on the cover of *Glycobiology*, and the graphics also won the Today's Life Sciences award for best modelling picture in 2000.

As conversion between two different conformations of PrP, “normal” PrP^C and “scrapie” PrP^{Sc}, is suggested to be part of the disease-causing mechanism, extended simulations (4 ns) were next conducted at elevated temperatures (effective 400 and 500 K, compared with previous 300 K) to investigate possible unfolding pathways [3]. Note that the simulation time is still much too short to simulate any transition state towards a PrP^{Sc}-like structure. Using α -distance matrices in combination with principal component analysis (PCA) we were able to identify regions of the protein which collapse first or, conversely, form fairly stable domains. Helix-C and Helix-A showed high intrinsic stability, even at 500 K, whereas Helix-B showed stability only in the glycosylated model, which might be attributed to its attached *N*-glycan.

The core of the stable structure at 500 K for the glycosylated model, is formed mainly by Helix-B and Helix-C; contact maps and PCA clearly identifies this region to be the last to unfold. Although the simulations showed apparently different unfolding pathways for the glycosylated and glycosylated models, multiple trajectory simulations are necessary to clarify whether this is really the case; some such simulations have been undertaken in conjunction with the following study.

In this most recent study we have combined the methods and experience of this project with those of project x11, which is developing the linear-scaling semi-empirical QM method, MOZYME [4]. This is apt as electrostatics appear important for PrP structure and function, and MOZYME offers the unique opportunity to calculate molecular electrostatic potentials (MEPs) of the whole protein at a QM level. Also, as another study on a protein ion channel recently undertaken at the ANU [Bliznyuk et al. (2002) J Phys Chem B 105, 12673] had found artifactual results for EPs calculated from MM force fields compared with those calculated by MOZYME, we were interested to compare these two calculational methods applied to a different, and more taxing, protein problem. An extensive simulation series for models of wild-type and 13 mutants of human PrP, 11 of them disease-associated, and including some with glycosylation, was undertaken. The results did not show differences in MEPs which would seem likely to contribute to the disease association of the mutants. This negative result is consistent with a now-huge amount of experimental structural and biophysical data on these mutants which has similarly failed to show a link between the mutations and likely properties, such as reduced general or local stability or gross changes in surface or electrostatic properties. However, the study did show large differences between MEPs calculated by the several methods. Although the implications of these results require further analysis, in general they sound a warning against drawing quantitative conclusions about protein-ligand or protein-protein interactions from calculations where electrostatic (i.e. long-range) components dominate.

Future work includes a detailed study of the water mobility from the glycosylated and non-glycosylated PrP simulations in order to gain further insight into factors influencing both protein and oligosaccharide flexibility. We are also starting to consider the effect of *N*-glycans on the structure and properties of another extracellular domain.

Computational Techniques Used

The SANDER module in Amber 5, with the all-atom AMBER94 force field for the protein and an adapted GLYCAM93 force field for the glycans, was used for the simulations, using periodic boundary conditions and constant volume. The PME (Particle Mesh Ewald) method was used to calculate long-range electrostatic interactions. Both unglycosylated and glycosylated models were immersed in a rectangular box of pre-equilibrated TIP3P water molecules (~5,000 and ~10,000, respectively) giving total system sizes of ~15,000 and 33,000 atoms, respectively, with simulation times of ~1.5 to 4 ns for studies 1-3, and 400-500 ps for study 4. Simulation analysis was done for: trajectories with the program CARNAL in Amber 5; secondary structure with the DSSP program; hydrophobic interactions using the VORONOI method; PCA using the fast NIPLAS algorithm; cluster analysis using the Average Linkage method; surface area and surface properties using the program NSC; model manipulation and visual analysis using INSIGHTII; picture generation using MOLSCRIPT, RASTER3D,

MOLMOL and POVRAY. MOZYME calculations were performed with MOPAC2000. For study 4, the MEP was calculated on a 65x65x65 grid, with spacing of 1.5 Å. For the AM1 method, a modified PMEP module in MOPAC2000 was used, while for AMBER charges the locally-written program GMIX was used. The modified PMEP module allows MEP calculation on a grid, as well as ability to add a contribution from the induced surface charges, for calculations on solvated molecules. Calculation of the MEP was also done with the DELPHI program, using default van der Waals radii and the same partial charges as for the Amber simulations, and also with partial charges obtained from the AM1 calculations in MOZYME.

The status of MOPAC on the VPP300, PC and APAC National Facility Compaq SC is given in the x11 report. Amber 5, particularly the SANDER module and PME option were highly vectorized (VPP300) or otherwise tuned for the VPP and PC. Some work has been done by National Facility staff to tune Amber (now) 6 for the Compaq SC, but further work is on hold pending release of a parallelized Amber 7 by its developers.

Publications, Awards and External Funding

1. J Zuegg, JE Gready. Molecular dynamics simulations of human prion protein: importance of electrostatic interactions for protein stability. *Biochemistry* 38, 1999, 13862-13876.
2. J Zuegg, JE Gready. Molecular dynamics simulations of human prion protein including the N-linked oligosaccharides and GPI anchor. *Glycobiology* 10, 2000, 959-974.
3. J Zuegg, JE Gready. Molecular dynamics simulations of human prion protein at high temperature. M/S in revision, 2002.
4. J Zuegg, AA Bliznyuk, JE Gready. Molecular electrostatic potential calculations on wild-type and mutants of human prion protein. M/S in preparation, 2002.

Study 4 was funded by a FECIT (Fujitsu European Centre for Information Technology) contract to ANUSF, as part of promotion activities for MOPAC2002 (demonstration of the capabilities and usefulness of MOZYME).

Evaluation of the Potential of MOZYME for the Study of Enzyme Reaction Mechanisms

Principal Investigator

Jill E. Gready

Computational
Proteomics Group
John Curtin School of
Medical Research
ANU

Co-Investigators

Stephen J. Titmuss

Computational
Proteomics Group
John Curtin School of
Medical Research
ANU

Alistair P. Rendell

Dept of Computer
Science
FEIT
ANU

Project

x11

Facilities Used

SC, VPP, PC

RFCD Codes

250106, 250601,
250699, 270108

Developments in applications software, coupled with increased HPC power which make very large calculations feasible, have created the expectation of a new era of large-scale biomolecular computation. One such development is MOZYME, a unique program module within the computational chemistry program MOPAC, which allows calculations to be undertaken at an approximate quantum mechanical (semiempirical QM) level on molecules with thousands of atoms. MOZYME is one of a few so-called linear-scaling (LC) QM methods, introduced recently, which achieve this capability by overcoming the bottle-neck of standard QM methods – the N^3 (or higher; N = no of orbitals) computational dependence. MOZYME LC is based on a localised molecular orbital method developed by J.J.P. Stewart, the originator of MOPAC. For conventional QM methods, the N^3 dependence limits calculations to 50-100 atoms (ab initio QM) or a few hundred atoms (semiempirical QM, as in MOPAC and MOZYME). This new capability to compute molecular properties and energies for such large molecules at an all-QM level, provides the opportunity to test the usefulness of MOZYME in real applications. We started with a problem for which MOZYME had been proposed, that is, studying enzyme reaction mechanisms, and compared its performance with another type of new method, hybrid quantum mechanical/mechanical methods (QM/MM) which overcomes the QM computational bottleneck in a completely different way. We also employed MOZYME to study protein properties in other contexts where an all-QM calculation would seem advantageous.

Significant Achievements, Anticipated Outcomes and Future Work

Comparison of results from MOZYME and QM/MM calculations on the reaction mechanism of the enzyme dihydrofolate reductase (DHFR) showed major and complex differences in components of the energies, as well as total energies in some cases, reflecting the different theoretical treatments of the interaction of the reactant subsystem with the rest of the enzyme; these have major theoretical ramifications (see below). In particular, the profiles showed close agreement in the total energies for fixed MM coordinates, but divergence when the MM geometry is allowed to relax during the reaction, largely due to the MM electrostatic energy [1]. After implementing energy decomposition into MOZYME to allow direct comparison with QM/MM components, the results showed a large change (~16 kcal/mol) in the MOZYME MM component of the fixed-geometry model due to polarization of the MM region surrounding

the active site QM region, which is not explicitly modelled by QM/MM. For the variable-geometry model the differences are even greater, with a 52 kcal/mol discrepancy in the relative reaction energies (product - reactant), largely due to MM energy differences [2]. Results so far have provided many surprises, not all of which have yet been explained. Future work will focus on further analysis of the different descriptions of the MM or MM-equivalent (i.e. in MOZYME) region, specifically polarization and other electrostatic contributions, which are especially large for the variable-geometry model. This will include investigation of relative energies as a function of distance from the reaction centre, and study of energy contributions for selected residues or residue-residue pairs. Also, although initially proposed as a method for studying, in particular, enzyme reactions (hence, the name) our studies suggested several other potential uses where an all-QM calculation of protein properties was desirable; we have used it for studying electrostatic potentials (see x04/d55) and it has been used by others at ANU for a similar purpose (Bliznyuk et al. (2002). *J Phys Chem B* 105, 12673). Ö

Other outcomes. Apart from these specific results, the project, and other concurrent work with MOZYME, provided several other significant insights; these findings could only be made once it became possible to do all-QM calculations on such large molecules. One unexpected outcome was the value of MOZYME as a quality control on QM/MM enzymic calculations, both on the quality of the parameterization of the QM/MM coupling terms, and on the risks of using static (i.e. without MD) QM/MM with optimization of MM-atom positions during the reaction, as is commonly done in the literature. Secondly, the results for the treatment of both short- and long-range electrostatics by MM, suggest that merely extending the size of the QM region in QM/MM calculations is not a universal solution to the MOZYME- and QM/MM-method differences; this is an important and timely result which conflicts with expectations in the literature and may provide a means for judicious choice of the QM boundary for large QM-region calculations now becoming practicable (e.g. see w05, u51/d52, u53/d55). Results from project x04/d55 which used both MOZYME and an MM-based method to calculate electrostatic potentials (EPs) of prion protein showed fundamental problems with the latter results. Finally, fundamental issues in the theoretical foundation of semiempirical theory and the meaning of energy components emerged; we have addressed one aspect so far [3], with other studies planned.

In summary the project offers an important conclusion. There has been a common assumption that application of current computational chemistry (CC) methods is limited largely by computing power, and, consequently, that if such power is now available, the methods would be able to deliver reliable results at the level now achievable by small-molecule CC. The MOZYME results demonstrate that this assumption is flawed. It cannot be assumed that the theoretical performance of current methods (including semiempirical QM theory), developed and extensively tested for small-molecule chemistry, will scale to large system size.

Computational Techniques Used

MOZYME is incorporated into MOPAC (Fujitsu Ltd Japan). Our application to real problems uncovered major deficiencies in the implementation and overall capability of the program. The project worked closely with MOPAC developer Dr A. Bliznyuk in ANUSF to remedy these. Major refinements, which also generated new R&D contracts with Fujitsu, were implementation of linear-scaling direct SCF to reduce memory limitations, and implementation

of linear-scaling COSMO to provide implicit solvation functionality. Within the project, an energy decomposition analysis routine (to complement that in the non-LC part of MOPAC) was implemented in the MOZYME module. The improved MOPAC program was tuned and available on the VPP300 and PC, and is now available on the APAC National Facility Compaq SC. The QM/MM calculations were performed using the locally-developed program MOPS (see u51/d52). MOPS was vectorised for the VPP300 and tuned for the PC; extensive efforts of parallelization for the National Facility Compaq SC are documented in the report for u51/d52. *ab initio* QM calculations required for parts of the study were done with GAUSSIAN98.

The MOZYME calculations were done on a DHFR model containing ~3200 atoms (3047 protein, 97 substrate + cofactor, 17 water) requiring ~8000 basis functions for the PM3 model. The QM/MM calculations were done with a QM region containing 48 atoms, with the remainder (i.e. ~3200 atoms) in the MM region. Several coordinate sets (e.g. with fixed or various relaxed MM regions) generated along a QM/MM reaction path were used as input to the MOZYME calculations. The main details of the QM/MM model were taken from ongoing QM/MM work (see u51/d52, u53/d55).

Publications, Awards and External Funding

ARC/DETYA SPIRT (APAI) Grant (1998-2001) with industry partner Fujitsu Ltd Japan to JE Gready, AP Rendell and R Nobes. "Application of linear scaling semi-empirical quantum chemical methods to the study of enzyme reaction mechanisms."

1. SJ Titmuss, PL Cummins, AA Bliznyuk, AP Rendell, JE Gready. Comparison of linear-scaling semiempirical methods and combined quantum mechanical/molecular mechanical methods applied to enzyme reactions. *Chem Phys Letters* 320, 2000, 169-176.
2. SJ Titmuss, PL Cummins, AP Rendell, AA Bliznyuk, JE Gready. Comparison of linear-scaling semiempirical methods and combined quantum mechanical/molecular mechanical methods for enzyme reactions II: an energy decomposition analysis. *J Comput Chem*, 2002, in press.
3. PL Cummins, SJ Titmuss, D Jayatilaka, AA Bliznyuk, AP Rendell, JE Gready. Comparison of semiempirical and *ab initio* QM decomposition analyses for the interaction energy between molecules. *Chem Phys Letters* 352, 2002, 245-251.

Hybrid QM/MM Studies of the Reaction Mechanism of Dihydrofolate Reductase

Principal Investigator

Jill E. Gready

Computational

Proteomics Group

John Curtin School of

Medical Research

ANU

Co-Investigators

Peter L. Cummins

Stephen P. Greatbanks

Computational

Proteomics Group

John Curtin School of

Medical Research

ANU

Ivan Rostov

Supercomputer Facility

ANU

Alistair P. Rendell

Dept of Computer

Science

FEIT

ANU

Kritsana Sagarik

School of Chemistry

Suranaree University of

Technology

Thailand

Project

w05

Facilities Used

SC, VPP, PC

RFCD Codes

250106, 250601,

250699, 270108

While much has been learned about enzyme mechanisms at the molecular level from x-ray crystallography, the picture is largely static, while other experimental techniques offer a more thermodynamic view but only at the macroscopic level. In principle, computer simulation provides a means to complement experimental data and provide an accurate energetic representation of the chemical events at the reaction centre, and of the energetic and dynamical roles of the enzyme and its components. In particular, it offers a means to study the “invisible” players, protons and water, not seen by crystallography. All these players and properties may play a role in enzyme catalysis: the problem is how to decipher for a particular enzyme of interest. Delineation of the key contributions for a molecular system of typical enzyme size (i.e. at least 3,000 protein atoms plus, at least, an equivalent number of solvent molecules) is beyond the feasibility of any current computational method; no current method can provide this reliably at the level now routinely achievable in small-molecule chemistry. Rather, current methods offer “windows” which can be divided into three main classes (QM level and type; MM/MD; and hybrid QM/MM+MD) depending on the quality of the electronic description of the molecule; the quality of the representation of its chemical environment; and the degree of dynamic sampling. We are pioneering a multi-method approach using all these methods to study the reaction mechanism of the enzyme dihydrofolate reductase (DHFR), and, by comparative evaluation, to define the ranges of applicability of the different methods to large complex protein systems. In addition to scientific interest and applied interest as a major drug target, DHFR is well recognized as a test system. This is because the catalytic mechanism and ligand-binding properties have proven so refractory to experimental and computational investigation over 30-40 years, that there is an expectation their elucidation will offer general insights into enzyme mechanisms and how to simulate them.

Significant Achievements, Anticipated Outcomes and Future Work

Background. DHFR catalyses the reduction of folate to dihydrofolate(DHF) and tetrahydrofolate (THF) *via* a two-step reaction of protonation, and hydride-ion transfer from the cofactor NADPH. Key aspects of the mechanism have long remained elusive, specifically: whether the carboxylate group of the conserved active-site Asp or Glu residue is protonated or ionized during the reaction; the source of the protons; and the role of active-site water molecules (especially W206, *E. coli* DHFR numbering). These issues have been difficult to define computationally because of DHFR complexities: it seems to involve many, if not all, protein complexities (proton shifts, active-site water, loop-closing mechanisms and dynamics, highly non-isotropic global and active-site electrostatics); it is an old essential and ubiquitous enzyme with an odd chemically-taxing role, catalysing *two* hydride-ion transfer reactions on *two different* carbons of a complex *N*-heterobicyclic π -system (pterin ring) in *two* oxidation states, and with an attendant requirement for protonations of *two different* ring nitrogens; and, finally, accurate electronic description of these large ring systems and the active moiety (nicotinamide ring) of the cofactor NADPH, and changes during the reaction, is non-trivial. There is no basis for assuming these effects are additive, so the challenge is to define how they interact, and to calculate their co-operative quantitative contributions.

Results. With respect to the main questions and complexities, we have established so far that:

(i) Polarization by DHFR of its substrates is not sufficient to favour hydride-ion transfer at C7 (folate) or C6 (DHF) without pre-protonation, as reported in the literature from earlier DFT calculations. The studies used both point-charge embedded *ab initio* QM (SCF, MP2 and various DFT models) and linear-scaling semiempirical QM (MOZYME) calculations [1,2]. Large substrate polarizations from the DFT results were shown to be artifacts of the DFT description of large anions, i.e. they represent a clear example of method breakdown. However, a new finding was that polarization at the active carbon (C4) of the nicotinamide ring *does* activate C4 for hydride-ion transfer. The MOZYME calculations [2] also allowed, uniquely, evaluation of charge-transfer effects to the enzyme; these were small in agreement with other MOZYME studies on the reaction path (see report for x11).

(ii) A combined *ab initio* QM and QM/MM+MD study [3] of the energetics of the H-bond interaction between the active-site carboxylate (i.e. ionized form) and a protonated substrate (a folate mimic, 8-methyl-pterin) predicted that the neutral-pair form (i.e. proton transferred from N3 of substrate to form the carboxylic acid) is more stable than the ion-pair form, the form universally assumed in DHFR discussions in the literature. This is the case both in vacuum and in the enzyme although this relative stability is reduced to ~ 5 kcal/mol by the more favourable electrostatic interactions made by the ion pair with the enzyme. Such a neutral-pair form does not activate the pterin ring for hydride-ion transfer, as required by the pre-protonation hypothesis for catalysis. However, the results were suggestive of a small free energy gap and the possibility of a low energy barrier for the proton transfer, a mechanism proposed in the literature as operative in many enzyme reactions. Followup *ab initio* SCRF calculations [4] indeed established a low-barrier hydrogen bond (LBHB): the ion pair corresponded to a shallow local minimum at the transition state (TS) between the two neutral-pair complexes. Thus, the H-bonded system has maximum probability of being found in the ion-pair form. As analysis of charge distributions indicates the TS is highly cationic, a

mechanism for activation of substrate towards hydride-ion transfer similar to that of the formal cation seemed possible.

(iii) However, using a combination of methods as for (ii), extension of these studies to consider both the role of protonation *and* a conserved active-site water molecule for the binding and reaction of folate and dihydrofolate provided another picture [6,7]. Results showed that a conserved water molecule (W206, *E. coli* DHFR) that is H bonded to both the OD2 oxygen atom of the carboxyl (Asp) sidechain and O4 of the pterin/dihydropterin ring, is critically important, and appears to determine preferred protonation sites for the enzyme-bound substrates [6]. Single protonation of the system does not lead to the N8 (folate) and N5 (DHF) protonated forms required for a pre-protonation activation mechanism, but these forms *are* the most stable after addition of two protons; the other proton resides on the carboxyl group. This result also appears to have indirect experimental support: calculated H-bond distances for W206 and protonated OD2 complexes (OD2–N3 distances < OD2–NA2 distances) correlate well with those observed in many x-ray structures. Analogous studies for the reaction at the enthalpic level (*ab initio* QM) showed that reduction takes place only if the active-site Asp27 is protonated, but that at the free-energy level (QM/MM +MD) the catalytic advantage of a protonated Asp27 residue is small [7]. For both ionized and neutral states of Asp27, analysis of the electrostatic interaction energies between the QM and MM regions in the MD simulations showed that the enzymic environment (MM region) plays a far greater role in stabilizing the final products than in stabilizing the TS complex. However, although for the larger QM region model (85 atoms), the range of activation free energies (10 - 12 kcal/mol) is similar to an experimental value (13 kcal/mol), in general quantitative agreement with experimental free energies was poor.

In summary using our mixed-method approach, we have established that a balanced description of the both the active-site and wider-enzyme environment needs to be built up. Provisionally we have established that: the conserved active-site water (W206, *E. coli* DHFR) is critically important for correct protonation of the enzyme-bound substrates, but plays little role in the actual hydride-ion step; Asp/Glu (Asp27, *E. coli* DHFR) seems to be needed to be protonated for both steps; and the enzymic environment plays a greater role in stabilizing the final products than the TS complex. The last result is novel and interesting as the reaction is rate-limited by product dissociation and is functionally irreversible. However, even the most recent calculations do not allow firm conclusions to be drawn.

Future work. Work along the following lines will be undertaken to test these results, to quantitate energetic components of the two reactions, and also to further develop and evaluate methods (see below and work in conjunction with project u51/d52). A large series of multiple-trajectory QM/MM calculations (200-500 ps each) will be performed on both protonation and reaction steps with the aim of testing the above provisional conclusions and the other mechanistic issues, as well as understanding the convergence of the results with respect to system size and composition. A systematic buildup of the QM region to include complete substrate and cofactor molecules as well as additional active-site residues is planned, particularly residues perpendicular to the pteridine-ring plane. This will allow estimation of long-range electrostatic effects from the cofactor, and possible π - π stacking interactions. As this would require QM regions of up to 250 atoms with 9 link atoms, concepts being developed in project x11 for choosing QM boundaries will be employed. These studies will be

complemented by very large *ab initio* fragment calculations on the QM regions. Our experience (and that of others in the literature) is that geometry optimizations of such systems with “untethered” ends (i.e. lacking constraints which would apply in the enzyme) are unpredictable; this will likely to much worse for larger QM-region analogues (i.e. up to 9 broken bonds). To date we have used fixed-atom constraints, but these are very artificial: we are now exploring the use of methods such as ONIOM in GAUSSIAN98, which allow imposition of softer enzyme-like potentials to restrict the conformational space of the fragment complexes.

Computational Techniques Used

Methods and programs. Three main methodologies are being used: (i) SE-QM/MM with MD simulation as implemented in our locally-developed MOPS program (see u51/d52). This includes recent parallelization under MPI in the APAC National Facility Compaq SC version of all compute-intensive terms except the QM SCF, and for running of multiple trajectories to improve the efficiency of free energy sampling. Other enhancements under development, which will be used and tested in this project, include: corrections to the SE-QM force fields to prevent both H-bond and covalent bond breaking, a significant problem for large QM regions; and use of system-specific parameters to recalibrate the SE-QM method for the specific enzymic reaction, to correct for the often large errors in reaction energies and activation energies from SE-QM. (ii) Linear-scaling MOZYME within MOPAC for all-enzyme QM calculations, as being developed and tested in project x11. (iii) GAUSSIAN98 for *ab initio* QM (including with implicit solvent methods, e.g. SCRF) and ONIOM (a type of QM/MM method) calculations, which was tuned and maintained for the VPP300 under the ANU-Fujitsu Area 3 contract, and is maintained for the APAC National Facility Compaq SC by National Facility staff.

Capabilities. Typical size-dependent limitations for these methods with current supercomputer resources are: 50-100 atoms for *ab initio* QM (e.g. SCF, DFT) calculations with full geometry optimization, including with an approximate environment (point-charge, implicit solvent); several hundred atoms for conventional semiempirical QM (e.g. AM1, PM3) calculations in MOPAC with full geometry optimization, including with implicit solvent (COSMO) model; 10,000 atoms or more for linear-scaling (MOZYME) SE-QM calculations with very limited geometry optimization, but including implicit solvent (COSMO) model; 30,000 atoms or more (enzyme + solvent) for multiple trajectory MD simulations (MOPS) of nanoseconds (ns) each; and 10,000-20,000 atoms for multiple trajectory simulations (MOPS) of ~500 ps each with a QM/MM potential (up to 200 atoms in the SE-QM region).

Publications, Awards and External Funding

1. SP Greatbanks, JE Gready, AC Limaye, AP Rendell. Enzyme polarization of substrates of dihydrofolate reductase by different theoretical methods. *Proteins* 37, 1999, 157-165.
2. SP Greatbanks, JE Gready, AC Limaye, AP Rendell. Comparison of enzyme polarization of ligands and charge-transfer effects for dihydrofolate reductase using point-charge embedded *ab initio* QM and linear-scaling semiempirical QM methods. *J Comput Chem* 21, 2000, 788-811.

3. PL Cummins, JE Gready. Combined quantum and molecular mechanics (QM/MM) study of the ionization state of 8-methylpterin substrate bound to dihydrofolate reductase. *J Phys Chem B* 104, 2000, 4503-4510.
4. PL Cummins, JE Gready. QM/MM and SCRF studies of the ionization state of 8-methylpterin substrate bound to dihydrofolate reductase: existence of a low-barrier hydrogen bond. *J Mol Graph Model* 18, 2000, 42-49.
5. JE Gready, PL Cummins. Mechanism-based substrates and inhibitors of dihydrofolate reductase. *in Free Energy Calculations in Rational Drug Design*, Reddy MR, Erion MD, eds., Kluwer Academic/Plenum Publishers, 2001, pp. 343-364.
6. PL Cummins, JE Gready. Energetically most likely substrate and active-site protonation sites and pathways in the catalytic mechanism of dihydrofolate reductase. *J Am Chem Soc* 123, 2001, 3418-3428.
7. PL Cummins, SP Greatbanks, AP Rendell, JE Gready. Computational methods for the study of enzymic reaction mechanisms I: application to the hydride transfer step in the catalysis of dihydrofolate reductase. *J Phys Chem B*, 2002, submitted.

The Density Matrix Renormalization Group

Principal Investigator**Miklos Gulacsi***Theoretical Physics**RSPHysSE**ANU***Co-Investigator****Ian McCulloch***Theoretical Physics**RSPHysSE**ANU*

The Density Matrix Renormalization Group (DMRG) method is used to obtain solutions of the many-body Schroedinger equation for a lattice Hamiltonian. It is mostly of use in strongly correlated electron systems where, due to the nature of the interactions, traditional techniques give very poor results. We have applied the technique to determine the ground state wavefunction for a variety of models. This is then used to determine various physical quantities, for example the ground-state phases, magnetization, susceptibility etc. The models we have studied have applications in several areas of condensed matter physics, including the theory of high-temperature superconductivity, heavy-fermion metals and the colossal-magnetoresistance materials.

Project

x18

Facilities Used

VPP, PC, SC

RFCD Codes

240203

Significant Achievements, Anticipated Outcomes and Future Work

The major technical achievement of the project has been the extension of the DMRG algorithm to conserve non-Abelian symmetries. This has a dramatic effect on the efficiency of the algorithm; improvements of more than two orders of magnitude have been seen in some cases. This arises because the conserved quantities associated with a non-Abelian symmetry allow the Hamiltonian matrix to be block-diagonalized while also removing what would otherwise appear as degenerate eigenvalues. In addition these conserved quantities have a direct physical interpretation, which allows us to calculate physical properties which were previously prohibitively difficult to obtain. This algorithm is the successor to and a major improvement on the earlier version described in the 1998 Annual Report for project x13.

The most significant achievement of the project to date is the discovery of a previously unrecognized ferromagnetic phase in the one-dimensional Kondo lattice model, and the subsequent discovery of the corresponding phase in the local-moment regime of the Anderson lattice model. This is a surprising result as prior to this work it was generally believed that the phase diagram of the 1D Kondo lattice had been completed. This new phase resolves some of the confusion regarding the ground-state phases of the Kondo lattice model, while also raising many new questions which are still to be resolved.

Some of the most interesting open questions in condensed matter physics revolve around the nature of strongly correlated electrons in two dimensions. This encompasses such effects as high-temperature superconductivity and the colossal-magnetoresistance materials. With the APAC National Facility machine, the application of the DMRG method to models of these

phenomena is becoming practical, so the major emphasis of the project for the future is to fully exploit the resources of the SC by parallelizing the code using MPI and/or multithreading.

Computational Techniques Used

The DMRG software has been written by I. McCulloch. The implementation makes heavy use of BLAS, LAPACK and GNUmp and ARPACK libraries. The SC has been invaluable by providing parametric parallelization, allowing the calculation of the properties of a model over a significant region of the parameter space of each model. Even using a single processor per job, the SC allowed us to make significant improvements in the accuracy of the calculations which were essential in obtaining the results described previously.

DMRG works by diagonalizing the Hamiltonian matrix of a quantum lattice model in a truncated basis set. The truncation is necessary because the dimension of the Hamiltonian matrix grows exponentially with the system size. This truncation is effected by dividing the lattice into two parts, and calculating which basis states in each part have the largest weight in the overall wavefunction. The basis is then truncated to keep only these states. The boundary of this partition is then shifted by one lattice site and the procedure is iterated until the energy and wavefunction has converged. The accuracy of the truncation depends strongly on the length of the interface between the two parts of the system, thus the algorithm performs substantially better for a one-dimensional chain than for higher dimensions.

Currently the bulk of the computational effort is used in obtaining the ground-state wavefunction and energy of the Hamiltonian. This is represented by the eigenvector and eigenvalue of largest absolute value of the Hamiltonian matrix. This matrix typically has a dimension of ~ 1 million by 1 million, and is stored in a block-sparse form. The diagonalization is done using the Davidson algorithm, which is near-ideal for this purpose as the matrix is usually strongly diagonally dominant. Also an extremely good initial guess-vector for the diagonalization is determined by a basis transformation from the previous DMRG step, thus only a small number of matrix-vector multiplies are required each step (usually less than 10).

Much of the complexity of the algorithm arises from the need to keep track of the basis at each iteration and perform the necessary basis truncations and transformations. While this represents a large proportion of the size of the code, the computational effort required here is relatively small.

Publications, Awards and External Funding

The algorithm developed by I. P. McCulloch and M. Gulacsi for use in this project is described in the following papers:

I. P. McCulloch and M. Gulacsi, Density matrix renormalization group method and symmetries of the Hamiltonian, *Aust. J. Phys.* 54, 2000, 4.

I. P. McCulloch and M. Gulacsi, Total Spin in DMRG, *Phil. Mag. Lett.* 81, 2001, 447.

I. P. McCulloch and M. Gulacsi, The Non-Abelian Density Matrix Renormalization Group Algorithm, *Europhys. Lett.*, 2002, in press.

Results obtained using the above algorithm on the PC and SC are published in the following papers:

I. P. McCulloch, M. Gulacsi, S. Caprara, A. Juozapavicius, A. Rosengren, Phase diagram of the 1D Kondo lattice model, *J. Low Temp. Phys.* 117, 1999, 323

I. P. McCulloch, A. Juozapavicius, A. Rosengren, M. Gulacsi, Localized spin ordering in Kondo lattice models, *Phys. Rev. B* 65, 2002, 052410

I. P. McCulloch, A. Juozapavicius, A. Rosengren, M. Gulacsi, Ferromagnetism in Kondo lattice models, *Phil. Mag. Lett.*, 2002, in press.

Results obtained using the PC and SC formed the major component of this PhD thesis:

I. P. McCulloch, *Collective Phenomena in Strongly Correlated Electron Systems*, PhD Thesis, Australian National University, 2001.

Fault Line and Boundary Estimation from Spatial Data

Principal Investigator

Peter Hall

Mathematics Research
Section
SMS
ANU

Co-Investigator

Christian Rau

Mathematics Research
Section
SMS
ANU

In many instances in applications, spatial data may be viewed as observations on an underlying regression surface. Estimation of jump curves or “fault lines” in such surfaces is then a task of primary interest. The developed methodology for estimating fault lines caters for features which are desirable both in theory and practice, such as irregularly-spaced design. An intuitively appealing but computationally demanding approach is taken, and asymptotically (as sample size tends to infinity) confidence envelopes for this estimator were derived theoretically. This approach consists of transforming the raw data into a smooth surface wherein the fault line appears as a ridge, which is defined to be its estimator.

Project

x21, d33

Facilities Used

PC, SC

RFCD Codes

230203

Significant Achievements, Anticipated Outcomes and Future Work

The projects explore the statistical and numerical aspects of the estimator, mainly in dependency on both the sample size, or more precisely the intensity of the random point process which generates the design points, and the local geometric properties of the fault line. Indeed, the radius of the confidence envelope is estimated using estimates of quantities such as curvature at each point of the fault line estimate, that is, the ridge.

The distribution which yields the percentiles for the asymptotically correct coverage level of the envelopes was studied via a Karhunen-Loeve expansion of a planar Gaussian field. Although the methodology has been established a considerable while ago, there is notably little existing work dealing with the practical aspects that were investigated in the projects.

Future work includes the investigation of the methodologically quite similar case of estimation of fault lines in bivariate densities, and the estimation of density support boundaries. In either case, methods for adaptive bandwidth selection are of interest. For regression surfaces, the investigation of the case where the distribution of the errors is not normal, but heavy-tailed, is also planned. Questions that arise in the context of simulating the distribution mentioned in the previous paragraph appear to be of independent interest, such as the expediency of sparse-grid techniques in calculating a four-dimensional Fourier Transform.

Computational Techniques Used

The code used in the projects was mainly developed in C++ and Matlab. Generally, the C++ algorithms served to generate the datasets used in the simulations, while the Matlab routines served to find the likelihood maximisers via the generalised Nelder-Mead method (or variants thereof), and to perform the geometrical construction of the confidence envelopes. In the latter task, estimated values of curvature of the fault line were produced using Fortran code from other authors. The computation of the likelihood, though only involving elementary operations, was strongly affected by sparseness difficulties that are common in spatial problems. These difficulties were compounded in estimating curvature. Karhunen-Loeve calculations involve the treatment of large matrix eigensystem problems, which posed inhibiting limitations, notably again in terms of memory, on smaller-scale facilities.

A Rigorous and Efficient Method for Inferring Phylogenetic Relationships

Principal Investigator

Alexander Isaev

Centre for

Bioinformation Science

JCSMR

ANU

The central challenge in biology is to understand how information contained in DNA sequences gives rise to the diverse physical properties of living systems. The differences observed between individuals of the same and/or different species stem largely from differences in their DNA. The research we seek to perform using the APAC National Facility is aimed at improving the biological realism of models of protein coding DNA sequences.

Co-Investigator

Gavin Huttley

Centre for

Bioinformation Science

JCSMR

ANU

Project

x36

Facilities Used

SC

RFCD Codes

270208

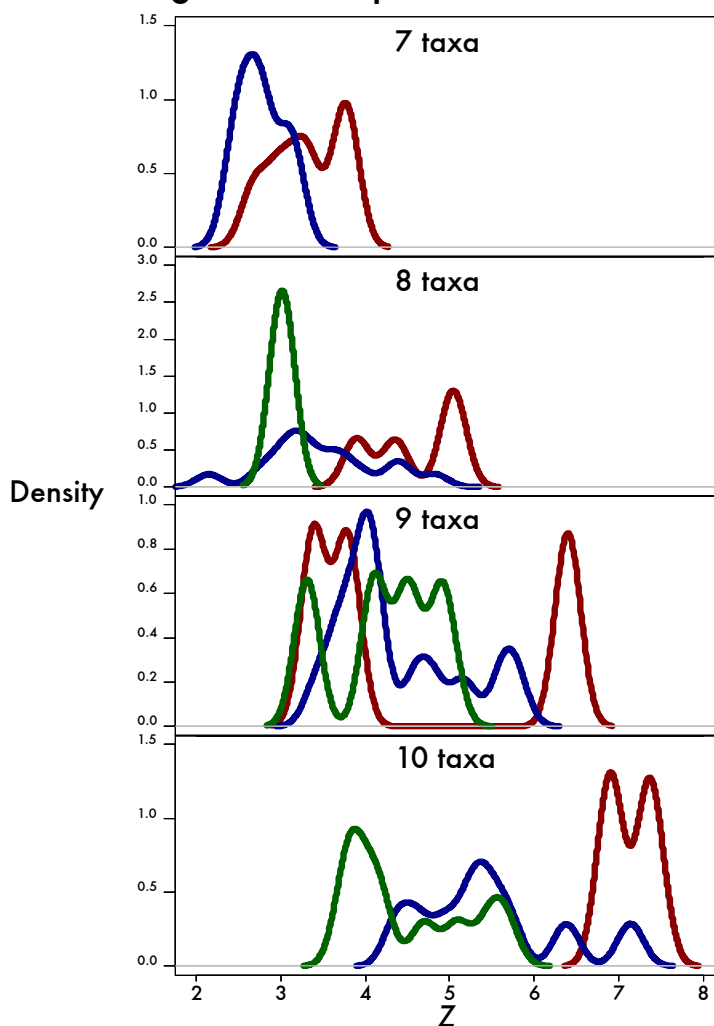
Significant Achievements, Anticipated Outcomes and Future Work

Since November 2001 when our project was first allocated time on the APAC National Facility, we have completed calculations for 7-10 OTU,s (2173509 trees) for 4 simulation rounds and obtained a preliminary distribution of the Z-score statistic (termed X in our original application). These calculations have been for project 1 from our original application.

Z is a measure of the ‘true-tree’ location, in units of standard deviations, relative to the mean weighted sums of squares determined from all possible tree topologies for a sample. Understanding its distribution is important for designing a fast and statistically rigorous method of phylogenetic reconstruction. From our first round of simulation we established that the number of taxa and the number of paired tips on the tree topology are two major factors influencing Z in a linear fashion (see figure). The latter effects were verified in simulations 2, 3, and 4 with minor deviations.

These results have led us to formulate what we call the ‘breathing tree’ model: tree topologies can be transformed into each other through simultaneous shrinking and growing of branches at different places. We anticipate constructing a numerical description of the ‘breathing tree’ that represents a parametrisation of tree topology. This parametrisation may allow us to numerically compare different evolutionary models, replacing the current parametric bootstrap procedure that is in many instances computationally prohibitive.

Number of taxa and tree topology – major factors effecting statistical power



Topology colour coding

— 4 paired tips

— 3 paired tips

— 2 paired tips

Project 2 will be started during the holiday period.

Computational Techniques Used

The sequence simulation program Seq-Gen combined with a custom C program, bushGen, were used to generate phylogenetic trees under a number of evolutionary models and to simulate replicate sequence data sets using these models. A custom application (wss) was written in C to calculate the weighted sum of squares for every simulated dataset and every tree topology for the corresponding number of OTUs. The mean and variance of the resulting distribution were used to measure the distance of the fit of the true-tree from the mean, as Z .

Parallelisation was achieved by submitting different sub-sets of the simulated data to the normal queue on the APAC National Facility SC and assembling the results after job completion using Python scripts. The number of CPUs available on the SC was critical in obtaining results in such a short time interval.

Highly Nonlinear Solitary Waves in Compressible Fluids

Principal Investigator

Brian Kennett
*Seismology and
Geomagnetism
RSES
ANU*

Co-Investigators

**Damien Bright
Douglas Christie**
*Seismology and
Geomagnetism
RSES
ANU*

This project is concerned with the numerical simulation of large amplitude trapped wave motions in atmospheric waveguides. Observations of such waves in the atmosphere indicate that the formation, evolution and decay of these waves are strongly influenced by the structure of the waveguide and the extent to which the overlying tropospheric layer absorb energy. The numerical simulation is based on a high resolution model of compressible fluid behaviour under different styles of forcing. A two layer system with a sharp change in the gradients of physical properties has been used to explore the development of strong disturbances from down drafts, e.g., due to thunderstorms.

Project

s52

Facilities Used

VPP

RFCD Codes

260201

Significant Achievements, Anticipated Outcomes and Future Work

The project was successful in generating stable numerical solutions for the fully compressible fluid models of the lower atmosphere subjected to large deformations. The way in which the nonlinear disturbances propagate depends strongly on the physical conditions and a wide variety of behaviour was explored in the PhD thesis written by Damien Bright in 2000. Nonlinear disturbances were generated by using down draft models of descending cold fluid, which tend to entrain their surroundings, and by careful transfer of results from incompressible flow followed by an adjustment period to allow the effects of compressibility to come into play. This study indicates that it is possible to achieve long-lived disturbances of high amplitudes whose effects can extend through much of the atmosphere.

Computational Techniques Used

The properties of the evolution on large amplitude wave phenomena need to be determination from integration of the fully nonlinear equations for compressible fluid flow. The conditions are nonhydrostatic and local dissipation below the scale of the numerical mesh is included through a second order operator. The boundary conditions are implemented using a combination of symmetry conditions and Fourier representations to allow for outward wave propagation at the top and sides of the model. The internal solution is carried out with a fine mesh second order finite-difference scheme including the full set of physical equations, e.g. diffusion of heat for the fluid.

Surface Wave Tomography

Principal Investigator

Brian Kennett

*Seismology and
Geomagnetism
RSES
ANU*

Using recording from portable and permanent seismic stations across Australia, the large amplitude surface waves from regional earthquakes are analysed to determine 3-dimensional structure in seismic wavespeed. For each path from a source to a receiver an average velocity model is determined and these constraints on the 3-D structure are finally combined to produce the 3-D image.

Co-Investigator

Eric Debayle

*Seismology and
Geomagnetism
RSES
ANU
(now) EOST— CNRS
Universite Louis
Pasteur
Strasbourg, France*

Project

x14

Facilities Used

SC, VPP

RFCD Codes

260206

Significant Achievements, Anticipated Outcomes and Future Work

We have been able to derive the first 3-D representation of both seismic wavespeed and anisotropy in the shear wave velocity beneath the Australian continent. We are able to demonstrate both the presence of high seismic wavespeeds beneath the Precambrian part of the continent, and a change in the nature of anisotropy with depth. Shallow anisotropy reflects past deformation, but the deep part which is nearly aligned with the motion of the Australian plate indicates current deformation at the bottom of the mechanical lithosphere. As time progressed we have incorporated larger volumes of data and improved the methods used to generate and assess the 3-D image.

Computational Techniques Used

The process requires both the inversion of seismic waveforms to extract a velocity model with depth and the combination of such models to retrieve the 3-D structure. The wavespeed inversion is nonlinear, whereas the second step is linear but involves a very large system of equations. We have used a continuous representation of the 3-D model and this requires

extensive computation but includes the influence of the neighbourhood of the propagation paths and thereby produces more reliable models.

Publications, Awards and External Funding

E. Debayle, B.L.N. Kennett , The Australian continental upper mantle - structure and deformation inferred from surface waves, *J. Geophys. Res.*, 2000, B'105, 25443-25540.

E. Debayle, B.L.N. Kennett, Anisotropy in the Australian upper mantle from Love and Rayleigh wave inversion, *Earth Planet. Sci. Lett.*, 184, 2001, 339-351.

Simulation of Seismic Wave Propagation Using a Wavelet-Based Method

Principal Investigator

Brian Kennett

Seismology
RSES
ANU

A wavelet-based method is developed for numerical simulation of seismic wave propagation in heterogeneous media. The method is not only accurate in obtaining seismic responses, but also stable during computation by treating spatial derivative terms in equations using a wavelet transform. We expand the method to seismic wave studies in global scale using parallel computing on a supercomputer.

Co-Investigator

Tae-Kyung Hong

Seismology
RSES
ANU

Project

x35

Facilities Used

SC

RFCD Codes

260206

Significant Achievements, Anticipated Outcomes and Future Work

The wavelet-based method has been introduced for numerical modeling of an elastic wave propagation in two dimensional media problems. The scheme represents spatial differentiation operators through wavelet bases and the resulting second-order differential equations for time are treated by a displacement-velocity formalism and a semigroup approach. The wavelet method for a spatial differentiation is not a grid-based scheme in the physical domain like a Fourier method although sampling is needed at given points. Therefore, we can maintain an accuracy of computation of spatial derivative uniformly throughout a whole domain in contrast to usual grid methods such as finite difference scheme that cumulates numerical errors during computation of derivative terms from grid to grid. For classical 2-D problems, the numerical results exhibit a high accuracy compared to analytic solutions. The method is not only stable during numerical computation, but also has achieved quite good results in various comparisons. Also, the method works well for both simple heterogeneous media and highly perturbed random velocity media. The wavelet-based method is very effective not only in the case of sudden variation of physical parameters at a boundary, but also for linear gradients where physical parameters are changing continuously. Moreover, the method provides stable time responses in a highly perturbed medium and obeys the energy conservation law. We expect the method can be extended to complex stochastic media problems where accurate treatments of spatial derivatives are essential for stable modeling. Also, the method can be used successfully in the seismic quantitative studies such as measurement of energy loss during wave propagation in attenuating media. Future work will concentrate on characterization of

the influence of different styles of heterogeneity and the extension of the technique to 3-D where parallel techniques are essential.

Computational Techniques Used

The principle of the wavelet-based method is described as follows. The governing elastic wave equations are transformed to a first-order differential equation system in time with a displacement-velocity formulation. Spatial derivatives are represented with a wavelet expansion using a semigroup approach. The evolution equations in time are derived from a Taylor expansion in terms of wavelet operators. The wavelet representation allows high accuracy for the spatial derivatives. Due to the use of a displacement-velocity scheme we reduce memory requirements by about 30% compared to the use of a velocity-stress scheme. Although the wavelet approach is highly accurate for the heterogeneous media, the method is still expensive in computational cost compared to other methods such as finite difference method or finite element method. With application of parallel computing using several supercomputers with high speed computational performance, we not only reduce the computational time, but also obtain good results from larger domains.

Non-perturbative Theory of Atomic Ionization with Two-electron Continuum

Principal Investigator

Anatoli Kheifets
Electron Physics Unit
RSPHysSE
ANU

A long-term objective of this project is to develop a general method for calculating atomic ionization leading to a final state with two-electron continuum. The theory of such processes should account for the inter-electron interaction between the two outgoing electrons and between each of these electrons with the ion core. A non-perturbative approach to the inter-electron interaction is to be taken by solving the Lippmann-Schwinger equation. The interaction with the core is described within the Hartree-Fock approximation.

Project

x01

RFCD Codes

240300

Significant Achievements, Anticipated Outcomes and Future Work

The close-coupling theory of the multiple atomic ionization has been further extended. The role of various correlation mechanisms has been uncovered in double photoionization of the helium atom. The cross-over between the shake-off and two-step mechanisms has been predicted theoretically and confirmed experimentally. The full parametrization of the symmetrized double photoionization amplitudes has been achieved offering a general description of the double photoionization process at a very wide range of photon energies and geometries of the two-electron escape. The second-order corrections to the electron impact ionization have been estimated in the double ionization of the helium atom.

Computational Techniques Used

The code has been ported from SGI to the Compaq AlphaServer SC. It was linked to the Compaq Extended Math Library (parallel routines). The code has been compiled using -omp multiprocessor option. With these modifications a two-fold gain in speed was achieved as compared to the SGI.

Publications, Awards and External Funding

A. S. Kheifets, I. Bray, J. Berakdar, and C. Dal Cappello. Comparative theoretical study of (e, 3e) on helium: Coulomb-waves versus close-coupling approach. *J. Phys. B*, 35(1):L15—L21, 2002.

A. Dorn, A. Kheifets, C. D. Schroeter, B. Najjari, C. Hoehr, R. Moshhammer, and J. Ullrich. Double ionization of helium by electron-impact in the impulsive regime. *Phys. Rev. A*, 65(2):115202, 2002.

G. Turri, L. Avaldi, P. Bolognesi, R. Camilloni, M. Coreno, G. Stefani, J. Berakdar, and A.S. Kheifets. Photodouble ionisation of He at 80 eV excess energy in equal energy sharing condition. *Phys. Rev. A*, 65(3):4702, 2002.

- A. S. Kheifets and I. Bray. Frozen core model of the double photoionization of beryllium. *Phys. Rev. A*, 65(1):2710, 2002.
- A. S. Kheifets and I. Bray. Symmetrized amplitudes of the helium atom double photoionization. *Phys. Rev. A*, 65(2):2708, 2002.
- C. Dawson, S. Cvejanovic, D. P. Seccombe, T. J. Reddish, F. Maulbetsch, A. Huetz, J. Mazeau, and A. S. Kheifets. Helium ($\gamma, 2e$) triple differential cross sections at an excess energy of 60 eV. *J. Phys. B*, 34(16):L525—L533, 2001.
- P. Bolognesi, R. Camilloni, M. Coreno, G. Turri, J. Berakdar, A.S. Kheifets, and L. Avaldi. Complementary TDCS for the photo-double ionisation of He at 40 eV above threshold in unequal energy sharing conditions. *J. Phys. B*, 34(15):3193—3203, 2001.
- A. Lahmam-Bennani, A. Duguet, M. N. Gaboriaud, I. Taouil, M. Lecas, A. S. Kheifets, J. Berakdar, and C. Dal Cappello. Complete experiments for the double ionisation of He: ($e, 3e$) cross sections at 1 keV impact energy and small momentum transfer. *J. Phys. B*, 34(15):3073—3087, 2001.
- A. S. Kheifets. On different mechanisms of the two-electron atomic photoionization. *J. Phys. B*, 34(8):L247—L252, 2001.
- A. Dorn, A. Kheifets, C. D. Schroeter, B. Najjari, C. Hoehr, R. Moshhammer, and J. Ullrich. Double ionization of helium by electron-impact: Complete pictures of the four-body break up dynamics. *Phys. Rev. Lett.*, 86(17):3755—3758, 2001.

Large-Scale Modeling of Multi-Phase Flows in Heterogeneous Porous Media

Principal Investigator

Mark Knackstedt

*Applied Mathematics
RSPHysSE
ANU*

Co-Investigators

Christopher Arns

Stephen Hyde

Viet Nguyen

Mohammad Saadatfar

Arthur Sakalleriou

Adrian Sheppard

*Applied Mathematics
RSPHysSE
ANU*

Ji-Youn Lee

Walid Mahmud

Val Pinczewski

Rob Sok

*Petroleum Engineering
University of NSW*

Lincoln Paterson

*CSIRO Petroleum
Resources*

*CSIRO and Bureau of
Meteorology*

Brent Lindquist

SUNY

New York, USA

Project

w09, d59

Facilities Used

VPP, PC, SC

RFCD Codes

260204, 240205,
240503, 290703,
240202, 291406

The purpose of this work is to study diffusion in a disordered medium which is characterized by a distribution of local conductances that contains long-range correlations. Our main interest in this problem is to understand whether diffusion of pollutants in soil can be modelled by the classical diffusion equation with a constant diffusivity, or whether the presence of the correlations gives rise to a nonlocal transport process at the macroscopic level that cannot be represented by the diffusion equation. In the latter case, one must develop the appropriate transport equation.

Significant Achievements, Anticipated Outcomes and Future Work

The correlations are generated by a fractional Brownian motion characterized by a Hurst exponent H . For $H=1/2$ the correlations appear to have no effect, and the transport process is diffusive. However, for $H<1/2$ and depending on the morphology of the medium, three distinct types of transport processes, namely, anomalous, Fickian, and superdiffusive transport may emerge. Moreover, if the medium is anisotropic and stratified, biased diffusion in it is characterized by power-law growth of the mean square displacements with the time in which the effective exponents characterizing the power law oscillates log-periodically with the time. This result cannot be predicted by any of the currently-available continuum theories of transport in disordered media.

Computational Techniques Used

We used the power spectrum method to generate the FBM distribution. All of the results presented in this paper were obtained with $128 \times 128 \times 128$ lattices. This size of the lattice gives rise to a conductance distribution the broadness of which is about 2-3 orders of magnitude variations in the bonds' conductances. Periodic boundary conditions were used in all the directions. The diffusion process was simulated by the random walk of a particle which is initially (at time $t=0$) inserted into the lattice at a randomly selected site. The particle executes a random walk between the nearest neighbor sites of the lattice. Each step of the walk from one site to another is taken with a transition probability proportional to the conductance of the bond between the two sites. After each step, the time t is increased by one unit. The mean square displacements (MSD) $R^2(t)$ of the walkers at time t are computed, where the averaging is taken over the initial positions of the walkers, and the different realizations of the lattice. Typically, we used 4,000 walkers (i.e., 4,000 initial positions) and 40 realizations of each of the conductance distributions. All the random walkers took 2×10^6 steps, (i.e., the MSD were computed up to time $t=2 \times 10^6$).

Publications, Awards and External Funding

M. Saadatfar M. Sahimi , Diffusion in Disordered Media with Long-Range Correlations: Anomalous, Fickian and Superdiffusive Transport and Log-Periodic Oscillations, Phys. Rev. E 65, 2002, 0261XX-1

Investigation of Continuum Approaches to Modelling of Membrane Channels

Principal Investigator

Serdar Kuyucak
Theoretical Physics
RSPHysSE
ANU

Co-Investigators

Shin-Ho Chung
Ben Corry
Scott Edwards
Chemistry
Faculty of Science
ANU

Project

x15, x16

Facilities Used

VPP

RFCD Codes

249901

Understanding structure and function of ion channels in membranes is one of the major problems in contemporary biophysics. During the last two decades, advances in experimental techniques have enabled accurate measurement of small channel currents, which are of the order of picoAmpere. Most recently, the crystal structure of a potassium channel has been determined from x-ray analysis for the first time. It is expected that similar techniques will be employed to reveal the structure of other channels. Thus the time is ripe for sophisticated modeling of channels so as to provide rigorous answers to some basic questions about ion channels such as their level of conductivity and its saturation with increasing concentration, and their selectivity towards a particular ion species. The simplest permeation models one can use for this purpose are the continuum theories based on the Nernst-Planck electro-diffusion equation. The project aims to address the validity of continuum theories in narrow pores and whether they can be used in studies of ion channels.

Significant Achievements, Anticipated Outcomes and Future Work

So far, most of the theoretical work on modeling of ion channels is based on methods borrowed from physical chemistry. For studies of ion conductivity, these range from the phenomenological reaction rate theories to solutions of the Poisson-Nernst-Planck equations. Similarly, the Poisson-Boltzman equation is routinely used to determine the electrostatic potential in channel studies. The radius of ion channels varies from a few angstroms in the narrow part to 10-15 Å in the mouth region. In a similar vein, the number of ions inside the channel volume at a given time is at most a few. Intuitively, one would expect the continuum approaches to break down under these conditions, and one should ideally use microscopic approaches such as Brownian dynamics and molecular dynamics. While the use of continuum techniques could have been justified a decade ago (because there was no computationally feasible alternative), with the current speed of computers this is no longer a valid excuse.

This project provides a systematic comparison of the continuum and microscopic approaches to ion channels, establishing the region of validity for the former. Typical channel shapes are parametrised with a radius, and the predictions of the Poisson-Nernst-Planck and Brownian dynamics approaches for the conductivity are compared as a function of this radius. It is found that the continuum model results deviate significantly from those of the Brownian dynamics at small radii (2-5 Å), and an agreement is recovered only when the radius is greater than two Debye lengths (16 Å for a 150 mM solution). The detailed comparisons of the two

theories show that macroscopic concepts such as shielding in fact play a much smaller role in a realistic channel environment than assumed in continuum models.†

Computational Techniques Used

Two algorithms used in the project involve solutions of the Poisson and Langevin equations. The first is achieved using a boundary element method and the second involves numerical integration of a second order stochastic differential equation. Both programs are about 90% vectorized and have been successfully implemented on the VPP. We have already used these codes in several Brownian dynamics investigations of membrane channels, to which we refer for details on the computation of the electric fields and the implementation of the Brownian dynamics algorithm (see the contributions by S.H. Chung). Computation of the electrostatic forces in arbitrary channel geometries and running of the Brownian dynamics code are CPU intensive jobs and require the use of a supercomputer.

Publications, Awards and External Funding

B. Corry, S. Kuyucak, S. H. Chung, Test of Poisson-Nernst-Planck theory in ion channels, *J. Gen. Physiol.* 114, 1999, 597-599.

B. Corry, S. Kuyucak, S. H. Chung, Invalidity of continuum theories of electrolytes in nanopores, *Chem. Phys. Lett.* 320, 2000, 35-41.

G. Moy, B. Corry, S. Kuyucak, S. H. Chung, Tests of continuum theories as models of ion channels: I. Poisson-Boltzmann theory versus Brownian dynamics, *Biophys. J.* 78, 2000, 2349-2363.

B. Corry, S. Kuyucak, S. H. Chung, Tests of continuum theories as models of ion channels: II. Poisson-Nernst-Planck theory versus Brownian dynamics, *Biophys. J.* 78, 2000, 2364-2381.

LIGO Gravitational Wave Data Analysis

Principal Investigator

David McClelland

*Department of Physics
Faculty of Science
ANU*

Co-Investigators

**Philip Charlton
Benedict Cusack
John Ringland
Susan Scott
Antony Searle
Bram Slagmolen**

*Department of Physics
Faculty of Science
ANU*

Ben Evans

*Supercomputer Facility
ANU*

Project

x19

Facilities Used

SC, PC, WYRD, MDSS

RFCD Codes

240101, 240201

In this decade, a number of interferometric gravitational wave observatories will come online at various sites around the world. General relativity predicts that violent astrophysical events will emit gravitational radiation - "ripples in space-time" - that will produce tiny but measurable changes in the relative path length of two perpendicular optical cavities. The difficulty of confidently detecting these small changes amid relatively large instrumental noise makes the combination of data from multiple observatories around the globe desirable. The relative locations and orientations of these detectors has implications for the sensitivity of the network as a whole; we seek to understand this interaction, and in particular how an Australian detector would improve the network.

Significant Achievements, Anticipated Outcomes and Future Work

The optimal placement and orientation for a new gravitational wave detector has been determined under several criteria, including a detection rate of binary inspiral events for implementations of two leading network analysis techniques. It has been demonstrated that under these criteria the siting of an observatory in Western Australia (Figure 1) is desirable. These results have been presented to the Fourth Edoardo Amaldi Conference and the 2001 Gravitational Wave Data Analysis Workshop. We will extend our analysis to consider the angular resolution of networks and detection rates of other source types.

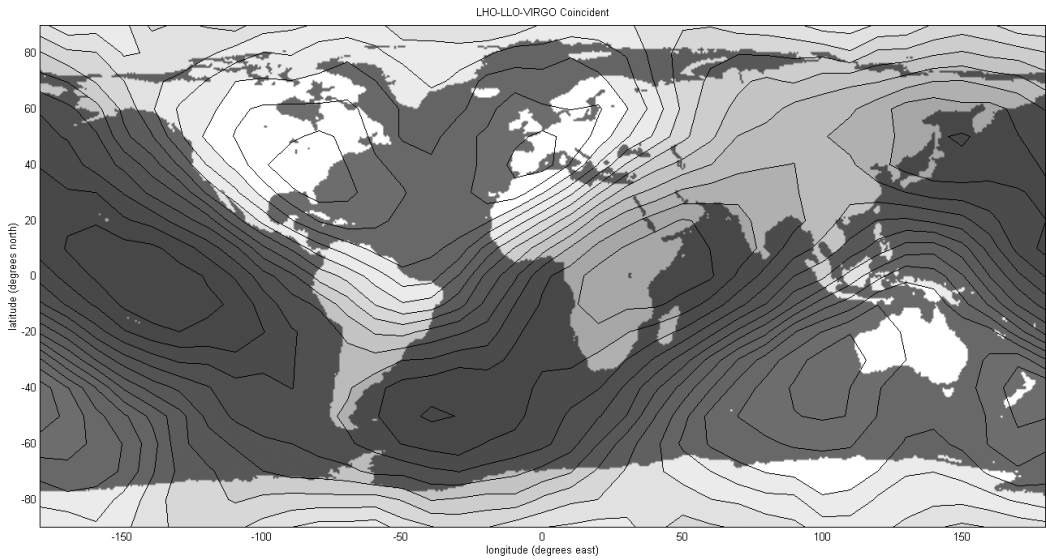


Figure 1. The relative merit of locations for a new observatory to augment the existing LIGO (North America) and VIRGO (Europe) observatories (lighter is better).

Computational Techniques Used

Monte-Carlo simulations of gravitational wave sources have been implemented in the MATLAB linear algebra package. The algorithm is amenable to parallelisation.

Publications, Awards and External Funding

A. C. Searle, S. M Scott, D. E. McClelland, Network Sensitivity to Geographical Configuration, *Classical and Quantum Gravity*, 2001 (accepted for publication).

Plume Zonation and Modelling with “Fluent”

Principal Investigator

Catherine A. Meriaux

Research School of
Earth Sciences
ANU

Project

x47

Facilities Used

SC

RFCD Codes

240502, 260200

Basalts from hotspot ocean islands (e.g. the Hawaiian and Galapagos Islands) are believed to be formed by melting of mantle plumes, which likely emanate from the core-mantle boundary layer. While it has long been recognized that lavas in these islands have different compositions, many geochemists believe that the variability within oceanic islands basalts reflects a source characteristic. The project is aimed at assessing whether a distribution of “heterogeneities” entrained in a plume conduit sheared by plate motion will maintain its identity. For transport in a vertical conduit, we expect simple stretching and thinning of heterogeneities. When horizontal shearing motion is superimposed on the buoyancy-driven vertical motion, the conduit is carried horizontally, becoming bent over. This leads to a complex recirculation in the conduit and entrainment of surrounding fluid. Thus “heterogeneities” at the source can be stirred, possibly to the extent that heterogeneities observed at the surface no longer reflect their distribution at the source. The amount of stirring depends on the buoyancy flux of the plume and the imposed shear, which together determine the angle of tilt of the conduit.

Significant Achievements, Anticipated Outcomes and Future Work

The initial testing phase of the project has been completed and confirmed that the entrainment of surrounding fluid by a thermal plume which is sheared by a background flow can be investigated using the Gambit and Fluent software. The mean flow structure and dynamics of the plume were successfully captured by our first three-dimensional model and entrainment of the surrounding fluid by the thermal plume was observed.

However, a detailed analysis of the results showed that the computation has to be optimized. First, the buoyancy flux of the plume was not constant although the computation had presumably converged to the steady-state solution. We think that the convergence criterion defined by default in Fluent may actually not be appropriate in our case. Thus, we are currently testing it. Alternatively, the problem could result from the use of a mesh too coarse for the problem.

The second issue which arose from these first calculations is that the horizontal non-uniform shear inherent to our geometry introduced a non-symmetric component in the characteristics of the vortices present in the plume conduit which is driven by the entrainment. This could not be a posteriori corrected in Fluent. Therefore, we are now inclined to change the geometry of our model such that, without altering the physics of the problem, it ensures that the shear motion is horizontally uniform.

As soon as the problem regarding the buoyancy flux has been identified and fixed, a new series of experiments will start in the new geometry. We expect that this new geometry will

allow for a better characterization of the entrainment into the plume conduit. The impact of stirring will be assessed by visualizing both the vorticity structure in the plume conduit and the pathlines of massless particles released in the problem domain. To complete our study, systematic calculations exploring a range of plume buoyancy fluxes and horizontal shearing motions will then be performed.

Computational Techniques Used

The calculations are performed by the computational fluid dynamics (CFD) packages Gambit and Fluent. The geometry and mesh of the model are initially built in Gambit while the flow and heat transfer are solved in Fluent. Fluent solvers utilize a finite-volume method based on fully unstructured meshes. Our application uses the segregated solver with a second-order implicit scheme (Gauss-Seidel-AMG) and the pressure-velocity coupling used is the PISO algorithm. The convection problem is solved within the Boussinesq approximation. Ultimately, we will take advantage of parallel processing which the Fluent solver allows for since our calculations are long and expensive.

Protein Refinement and Engineering

Principal Investigator

David Ollis

*Research School of
Chemistry
ANU*

Co-Investigator

Paul Carr

*Research School of
Chemistry
ANU*

Our research is focussed on the exciting interface between chemistry and biology. This area of research is crucial for understanding both the molecular basis of biological phenomena and for the utilisation of biological materials. Projects usually start with structure determination using diffraction techniques. Following this, structures are used to rationalise protein behaviour and to design experiments that further probe the intimate relationship between sequence, structure and function. Our current work has focussed on the following systems: 1) signal transduction: the GlnK and PII protein; 2) cell surface receptors: the β subunit of the IL5 receptor; and 3) replication protein: the ϵ subunit of DNA polymerase III.

Project

s06

Facilities Used

SC, PC

RFCD Codes

270108, 270199,
270805

Significant Achievements, Anticipated Outcomes and Future Work

The main experimental technique used by the group is X-ray diffraction. The object of our research is to firstly determine the structures of large bio-molecules and to use this information to understand how the molecules function.

The structure of the extra-cellular component of the β common receptor of the IL5 cytokine was published during 2001. IL5 plays a role in the generation of eosinophilia characteristic of helminth infections and allergic diseases such as asthma. The receptor is vital to understanding how IL5 stimulates cells. The arrangement of domains within the receptor is novel and allowed interaction sites with IL5 to be predicted. Experiments to verify these predictions are in progress.

The structure of the PII protein as a complex with ATP was also published during the year. Previous papers have described the protein structure while the latest manuscript describes the interaction between the protein and ATP, a co-factor that is necessary for biological activity.

Computational Techniques Used

Electron density maps or model construction are produced by Fourier techniques. Models are refined with non-linear least squares programs that minimise the difference between the observed and calculated structure amplitudes.

Publications, Awards and External Funding

Carr, P.D., Verger, D., Ashton, A.R. and Ollis, D.L. Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. *Structure with folding and design* (1999) **7** 461- 475.

Lappanen, V.L., Merckel, M., Ollis, D.L., Wong, K.K., Kozarach, J.W, and Goldman, A. Pyruvate formate lyase is structurally homologous, to type I ribonucleotide reductase. *Structure with folding and design* (1999) **7** 733-744.

Ashton, A.R., Trevanion, S.J. , Carr, P.D., Verger D. and Ollis, D.L. Structural basis for the light regulation of chloroplast NADP malate dehydrogenase *Physiologia. Plantarum*, (2000) **110**, 314-321.

Robinson, A., Edwards, K., Carr, P.D., Barton, J.D. Gary. D. Ewart and Ollis, D.L. Structure of the C123S mutant of diene lactone hydrolase (DLH) bound with the PMS moiety of the protease inhibitor phenylmethylsulfonyl fluoride (PMSF). *Acta Cryst D* (2000) **D56** 1376-1384.

Carr, P.D., Gustin, S.E., Church, A.P., Murphy, J.M., Ford, S.C., Mann, D.A., Woltring, D.M., Walker, I., Ollis, D.L. and Young, I.G. Structure of the complete extracellular domain of the common β subunit of the human GM-CSF, IL-3 and IL-5 receptors reveals a novel dimer configuration. *Cell* (2001) **104**, 291-300.

Xu, Y., Carr, P.D., Huber, T., Vasudevan, S.G. and Ollis, D.L. The structure of the PII - ATP complex. *European Journal of Biochemistry* (2001) **268**, 2028 -2037.

Hamdan, S., Carr, P.D., Brown, S.E., Ollis, D.L., Dixon, N.E. Structural basis for proofreading during replication of the *Escheria coli* chromosome. *Structure with folding and design* (in press).

Transport Properties of Systems with Long Range Interactions

Principal Investigator**Janka Petravc**

Physical and
Theoretical Chemistry
RSC
ANU

Project

x31

Facilities Used

SC

RFCD Codes

250602, 250699

The aim of this project is to develop algorithms for equilibrium and non-equilibrium molecular dynamics simulation of molecular systems with electrostatic interactions, in order to apply them to three different problems. The first is related to the intrinsic time-periodicity of phase variables occurring in the simulation of shear flow using the standard Sllod algorithm. Periodic time dependence of the response is of practical interest because it can cause wrong estimates of viscosity in sheared fluids, and of fundamental interest because it is not clear whether its origin is in the change of structure with lattice strain in equilibrium, or if it is purely dynamical. The second aims to relate the structural changes in associated liquids and electrolytes to anomalous behaviour of their transport properties. In particular, the aim is to find the structural reasons for the sign reversal in the Soret coefficient of salt solution in water with increase in salt concentration. The third concerns dynamical properties of a solution of ions in water confined between charged walls and forming an electrical double layer, a problem which has applications in biological systems.

Significant Achievements, Anticipated Outcomes and Future Work

The first year of this project has been devoted to development and testing of programs. The basic equilibrium programs for rigid and (semi-) flexible molecules with partial charges were tested on models for xylene isomers and ethanol. Xylene isomers are aromatics occurring in crude oil and the possibility of estimating their transport properties using molecular simulation is of interest to oil industry. It was shown that the available force field gives a sufficiently good description of intermolecular interactions to reproduce not only densities and viscosities for a wide range of temperatures and pressures, but also their peculiar counter-intuitive order. The equilibrium simulation of ethanol verified that the code was able to reproduce published simulation results for the dependence of diffusion coefficient on temperature, and was subsequently used to determine the influence of pressure on static structure and stability of hydrogen bonds and consequently on transport. The way in which different internal degrees of freedom (such as torsion and bending) influence diffusion was also explored.

The Sllod program for molecular systems with electrostatic interactions applied to liquid ethanol gave excellent agreement when extrapolated to equilibrium Green-Kubo results. It showed that in ethanol under strong shear, hydrogen bonded chains would become longer, more aligned with the x-axis and have less branching, and hydrogen bonds would become less stable. This results in faster decrease of viscosity with shear rate in ethanol than in say, butane.

Another test system was molten sodium chloride. After verifying that shear viscosity from Sllod calculation extrapolates to Green-Kubo zero-shear result, we determined maximum shear rate for which the standard kinetic temperature control is applicable. For higher shear rates, if linear streaming velocity profile is assumed with kinetic thermostat, one obtains wrong structural characteristics of the mixture. This can be avoided by controlling configurational temperature, which has particularly simple form for ionic systems. This system will be used to show that, when infinitely-ranged interactions are present, the amplitude of Sllod oscillations of shear viscosity is much larger than for Lennard-Jones systems and is present even for systems with very large number of particles.

Thermal conductivity of water using was calculated using a newly developed expression for heat flux in systems with electrostatic interactions, and was found to give good agreement with experimental data. This algorithm will be used in the future to find if the dependence of thermal conductivity on concentration of dissolved sodium chloride ions is well represented, and to calculate thermal diffusion coefficient of this mixture depending on concentration. It will also be applied to calculation of diffusion coefficients in the electrical double layer problem.

Computational Techniques Used

In all cases, molecular dynamics programs with Ewald summation were used, with special adaptations made for the case of Sllod algorithm and its peculiar periodic boundary conditions.

Publications, Awards and External Funding

There are 3 publications submitted, but they have not yet been accepted.

Structural and Mechanistic Chemistry

Principal Investigator

Leo Radom

*Computational
Quantum Chemistry
RSC
ANU*

Co-Investigators

Justine Bennett

Thomas Buesgen

Michelle Coote

Michael Hartmann

David Henry

Rebecca Jacob

Christopher Parkinson

Danne Rasmussen

Robyn Schofield

Anthony Scott

Stefan Senger

David Smith

Michael Sullivan

Stacey Wetmore

Geff Wood

*Computational
Quantum Chemistry
RSC
ANU*

Brian Smith

*Structural Biology
Walter & Eliza Hall
Institute*

Mark Freitag

Mark Gordon

*Chemistry
Iowa State University
USA*

Theis Solling

*Chemistry
Caltech
USA*

Project

r54, k29, q07, q08, v56,
v55, v01, s08, p03, d39

*Chemistry is traditionally an experimental science. However, recent advances in computer technology and the development of highly efficient computer algorithms have opened the way for a viable alternative approach to chemistry: chemistry by computer. We use such computer calculations to determine the structures of molecules and to help understand how molecules react with one another. The procedure employed is called *ab initio* molecular orbital theory, the term *ab initio* signifying that the calculations are carried out from first principles using the laws of quantum mechanics. No experimental data other than values of fundamental physical constants are used. An important feature is that the calculations can be carried out as readily for reactive or hazardous species as for normal, stable molecules. They are therefore particularly useful in cases where experimental studies might be difficult or impossible. The calculations are highly computationally intensive and the APAC National Facility has been and will continue to be extremely valuable to us. One of the major areas of current activity is to try to understand how enzymes make certain reactions go faster. Another focus is to try to better understand the chemistry of free radicals since these are of widespread importance in chemistry, biology and polymers. In addition, we are developing and assessing better procedures for obtaining accurate chemical information from such quantum chemistry computations.*

Facilities Used

PC, VPP, SC, MDSS

RFCD Codes

250699

Significant Achievements, Anticipated Outcomes and Future Work

Reactions Mediated by Coenzyme B12

Although these reactions have been extensively studied experimentally, there is certainly no consensus as to how they proceed. Our *ab initio* calculations show that the reactions are facilitated by partial-proton-transfer. For the reaction catalyzed by the enzyme glutamate mutase, we predicted that a specific amino acid moiety within the enzyme (Glu171) is likely to play a major role in this regard. If Glu171 is removed, the enzyme should be much less effective. Precisely such experiments have very recently been carried out in the USA and have confirmed this prediction. The reaction is slowed down by a factor of 50 when Glu171 is “mutated out”. Such examples provide strong encouragement for the use of computer calculations in a predictive manner in the study of enzyme reactions.

Free Radicals

Radicals are ubiquitous in chemistry, biology and polymer science. Because they are reactive species, they are often difficult to study experimentally and therefore theory has a potentially useful role to play in their characterization. It turns out that the theoretical study of radicals is also not entirely straightforward. Therefore it is important to carefully assess the various theoretical methods that are to be used and make improvements where necessary before embarking on extensive application. This is a key part of our research. We have been using theory to determine radical stabilization energies, with the important aim of seeing how individual substituents stabilize or destabilize a radical centre. We have also been examining the details of radical addition and abstraction reactions, both of which are very important in biological chemistry and polymer chemistry.

Development of Improved Theoretical Procedures

The ability to predict reliable thermochemistry represents a very important application of *ab initio* molecular orbital theory. Much work has been done with this aim in mind. Our most recent contributions have involved developing analogues of the successful G3 technique that are capable of adequately describing systems with biradical character. The most successful of these to date is called MRG3(MP2).

Hydrogen Bonding

Hydrogen bonding is of great importance in chemical and biological systems. Previous studies have mostly focussed on hydrogen bonds involving electronegative elements, e.g., O–H•••N. We have examined instead the weaker hydrogen bonds to carbon, e.g., C–H•••N. We have carried out systematic studies aimed at identifying which types of systems will exhibit the strongest C–H•••X hydrogen bonds. The effect of electronegative substitution has been particularly targeted.

Computational Techniques Used

We use mainstream computational quantum chemistry programs for our work. These have generally been refined by ANUSF programmers and are also tested within my research group. The main programs that we use are GAUSSIAN 98, MOLPRO 2000, ACES2 and QCHEM. The greater computing power of the new APAC facility has made possible calculations on larger molecules and using more reliable theoretical procedures.

Publications, Awards and External Funding

C. J. Parkinson, P. M. Mayer and L. Radom, Cyanovinylyl Radical: An Illustration of Poor Performance of Unrestricted Perturbation Theory and Density Functional Theory Procedures in Calculating Radical Stabilization Energies, *Theoretical Chemistry Accounts*, 102, 1999, 92-96.

T. I. Sølling and L. Radom, Exchange Reactions of Chloriranium Ions and Chlorirenium Ions: A G2 Investigation, *International Journal of Mass Spectrometry*, 185/186/187, 1999, 263-270.

M. L. Coote, T. P. Davis and L. Radom, The Effect of Remote Substituents in Free Radical Addition Reactions: New Evidence for the Penultimate Unit Effect, *Journal of Molecular Structure, Theochem*, 461-462, 1999, 91-96.

D. M. Smith, B. T. Golding and L. Radom, On the Mechanism of Action of Vitamin B12: Theoretical Studies of the 2-Methyleneglutarate-Mutase-Catalyzed Rearrangement, *Journal of the American Chemical Society*, 121, 1999, 1037-1044.

T. I. Sølling, S. B. Wild and L. Radom, Are Pi-Ligand Exchange Reactions of Thiirenium and Thiranium Ions Feasible? An Ab Initio Investigation, *Chemistry - A European Journal*, 5, 1999, 509-514.

D. M. Smith, B. T. Golding and L. Radom, Facilitation of Enzyme-Catalyzed Reactions by Partial Proton Transfer: Application to Coenzyme B12-Dependent Methylmalonyl-CoA Mutase, *Journal of the American Chemical Society*, 121, 1999, 1383-1384.

A. J. Chalk and L. Radom, Ion-Transport Catalysis: Catalyzed Isomerizations of NNH⁺ and NNCH₃⁺, *Journal of the American Chemical Society*, 121, 1999, 1574-1581.

T. I. Sølling, S. B. Wild and L. Radom, Exchange and Insertion Reactions Involving Borane Adducts of Phosphirane and Phosphirene: A G2(MP2) Ab Initio Investigation, *Journal of Organometallic Chemistry*, 580, 1999, 320-327.

M. L. Coote, T. P. Davis and L. Radom, Effect of the Penultimate Unit on Radical Stability and Reactivity in Free-Radical Polymerization, *Macromolecules*, 32, 1999, 2935-2940.

M. L. Coote, T. P. Davis and L. Radom, The Conformational Dependence of the Penultimate Unit Effect in Free-Radical Copolymerization, *Macromolecules*, 32, 1999, 5270-5276.

D. M. Smith, B. T. Golding and L. Radom, Towards a Consistent Mechanism for Diol-Dehydratase-Catalyzed Reactions: An Application of the Partial-Proton-Transfer Concept, *Journal of the American Chemical Society*, 121, 1999, 5700-5704.

- D. R. Rasmussen and L. Radom, Planar Tetracoordinate Carbon in a Neutral Saturated Hydrocarbon: Theoretical Design and Characterization, *Angewandte Chemie, International Edition in English*, 38, 1999, 2876-2878.
- S. Petrie and L. Radom, Magnesium- and Calcium- Containing Molecular Dications: A High Level Theoretical Study, *International Journal of Mass Spectrometry*, 192, 1999, 173-183.
- A. Schulz, B. J. Smith and L. Radom, Heats of Formation of Alkali and Alkaline Earth Oxides and Hydroxides: Some Dramatic Failures of the G2 Method, *Journal of Physical Chemistry A*, 103, 1999, 7522-7527.
- D. M. Smith, B. T. Golding and L. Radom, Understanding the Mechanism of B12-Dependent Methylmalonyl-CoA Mutase: Partial Proton-Transfer in Action, *Journal of the American Chemical Society*, 121, 1999, 9388-9399.
- C. J. Parkinson, P. M. Mayer and L. Radom, An Assessment of Theoretical Procedures for the Calculation of Reliable Radical Stabilization Energies, *Journal of the Chemical Society, Perkin Transactions. 2*, 1999, 2305-2314.
- T. I. Sølling, S. B. Wild and L. Radom, A G2 Ab Initio Investigation of Ligand-Exchange Reactions Involving Mono- and Bis⁺ Adducts of the Phosphenium Ion, *Inorganic Chemistry*, 38, 1999, 6049-6054.
- A. J. Chalk, P. M. Mayer and L. Radom, Rearrangement and Fragmentation Pathways of [C3H7Z]⁺ Ions (Z = NH and S): Are Ion-Neutral Complexes Important? *International Journal of Mass Spectrometry*, 194, 2000, 181-186.
- M. L. Coote, T. P. Davis and L. Radom, Models for Free-Radical Copolymerization Propagation Kinetics, In *Controlled/Living Radical Polymerization*, ACS Symposium Series 768, K. Matyjaszewski, Editor, pp 82-92, American Chemical Society, Washington, DC, 2000.
- M. Hartmann and L. Radom, The Acetylene-Ammonia Dimer as a Prototypical C⁻HooN Hydrogen-Bonded System: An Assessment of Theoretical Procedures, *Journal of Physical Chemistry A*, 104, 2000, 968-973.
- S. Senger and L. Radom, Zeolites as Transition-Metal-Free Hydrogenation Catalysts: A Theoretical Mechanistic Study, *Journal of the American Chemical Society*, 122, 2000, 2613-2620.
- A. J. Chalk and L. Radom, The Involvement of Ion-Neutral Complexes in Ethylene Loss From [PhC(CH₃)₂]⁺ and its Isomers, *International Journal of Mass Spectrometry*, 199, 2000, 29-40.
- T. I. Sølling, S. B. Wild and L. Radom, Formation of Three-Membered Phosphorus Heterocycles via Ligand-Exchange Reactions in Mono-Adducts of the Phosphenium Ion: An Ab Initio Investigation, *International Journal of Mass Spectrometry*, 201, 2000, 205-213.
- A.P. Scott and L. Radom, Are Cumulenones Kinked? A Systematic High-Level Ab Initio Study of H₂CCCO, H₂CCCCO and H₂CCCCCO, *Journal of Molecular structure*, 556, 2000, 253-261.

M. A. Collins, S. Petrie, A. J. Chalk and L. Radom, Proton-Transport Catalysis and Proton-Abstraction Reactions: An Ab Initio Dynamical Study of $X + \text{HOC}^+$ and $\text{XH}^+ + \text{CO}$ ($X = \text{Ne}, \text{Ar}$ and Kr), *Journal of Chemical Physics*, 112, 2000, 6625-6634.

D. R. Rasmussen and L. Radom, Hemispiroalkaplanes: Hydrocarbon Cage Systems with a Pyramidal-Tetracoordinate Carbon Atom and Remarkable Basicity, *Chemistry ^ A European Journal*, 6, 2000, 2470-2483.

T. I. Sølling, S. B. Wild and L. Radom, Are the Approach Directions of s and p Nucleophiles to the Sulfur Atom of Thiiranium and Thiirenium Ions Different? *Chemistry ^ A European Journal*, 6, 2000, 590-591.

S. Senger and L. Radom, Towards a Low-Barrier Transition-Metal-Free Catalysis of Hydrogenation Reactions: A Theoretical Mechanistic Study of HAIX_4 -Catalyzed Hydrogenations of Ethene ($X = \text{F}, \text{Cl}$ and Br), *Journal of Physical Chemistry A*, 104, 2000, 7375-7385.

T. I. Sølling and L. Radom, Exchange of Cl^+ Between Lone-Pair Donors and p-Donors: A Computational Study, *European Mass Spectrometry*, 6, 2000, 153-160.

S. D. Wetmore, D. M. Smith and L. Radom, How B6 Helps B12: The Roles of B6, B12 and the Enzymes in Aminomutase-Catalyzed Reactions, *Journal of the American Chemical Society*, 122, 2000, 10208-10209.

quD. M. Smith, B. T. Golding and L. Radom, , Understanding the Mechanism of B12-Dependent Diol Dehydratase: A Synergistic Push-Pull Proposal, *Journal of the American Chemical Society*, 123, 2001, 1664-1675.

M. Hartmann, S.D. Wetmore and L. Radom, $\text{C}^{\text{H}}\text{HooX}$ Hydrogen Bonds of Acetylene, Ethylene and Ethane with First- and Second-Row Hydrides, *Journal of Physical Chemistry A*, 105, 2001, 4470-4479.

D. M. Smith, S. D. Wetmore and L. Radom, Theoretical Studies of Coenzyme B12-Dependent Carbon-Skeleton Rearrangements, In *Theoretical Biochemistry - Processes and Properties of Biological Systems*, L. Ericsson Editor, pp. 183-214, Elsevier, Amsterdam, 2001.

T. I. Sølling and L. Radom, A G2 Study of SH^+ Exchange Reactions Involving Lone-Pair Donors and Unsaturated Hydrocarbons, *Chemistry ^ A European Journal*, 7, 2001, 1516-1524.

H. Fischer and L. Radom, Factors Controlling the Addition of Carbon-Centered Radicals to Alkenes ^ An Experimental and Theoretical Perspective, *Angewandte Chemie, International Edition in English*, 40, 2001, 1340-1371.

S. D. Wetmore, D. M. Smith, B. T. Golding and L. Radom, , Interconversion of (S)-Glutamate and (2S,3S)-3-Methylaspartate: A Distinctive B12-Dependent Carbon-Skeleton Rearrangement, *Journal of the American Chemical Society*, 123, 2001, 7963-7972.

T.I. Sølling, A. Pross and L. Radom, A High-Level Ab Initio Investigation of Identity and Non-Identity Gas-Phase SN_2 Reactions of Halide Ions with Halophosphines, *International Journal of Mass Spectrometry*, 210/211, 2001, 1-11.

D. J. Henry, C. J. Parkinson, P. M. Mayer and L. Radom, Bond Dissociation Energies and Radical Stabilization Energies Associated With Substituted Methyl Radicals, *Journal of Physical Chemistry A*, 105, 2001, 6750-6756.

S. D. Wetmore, D. M. Smith and L. Radom, The Enzyme Catalysis of 1,2-Amino Shifts: The Cooperative Action of B6, B12 and Aminomutases, *Journal of the American Chemical Society*, 123, 2001, 8678-8689.

A. P. Scott, M. S. Platz and L. Radom, Singlet-Triplet Splittings and Barriers to Wolff Rearrangement for Carbonyl Carbenes, *Journal of the American Chemical Society*, 123, 2001, 6069-6076.

S. D. Wetmore, D. M. Smith, R. Schofield and L. Radom, A Theoretical Investigation of the Effects of Electronegative Substitution on the Strength of C[∞]H[∞]∞N Hydrogen Bonds, *Journal of Physical Chemistry A*, 105, 2001, 8718-8726.

D. J. Henry and L. Radom, Theoretical Thermochemistry of Radicals, In *Theoretical Thermochemistry*, J. Cioslowski, Editor, pp 161-197, Kluwer, Dordrecht, 2001.

T. I. Sølling, D. M. Smith, L. Radom, M. A. Freitag and M. S. Gordon, Multi-Reference Equivalents of the G2 and G3 Methods, *Journal of Chemical Physics*, 115, 2001, 8758-8772.

S. D. Wetmore, D. M. Smith and L. Radom, Catalysis by a Mutant of Methylmalonyl-CoA Mutase: A Theoretical Rationalization for a Change in the Rate-Determining Step, *ChemBioChem*, 2, 2001, 919-922.

Development of Computational Chemistry Methods for Parallel Processors

Principal Investigator

Alistair Rendell
Computer Science
Faculty of Science
ANU

Co-Investigator

Stephen Titmuss
Computational
Molecular Biology
JCSMR
ANU

Computational chemistry aims to model the behaviour of atoms and molecules using computers. There are many parts to computational chemistry, but our specific interest is in electronic structure theory. These methods seek to solve, with minimal approximation, the fundamental equations of quantum mechanics for a system of electrons and nuclei. The methods used are complex having been developed over many decades. In part because of this historical legacy many of the computer codes that are currently in use are poorly adapted to massively parallel processors. This project aims to address this issue.

Project

x32

RFCD Codes

250601

Significant Achievements, Anticipated Outcomes and Future Work

Our initial work has been exploratory, assessing the parallel capability and performance of some current electronic structure codes. To this end we have considered the GAMESS (US Version), Gaussian 98, and NWChem codes. Ultimately our aim is to perform routinely simple (i.e. Hartree-Fock and density functional) energy and gradient calculations on systems containing several hundred atoms, while making more complicated perturbation theory and coupled cluster calculations routine on systems with fifty or more atoms. To this end our future work will be focused on the Gaussian code, working closely with the developers to improve the parallel performance of the various linear scaling algorithms.

Computational Techniques Used

Essentially, electronic structure codes use spectral methods to solve a system of differential equations. The behaviour of each electron is described by a series of basis functions that are centered on each nuclear center. The formation and manipulation of integrals involving these functions and a variety of operators form the core of all electronic structure methods. The specific codes used here are GAMESS (US Version), Gaussian 98 and NWChem. All three codes use a multiple instruction multiple data (MIMD) approach to parallelisation, but differ in the method of parallelisation (e.g. message passing, Linda, and Global Arrays) and degree to which different quantities have been distributed across the memory associated with each process. The APAC National Facility SC provides us with access to significantly more processors than would be available on any other resource.

Publications, Awards and External Funding

This work has been supported, in part, by funds from the APAC / ANU computational chemistry expertise program.

Atom Lasers and Bose-Einstein Condensates

Principal Investigator

Craig Savage

*Physics and Theoretical
Physics
Faculty of Science
ANU*

This project investigates the quantum dynamics of atom lasers and dilute gas Bose-Einstein Condensates (BECs). These are highly quantum mechanical systems in which atoms behave like waves. The 2001 Nobel Prize in Physics was awarded for pioneering experimental work in this area.

Co-Investigators

Simon Haine

Joseph Hope

Pearl Louis

Adele Morrison

Nicholas Robins

Anthony Searle

*Physics and Theoretical
Physics
Faculty of Science
ANU*

Glenn Moy

*Chemistry
Faculty of Science
ANU*

Dan Gordon

*Engineering
FEIT
ANU*

Project

r62, x27, d66

Facilities Used

PC, SC, VPP, MDSS

RFCD Codes

240201, 240204

Significant Achievements, Anticipated Outcomes and Future Work

Our first publication for 2001 was a clarification of the limitations on producing macroscopic quantum superpositions (Schrodinger Cats) in BECs. Our next publication concerned the quantum statistics of coupled atomic and molecular condensates in three dimensions.

A paper has been submitted describing a previously unknown type of instability in atom lasers. We have developed new numerical techniques for computational modelling of the full

quantum field theory of atom lasers. Modelling of the ANU BEC and atom laser experiments is ongoing.

A major new direction is the modelling of the “Bosenova” BEC explosions observed in experiments this year. Once we have developed an adequate model of the phenomenon we anticipate using it, in conjunction with experiments, to explore important new atomic physics.

Computational Techniques Used

We solve nonlinear Schrodinger type partial differential equations in up to three dimensions. We use a pseudo-spectral method which takes advantage of the efficiently parallelised FFT algorithm. Our variant, developed at the University of Otago, is particularly effective, as it uses a fourth order Runge-Kutta time step.

Our quantum field work adds stochastic terms, representing quantum noise, to the nonlinear partial differential equations. Useful results are obtained by averaging over ensembles of solutions.

The APAC National Facility has made it possible for us to model actual experimental conditions. This enables us to make definitive statements about the physics involved in particular cases. Without the National Facility only very limited ensembles could be used in our quantum field calculations, limiting their relevance.

Publications, Awards and External Funding

C.M. Savage, ARC Large Grant, The quantum theory of atom lasers, 2000-2002.

P. J. Y. Louis, P. M. R. Brydon, C. M. Savage, Macroscopic quantum superposition states in Bose-Einstein condensates: Decoherence and many modes, *Physical Review A*, 64, 2001, 053613.

J.J. Hope, Quantum effects in coupled atomic and molecular Bose-Einstein condensates, *Physical Review A*, 64, 2001, 053608.

Filamentary Structure in the Lobes of Radio Galaxies

Principal Investigator

Curtis J. Saxton

Mt Stromlo Observatory
RSAA and
Dept. of Physics and
Theoretical Physics
Faculty of Science
ANU

Co-Investigators

Geoffrey V. Bicknell

Mt Stromlo Observatory
RSAA and
Dept. of Physics and
Theoretical Physics
Faculty of Science
ANU

Ralph S. Sutherland

Mt Stromlo Observatory
RSAA
ANU

Project

x34

Facilities Used

SC

RFCD Codes

240101, 240502

Significant Achievements, Anticipated Outcomes and Future Work

We have investigated the structure and behaviour of the termination shock and backflow cocoon in extragalactic jets ranging over two orders of magnitude in density. We find that the distribution of radio emitting plasma in the cocoon depends substantially on the jet density. Moreover, the density dependent turbulence in the jet backflow also alters the surging action of the jet through compression and decompression. This affects the structure of luminous shocks. We have also found that boundary conditions of the system have a significant effect on the width and mixing of the jet cocoon.

We have reproduced images of radio surface brightness similar to the hot-spot and bar-like filament revealed in radio and optical observations of the western radio lobe of Pictor†A. These structures are well explained in terms of the surging termination shock of the jet, surrounded by bright transitory ring-like shocks in the jet's backflow. The transient simulated

structures match observation in episodes amounting to approximately 4% of the time, implying that such morphologies would be uncommon in radio galaxies, yet could be observed in a small fraction. Our studies suggest that the jet of Pictor A is almost in the plane of the sky; the observed morphology is not reproduced by a jet that is rendered at inclinations of 45° or less. Furthermore, we predict that these bright structures have dynamical lifetimes that are shorter than the light travel time across the source. The shock luminosity can vary in time by over an order of magnitude even though the power and thrust of the jet source is constant. This provides an important qualification to comparisons between the brightnesses of the independent eastern and western hot-spots in the radio source.

The ring-like structures observed in the radio lobes of Hercules A are well explained as shocks propagating in the backflow surrounding the jet. In this case, we infer that the jet is oriented $\sim 45^\circ$ to the line of sight. The lack of bright hot-spots in these radio lobes can be understood if the head of the Hercules A jet is presently in one of the low-luminosity phases seen in our simulation, during the reformation or surging of the termination shock.

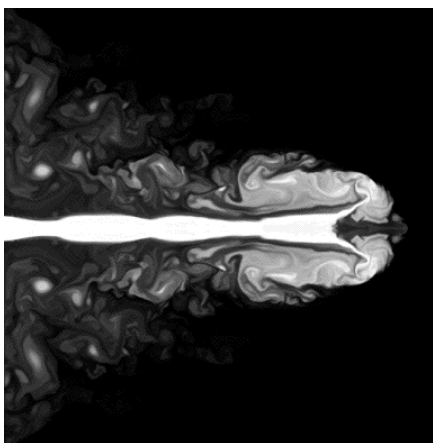


Figure 1.

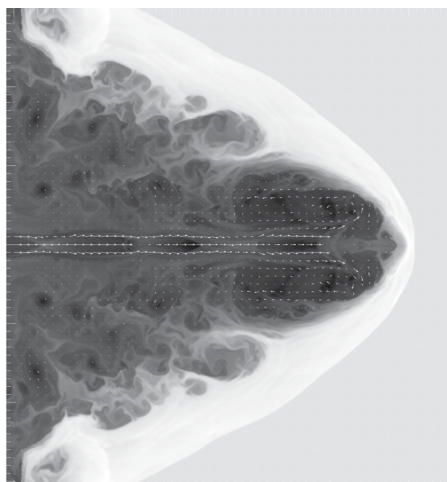


Figure 2.

Figure 1. Cross-sectional map of the concentration of radio plasma, derived from the jet, in a snapshot from a simulation with $(\eta, M) = (10^{-4}, 10)$. The jet is the white horizontal feature. It propagates to the right through the denser background medium.

Figure 2. Log density map with velocity vectors superimposed, corresponding to the previous image. Within the bow shock, a turbulent cocoon surrounds the jet. The compression and decompression of the jet by the turbulence in the cocoon is evident through the velocity vectors.

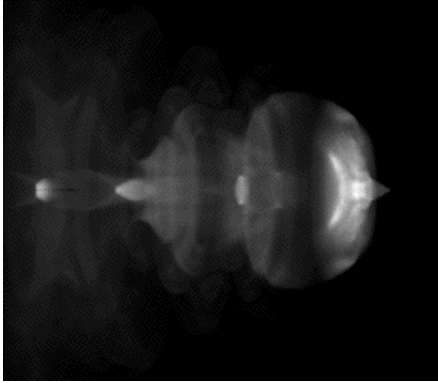


Figure 3.

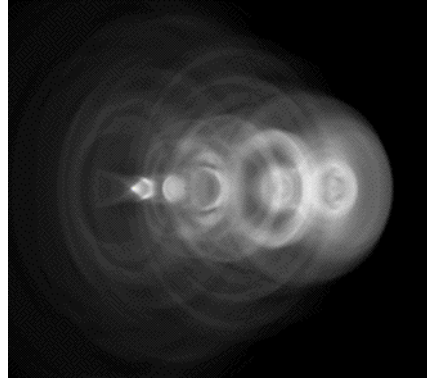


Figure 4.

Figure 3. Raytraced image of a snapshot from a jet with $(\eta, M) = (10^{-4}, 5)$ viewed at an inclination of 80° .

Figure 4. Raytraced image of a snapshot from a jet with $(\eta, M) = (10^{-4}, 5)$ viewed at an inclination of 45° .

Computational Techniques Used

Our hydrodynamic simulations were conducted using the VH-1 code (Blondin and Lukfin 1993) which is an implementation of the Piecewise Parabolic Method (PPM) (Colella and Woodward 1984). An advantage of PPM for this type of simulation is its excellent resolution of shocks. We have enhanced the code to achieve greater efficiency, and have added a scalar tracer to distinguish and follow the evolution of various constituents of the physical system.

Our images of simulated radio surface brightness were rendered by a special-purpose ray-tracing program that projects three-dimensional structures obtained from the PPM output. At a given orientation, the volumetric emissivity (a function of the PPM output variables) is integrated along the line of sight for each pixel of the sky/screen image.

Publications, Awards and External Funding

This project is connected with the ARC large grant A69905341: “Jet-cloud interactions in active galactic nuclei.”

C.J. Saxton, R.S. Sutherland, G.V. Bicknell, Evidence of time-dependent ultra-light jet behaviour in Pictor A, *Astronomy & Astrophysics*, 2002, submitted.

C.J. Saxton, R.S. Sutherland, G.V. Bicknell, Production of ring-like structure in the cocoon of Hercules A, *Astrophysical Journal*, 2002, submitted.

Efficient Calculation of Statistical and Dynamical Reaction Rates for Large Dimensional Molecular Systems

Principal Investigator

Harold Schranz

*Physical and
Theoretical Chemistry
RSC
ANU*

Knowledge of how quickly chemical reactions occur is an essential ingredient in the rigorous modelling of combustion, industrial and atmospheric reaction systems. This project focuses on the development of new methods for the accurate prediction of rate constants for chemical reaction. Crucial to this development is a more complete understanding of how and on what timescale energy moves about a molecule (IVR) and between molecules (collisional energy transfer).

Co-Investigator

Terry Frankcombe

*Chemistry
Molecular & Microbial
Sciences
University of
Queensland*

The current focus is in two areas. Firstly, on the role of the intermolecular potential energy and quantum effects on collisional energy transfer in highly excited molecules. Secondly, on the modeling of complex chemical reactions in the gas phase, which involves describing the molecular species and their transformation by reaction at a detailed level.

Project

s10

Facilities Used

VPP, PC, MDSS

RFCD Codes

250101, 250601,
250602, 250603

Significant Achievements, Anticipated Outcomes and Future Work

A. Effect of Potential Energy on Collisional Energy Transfer: A quantitative measure of the average steepness of an intermolecular potential energy surface between a polyatomic molecule and a monotonic gas atom is proposed. Classical trajectory studies of collisions between rovibrationally excited carbon dioxide molecules and argon atoms selected from a thermal bath show that there is a strong correlation between this average steepness and the first moment of intermolecular energy transfer over a wide range of intermolecular potentials describing this CO₂-Ar interaction. The strong correlation, as illustrated, has the potential to allow insight into the effect that more subtle features of the intermolecular potential energy surface have on collisional energy transfer.

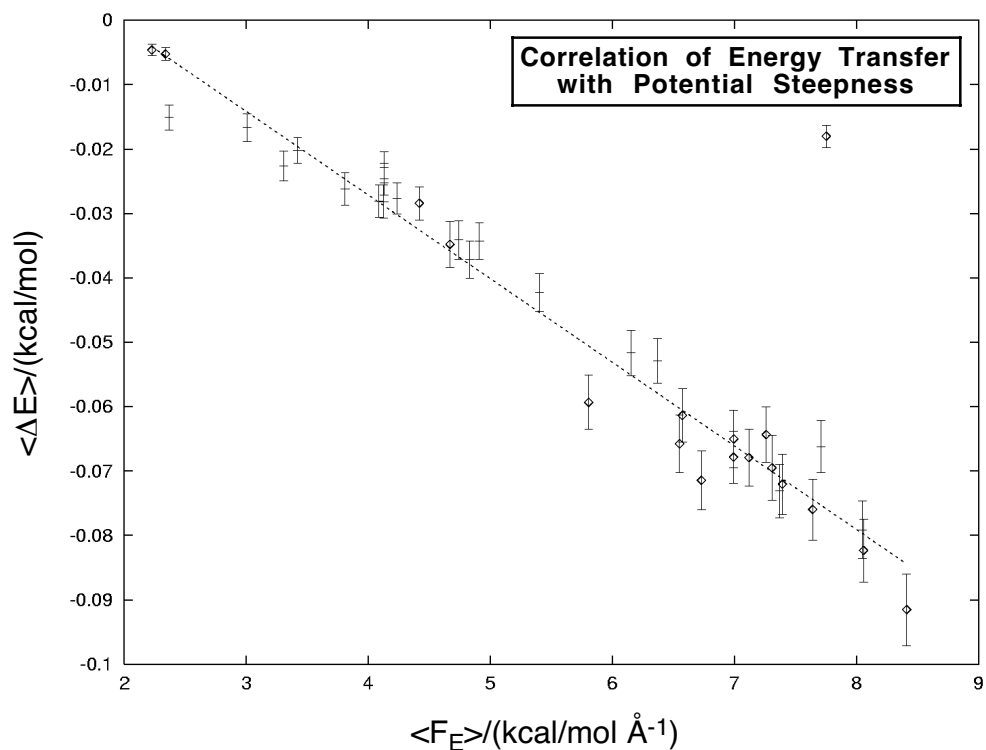


Fig. 1. Correlation of average energy transfer per collision with a intermolecular potential energy steepness parameter for highly excited CO_2 with internal energy $E=75$ kcal/mol colliding with Ar at $T=300\text{K}$.

B. Prediction of Absolute Rate Coefficients and Product Branching Ratios: Complex chemical reactions in the gas phase can be decomposed into a network of elementary (e.g. unimolecular and bimolecular) steps which may involve multiple reactant channels, multiple intermediates, and multiple products. Here we focus on a detailed modeling of the $\text{C}(^3\text{P})+\text{Allene}$ (C_3H_4) reaction, for which molecular beam experiments and theoretical calculations have previously been performed.

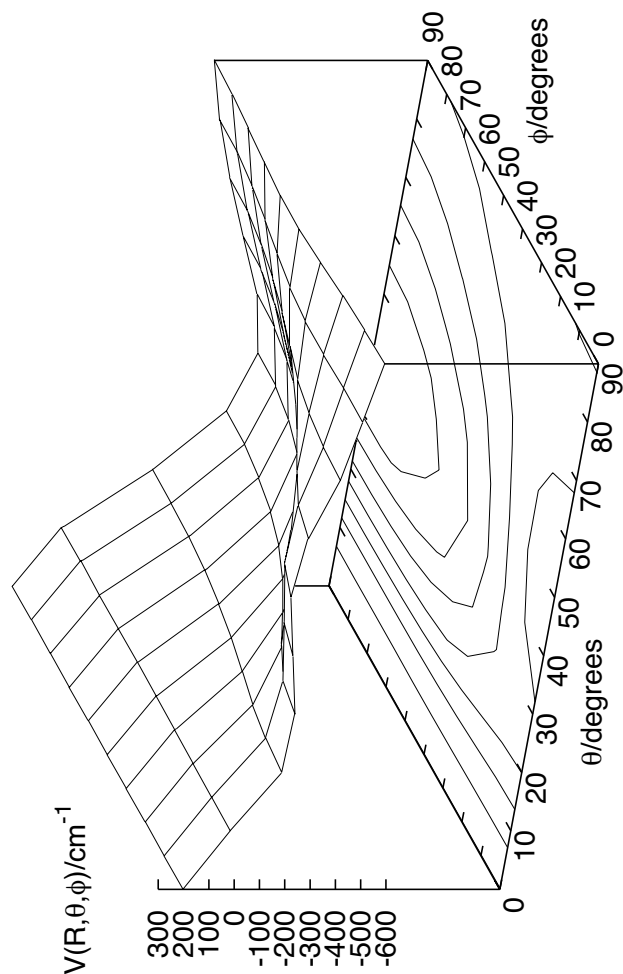


Fig. 2. The potential energy surface for $C(^3P)+\text{Allene } (C_3H_4)$ at a separation of 4 \AA based on ab initio calculations generated with Gaussian 98.

We predict absolute unimolecular rate coefficients and branching ratios using microcanonical variational transition state theory (μ -VTST) with full energy and angular momentum resolution. Our calculation of the initial capture rate is facilitated by systematic ab initio potential energy surface calculations that describe the interaction potential between carbon and allene as a function of the angle of attack. Furthermore, the chemical kinetic scheme is enhanced to explicitly treat the entrance channels in terms of a predicted overall input flux and also to allow for the possibility of redissociation via the entrance channels. Thus, the computation of overall capture rates and partial capture rates is now possible.

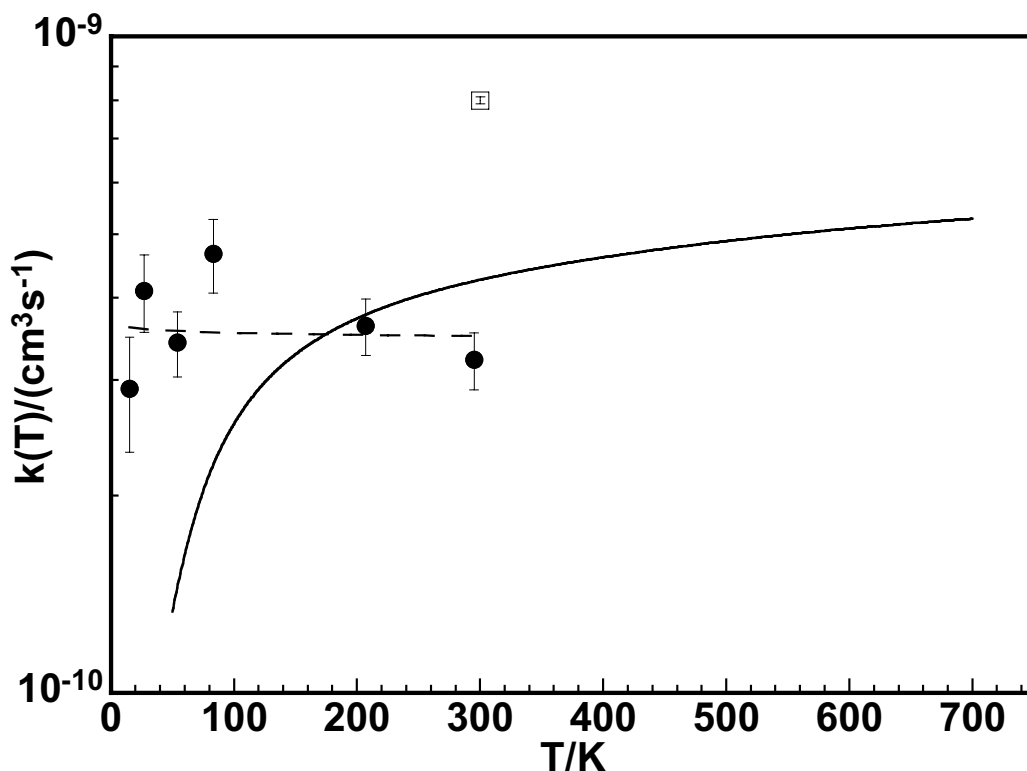


Fig. 3. Temperature dependence of the capture rate constant $k(T)$ predicted by this work (solid curve) compared to experimental results (points) for $C(^3P) + \text{allene}$.

Computational Techniques Used

Quantum chemical ab initio packages (GAUSSIAN) and density functional theory packages (ADF) were used for calculating points for the intermolecular and intramolecular potential energy of the molecular systems under study.

For the collisional energy transfer studies, a model of the colliding system (e.g. an excited target molecule and a thermal projectile molecule) was described by a global potential energy surface including intra- and intermolecular parts, and the classical equations of motion were solved for a specified ensemble of initial conditions before the collision, thereby ending up with a set of final states after the collision is over. Home grown efficient vectorised trajectory codes were employed for this portion of the project and production runs were performed using the VPP and SGI Power Challenge supercomputers. A variety of detailed and averaged quantities were extracted from such studies.

For the calculation of overall capture rates and partial capture rates for complex chemical reactions, a statistical treatment was employed which involved a combination of Monte Carlo sampling over configuration space and convolution techniques.

Kernel Methods for Bioinformatics

Principal Investigator

Alexander Smola

Computer Sciences
Laboratory
RSISE
ANU

Co-Investigators

James Macnicol

Gunnar Raetsch

Computer Sciences
Laboratory
RSISE
ANU

Soeren Sonnenburg

FIRST
Fraunhofer Institute
Germany

*Splice sites are locations in DNA at the boundaries of exons (which code for proteins) and introns (which do not). The more accurately splice sites can be located, the easier it becomes to identify the coding regions in DNA. Accurate splice site sensors thus form critical components of computational gene finders. Since ever-larger chunks of DNA are to be analyzed by gene finders the problem of accurate splice site recognition has never been more important. We pose splice site recognition as a classification problem with the classifier learned from a labeled data set consisting of only local information around the potential splice site. Finding the correct position of splice sites without using global information is assumed to be a difficult task. We apply Support Vector Machines - a state-of-the-art previously untried method - to the splice site recognition problem. We analyze the genomes of *C. elegans* and of humans using specifically designed support vector kernels. One of the kernels is adapted from our work on detecting translation initiation sites in vertebrates and another uses a Fisher-kernel like idea. We find that both approaches perform very well on this task.*

Project

x46

Facilities Used

SC

RFCD Codes

280213

Significant Achievements, Anticipated Outcomes and Future Work

During the first month of the project we developed more optimization code for the APAC National Facility Compaq platform and we are working on tuning the performance to exploit the potential of the machines available.

Despite the short period of time, we have two publications on work done on the APAC National Facility that have been submitted or will be submitted to machine learning conferences. Further details and full paper versions will be available in due course.

Computational Techniques Used

The algorithm to be implemented consists of two stages. The first involves the training of a Hidden Markov Model (HMM), which means that we will have to test approximately 500 different settings in order to accommodate for two genomes, deal with acceptor and donor sites, optimize over various HMM architectures, etc. Training of each of the models (we have approximately 100,000 training samples but better results can be obtained for larger datasets)

can take several days on an AMD 1.4GHz platform (the latter has roughly 70% of the performance of a single Alpha 1GHz CPU). Optimization can be parallelized on up to 4 CPUs (with shared memory architecture). We estimate the usage of $500 \times 3 \times 24 = 36,000$ CPU hours.

In the second stage we use an interior point algorithm for the SV optimization problem (the specific version is a so-called ν -SVM, as described in [1]) with a Sherman-Morrison-Woodbury decomposition to deal with large dense matrices. Each interior point algorithm requires roughly 40 large matrix multiplications of $M \times N$ matrices where $M \sim 10^6$ and $N \sim 10^4$ which leads to approximately $4 \times 10^{15} = 40 \times M \times N^2$ operations, or in other words, an estimated 850 CPU hours, assuming 1.2GFlops sustained performance per CPU. Including overhead, 1,000 CPU hours per training run is a reasonable estimate. For model selection we will need to train 10-15 models per genome, which leads to another 30,000 CPU hours.

Communication between the nodes will be mainly by using MPI to scatter and gather results. One node will play the role of a master-node summing up the partial terms of an $N \times M$ matrix multiplication with an $M \times N$ matrix (MN).

[1] Scholkopf B, Smola AJ, Williamson RC, Bartlett PL, New support vector algorithms, *Neural Computation* 12(5):1207-1245, 2000.

Magnetic Interactions in Transition Metals

Principal Investigator

Robert Stranger

Chemistry

ANU

Co-Investigator

Christopher Delfs

Chemistry

ANU

Project

x24

Facilities Used

PC, SC

RFCD Codes

250601, 250201

Significant Achievements, Anticipated Outcomes and Future Work

Broken-symmetry, density functional calculations on the mixed oxo- and carboxylato-bridged complexes $[(\text{NH}_3)_4\text{Mn}(\mu\text{-O})_2(\mu\text{-O}_2\text{CH})\text{Mn}(\text{NH}_3)_4]^{n+}$ and $[(\text{NH}_3)_4\text{Mn}(\mu\text{-O})(\mu\text{-O}_2\text{CH})_2\text{Mn}(\text{NH}_3)_4]^{n+}$ which serve as model systems for several oxomanganese enzymes, have been completed for the $\text{Mn}^{\text{IV}}\text{Mn}^{\text{IV}}$, $\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}$ and $\text{Mn}^{\text{III}}\text{Mn}^{\text{III}}$ oxidation states. A comparison with our earlier work on the planar di-oxo system $[(\text{NH}_3)_4\text{MnO}_2\text{Mn}(\text{NH}_3)_4]^{n+}$ has yielded valuable insights into how the interactions in these magnetically-coupled manganese clusters depend on the number and type of bridging ligands, and the oxidation states of the metal centres. The principal magnetic interactions in $[\text{Mn}_2(\mu\text{-O})_2(\text{HCO}_2)(\text{NH}_3)_6]^{n+}$ were found to involve the $J_{xz/xz}$ and $J_{yz/yz}$ pathways but due to the tilt of the Mn_2O_2 core, they are less efficient than in the planar di- μ -oxo structure. In both the $\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}$ and $\text{Mn}^{\text{III}}\text{Mn}^{\text{III}}$ dimers, the Mn^{III} centres are high spin and the Jahn-Teller effect gives rise to an axially elongated Mn^{III} geometry. In the mixed-valence $\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}$ dimer, the crossed exchange J_{x-y^2/z^2} pathway is quite efficient, leading to significant delocalisation of the odd electron. Since this delocalisation pathway partially converts the Mn^{IV} ion into low-spin Mn^{III} , the coupling can be considered to arise from two interacting spin ladders, one the result of coupling between Mn^{IV} ($S=3/2$) and high-spin Mn^{III} ($S=2$), the other between Mn^{IV} ($S=3/2$) and low-spin Mn^{III} ($S=1$). In $[\text{Mn}_2(\mu\text{-O})(\text{HCO}_2)_2(\text{NH}_3)_6]^{n+}$, the $\text{Mn}^{\text{III}}\text{Mn}^{\text{III}}$ dimer and the lowest energy structure for the $\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}$ dimer both involve high-spin Mn^{III} but the Jahn-Teller effect gives rise to an axially compressed Mn^{III} geometry. In the $\text{Mn}^{\text{IV}}\text{Mn}^{\text{IV}}$ dimer, the ferromagnetic J_{yz/z^2} pathway partially cancels $J_{yz/yz}$ and as a consequence, the antiferromagnetic $J_{xz/xz}$ pathway dominates the magnetic coupling. In the $\text{Mn}^{\text{III}}\text{Mn}^{\text{III}}$ dimer, the competition between the ferromagnetic J_{yz/z^2} and antiferromagnetic $J_{yz/yz}$ and $J_{xz/xz}$ pathways results in relatively weak overall antiferromagnetic coupling. In the $\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}$ dimer, the structures involving high-spin and low-spin Mn^{III} are almost degenerate.

In the high-spin case, the odd electron is localised on the Mn^{III} centre and the resulting antiferromagnetic coupling is similar to the Mn^{IV}Mn^{IV} dimer. In the alternative low-spin structure, the odd electron is significantly delocalised due to the crossed $J_{yz/z}2$ pathway, and cancellation between ferromagnetic and antiferromagnetic pathways leads to overall weak magnetic coupling. The delocalisation partially converts the Mn^{IV} ion into high-spin Mn^{III} and consequently the spin ladders arising from coupling of Mn^{IV}(S=3/2) with high-spin (S=2) and low-spin (S=1) Mn^{III} are configurationally mixed. Thus, in principle the ground state magnetic coupling in the mixed valence dimer involves contributions from three spin-ladders, two associated with the delocalised low-spin structure and the third arising from the localised high-spin structure.

We have also undertaken density functional calculations on the μ -oxo, μ -peroxo bridged model complexes $[\text{Mn}_2^{\text{III}}(\mu\text{-O})(\mu\text{-O}_2)(\text{NH}_3)_8]^{2+}$ and $[\text{Mn}_2^{\text{III}}(\mu\text{-O})(\mu\text{-O}_2)(\text{NH}_3)_6(\text{H}_2\text{O})_2]^{2+}$ in order to rationalize the unusual stability of the μ -peroxo bridge in the structurally characterised trimer $[\text{Mn}_3^{\text{III}}(\mu_3\text{-O})(\mu\text{-O}_2)(\text{AcO})_2(\text{dien})_3]^{2+}$. The calculations reveal that the stability of the peroxide bridge can be attributed to a Jahn-Teller elongation of the axial Mn-N bonds perpendicular to the $\text{Mn}_2(\mu\text{-O})(\mu\text{-O}_2)$ plane which results in a stabilization of the high-spin Mn^{III} oxidation state. However, the difference between the relative energies of the bridged and cleaved peroxide structures is small (ca 0.5 eV), the lowest energy structure depending on the nature of the terminal ligands. Calculations on the model trimer $[\text{Mn}_3^{\text{III}}(\mu_3\text{-O})(\mu\text{-O}_2)(\text{HCO}_2)_2(\text{NH}_3)_9]^{2+}$ indicate that the energetic differences between the cleaved and uncleaved structures is even smaller (ca 0.2 eV) and although the peroxo-bridge still remains intact, it is likely to be quite fuscile.

Computational Techniques Used

The Amsterdam Density Functional (ADF) computer package was used for all calculations.

Publications, Awards and External Funding

External Funding:

ARC Large Grant (2000-2002): A density functional study of oxomanganese enzymes.

ARC Large Grant (2001-2003): Periodic and redox-induced trends in metal-metal bonding.

Publications:

McGrady, J.E. and Stranger, R. Redox-Induced Formation and Cleavage of O-O Bonds in a Peroxo-Bridged Manganese Dimer: a Density Functional Study. *Inorganic Chemistry*, 38, 1999, 550-558.

Delfs, C.D. and Stranger, R. Magnetic Exchange in $[\text{Mn}_2(\mu\text{-O})_3(\text{tmtacn})_2]^{2+}$: Metal-Metal Bonding or Superexchange? *Inorganic Chemistry*, 39, 2000, 491-495.

Delfs, C.D. and Stranger, R. Oxidation State Dependence of the Geometry, Electronic Structure, and Magnetic Coupling in Mixed Oxo- and Carboxylato-Bridged Manganese Dimers. *Inorganic Chemistry*, 40, 2001, 3061-3076.

Delfs, C.D. and Stranger, R. Investigating the Stability of the Peroxide Bridge in μ -oxo and Di- μ -oxo Manganese Clusters. *Inorganic Chemistry*, 2002, submitted for publication.

Astrophysical Radiative Shock Simulations

Principal Investigator

Ralph Sutherland

*Mt Stromlo Observatory
Research School of
Astronomy and
Astrophysics
ANU*

We are undertaking very high resolution hydrodynamical simulations of hypersonic radiative wall shocks in 2 dimensions, using computed non-equilibrium cooling. A Mach 15-20 shockwave is propagated into media with density fluctuations with both power-law and gaussian spectra, and the development of the post shock thermal instabilities and turbulence is examined. We show that due to the intrinsic scales introduced by the cooling process, extremely high resolutions are required to achieve asymptotic solutions. The post shock turbulence is no longer a simple power-law, with strong dissipation occurring at intermediate scales. Finally, a fractal dimensional treatment of the resulting structures gives useful means of estimating the cooling efficiencies of these strong shocks compared to uniform or even steady models.

Co-Investigator

Geoff Bicknell

*Mt Stromlo Observatory
Research School of
Astronomy and
Astrophysics
ANU*

Project

v60

Facilities Used

VPP

RFCD Codes

240101, 240502,
240304

Significant Achievements, Anticipated Outcomes and Future Work

We are also addressing the numerical instability that arises in the directionally split computation of hydrodynamic flows when shock fronts are parallel to a grid plane. Transverse oscillations in pressure, density and temperature are produced that are exacerbated by thermal instability when cooling is present. We have developed a spatial oscillation filter method that removes stripes and permits other high velocity gradient regions of the flow to evolve in a physically acceptable manner. It also has the advantage of only acting on a small fraction of the cells in a two or three dimensional simulation compared to conventional methods, and does not significantly affect performance. These results were presented at the High Energy Density Laboratory Astrophysics 2002 conference in Ann Arbor (U. Michigan), and at workshops here in Australia and at the Lawrence Livermore National Laboratory, California.

The 2D inhomogeneous simulations have demonstrated the true resolution required for accurate modelling of thermal instabilities for the first time. When fully resolved, the fractal nature of the cooling plasma and the large fluctuations in density and pressure at many scales required extremely high resolution. The self-similar nature of the structures means that adaptive mesh techniques fail, and resort to the finest resolution over most of the grid. The implication of this is that fully 3D simulations, required to correctly simulate the internal turbulence of the

post shock flows, will require approximately $10^{(9-10)}$ 3D grid cells. These 2D calculations on the VPP represent the preliminary tests for large scale calculations on the new APAC National Facility.

In the near future fully 3 dimensional simulations will be undertaken, with the aim of producing emission spectra that can be compared to observations of real astrophysical shockwaves.

Computational Techniques Used

Standard ways of dealing with this type of hydrodynamical problem include the introduction of artificial viscosity or grid “jittering”. Artificial viscosity smoothes the shock over more than two zones, thereby reducing the fluctuations between zones parallel to the shock front. Grid-jittering involves displacement of the grid in an oscillatory fashion thereby smoothing fluctuations between cells in the additional grid resampling process. Both of these techniques can be useful. However, they have the disadvantage that they are applied to the entire flow. This has the consequence that otherwise sharp features, e.g. tangential discontinuities in velocity, are smoothed unnecessarily, degrading the resolution. We have developed a new approach involving a “localized oscillation filter”. This filter applies a light smoothing to cells parallel to the shock front. However, the smoothing is restricted to cells local to the shock, thereby avoiding the undesirable side-effects of other methods.

We have utilized a version of the VH-1 code, made available by J. Blondin et al. via their website (<http://wonka.physics.ncsu.edu/pub/VH-1/>). We have extensively reorganized the basic code for vectorization and overall efficiency on the ANU Fujitsu VPP300 supercomputer as well as adding subroutines to update the energy density when optically thin radiative cooling is operating. The new code is now called, simply, ppmr, to distinguish it from VH-1 and to refer to the method used.

The code achieved 98% vector efficiency, processing cells at up to 300,000 per second, and at the time this was the fastest available computing, which allowed the high resolution 2D models to be computed for the first time.

Publications, Awards and External Funding

Grants:

ARC Large Grant: Jet-Cloud Interactions in Active Galactic Nuclei (related to v60 and the newer x34 project)

Publications:

R. Sutherland, D. K. Bisset, G.V. Bicknell, The Numerical Simulation of Radiative Shocks I: The elimination of numerical shock instabilities using a localized oscillation filter, ApJ, 2002, submitted

R. Sutherland, G.V. Bicknell, The Numerical Simulation of Radiative Shocks II: 2D inhomogeneous radiative shocks, ApJ, 2002, in prep

NMR Spectral Assignment Using Structural Information

Principal Investigator

Andrew Torda

Chemistry

RSC

ANU

Co-Investigators

Mark Abraham

Zsuzsanna Dosztanyi

Oliver Martin

James Procter

Regula Walser

Chemistry

RSC

ANU

Project

x08, v04, w51, d97

Facilities Used

PC, SC

RFCD Codes

250503, 250699,

230204, 239901

Significant Achievements, Anticipated Outcomes and Future Work

Some people think force fields are uninteresting sets of formulae and parameters which tell you about the energy of a system. In fact, force fields are fun to build, take apart and use in unusual ways. In previous years, we have spent much time automatically building score functions (like force fields) which tell us how much a protein sequence likes to sit on a particular structure. This is not fanciful. It lets us numerically test the compatibility of a new sequence with a known structure. While we continue the search for better score functions, we are now particularly interested in unexpected uses of these creations.

Firstly, score functions tell us about the interactions between different kinds of units within proteins. Some of these units (amino acids) behave in similar ways, while others are quite different. For example, some like to interact with water molecules, while others are happiest when hidden from the surface of a protein. These similarities tell us how compatible different residues are with each other. In other words, how likely one amino acid is to fit in the place of another during the course of evolution. This means we can build tables of compatibility based on our score functions. To a biologist, this would sound like the definition of an amino acid substitution matrix, a vital table used in all protein sequence comparisons. From our

work, we can generate these tables without relying on an evolutionary model. Aside from the elegance of the method, our home-built substitution matrices have turned out to be useful for analysing the important properties of force fields.

Recognising protein structures also requires some classification of known structures. This is a nightmarishly difficult task since even similar proteins will differ in size and detail. There is no known, deterministic fail-proof approach to comparing protein structures, but we have been developing a graph theory based method. This works, but is somewhat slow. New approaches will use compatibility as calculated from our score functions.

Lastly, our score functions are a direct numerical way to fit protein sequences to protein structures. We are now working on approaches to classifying and predicting protein properties which will merge information from protein sequence comparisons, structure comparisons and sequence to structure comparisons using our score functions.

Computational Techniques Used

All code was home-built and lovingly crafted, except for the pieces which were recklessly slapped together. Across the different parts of the work, a smorgasbord of algorithms was used. This included various numerical optimisation methods for optimising force fields and parameters and a fast dynamic programming method for placing pieces of amino acid sequences on candidate structures. A selection of exotic clique-detection algorithms was used in the protein structure comparison work.

Publications, Awards and External Funding

Zs. Dosztányi, A.E. Torda, Amino acid similarity matrices based on force fields, *Bioinformatics*, 17, 2001, 686-699.

J.B. Procter, A.J. Perry, A.E. Torda, Comparing objects of different sizes: treating proteins as strings, *Aust. J. Chem.*, 54, 2001, 367-373.

D. Reith, T. Huber, F. Müller-Plathe, A.E. Torda, Free energy approximations in simple lattice proteins, *J. Chem. Phys.* 114, 2001, 4998-5005.

A.J. Russell, A.E. Torda, Protein sequence threading - averaging over structures, *Proteins*, 2002, in press.

Computation of X-ray Diffraction Patterns for 3D Model Systems

Principal Investigator

T Richard Welberry

Research School of
Chemistry
ANU

Co-Investigators

Darren Goossens

Matthias Honal

Research School of
Chemistry
ANU

The aim of our project is modelling the disorder that occurs in crystals of some organic molecules, inorganic materials and mineral systems, which we observe in our diffuse X-ray diffraction experiments. We address the question: can we, by using a detailed potential model of the systems under investigation, describe the short-range order properties of the materials sufficiently well that we may obtain computed diffuse diffraction patterns which are in substantive agreement with observed X-ray diffraction patterns? The process is an iterative one involving several stages of computation.

Project

p05

Facilities Used

SC, (VPP, PC)

RFCO Codes

240202, 250105,
260101, 291302

Significant Achievements, Anticipated Outcomes and Future Work

The method has been used to study disorder in a number of quite diverse systems. Major projects have been trying to understand the disorder in cubic stabilized zirconias, CSZ's, (these have commercial importance as "cubic zirconia" gems); in Mullite (a major component of nearly all aluminosilicate ceramics); and the non-stoichiometric iron oxide, wustite Fe_{1-x}O (a major constituent of the Earth's lower mantle). For each of these systems three dimensional models of the way in which vacancies order, and the way in which the rest of the structure relaxes around the defects, have been established. Systems of more recent interest have been various guest/host systems such as the family of urea inclusion compounds; thallium antimonyl germanate which is a non-linear optical material; B8-type alloys which involve interstitial transition metal ordering and the transition metal compound $\text{Fe}_3(\text{CO})_{12}$ in which the Fe_3 moiety is disordered.

In 1997 we developed algorithms to perform this iterative MC methodology solely by computer, using quantitative comparison of observed and calculated diffraction patterns and automatic updating of model parameters, using a least-squares algorithm. This represents a formidable computational task, which is only feasible with state-of-the-art computational facilities. We believe this methodology will become increasingly powerful and more widely used as computers become even faster, allowing increasingly complex systems to be studied.

Computational Techniques Used

- (1) A simulation model is first set-up in terms of basic inter-atomic or inter-molecular interactions.
- (2) A realisation of the model is obtained via simulation, usually Monte Carlo (MC).
- (3) The diffraction pattern of the model system is calculated and compared to the observed data.
- (4) The model is adjusted as a result of the findings in step (3) and the process is repeated from step (1).

Approximately equal amounts of time are spent on steps (2) and (3) above. The latter calculation uses the software algorithm DIFFUSE developed by Dr. Brent Butler at RSC some years ago (this was highly vectorised code). The newly developed methodology for automatic refinement of a MC model requires long production runs. A substantial proportion of the work involves parallel computations in which basically the same calculation is carried out with different sets of model parameters. The methodology is thus ideally suited to a parallel processing environment. Since 1999 and before the advent of the SC machine, virtually all our work has been carried out on a local cluster of 12 Pentium II/III workstations. At present we are carrying out work to implement the methodology on the APAC National Facility SC machine.

Publications, Awards and External Funding

M. Honal, T.R. Welberry, Monte Carlo Study of the Quasicrystal-to-Crystal Transformation Using an Approach Based on the Gummelt Covering, *Zeitschrift für Kristallographie* (in press).

Collapse and Folding of Block Copolymers

Principal Investigator**David Williams***Applied Mathematics**RSPHysSE**ANU***Co-Investigator****Ira Cooke***Applied Mathematics**RSPHysSE**ANU***Project**

x20

Facilities Used

SC

RFCD Codes

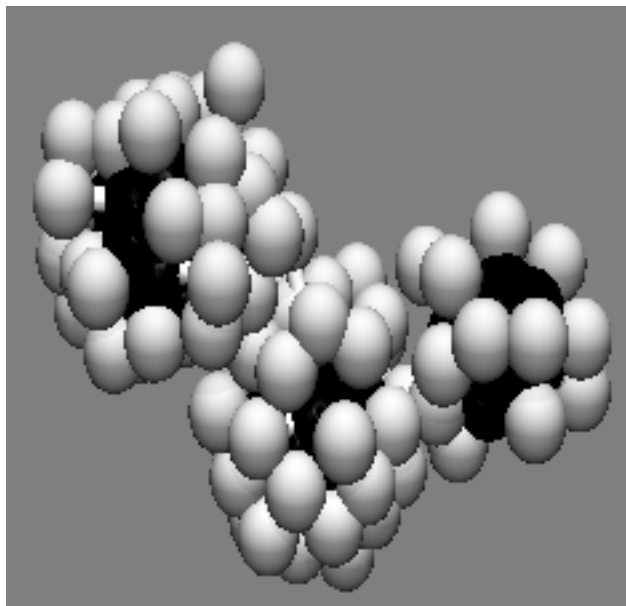
240202, 249901

The protein folding problem involves an attempt to elucidate the tertiary structure of biologically active proteins from their amino acid sequence. One approach to this problem is to strip the protein of all its complexity and to build up knowledge of folding, starting from the very simplest systems. The aim of this approach is to bring to light some of the very general laws of polymer/protein folding. In our research we use a model polymer chain consisting of alternating sequences of hydrophilic and hydrophobic segments. At sufficiently low temperatures this chain undergoes a “collapse” transition, in which the hydrophobic regions micro phase separate from the hydrophilic ones. Our aim is to elucidate general dynamical laws governing this type of “collapse” using Brownian dynamics simulations. Even for a relatively simple system such as this, the computer time required to obtain good results is very large. Where previous studies were restricted to equilibrium conditions, the new APAC National Facility will allow us to examine the much more interesting and physically relevant problem of dynamics.

Significant Achievements, Anticipated Outcomes and Future Work

Our research looks at the collapse of copolymers where one monomer type is in a good solvent and the other is in a poor solvent. One of the key differences between such copolymer systems and those of homopolymers (one monomer type only) is the additional variability afforded by different backbone chain sequences. The influence of backbone chain sequence on polymer collapse is clearly of relevance to the important question of bio polymer folding but is presently not well understood. As a beginning to this type of work we have made a study of the effect of sequence “blockyness” on the collapse of polymers from random coil to compact globule. In this work, polymers were represented by simple bead and spring models, and the collapse process was studied using Langevin dynamics simulations over a range of temperatures. We found that “blockyness” has a strong influence on the thermodynamics and kinetics of collapse and also on the nature of the collapsed state.

Having completed an initial study of the effect of blockyness on collapse, our attention in the coming year will turn towards a more detailed study of the nature of collapsed states. Of particular interest is the formation of intra-chain micelles (see illustration), which show promise as drug delivery agents. Our aim will be to determine the mechanisms underlying their formation and stability, as well as the number that form along a given polymer chain.



Collapsed state of a polymer in aqueous solution with alternating hydrophilic (light) and hydrophobic (dark) monomers. Note the presence of micelle like structures.

Computational Techniques Used

The primary computational task in our work is the calculation of forces for Langevin dynamics simulations. We make use of standard algorithms for this type of work such as nesting lists and lookup tables.