# ENERGETIC AND CONFORMATIONAL STUDIES OF NONSPECIFIC ADSORPTION OF SIMPLE PROTEIN-LIKE CHAIN MOLECULES USING DYNAMIC MONTE CARLO SIMULATIONS 

by
Susan Marisa Liu
B.Eng., McGill University, 1993
M.A.Sc., The University of British Columbia, 1997

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in THE FACULTY OF GRADUATE STUDIES<br>The Department of Chemical and Biological Engineering and<br>The Michael Smith Laboratories<br>We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
October 2004
© Susan Marisa Liu, 2004


#### Abstract

Dynamic Monte Carlo simulations of short HP (hydrophobic-polar) protein-like chains to solid-liquid surfaces are used to probe thermodynamic and dynamic aspects of protein adsorption. The HP model enables the enumeration of all chain conformations, thereby aiding understanding of the relation between adsorption thermodynamics and changes in accessible chain conformations resulting from the sorption process. Simulation results indicate that HP chains having a single conformation at their lowest energy in solution adsorb such that the new lowest energy state of the system is conformationally degenerate. As a result, adsorption can lead to an increase in chain entropy. Entropically-driven adsorption is found to be likely when the interaction energy between the hydrophobic segments of the chain and the sorbent is weak and equals the contact energy between two hydrophobic units within the chain.

Chain sequence and sorbent properties are shown to profoundly influence adsorption thermodynamics. Simulations are carried out where intra- and intermolecular hydrophobic interaction energies are varied to examine the influence of the stability of the native-state conformation on adsorption thermodynamics over a range of sorbent hydrophobicities. Lower stability chains tend to adsorb more readily on hydrophilic sorbents and experience greater average changes in conformation, usually accompanied by a loss in entropy. Adsorption to more hydrophobic sorbents leads to a loss in chain conformational entropy, irrespective of the stability of the native state.

Lateral confinement on the sorbent surface is shown to greatly reduce the degrees of freedom in the chain, thereby resulting in a strong stabilization of the native-state conformation of the chain in its adsorbed state. This effect is compared to experimental


data for nonspecific adsorption of hen egg-white lysozyme to silica to explain the increase in adsorbed enzyme activity as a function of surface loading and geometry. Studies of run-averaged energy trajectories for chain adsorption indicate that the process follows a basic energy path characterized by well-defined energy levels, suggesting the presence of natural kinetic barriers.

This thesis demonstrates the value of simple mesoscopic protein-like chain models and dynamic Monte Carlo simulations of their adsorption behavior in understanding better the mechanisms and forces driving nonspecific protein adsorption.

## Table of contents

Abstract ..... ii
Table of contents ..... iv
List of tables ..... vii
List of figures ..... viii
List of nomenclature ..... xi
Main symbols ..... xi
Superscripts ..... xii
Subscripts ..... xiii
Acknowledgements ..... xiv
1 Introduction, background and objectives ..... 1
1.1 Literature review ..... 5
1.1.1 General aspects of proteins ..... 5
1.1.2 Driving forces for nonspecific protein adsorption ..... 12
1.1.3 Kinetics of nonspecific protein adsorption ..... 16
1.1.4 Simulating and modeling protein adsorption ..... 17
1.2 Thesis objectives ..... 21
1.3 References ..... 23
1.4 Tables ..... 31
1.5 Figures ..... 32
2 Mesoscopic analysis of conformational and entropic contributions to nonspecific adsorption of HP copolymer chains using dynamic Monte Carlo simulations ..... 36
2.1 Introduction. ..... 36
2.2 Theory ..... 42
2.2.1 Dynamic Monte Carlo simulations of HP-model polymers ..... 42
2.2.2 The model system ..... 45
2.2.3 Calculation of thermodynamic parameters ..... 47
2.3 Results ..... 49
2.3.1 Dynamics and thermal unfolding behaviour of model sequences ..... 49
2.3.2 Chain adsorption to the surface ..... 51
2.3.3 Thermodynamics of chain adsorption ..... 54
2.3.4 Influence of sorbent surface geometry ..... 56
2.3.5 Dependence on the total hydrophobicity ..... 57
2.4 Discussion ..... 58
2.4.1 Protein adsorption kinetics are linked to energetic barriers that frustrate conformational trajectories of the peptide chain ..... 62
2.5 Summary ..... 65
2.6 References ..... 67
2.7 Tables ..... 72
2.8 Figures ..... 77
3 Mesoscopic dynamic Monte Carlo simulations of the adsorption of protein-like HP chains within laterally constricted spaces ..... 85
3.1 Introduction ..... 85
3.2 Dynamic Monte Carlo simulations of chain adsorption ..... 88
3.3 Materials and Methods ..... 92
3.3.1 Reagents ..... 92
3.3.2 Measurement of Adsorption Isotherms ..... 93
3.3.3 Lysozyme activity measurements ..... 95
3.3.4 Isothermal Titration Calorimetry ..... 96
3.4 Results and discussion ..... 97
3.4.1 dMC simulations of HP chain adsorption within confined spaces ..... 97
3.4.2 Influence of lateral confinement on the thermodynamics of HP chain adsorption ..... 103
3.5 Summary ..... 107
3.6 References ..... 109
3.7 Tables ..... 112
3.8 Figures ..... 113
4 Energy landscapes for adsorption of protein-like HP chains as a function of native- state stability ..... 122
4.1 Introduction ..... 122
4.2 Protein-like HP chain and dMC simulation algorithm ..... 125
4.3 Results and discussion ..... 128
4.3.1 HP chain adsorption thermodynamics ..... 128
4.3.2 Thermally-averaged energy landscape analysis of HP chain adsorption ..... 133
4.3.3 The deformation entropy for HP chain adsorption ..... 137
4.4 Summary ..... 142
4.5 References ..... 144
4.6 Tables ..... 147
4.7 Figures ..... 152
5 Conclusion ..... 157
5.1 References ..... 161
Appendix ..... 162
Description of dynamic Monte Carlo Simulations and related program code ..... 162
The lattice grid ..... 162
The HP chain. ..... 163
The algorithm. ..... 164
Ending the program ..... 165
Program files ..... 166
Main program <*.for> ..... 166
Input file $<^{*}$.dat $>$ ..... 167
Output files $<^{*}$.out $>,<$ Screen $>$, etc ..... 167
Editing, compiling and running the program ..... 168
Running a single program ..... 168
Running a collection of programs in series ..... 169
Example program for dynamic Monte Carlo simulation, CONTACT9 ..... 170
Main program, <contact9.for> ..... 171
Input data, <test.dat> ..... 219
Output files ..... 228

## List of tables

Table 1.1: Interactions governing the native-state structural stability of globular proteins. $\Delta_{N-D} G$ refers to the Gibbs energy of denaturation for the protein.............................. 31
Table 2.1: Thermodynamic changes for the native to denatured state transition of sequence I at reduced temperatures $T^{*}=0.44$ and $T^{*}=0.71 \ldots . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . ~ 7 ~ 72 ~$

Table 2.2: Thermodynamic changes for the adsorption of sequence I to a relatively hydrophilic surface $\left(\chi_{H W}=-1\right)$ as a function of chain stability $\left(\chi_{H H}\right)$...................... 73

Table 2.3: Thermodynamic changes for the adsorption of denatured sequence I $\left(\chi_{H H}=-\right.$
Table 2.4: Thermodynamic changes for the adsorption of sequence I on surfaces of varying geometry with $\chi_{H H}=\chi_{H W}=-4$.................................................................. 75

Table 2.5: Thermodynamic changes for the adsorption of sequence I to a planar surface when $\chi_{H H}=\chi_{H W}$. 76

Table 3.1: Molar enthalpy change $\Delta H_{a d s}$ as a function of percent sorbent surface coverage for the adsorption of HEWL to particulate silica in $50-\mathrm{mM} \mathrm{KCl}(\mathrm{pH} 7)$ at $37^{\circ} \mathrm{C} . \Delta H_{\text {ads }}$ is expressed on a per mole of HEWL adsorbed basis........................ 112

Table 4.1: Adsorption thermodynamics for HP chain sequence I as a function of native-
state stability and sorbent surface affinity. ....................................................... 147
Table 4.2: The average number of favorable intramolecular and intermolecular contacts formed by sequence I in solution and when adsorbed onto a planar surface.......... 148

Table 4.3: Adsorption thermodynamics for sequence II as a function of native-state
stability and sorbent surface affinity.............................................................. 149
Table 4.4: The average number of favorable intramolecular and intermolecular contacts formed by sequence II in solution and when adsorbed onto a planar surface. ....... 150

Table 4.5: The density of unique conformational states for the adsorbed chain at the global energy minimum energy state as a function of $\chi_{H H} / \chi_{H W} \ldots . . . . . . . . . . . . . . . . . . . . . . . . . . ~ 151 ~$

## List of figures

Figure 1.1: The basic chemical structure of an amino acid and schematic diagrams of the 20 naturally-occuring amino acids 32

Figure 1.2: The condensation reaction forming the peptide link between two amino acids 34

Figure 1.3: Isotherm for the adsorption of hen egg-white lysozyme ( $\mathrm{pI} \sim 11$ ) to particulate silica at pH 7 and $37^{\circ} \mathrm{C}$. Shown is the concentration of protein on the surface as a function of protein in bulk solution. Data points collected from two separate trials are indicated. The solid line represents the ascending isotherm, while dotted lines represent descending isotherms. The vertical dashed line extended from zero indicates that to the lowest possible detectable levels, no desorbed protein was detected. This data is the author's own from reference (43) 35

Figure 2.1: Examples of Verdier-Stockmeyer moves used to manipulate the protein-like chains during simulations. Shown are a 2-bead crankshaft move, a 1-bead flip and an example of an end-bead turn. 77

Figure 2.2: Schematic diagrams of the two model chains: sequence I and sequence II. Filled circles represent hydrophobic (H) units while open circles represent polar (P) units 78

Figure 2.3: Simulation results demonstrating the conformational dependence of sequence I on temperature. Shown is the ensemble-averaged number of intramolecular hydrophobic contacts as a function of reduced temperature, $T^{*}$. The line drawn indicates the trend of the data.

Figure 2.4: System energy probability histograms for sequence I: filled grey bars - chain in solution when $\chi_{H H}=-4\left(T^{*}=0.71\right)$; filled black bars - chain adsorbed on a relatively hydrophobic surface $\left(\chi_{H W}=-4\right)$ when $\chi_{H H}=-4$. The error bars refer to the standard deviation of 5 runs. 80

Figure 2.5: Energy trajectory for the adsorption of sequence I to a planar hydrophobic surface when $\chi_{H H}=\chi_{H W}=-4$. 81

Figure 2.6: Representative conformational states of sequence I shortly after adsorbing to a planar hydrophobic surface when $\chi_{H H}=\chi_{H W}=-4$. All conformations correspond to system energies $E_{i} \geq-48 k T$. 82

Figure 2.7: Representative examples of lowest energy states of sequence I adsorbed to a hydrophobic surface when $\chi_{H H}=\chi_{H W}=-4$. At the global minimum, $E_{i}=-56 \mathrm{kT} . .83$

Figure 2.8: The 6 lowest energy states for sequence I when $\chi_{H H}=-3$ and $\chi_{H W}=-1$. The structure in the centre is the adsorbed chain in its native state.84

Figure 3.1: Schematic diagram of the model HP chain. Filled circles represent hydrophobic (H) chain segments while open circles represent polar (P) chain segments.

113
Figure 3.2: Energy trajectory for adsorption of the HP chain to a planar sorbent surface when $\chi_{H H}=-1$ and $\chi_{H W}=-4$. All other segment-segment interaction energies set equal to zero. 114

Figure 3.3: Dependence of the degeneracy of the lowest energy state (i.e., the total number of unique chain conformations) on accessible sorbent surface area and volumetric space above it. The $x$-axis indicates the number of lattice sites on the sorbent surface available for binding. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-1$

Figure 3.4: Energy probability histograms for the HP chain adsorbed to sorbent surfaces of two different widths: etched grey bars - width of sorbent surface, $x_{S}$, is 4 lattice units; solid black bars $-x_{S}$ is 18 lattice units. The error bars refer to the standard deviation of 5 runs. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}$ $=-1$. 116

Figure 3.5: Ratio of the probability of the HP chain being in its native-state conformation when adsorbed, $P_{N S}{ }^{\text {adsorbed }}$, to that in solution, $P_{N S}{ }^{\text {solution }}$, as a function of the width of the available adsorption site, $x_{S}$. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-4$ (open squares), and $\chi_{H H}=-4, \chi_{H W}=-1$ (open circles). The line drawn indicates the trend of the data

117
Figure 3.6: HEWL-catalyzed pNP-C5 hydrolysis kinetics when HEWL is dissolved in aqueous solution ( $\mathrm{pH} 6,22{ }^{\circ} \mathrm{C}$ ) and when HEWL is nonspecifically adsorbed to particulate silica at different levels of surface coverage. In all experiments, the amount of HEWL and the initial concentration of pNP-C5 are held constant at $50 \mu \mathrm{~g}$ and 0.20 mM , respectively. Shown are measurements taken for HEWL in solution with no surface (squares), HEWL adsorbed on silica at monolayer coverage (circles), and HEWL adsorbed at 20\% monolayer coverage (triangles). 118

Figure 3.7: Thermodynamics of adsorption of the HP chain to a planar sorbent surface as a function of accessible sorbent surface area and volumetric space above it. The $x$ axis indicates the number of lattice sites on the sorbent surface available for binding. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-1$. Trends indicated are $\Delta U / k T$ (solid line), $\Delta A / k$ (dashed line) and $\Delta S / k$ (dash-dotted line). To avoid excessive clutter, simulation points are shown for $\Delta U / k T$ only. Error bars are calculated from the standard deviation of 5 runs. 119

Figure 3.8: Thermodynamics of adsorption of the HP chain to a planar sorbent surface as a function of accessible sorbent surface area and volumetric space above it. Nonzero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-4$. Symbology is the same as in Figure 3.7. 120

Figure 3.9: Thermodynamics of adsorption of the HP chain to a planar sorbent surface as a function of accessible sorbent surface area and volumetric space above it. Nonzero segment-segment interaction energies are $\chi_{H H}=-1, \chi_{H W}=-2$. Symbology is the same as in Figure 3.7

Figure 4.1: Schematic of HP chain sequences used in these simulations: sequence I and sequence II. Hydrophobic ( H ) segments are filled and polar segments ( P ) units are unfilled.

Figure 4.2: Helmholtz energy ( $\Delta_{a d s} A / k T$, squares) and entropy ( $\Delta_{a d s} S / k$, circles) of adsorption for the HP chain sequence I adsorbing on a weakly attractive surface $\left(\chi_{H W}=-1\right)$ as a function of $\Delta_{D-N} A / k T$, the stability of the native-state of the chain relative to its fully denatured state. Lines drawn indicate data trends. Error is within $11 \%$ and $22 \%$ for values of $\Delta_{a d s} A / k T$ and $\Delta_{a d s} S / k$, respectively. 153

Figure 4.3: Thermally-averaged energy landscapes for the HP chain sequence I in solution (squares) and adsorbed (circles) to a weakly attractive sorbent surface ( $\chi_{H W}$ $=-1$ ) under conditions where the stability of the native-state conformation is low $\left(\chi_{H H}=-1\right)$. Lines drawn indicate data trends. 154

Figure 4.4: The average length, $\overline{l_{x}}$, in the $x$ dimension of each bond vector in HP chain sequence I as a function of $\chi_{H H} / \chi_{H W}$, the ratio of the average intramolecular to intermolecular contact energies. The dotted line represents the average length of each bond vector in the $x$ dimension when the chain is in its native-state conformation. The line drawn indicates the data trend. 155

Figure 4.5: Comparison of $\Delta_{\text {def }} S / k$ values calculated with Eq. [4.12] to dMC simulation data values over a range of chain deformations, $\overline{l_{x}} / \overline{l_{x o}}$ : dashed line $a=2$, solid line $a=2.8$. The dMC data (open squares) shown correspond to adsorption conditions where $\chi_{H H} / \chi_{H W}=0.25\left(\overline{l_{x}} / \overline{l_{x o}}=2.2\right), 1\left(\overline{l_{x}} / \overline{l_{x o}}=1\right)$, and $4\left(\overline{l_{x}} / \overline{l_{x o}}=0.85\right)$, respectively

## List of nomenclature

## Main symbols

$\alpha_{x} \quad=$ ratio of projection of bond vectors $=\overline{l_{x}} / \overline{l_{x o}}$
$\beta=1 / \sqrt{2 n l_{x}^{2}}$, relation from Eq.[4.7]
$\chi_{i j} \quad=$ Flory energy parameter between components/residues $i$ and $j$
$\omega i j \quad=$ interaction energy between components $i$ and $j$
$\Omega(x)=$ density of states having value $x$
$a \quad=$ length of one lattice unit
$a \quad=$ dimensionality of system [Eq. 4.12]
$A \quad=$ Helmholtz energy
$E \quad=$ energy
$E_{i} \quad=$ energy of conformation $i$
$\delta E \quad=$ energy difference between start and end of Monte Carlo move
$h \quad=$ number of intramolecular hydrophobic contacts
$h_{N} \quad=$ number of intramolecular hydrophobic contacts in the native state conformation
$k \quad=$ Boltzmann constant
$l_{x} \quad=$ projection of bond vector on $x$-axis
$\bar{l}_{x} \quad=$ mean projection of bond vector on $x$-axis
$L \quad=$ chain length, number of residues in a chain
$n \quad=$ number of bonds within the chain Eq.[4.7]
$n \quad=$ random number, chosen for comparison in Monte Carlo algorithm
$n_{i} \quad=$ number within sample having end-to-end distances within $x_{i}$ and $x_{i}+d x$ Eq.[4.11]
$N \quad=$ total number of end-to-end distances sampled (Chapter 4)
$N \quad=$ number of lattice sites in the simulation grid
$p_{i}\left(E_{i}\right)=$ probability that the energy of the lattice is $E_{i}$
$P \quad=$ calculated probability for comparison to randomly chosen number in the Monte Carlo algorithm
$Q \quad=$ partition function
$r_{i} \quad=$ coordinates of lattice site $i$
$\delta\left(r_{i}-r_{j}\right)=$ function based on relation of positions of lattice sites $i$ and $j$
$S \quad=$ entropy
$T$ = temperature
$U \quad=$ internal energy
$T_{m}=$ "melting" temperature
$T^{*} \quad=$ reduced temperature
$V \quad=$ volume
$W(x)=$ distribution of chain end-to-end distances in $x$-dimension
$\Delta_{Y}<X>=$ difference of ensemble averaged values of $X$ for transition $Y$
$\langle X\rangle=$ ensemble average of value $X$
$\left(x_{i}, y_{i}\right)=$ coordinates of lattice site $i$

## Superscripts

$o \quad=$ in maximum entropy conditions (Chapter 4)

## Subscripts

```
A = wall (athermal)
ads = adsorption (non-adsorbed to adsorbed transition)
conf = conformational
D = denatured state
def = deformation
free = in solution, non-adsorbed
H = hydrophobic residue of chain
N = native state
o = at maximum entropy conditions (Chapter 4)
P = polar residue of chain
S = solvent
W = wall (active)
x = with respect to the }x\mathrm{ -axis
```


## Acknowledgements

I would first like to express my heartfelt gratitude to my research supervisor, Chip Haynes. His scientific guidance and unwavering support during my time in UBC will always be appreciated and remembered.

I am grateful to the Institute of Applied Mathematics for the generous use of their computer facilities. Grants from the National Science and Engineering Research Council and the Protein Engineering Network of Centres of Excellence enabled this work.

I would also like to acknowledge fellow members of the Michael Smith Laboratories (formerly the Biotechnology Laboratory) past and present, who contributed to my work and learning there. I'd especially like to thank Louise Creagh, for her tutelage and sound advice regarding all matters practical. My appreciation goes out to members of the Haynes Lab and also to the labs of Robin Turner, Jamie Piret, Douglas Kilburn and Donald Brooks for frequent sharing of their resources.

The richness of my graduate student experience can be attributed in large part to Green College, UBC. I feel fortunate to be a part of this special and inspiring community.

Finally, I am indebted to many kind friends, my father, Dirkson, and dear sisters, Dora and Jean for their encouragement and support. It is with great sadness that my late mother, Esther Liu, is not able to share this personal milestone with me.

## 1 Introduction, background and objectives

The need to understand and control protein adsorption at solid-liquid interfaces is driven in large part by current limitations in synthetic biomaterials for human implantation and in bioprocessing equipment, particularly chromatography resins (1-3).

A substantial effort is currently being made to tailor properties of artificial materials to minimize or eliminate a negative immunological host response, such as the formation of thrombi at or near the blood-implant interface (1,4). As adsorption of plasma proteins to a poorly designed foreign surface is known to trigger a biochemical reaction cascade leading to the formation of thrombi $(5,6)$, much of this work is devoted to identifying surface chemistries which properly control the density, composition and conformation of the adsorbed protein layer ( 1,7 ). Similarly, a myriad of empirical strategies are now being tested to improve the performance of other synthetic body-fluid-contacting materials, including contact lenses $(8,9)$, dental fixtures $(10,11)$, and hemodialysis equipment (7).

An increased dependence on biomolecule-based pharmaceuticals has further intensified our need to better understand and control protein adsorption. FDA-approved recombinant-protein, DNA and viral products have increased by more than $1000 \%$ over the past decade, and are now a preferred strategy to treat a number of life-threatening cancers and auto-immune diseases (12-14). Production costs for protein and DNA-based therapeutics are substantive, and the need to deliver a product free of contaminants is
increasing due to growing concerns over product safety (i.e. the avoidance of transmission of Creutzfeldt-Jakob disease, hepatitis, etc.). Identifying ways to reduce costs and increase safety by effectively streamlining the production of these valuable materials is therefore crucial. The separation of complex protein mixtures is typically carried out using large-scale chromatographic processes where the proteins are purified by preferential adsorption or partitioning from a mobile liquid carrier to a solid matrix. In general, the protein's specific affinity for and interaction with the stationary phase dictates the quality of the separation. Improvements in column performance are therefore intimately linked with the ability to carefully control the sorbent chemistry and geometry.

Over the past half century, a significant body of literature has been devoted to adsorption of chain molecules at solid-liquid interfaces. Adsorption phenomena in such systems are complex due to the unique properties of chains. Monodisperse linear homopolymer chains (possibly the simplest member of this adsorption group) can fluctuate among a large number of chain configurations at the solid-solution interface. Characterizing the adsorption behavior of the chain is therefore difficult, as it generally takes several parameters to describe the state of the polymer at the interface. Representative parameters include the average number of points of attachment, the horizontal spread (often defined in terms of the average chain radius $\sqrt{\overline{r^{2}}}$ ), and the average chain thickness $\Delta r$. There are in fact more parameters than can be uniquely specified from fitting of adsorption data, and this has made it difficult to confirm or deny proposed models for homopolymer adsorption. Advances in our understanding of homopolymer adsorption have therefore come largely through experiments that specifically probe conformations of
adsorbed chains (x-ray diffraction, neutron scattering, etc.) or, more commonly, through adsorption data generated for model systems using molecular simulation techniques (15, 16). On and off-lattice Monte Carlo simulations, other random-walk approaches, and various molecular dynamics and statistical approaches have all been used to generate useful model adsorption data for understanding homopolymer adsorption phenomena (17-23).

Although it shares many of the classic features associated with flexible homopolymer adsorption, globular protein adsorption has proven considerably more difficult to understand at the molecular and thermodynamic levels. For example, in aqueous solutions, both linear flexible homopolymer adsorption and globular protein adsorption can be endothermic, indicating that adsorption is entropy rather than energy favored. However, as flexible polymers are known to lose conformational entropy upon adsorption, the gain in system entropy can be precisely ascribed to a gain in solvent entropy. The origin of the entropy gain in globular protein adsorption systems is more complex and therefore far less well understood. Flexible polymer adsorption and globular protein adsorption also share a tendency for surprisingly slow adsorption kinetics (24). Adsorption that appears to have reached equilibrium can often display a slow but continuous drift for days $(25,26)$. For the case of flexible polymer adsorption, Cohen-Stuart et al. $(27,28)$ have argued that the slow adsorption kinetics are due to polymer polydispersity effects. The same argument cannot hold for globular proteins, which are uniform in size and chemistry.

This thesis is founded on the hypothesis that simulation of the adsorption of simple protein-like chains using a dynamic Monte Carlo (dMC) method will serve to improve our fundamental understanding of nonspecific protein adsorption. The argument is supported by the successful use of molecular simulations to greatly improve our understanding of the adsorption of flexible linear polymers to solid-liquid interfaces (18,21,23), and by recent dMC simulations performed by Anderson et al. (29) on the adsorption of simple mesoscopic protein-like chains to a liquid-liquid interface which gave important insights into the dynamics and thermodynamics of the process.

Currently, full atomistic simulations of a related problem, protein folding, can only be carried out over a millisecond time period (30). This duration limit is not practical for the simulation of protein adsorption processes, since they are known to occur over periods of minutes to days (26). We therefore choose to restrict our studies to the simulation of simple heteropolymer chains composed of two types of units, hydrophobic $(\mathrm{H})$ and polar (P), known commonly as the HP model. By proper sequencing of the residues, HP chains will fold into unique protein-like conformations. Despite the obvious simplicity of the HP model, Onuchic (31), Dill $(32,33)$, and many others $(34,35)$ have shown it to be useful in understanding the mechanism of protein folding, the forces that stabilize the native-state conformation, and the sensitivity of native-state stability on system conditions.

### 1.1 Literature review

### 1.1.1 General aspects of proteins

Proteins are biological macromolecules. Diverse in structure and function, they are abundant in all living beings. Proteins often function as biological catalysts, but are also known to be important in ligand transport and cell signaling and as structural materials within biological systems. The building blocks of proteins are L-amino acids. The general structure of an amino acid is $\mathrm{RCH}\left(\mathrm{NH}_{3}{ }^{+}\right) \mathrm{CO}_{2}^{-}$, where R represents its characteristic side chain (see Figure 1.1). Twenty naturally-occurring amino acids exist, each having characteristics defined by the chemical properties of their respective side chain.

Within a protein, amino acids are joined through a condensation reaction between the carbonyl and amino groups (Figure 1.2). The peptide link formed is highly rigid due to its partial double bond character. The sequence in which amino acids connect, referred to as the primary structure of the protein, is significant in that it is the interactions of the peptide units and their respective side chains which impart the principal characteristics and eventual function to the molecule. The primary sequence determines the formation of highly organized substructures such as $\alpha$-helices or $\beta$-sheets, whose content define the secondary structure of the protein. Similarly, its composition, as determined by how secondary structures and other less ordered forms are placed, determines its overall threedimensional arrangement (its tertiary structure). The chain's noncovalent association
with like or unlike chains (its quaternary structure) is also ultimately defined by its amino acid sequence.

Protein structures are varied and complex. Unlike many other types of linear chain molecules, proteins in an aqueous solution of moderate temperature and pH usually display a specific conformation or a limited set of similar conformations. Maintaining this state, known as the native-state conformation, is essential to the function of the protein. The native state is stabilized relative to the large ensemble of denatured states through a closely-balanced combination of energies. Perturbations, whether originating from changes in the protein's solvent environment (alterations in temperature, ionic strength, or pH for instance), or the presence of external factors (e.g. a solid-liquid interface nearby) can easily shift the balance of these stabilizing energies and cause the protein to unfold and therefore cease to function.

Proteins comprised of a single polypeptide chain can be divided into three broad categories according to their tertiary structure:
a) expanded coil structures: flexible and highly solvated,
b) fibrillar proteins: mainly consisting of regular secondary structures such as $\alpha$ helices and $\beta$-sheets, and
c) globular proteins: compact proteins that are made up of both random and structured parts, all folded into a roughly spherical configuration.

Most proteins of interest, such as enzymes and antibodies, are globular, and they are the type of proteins of interest in this work.

### 1.1.1.1 Globular proteins

Globular proteins in an aqueous solution have a number of general characteristics (36), some of which are described here.

1) Globular proteins are roughly spherical in shape with average diameters of the order of angstroms to nanometers.
2) Hydrophobic side groups have a tendency to reside in the interior of globular proteins to avoid contact with water. This does not necessarily mean that all hydrophobic residues are sheltered from the solvent or that the interior is composed entirely of hydrophobic groups. Internal hydrophobicity is limited, for instance, by the presence of the relatively hydrophilic polypeptide backbone.
3) Charged groups are found predominantly on the exterior of the protein, while the very few charged groups in the interior are almost always found in ion pairs.
4) Globular proteins are very densely packed, i.e. comparable to densities for polymer glasses. The atomic packing fraction of a protein is about $75 \%$, significantly higher than the packing fraction of liquid water which is $58 \%$ at $25^{\circ} \mathrm{C}$ and 1 atm .

### 1.1.1.2 Factors affecting native-state stability in aqueous solution

A folded protein is a densely packed molecule stabilized by an intricate heterogeneous network of intra- and intermolecular forces. Thermodynamic investigations carried out by Privalov (37) indicate the Gibbs energy stabilizing the native state of a globular
protein at physiological conditions typically lies between 30 and $70 \mathrm{~kJ} / \mathrm{mol}$, roughly equivalent to the energy of 4 to 12 hydrogen bonds.

Table 1.1 summarizes favorable and unfavorable interactions and forces known or at least thought to affect the stabilities of proteins dissolved in aqueous solutions. The table shows that hydrophobic dehydration, dispersion forces and hydrogen bonding can drive folding of the protein into its native state. Compensating those forces are the associated loss of conformational entropy and distortions of bond lengths and bond angles. Coulombic forces can be favorable or unfavorable to the stability of the native structure, depending on the overall pH of the system relative to the isoelectric point of the protein.

Hydrophobic dehydration is generally believed to be a major driving force for protein folding in aqueous solutions (38-40). It refers to the change in solvation of hydrophobic amino acid side chains of a protein in its folded (native) state relative to its unfolded (denatured) state. When a protein is fully denatured, most if not all of its side chains are exposed to the aqueous solvent. This results in a high degree of solvation which requires the water molecules to locally arrange themselves around the apolar solute (in this case, the denatured protein) in a relatively ordered shell which seeks to maximize solvent hydrogen bonding at the solvent-solute interface. In general, the native-state structure has considerably fewer hydrophobic side chains exposed to the solvent since a majority of the apolar groups fold into the interior of the protein. Solvent entropy is therefore gained from dehydration of these apolar groups.

Intramolecular hydrogen bonding is also thought to make a substantial contribution to the stability of the native state. Creighton (36) has argued that hydrogen bonding contributes $c a .30-45 \%$ of the energy driving folding, with less significant contributions from dehydration, van der Waals and electrostatic interactions.

The significance of electrostatic forces to the state of the protein can be estimated by observing the dependence of protein stability on changes in pH and ionic strength. For example, at extreme pH relative to the isoelectric point, the charge density of a folded protein becomes high, and there is a tendency for the protein to unfold. Specific electrostatic interactions such as ion pairing within a protein have an opposite effect in that they usually lead to the stabilization of the native protein. A rough gauge of the degree of electrostatic interaction occurring at the surface of the protein is given by the local water density adjacent to the protein surface: an ion of significant charge density increases the molar density of water directly surrounding it (41).

Dispersion forces, which are highly dependent on the distance between atoms ( $\Gamma \alpha r^{-6}$ ), are likely important for local protein structure due to the dense atomic packing in a typical protein. The total magnitude of stabilization energy from dispersion effects, however, is thought to be less than that due to the effects of hydrophobic dehydration and hydrogen bonding (39).

Counteracting all the positive stabilization forces is one main destabilizing force: the loss of conformational entropy resulting from the folding of the polypeptide chain $(39,40)$.

Creighton has estimated that for a polypeptide chain in a random coil, approximately 4 distinct backbone conformations exist per peptide unit (36). Assuming that a peptide unit has only one backbone conformation when involved in an $\alpha$-helix or $\beta$-sheet, the loss of entropy per peptide unit is $R \ln 4=11.53 \mathrm{~J} \mathrm{~mol}^{-1} \mathrm{~K}^{-1}$. For a protein consisting of 100 residues, the loss in entropy is approximately $1200 \mathrm{~J} / \mathrm{K}$ per mole of protein due to freezing of the backbone structure. This results in a Gibbs free energy gain of $350 \mathrm{~kJ} / \mathrm{mol}$ at 300 K . Additional entropy losses occur from reductions in conformational freedom of side chains within the interior of the folded protein.

A somewhat less significant force opposing the native state is the distortion of covalent bond lengths and bond angles as determined by energy minimization calculations (36, 42). These distortions, which add approximately 4 to 8 kJ per distorted bond to the native state energy, are believed to be necessary to optimize the various interactions (hydrophobic, dispersive and peptide-peptide hydrogen bonding) required for a tightly packed, compact molecule.

### 1.1.1.3 Macroscopic properties of protein adsorption systems

Steady-state behavior of a protein adsorption system is most often illustrated by an adsorption isotherm, where the surface concentration of protein is measured against the concentration of free protein in bulk solution. Figure 1.3 shows an example isotherm for the adsorption of hen egg-white lysozyme to silica at pH 7 and $37^{\circ} \mathrm{C}$ (43). The direction of the arrows within the figure indicate the type of isotherm being represented. Arrows
up and to the right indicate the ascending isotherm, where the total protein concentration is progressively increased. Arrows to the left indicate the descending isotherm where the free protein concentration is diluted at otherwise constant conditions.

The isotherm shown in Figure 1.3 demonstrates some general features typical of many globular protein adsorption systems. First, the behavior of proteins on solid surfaces is complex. Initial slopes of ascending protein adsorption isotherms are usually steep, indicating a strong affinity of the protein for the sorbent surface, but not infinite, suggesting establishment of a quasi-equilibrium between the sorbate and protein in the bulk phase. Second, ascending and descending isotherms differ in an apparent timeindependent manner, indicating that at a given free protein concentration, the system can exist in more than one state. As a result, a protein adsorption process shows some classic features of irreversible thermodynamics. What is also true for this particular system and many others is that for a given ascending isotherm there are an infinite number of descending isotherms, each of which is defined by the departure point from the ascending isotherm. In most cases, there is no evidence that the descending isotherms rejoin the ascending isotherm.

Despite the apparent irreversible nature of their binding interaction, adsorbed proteins are dynamic on a sorbent surface. For example, fluorescence recovery after photobleaching (FRAP) experiments demonstrate that nonspecifically adsorbed proteins are mobile on the sorbent surface and therefore are able to reposition themselves in response to a concentration gradient (44). Limited exchange of the protein on and off the surface has
also been shown in radiotracer experiments, where labelled proteins are displaced from the surface by non-labelled proteins $(45,46)$. These results suggest that although nonspecific protein adsorption appears macroscopically irreversible, atomic intermolecular contacts between sorbent and protein are in a constant state of flux, indicating some level of reversibility at the microscopic level.

### 1.1.2 Driving forces for nonspecific protein adsorption

Isothermal titration calorimetry has allowed for the direct quantification of the heat associated with protein adsorption. Often, as one might expect, enthalpy is found to drive adsorption. But in many cases, the process is endothermic, indicating that adsorption occurs through an increase in system entropy (47, 48). This supports growing evidence that protein adsorption is not the result of a simple single-step reaction process, but a combination of multiple, perhaps synergistic subprocesses occurring at different rates (47). Of those known to occur, three subprocesses are routinely observed: 1) changes in protein conformation, 2) dehydration of parts of the sorbent and protein surfaces, and 3) redistribution of charged groups at the protein-sorbent interface $(40,49)$.

### 1.1.2.4 Protein structural changes.

The introduction of a sorbent interface can influence the stability of the native-state conformation of a protein by interacting directly with the protein (40). Structural rearrangement of adsorbed proteins has been observed using a wide range of analytical
techniques, including many spectroscopic techniques such as total internal reflection fluorescence (TIRF) (50), microscopic techniques such as atomic force microscopy (51, 52 ), probing of the charge profile of the protein using proton titrations (53,54), and thermodynamically, using calorimetry $(47,55)$. The severity of the conformational change varies. In some instances, the structural perturbation away from the native state is small. Ellipsometry and reflectometry experiments, for instance, have shown thicknesses of adsorbed layers which are comparable to the native protein adsorbed in a side-on or end-on orientation $(56,57)$. Activity measurements for many structurally stable enzymes adsorbed onto hydrophilic surfaces have also indicated minimal structural changes in the adsorbed state (58). In most cases, however, the conformational changes are more severe, and several analytical studies have suggested multiple conformational states of adsorbed proteins (59-61).

Structural denaturation has been shown to increase with certain properties of the system. In particular, high sorbent hydrophobicity, low surface coverage, and low protein stability are correlated with increases in conformational change (43, 50, 62). These aspects will be discussed in detail in upcoming chapters, and therefore will not be discussed further here.

### 1.1.2.5 Hydrophobic interactions

The major contributing effect to a positive entropy of adsorption is thought to be the hydrophobic effect (40), where, in an aqueous environment, the presence of a
hydrophobic solute causes the solvating water molecules to arrange themselves in order to minimize unfavorable polar-nonpolar contacts with the surface while maximizing preferred hydrogen bonding amongst themselves (38, 39). This local enhancement in solvent structure is characterized by a significant decrease in entropy and a large positive change in heat capacity, as demonstrated by the $\Delta S$ and $\Delta C_{p}$ for transfer of non-polar solutes into dilute aqueous solutions (63), and by differential scanning calorimetry (DSC) studies of protein unfolding (38).

The significance of the hydrophobic effect has been investigated from the perspective of both the protein and the sorbent surface. Greater hydrophobicity, either at the sorbent or protein surface, often lead to increased adsorption and energies of adhesion (50, 64-66). Some studies however show that moderate degrees of hydrophobicity maximize adsorption affinity ( 67,68 ), indicating that although hydrophobic interactions are significant, they do not necessarily dominate the driving force for protein adsorption.

### 1.1.2.6 Electrostatics

Protein adsorption electrostatics are complex, involving the overlap of the electric double layers of the protein and the sorbent surface, charge interactions between the protein macromolecules, and, especially in the case of repelling charge, involvement of low-molecular-weight ions present in solution (40). As a result, protein adsorption is often strongly dependent on the pH of the system. For example, when electrostatic interactions between neighbouring proteins dominate, adsorption levels are typically maximized when
the solution pH is near the isoelectric point of the protein, a condition where the protein surface charge is neutralized, thus allowing proteins to pack closely together on the surface $(40,64,69)$.

Alternatively, for cases when electrostatic forces between the protein and the interface dominate, proteins tend to show maximum affinity for the sorbent at the condition of charge complementarity given by the pH where the protein charge is equal and opposite to that of the sorbent surface (40). In the case where charge complementarity does not occur, ions from solution must be incorporated into the protein-sorbent interface to achieve a neutral situation. Changes in the pH of the system prior to and after adsorption have therefore been monitored to determine the net transfer of charge to the interface (53).

### 1.1.2.7 Other driving forces

Other subprocesses can also contribute to the driving force for adsorption, including formation of specific intermolecular ion pairs or hydrogen bonds, intra- and intermolecular van der Waals interactions, and those forces which scale with $1 /(\text { distance })^{3}$ such as dipole-dipole and dipole-induced dipole interactions (70, 71).

### 1.1.3 Kinetics of nonspecific protein adsorption

The apparent irreversible nature of nonspecific protein adsorption has motivated a large number of fundamental investigations of the kinetics of the process, both for singleprotein and competitive adsorption systems ( $45,46,72,73$ ). Some of the central goals of these kinetic studies are 1) to identify time constants for the various steps (subprocesses) along the reaction pathway, 2) to identify that step, and associated energy barrier, which allows limited protein exchange (i.e. the desorbed protein is replaced by a protein of the same or different kind) but prevents spontaneous desorption upon dilution, and 3) to obtain the necessary database to test proposed kinetic models for the adsorption process.

The protein adsorption reaction involves a number of time-dependent steps, including 1) diffusion of the protein to the surface or corresponding boundary layer area, 2) attachment (and detachment) of the protein to (from) the surface, and 3) reorientation, rearrangement and conformational changes to the protein once adsorbed. Due to the lack of sensitivity in available analytical techniques, these reaction steps cannot in general be directly probed, but only inferred from the nature of the total signal.

Additional subprocesses have been considered in kinetic theories for nonspecific protein adsorption. For example, Lundstrom and Elwing (74) include the possibility of protein exchange in their simple theoretical model. They hypothesize that the state of the protein on the surface determines its exchange rate with protein in solution; specific rates exist for self-exchange reactions, when the adsorbed protein remains in a native conformation,
and for alternate-exchange reactions, when the adsorbed protein has undergone a conformational change. Competitive and multiple layer adsorption have also been considered (45, 46, 75 ).

In general, model parameters are not determined directly, but are simultaneously regressed from global kinetic data obtained from ellipsometry $(59,76)$ or surface pressure experiments (72, 77-79). As a result, development of kinetic models for nonspecific protein adsorption remains a largely empirical science.

### 1.1.4 Simulating and modeling protein adsorption

Although the body of knowledge regarding protein structure is substantial, practical limitations (in computational power, for example) have so far restricted the development of detailed molecular simulations of protein adsorption. Simulations that have been carried out generally treat the protein as a hard sphere or conformationally static macromolecule and focus primarily on describing electrostatic interactions (80), dispersive effects $(81,82)$ and solvent interactions $(83,84)$. However, a small number of more advanced and realistic simulations have recently been reported, and results from these studies are discussed in detail in subsequent chapters of this thesis.

In general, analytical models describing protein adsorption isotherms and kinetics are largely based on comparison with experimental data and tend to be highly empirical in nature. The simplest models are those which assume the protein to be a hard particle
which adsorbs reversibly to the sorbent surface (85-87). For example, the Langmuir equation has often been used to describe protein adsorption isotherms (87) At first glance, this approach seems reasonable as proteins often adsorb at monolayer coverage $(47,85,88)$, and as a consequence result in isotherms exhibiting a shape similar to that of a Langmuir curve. However, mechanistic inferences drawn from the good fit of the Langmuir equation are clearly misleading $(40,89)$, as protein structural changes and adsorption irreversibilities are not accounted for in the model.

The random sequential adsorption (RSA) model represents proteins as hard particles, while acknowledging the irreversible nature of protein adsorption (90-93). In a classic RSA-type model, hard particles (most commonly a round disk or sphere) adsorb randomly and are not allowed to overlap. Once adsorbed, the particles are immobile and eventually reach a "jamming limit" where surface coverage cannot increase. Tarjus et al. (90) modified the RSA model to investigate the more realistic situation where desorption and surface diffusion take place. They presented a generalized description of how a process can reach an equilibrium-like steady state characterized by the relative rates of adsorption, desorption and diffusion. Van Tassel et al. (91) introduced conformational change into the RSA model by including an expansion factor for the sorbate. Upon contact with the surface, the particles (in this case, circular disks representing proteins) are allowed to symmetrically expand to a greater diameter, representing the tendency for a globular protein to flatten out when adsorbed. A later paper by the same group which incorporate partial reversibility in the model (92) showed that greater saturation occurs
when sorbate expansion is allowed and that conformational change is dependent on bulk concentration.

In an attempt to advance the RSA model beyond steric effects, electrostatic forces were added by Adamczyk et al. (94) and Lenhoff et al. (95, 96). Both groups incorporated charged particles and solution electrolytes into the model, resulting in particle-particle and particle-surface energy considerations. Oberholtzer et al. (95) also incorporated a perpendicular force imposed on approaching particles to represent a barrier to adsorption taking into account the opposing forces of surface attraction and electrostatic repulsion of already-adsorbed particles. The latter 3D model was more successful in describing experimental results and demonstrated the importance of including an energy barrier prior to contact in the process. Changes in protein conformation were not addressed in these models.

Electrostatic and dispersion forces were incorporated in a protein adsorption model in the case of charged proteins (represented by spheres) adsorbing to a similarly charged surface. (97, 98). In a later study, Stahlberg et al. (99) solved for the electrostatic and van der Waals interaction energies between two parallel plates to determine a capacity factor for ion-exchange chromatography. Despite the coarse geometry of the model, their results correlated well with experimental data over a large range of ionic strengths. Roth and Lenhoff (100) calculated equilibrium constants for proteins adsorbing to oppositely charged surfaces at low coverage where proteins were represented as a low dielectric spheres each containing a single central charge. Interactions with the surface were
determined to be functions of the size and charge of the protein, the charge density on the surface and solvent characteristics. Later papers by the same group indicated that success of the model is highly dependent on proper estimation of particle shape and the charge distribution on the surface of the protein $(71,101)$.

Models which attempt to take into account protein conformational change are few in number. Combining experimental and theoretical estimates, Haynes et al. (43, 47) developed a model based on the enthalpic and entropic contributions of six adsorption subprocesses: changes in the protein structure, hydration effects, protonation or deprotonation of titratable residues on the protein and sorbent surfaces, and three electrostatic effects originating from overlapping of the protein and sorbent electric fields, i) coulombic interactions, ii) specific ion pairing between oppositely charged residues, and iii) ion co-adsorption from the solvent into the interfacial layer. Estimates of protein structural changes in the model are derived from thermodynamic parameters determined using microcalorimetry.

A mean-field approach was taken by Fang and Szleifer (102) who incorporated conformational change and competitive adsorption to their model. A general diffusion equation was used to define the movement of the protein to the surface while the protein's chemical potential gradient provided a measure of the overall driving force for adsorption. The study of both kinetic and thermodynamic aspects allowed for the examination of the initial adsorption sequence as well as transition of the adsorbed layer to its eventual equilibrium state. They investigated the situation where particles undergo
a surface-induced conformational change upon adsorption. Their results revealed that the composition of the adsorbed layer is dependent on the bulk concentration of the protein and the degree of intermolecular interaction at the surface. In their study of the adsorption of protein mixtures, they were able to mimic some aspects of the Vroman effect, where, due to differences in diffusion rates and degrees of attraction to the surface, larger particles eventually displace smaller particles from the sorbent surface.

### 1.2 Thesis objectives

The objective of this thesis is to study the adsorption behavior of simple protein-like HP chains using dynamic Monte Carlo simulations with the aim of understanding how the tendency of such chains to form unique low-energy conformations in solution alters their adsorption thermodynamics relative to the adsorption behavior of a random-coil homopolymer. A two-dimensional lattice is used to simplify the problem such that all chain conformations can be observed. The model is used to:

1) Determine how system variables such as sequence order, structural stability and surface hydrophobicity affect adsorption thermodynamics, particularly their effect on the change in the conformational entropy of the chain.
2) Study average adsorption trajectories of protein-like chains to better understand the dynamic behavior of nonspecific protein adsorption systems.
3) Determine the effects of sorbent geometry and available surface area on adsorption thermodynamics through simulations involving spatial restrictions of an adsorbing protein-like chain.
4) Identify the driving forces and mechanisms leading chains of varying conformational stability to adsorb onto hydrophobic and hydrophilic surfaces.
5) Identify conditions where adsorption is entropically favored and then use the model to obtain a clear understanding of the origin of the gain in entropy.
6) Evaluate and compare energy landscapes for the HP chains in solution and adsorbed to a solid-liquid interface to assess how the multiplicity of states are distributed in both systems and how these differences determine the accessible conformation(s) (and energies) of the chain.

These results are then compared with experimental observations for nonspecific protein adsorption with the aim of better understanding how the inherent conformational degrees in the chain influence its adsorption thermodynamics.

What follows are three research chapters composing the main body of this thesis. Objectives 1 and 2 are the focus of Chapter 2, entitled "Analysis of conformational and entropic contributions to nonspecific protein adsorption at solid-liquid interfaces using dynamic Monte Carlo simulations". Chapter 3, entitled "Mesoscopic dynamic Monte

Carlo simulations of the adsorption of protein-like HP chains within volumetrically constricted spaces", concentrates on Objective 3. Finally, Objectives 4 to 6 are addressed in Chapter 4, entitled "Energy landscapes for adsorption of a protein-like HP chain as a function of native-state stability".

### 1.3 References

1. Elbert, D.L. and Hubbell, J.A., Surface treatments of polymers for biocompatibility. Ann. Rev. Mater. Sci., 26 (1996) 365-394.
2. Klein, E., Affinity membranes: a 10-year review. J. Membr. Sci., 179 (2000) 1-27.
3. Moses, J.W., Kipshidze, N. and Leon, M.B., Perspectives of drug-eluting stents: the next revolution. J. Cardiovascular Drugs, 2 (2002) 163-172.
4. Shen, M.C., Wagner, M.S., Castner, D.G., Ratner, B.D. and Horbett, T.A., Multivariate surface analysis of plasma-deposited tetraglyme for reduction of protein adsorption and monocyte adhesion. Langmuir, 19 (2003) 1692-1699.
5. Tang, L.P. and Eaton, J.W., Fibrin(ogen) mediates acute inflammatory responses to biomaterials. J. Exp. Med., 178 (1993) 2147-2156.
6. Balasubramanian, V., Grusin, N.K., Bucher, R.W., Turitto, V.T. and Slack, S.M., Residence-time dependent changes in fibrinogen adsorbed to polymeric biomaterials. J. Biomed. Mater. Res., 44 (1999) 253-260.
7. Lin, W.C., Liu, T.Y. and Yang, M.C., Hemocompatibility of polyacrylonitrile dialysis membrane immobilized with chitosan and heparin conjugate. Biomaterials, 25 (2004) 1947-1957.
8. Bruinsma, G.M., van der Mei, H.C. and Busscher, H.J., Bacterial adhesion to surface hydrophilic and hydrophobic contact lenses. Biomaterials, 22 (2001) 3217-3224.
9. Cheng, L., Muller, S.J. and Radke, C.J., Wettability of silicone-hydrogel contact lenses in the presence of tear-film components. Current Eye Research, 28 (2004) 93108.
10. Steinberg, D., Klinger, A., Kohavi, D. and Sela, M.N., Adsorption of Human Salivary Proteins to Titanium Powder .l. Adsorption of Human Salivary Albumin. Biomaterials, 16 (1995) 1339-1343.
11. Hayakawa, T., Yoshinari, M. and Nemoto, K., Direct attachment of fibronectin to tresyl chloride-activated titanium. Journal of Biomedical Materials Research Part A, 67A (2003) 684-688.
12. Hudson, P.J., Recombinant antibodies: a novel approach to cancer diagnosis and therapy. Expert Opin. Invest. Drugs, 9 (2000) 1231-1242.
13. Opalinska, J.B. and Gewirtz, A.M., Nucleic-acid therapeutics: Basic principles and recent applications. Nat. Rev. Drug Discovery, 1 (2002) 503-514.
14. Souriau, C. and Hudson, P.J., Recombinant antibodies for cancer diagnosis and therapy. Expert Opin. Biol. Ther., 3 (2003) 305-318.
15. Cohen-Stuart, M.A., Cosgrove, T. and Vincent, B., Experimental aspects of polymer adsorption at solid-solution interfaces. Adv. Colloid Interface Sci., 24 (1986) 143239.
16. Kawaguchi, M. and Takahashi, A., Polymer adsorption at solid liquid interfaces. Adv. Colloid Interface Sci., 37 (1992) 219-317.
17. Silberberg, A., J. Phys. Chem., 66 (1962) 1872.
18. Scheutjens, J.M.H.M. and Fleer, G.J., Statistical theory of the adsorption of interaction chain molecules. 1. Partition function, segment density distribution, and adsorption isotherms. J. Phys. Chem., 83 (1979) 1619-1635.
19. Scheutjens, J.M.H.M. and Fleer, G.J., Statistical theory of the adsorption of interacting chain molecules. 2. Train, loop, and tail size distribution. J. Phys. Chem., 84 (1980) 178-190.
20. Wang, Y.M. and Mattice, W.L., Adsorption of homopolymers on a solid-surface - a comparison between Monte-Carlo simulation and the Scheutjens-Fleer mean-field lattice theory. Langmuir, 10 (1994) 2281-2288.
21. Zajac, R. and Chakrabarti, A., Statics and dynamics of homopolymer adsorption and desorption: A Monte Carlo study. J. Chem. Phys., 104 (1996) 2418-2437.
22. Liu, H. and Chakrabarti, A., Molecular dynamics study of adsorption and spreading of a polymer chain onto a flat surface. Polymer, 40 (1999) 7285-7293.
23. Binder, K., Milchev, A. and Baschnagel, J., Simulation studies on the dynamics of polymers at interfaces. Ann. Rev. Mater. Sci., 26 (1996) 107-134.
24. Clark, A.J., Whitehead, L.A., Haynes, C.A. and Kotlicki, A., Novel resonantfrequency sensor to detect the kinetics of protein adsorption. Rev. Sci. Instrum., 73 (2002) 4339-4346.
25. Bentaleb, A., Abele, A., Haikel, Y., Schaaf, P. and Voegel, J.C., FTIR-ATR and radiolabeling study of the adsorption of ribonuclease A onto hydrophilic surfaces: Correlation between the exchange rate and he interfacial denaturation. Langmuir, 14 (1998) 6493-6500.
26. Giacomelli, C.E., Esplandiu, M.J., Ortiz, P.I., Avena, M.J. and De Pauli, C.P., Ellipsometric study of bovine serum albumin adsorbed onto $\mathrm{Ti} / \mathrm{TiO}_{2}$ electrodes. $J$. Colloid Interface Sci., 218 (1999) 404-411.
27. Cohen-Stuart, M.A., Scheutjens; J.M.H.M. and Fleer, G.J., Polydispersity effects and the interpretation of polymer adsorption-isotherms. J. Polym. Sci., Part B: Polym. Phys., 18 (1980) 559-573.
28. Cohen-Stuart, M.A., Fleer, G.J. and Bijsterbosch, B.H., The adsorption of polyvinyl pyrrolidone onto silica. 1. Adsorbed amount. J. Colloid Interface Sci., 90 (1982) 310-320.
29. Anderson, R.E., Pande, V.S، and Radke, C.J., Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. J. Chem. Phys., 112 (2000) 9167-85.
30. Fersht, A.R., On the simulation of protein folding by short time scale molecular dynamics and distributed computing. PNAS, 99 (2002) 14122-14125.
31. Socci, N.D. and Onuchic, J.N., Folding kinetics of proteinlike heteropolymers. J. Chem. Phys., 101 (1994) 1519-1528.
32. Lau, K.F. and Dill, K.A., A lattice statistical mechanics model of the conformational and sequence spaces of proteins. Macromolecules, 22 (1989) 3986-3997.
33. Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D. and Chan, H.S., Principles of protein folding - A perspective from simple exact models. Protein Sci., 4 (1995) 561-602.
34. Camacho, C.J. and Thirumalai, D., Kinetics and thermodynamics of folding in model proteins. PNAS, 90 (1993) 6369-6372.
35. Gupta, P. and Hall, C.K., Effect of solvent conditions upon refolding of pathways and intermediates for a simple lattice protein. Biopolymers, 42 (1997) 399-409.
36. Creighton, T.E., Proteins, structures and molecular properties (2nd Edition). W.H. Freeman, New York, N.Y., 1993.
37. Privalov, P.L., Thermodynamic problems of protein-structure. Annu. Rev. Biophys. Biophys. Chem., 18 (1989) 47-69.
38. Privalov, P.L. and Gill, S.J., Stability of protein-structure and hydrophobic interaction. Adv. Protein Chem., 39 (1988) 191-234.
39. Dill, K.A., Dominant forces in protein folding. Biochemistry, 29 (1990) 7133-7155.
40. Haynes, C.A. and Norde, W., Globular protein at solid/liquid interfaces. Colloids Surfaces B, 2 (1994) 517-566.
41. Burgess, J., Ions in solution. Halsted Press, New York, NY, 1988.
42. Levitt, M., Conformational preferences of amino-acids in globular pProteins. Biochemistry, 17 (1978) 4277-4284.
43. Liu, S.M., The thermodynamics of irreversible nonspecific adsorption at a solidliquid interface. University of British Columbia, Vancouver, BC, 1997.
44. Tilton, R.D., Gast, A.P. and Robertson, C.R., Surface-diffusion of interacting proteins - Effect of concentration on the lateral mobility of adsorbed bovine serumalbumin. Biophys. J., 58 (1990) 1321-1326.
45. Hunter, J.R., Carbonell, R.G. and Kilpatrick, P.K., Coadsorption and exchange of lysozyme beta-casein mixtures at the air-water-interface. J. Colloid Interface Sci., 143 (1991) 37-53.
46. Lee, W.K., McGuire, J. and Bothwell, M.K., Competitive adsorption of bacteriophage T4 lysozyme stability variants at hydrophilic glass surfaces. J. Colloid Interface Sci., 269 (2004) 251-254.
47. Haynes, C.A. and Norde, W., Structures and stabilities of adsorbed proteins. J. Colloid Interface Sci., 169 (1995) 313-328.
48. Omanovic, S. and Roscoe, S.G., Interfacial behaviour of beta-lactoglobulin at a stainless steel surface: An electrochemical impedance spectroscopy study. J. Colloid Interface Sci., 227 (2000) 452-460.
49. Norde, W. and Lyklema, J., Why proteins prefer interfaces. J. Biomater. Sci. Polymer Edn., 2 (1991) 183-202.
50. Buijs, J. and Hlady, V., Adsorption kinetics, conformation, and mobility of the growth hormone and lysozyme on solid surfaces, studied with TIRF. J. Colloid Interface Sci., 190 (1997) 171-181.
51. Chen, X., Davies, M.C., Roberts, C.J., Tendler, S.J.B., Williams, P.M., Davies, J., Dawkes, A.C. and Edwards, J.C., Recognition of protein adsorption onto polymer surfaces by scanning force microscopy and probe-surface adhesion measurements with protein-coated probes. Langmuir, 13 (1997) 4106-4111.
52. Kim, D.T., Blanch, H.W. and Radke, C.J., Direct imaging of lysozyme adsorption onto mica by atomic force microscopy. Langmuir, 18 (2002) 5841-5850.
53. Haynes, C.A., Sliwinsky, E. and Norde, W., Structural and electrostatic properties of globular proteins at a polystyrene-water interface. J. Colloid Interface Sci., 164 (1994) 394-409.
54. Galisteo, F. and Norde, W., Adsorption of lysoyzme and alpha-lactalbumin on poly(styrenesulphonate) latices 2. Proton titrations. Colloid. Surface. B, 4 (1995) 389-400.
55. Welzel, P.B., Investigation of adsorption-induced structural changes of proteins at solid/liquid interfaces by differential scanning calorimetry. Thermochim. Acta, $\mathbf{3 8 2}$ (2002) 175-188.
56. Malmsten, M. and Lassen, B.,Ellipsometry studies of protein adsorption at hydrophobic surfaces. In "Proteins at Interfaces II: Fundamentals and Applications" (T.A. Horbett and J.L. Brash, Eds.), p.228-238. American Chemical Society, Washington, DC, 1995.
57. Froberg, J.C., Arnebrant, T., McGuire, J. and Claesson, P.M., Effect of structural stability on the characteristics of adsorbed layers of T4 lysozyme. Langmuir, 14 (1998) 456-62.
58. Kondo, A., Urabe, T. and Yoshinaga, K., Adsorption activity and conformation of alpha-amylase on various ultrafine silica particles modified with polymer silane coupling agents. Colloids Surfaces A, 109 (1996) 129-36.
59. Krisdhasima, V., McGuire, J. and Sproull, R., Surface hydrophobic influences on $\beta$ lactoglobulin adsorption kinetics. J. Colloid Interface Sci., 154 (1992) 337-350.
60. Zoungrana, T. and Norde, W., Thermal stability and enzymatic activity of alphachymotrypsin adsorbed on polystyrene surfaces. Colloids Surfaces B, 9 (1997) 157167.
61. Vermonden, Giacomelli and Norde, W., Reversibility of structural rearrangements in bovine serum albumin during homomolecular exchange from AgI particles. Langmuir, 17 (2001) 3734-3740.
62. Billsten, P., Wahlgren, M., Arnebrant, T., McGuire, J. and Elwing, H., Structural changes of T4 lysozyme upon adsorption to silica nanoparticles measured by circular dichroism. J. Colloid Interface Sci., 175 (1995) 77-82.
63. Cabani, S., Gianni, P., Mollica, V. and Lepori, L., Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous-solution. J. Solution Chem., 10 (1981) 563-595.
64. Norde, W., Adsorption of proteins from solution at the solid-liquid interface. $A d v$. Colloid Interface Sci., 25 (1986) 267-340.
65. Kull, T., Nylander, T., Tiberg, F. and Wahlgren, N.M., Effect of surface properties and added electrolyte on the structure of $\beta$-casein layers adsorbed at the solid/aqueous interface. Langmuir, 13 (1997) 5141-5147.
66. Malmsten, M., Formation of adsorbed protein layers. J. Colloid Interface Sci., 207 (1998) 186-99.
67. Baszkin, A. and Lyman, D.J., The interaction of plasma-proteins with polymers .1. Relationship between polymer surface-energy and protein adsorption-desorption. J. Biomed. Mater. Res., 14 (1980) 393-403.
68. Poncet, C., Tiberg, F. and Audebert, R., Ellipsometric study of the adsorption of hydrophobically modified polyacrylates at hydrophobic surfaces. Langmuir, 14 (1998) 1697-1704.
69. Suzawa, T. and Shirahama, H., Adsorption of plasma-proteins onto polymer lattices. Adv. Colloid Interf. Sci., 35 (1991) 139-172.
70. Roth, C.M. and Lenhoff, A.M., Electrostatic and van der Waals contributions to protein adsorption: Comparison of theory and experiment. Langmuir, 11 (1995) 3500-3509.
71. Roth, C.M., Neal, B.L. and Lenhoff, A.M., Van der Waals interactions involving proteins. Biophys. J., 70 (1996) 977-987.
72. Wang, J. and McGuire, J., Surface tension kinetics of the wild type and four synthetic stability mutants of T4 phage lysozyme at the air-water interface. J. Colloid Interface Sci., 185 (1997) 317-23.
73. Wertz, C.F. and Santore, M.M., Effect of surface hydrophobicity on adsorption and relaxation kinetics of albumin and fibrinogen: Single-species and competitive behaviour. Langmuir, 17 (2001) 3006-3016.
74. Lundstrom, I. and Elwing, H., Simple kinetic-models for protein exchange-reactions on solid-surfaces. J. Colloid Interface Sci., 136 (1990) 68-84.
75. Xu, S.Q. and Damodaran, S., Kinetics of adsorption of proteins at the air-waterinterface from a binary mixture. Langmuir, 10 (1994) 472-480.
76. Elwing, H., Protein adsorption and ellipsometry in biomaterial research. Biomaterials, 19 (1998) 397-406.
77. Suttiprasit, P., Krisdhasima, V. and McGuire, J., The surface-activity of alphalactalbumin, beta-lactoglobulin, and bovine serum-albumin .1. surface-tension measurements with single-component and mixed-solutions. J. Colloid Interface Sci., 154 (1992) 316-326.
78. Razumas, V., Nylander, T. and MacRitchie, F., Surface pressure study of hemin, microperoxidase-8, -11 , and cytochrome c adsorption at the air-water interface. $J$. Colloid Interface Sci., 178 (1996) 303-308.
79. Ybert, C. and di Meglio, J.M., Study of protein adsorption by dynamic surface tension measurements: Diffusive regime. Langmuir, 14 (1998) 471-475.
80. Juffer, A.H., Argos, P. and DeVlieg, J., Adsorption of proteins onto charged surfaces: A Monte Carlo approach with explicit ions. J. Comput. Chem., 17 (1996) 1783-1803.
81. Lu, D.R. and Park, K., J. Biomater. Sci. Polymer Edn., 1 (1990) 243-.
82. Noinville, V., Vidalmadjar, C. and Sebille, B., Modeling of protein adsorption on polymer surfaces - Computation of adsorption potential. J. Phys. Chem., 99 (1995) 1516-1522.
83. Lu, D.R., Lee, S.J. and Park, K., Calculation of solvation interaction energies for protein adsorption on polymer surfaces. J. Biomater. Sci. Polymer Edn., 3 (1991) 127-147.
84. Latour, R.A.J. and Rini, C.J., Theoretical analysis of adsorption thermodynamics for hydrophobic peptide residues on SAM surfaces of varying functionality. (2002) 564 77.
85. Almalah, K., McGuire, J. and Sproull, R., A macroscopic model for the singlecomponent protein adsorption-isotherm. J. Colloid Interface Sci., 170 (1995) 261268.
86. Johnson, R.D., Wang, Z.G. and Arnold, F.H., Surface site heterogeneity and lateral interactions in multipoint protein adsorption. Journal of Physical Chemistry, 100 (1996) 5134-5139.
87. Perrin, A., Elaissari, A., Theretz, A. and Chapot, A., Atomic force microscopy as a quantitative technique: correlation between network model approach and experimental study. Colloids and Surfaces B-Biointerfaces, 11 (1998) 103-112.
88. Arai, T. and Norde, W., The behavior of some model proteins at solid-liquid interfaces 1. Adsorption from single protein solutions. Colloid. Surface., 51 (1990) 1-15.
89. Martensson, J., Arwin, H., Lundstrom, I. and Ericson, T., Adsorption of lactoperoxidase on hydrophilic and hydrophobic silicon dioxide surfaces - an ellipsometric study. J. Colloid Interface Sci., 155 (1993) 30-36.
90. Tarjus, G., Schaaf, P. and Talbot, J., Generalized random sequential adsorption. J. Chem. Phys., 93 (1990) 8352-8360.
91. Van Tassel, P.R., Viot, P., Tarjus, G. and Talbot, J., Irreversible adsorption of macromolecules at a liquid-solid interface - Theoretical-studies of the effects of conformational change. J. Chem. Phys., 101 (1994) 7064-7073.
92. Van Tassel, P.R., Viot, P. and Tarjus, G., A kinetic model of partially reversible protein adsorption. J. Chem. Phys., 106 (1997) 761-770.
93. Van Tassel, P.R., Guermouri, L., Ramsden, J.J., Tarjus, G., Viot, P. and Talbot, J., A particle-level model of irreversible protein adsorption with a postadsorption transition. J. Colloid Interface Sci., 207 (1998) 317-23.
94. Adamczyk, Z., Zembala, M., Siwek, B. and Warszynski, P., Structure and ordering in localized adsorption of particles. J. Colloid Interface Sci., 140 (1990) 123-137.
95. Oberholzer, M.R., Stankovich, J.M., Carnie, S.L., Chan, D.Y.C. and Lenhoff, A.M., 2-D and 3-D interactions in random sequential adsorption of charged particles. $J$. Colloid Interface Sci., 194 (1997) 138-153.
96. Yuan, Y., Oberholzer, M.R. and Lenhoff, A.M., Size does matter: electrostatically determined surface coverage trends in protein and colloid adsorption. Colloid. Surface. A, 165 (2000) 125-141,
97. Ruckenstein, E. and Prieve, D.C., Adsorption and desorption of particles and their chromatographic-separation. AIChE J., 22 (1976) 276-283.
98. Prieve, D.C. and Ruckenstein, E.,In "Colloid and Interface Science" (M. Kerker, Eds.), p.73. Academic Press, New York, NY, 1976.
99. Stahlberg, J., Jonsson, B. and Horvath, C., Combined effect of coulombic and van der Waals interactions in the chromatography of proteins. Anal. Chem., 64 (1992) 3118-3124.
100. Roth, C.M. and Lenhoff, A.M., Electrostatic and van der Waals contributions to protein adsorption: Computation of equilibrium constants. Langmuir, 9 (1993) 962972.
101. Asthagiri, D. and Lenhoff, A.M., Influence of structural details in modeling electrostatically driven protein adsorption. Langmuir, 13 (1997) 6761-6768.
102. Fang, F. and Szleifer, I., Kinetics and thermodynamics of protein adsorption: A generalized molecular theoretical approach. Biophys. J., 80 (2001) 2568-2589.

### 1.4 Tables

Table 1.1: Interactions governing the native-state structural stability of globular proteins. $\Delta_{N-D} G$ refers to the Gibbs energy of denaturation for the protein.

| Type of Interaction | Contribution to <br> $\Delta_{N-D} G$ | Comments |
| :--- | :--- | :--- |
| Hydrophobic <br> dehydration | $\ll 0$ | An increase in entropy results from the <br> release of water molecules contacting <br> hydrophobic residues. |
| Hydrogen bonding | $<0(?)$ | Intramolecular hydrogen bonding, <br> especially in ordered secondary structures, <br> may contribute to stability. |
| Dehydration of polar <br> groups | $<0$ | Dehydration of polar groups may <br> contribute up to $5 \mathrm{~kJ} /($ mol amino acid) |
| Electrostatic forces | $>$ or $<0$ | Contribution is dependent on the pH of the <br> system relative to the pI of the protein. |
| Dispersion forces | $\leq 0$ | Favorable due to the dense packing of the <br> atoms in a protein structure. |
| Conformational <br> entropy | $\gg 0$ | A substantial loss of conformational <br> freedom from folding and the formation of <br> highly ordered secondary structures. |
| Distortion of covalent <br> bond lengths and <br> bond angles | $>0$ | Unfavorable strains existing to |
| accommodate other, more dominant |  |  |
| interactions. |  |  |

### 1.5 Figures

Figure 1.1: The basic chemical structure of an amino acid and schematic diagrams of the 20 naturally-occuring amino acids.


[Ala]



Methionyl
[Met]


Threonyl
[Thr]

[Arg]


Glutaminyl [GIn]




Tryptophanyl
[Trp]



Glutamyl [Glu]

[Asn]
[Asp]

Glycyl [Gly]




Tyrosyl
[Tyr]




Valyl
[Val]

Figure 1.2: The condensation reaction forming the peptide link between two amino acids.


Figure 1.3: Isotherm for the adsorption of hen egg-white lysozyme ( $\mathrm{pI} \sim 11$ ) to particulate silica at pH 7 and $37^{\circ} \mathrm{C}$. Shown is the concentration of protein on the surface as a function of protein in bulk solution. Data points collected from two separate trials are indicated. The solid line represents the ascending isotherm, while dotted lines represent descending isotherms. The vertical dashed line extended from zero indicates that to the lowest possible detectable levels, no desorbed protein was detected. This data is the author's own from reference (43).


## 2 Mesoscopic analysis of conformational and entropic contributions to nonspecific adsorption of HP copolymer chains using dynamic Monte Carlo simulations*

### 2.1 Introduction

The success of invasive and extracorporeal medical devices is typically limited by the incompatibility of the materials of construction of the device with the tissue and blood with which they come into contact. Despite improvements over the past several decades in the physical and chemical properties of artificial biomaterials, inflammatory reactions against the foreign substances of the devices have not been eliminated. In critical therapeutic interventions such as major surgery, the activation of the immune and inflammatory system within a patient, as well as the coagulation process, contribute to a slower recovery and increased susceptibility to infections and post-operative complications. Moreover, when the use of a device is required in long-term care (e.g. implantable devices and hemodialysis), the body's chronic response against the foreign materials of construction requires lifelong medication.

[^0]Regulating the behaviour of cells and tissues at a biomaterial interface requires strict control over the surface properties of the material and an ability to impart to the material a defined biological response. Among the greatest challenges for meeting these requirements are controlling protein adsorption, retaining protein activity following adsorption, and tailoring protein distributions on the artificial surface to elicit a desired cellular response. The hydrophobicity, charge, and chemical makeup of the surface and the contacting protein have all been shown to impact the energetics and kinetics of the protein adsorption process, and can affect both the stability and orientation of a protein at a surface (1-3). Most protein-material interactions therefore result in decreased protein activity.

Improving the biocompatibility of synthetic materials for traditional device-based therapies will require a better fundamental understanding of the kinetics and energetics of protein adsorption. Considerable attention has therefore been given to understanding how and why proteins adsorb at interfaces (1,3). Regrettably, the problem has proven recalcitrant, due in part to the inherent complexity of the process and to the lack of experimental methods capable of accurately measuring molecular contributions to overall adsorption energetics and of visualizing the protein adsorption process at the molecular level.

Numerous experimental studies have shown that electrostatics, dehydration forces, interactions between neighboring adsorbed proteins, and protein conformation all contribute to the adsorption process (1,2). Protein and sorbent surface geometry,
including surface roughness, are also thought to be important. Globular proteins appear to prefer surfaces of greater hydrophobicity and greater roughness (3). For example, Kondo et al. (4) showed that for identical total protein loads, the surface density of adsorbed $\alpha$-amylase was lowest on silica particles, the least hydrophobic (most polar) surface tested, and increased with increasing hydrophobicity of the sorbent surface, indicating a strong correlation between the affinity of the protein for the surface and sorbent surface polarity. A dependence of adsorption rates on the sorbent surface hydrophobicity has also been shown, with increasing sorbent hydrophobicity leading to an increase in the forward rate constant for adsorption and a decrease in the off-rate (5). Kull et al. (6), for example, showed that $\beta$-casein adsorbs quickly onto hydrophobic silica, while the adsorption kinetics appear to be much slower on hydrophilic silica under similar conditions.

Sorbent hydrophobicity has also been shown to influence the extent of differences between a protein's adsorbed and native-state conformations. Circular dichroism (7, 8), ellipsometry (9), NMR (10), TIRF $(11,12)$, and calorimetry $(7,13)$ data have all shown evidence of significant conformational changes in proteins adsorbed at hydrophobic surfaces. Activity assays, an indirect measurement of the extent of protein structural change, have also shown that the tendency for a protein to change conformation during adsorption can be correlated to sorbent surface hydrophobicity $(4,14)$

The extent of conformational change in proteins upon adsorption is influenced by other factors, most notably by protein stability (1). McGuire and coworkers (15-17) conducted
a number of studies using wild-type T4 bacteriophage lysozyme and a set of mutants of varying thermal stability. They concluded that T4 variants of greater thermodynamic stability retain greater amounts of both secondary and tertiary structure upon adsorption, and that the less stable mutants adhere more strongly to the surface. The rate of adsorption was also affected, with less stable proteins adsorbing more quickly. Similar conclusions were made by Billsten et al. $(18,19)$, who studied adsorption of variants of human carbonic anhydrase II on silica particles. Based on analysis of differential scanning calorimetry, circular dichroism, and fluorescence spectroscopy data, they concluded that adsorption of the more-stable pseudo-wild-type protein resulted in the smallest perturbation in protein conformation relative to the native state.

Efforts have been made to model the thermodynamics and kinetics of globular protein adsorption $(5,20)$ by assuming that the protein first adsorbs to the surface in its native state, then undergoes structural rearrangements. Van Tassel et al. (21), for instance, represented proteins as disk-shaped particles that symmetrically spread on the surface once adsorbed. While validation of these rather coarse models is generally restricted to successful correlation with adsorption isotherm and binding kinetics data, their predictions suggest a relation between protein conformational change and adsorbed protein concentration that has been observed experimentally. Thus, despite their rather simple nature, the models indicate that an accurate description of protein conformational changes during adsorption is essential to understanding the thermodynamics and kinetics of protein adsorption.

More detailed descriptions of conformational changes in a protein during adsorption have been elucidated from thermodynamic studies. Haynes and Norde utilized a combination of calorimetry, proton titrations and adsorption isotherm measurements to quantify the contribution of changes in protein structure to the thermodynamics of the adsorption process (1, 13). A number of general conclusions can be drawn from these investigations. First, the degree of protein structural change during adsorption is often significant, and consequently can have a large effect on the overall energy change accompanying the adsorption process. For example, in a study of the adsorption of $\alpha$ lactalbumin to negatively-charge polystyrene surfaces at pH 10 , protein structural changes were estimated to make the dominant contribution to the overall enthalpy of adsorption (13).

Conformational changes in the protein must also result in a change in entropy, but data defining the magnitude and sign of entropy changes accompanying protein adsorption are limited and generally indirect. In their seminal early work, Norde and Lyklema (22, 23) argued that protein secondary structure losses upon adsorption led to a net increase in the rotational mobility of the backbone and side chains of the protein. As a consequence, the conformational entropy of the protein was predicted to increase. They argued that the resulting entropy gain would be large enough to compensate for the unfavorable decrease in enthalpy associated with adsorption. Losses in protein secondary structure upon adsorption have been observed. In particular, circular dichroism has been used in several instances to measure decreases in $\alpha$-helix content during nonspecific protein adsorption to relatively hydrophilic sorbents such as silica $(7,17)$. While these studies provide no
direct evidence that losses in protein secondary structure lead to an increase in the conformational entropy of the adsorbed protein relative to its native state structure in solution, the overall entropy change for the adsorption process is often favorable. In contrast, Giacomelli et al. (24) and Sane et al. (25) and others have recently shown that changes in protein conformation upon adsorption at hydrophobic surfaces often involves an increase in ordered structure, and therefore a decrease in conformational entropy. This is consistent with recent modeling results by Ben-Tal et al. (26) which predict a net loss in chain entropy of approximately $1.7 k T$ for a short peptide, pentalysine, during adsorption to a hydrophobic lipid membrane.

The aim of this paper is to examine more closely the dependence of the energy and entropy of adsorption on associated changes in the conformation of a chain molecule that folds into a unique compact lowest-energy state, such as the native state of a globular protein. Regrettably, experimental methods capable of directly quantifying changes in chain conformation and entropy and then relating them to system energy and entropy are not available. However, such connections can be made within much simpler model systems, such as the well-known HP chain model of Dill, that involve linear copolymer chains that fold in solution into a unique lowest-energy conformation and therefore share with single-domain globular proteins important conformational and energetic properties. Here, we ask what can be learned from mesoscopic dynamic Monte Carlo simulations of the adsorption of protein-like HP model chains about the dependence of adsorbed-state chain conformations and entropy on both the stability of the lowest-energy ("native state") conformation of the chain in solution and the topography and hydrophobicity of
the sorbent surface? Throughout the paper, connections are made with general experimental observations of protein adsorption processes to identify where results from our simple model are in qualitative agreement and may therefore provide insight into the contributions of chain energy and entropy to the overall thermodynamics of protein adsorption.

The complexity of the folding behaviour of proteins, combined with limitations in computing capabilities, set limits on the class of protein simulation problems that can currently be solved at the atomistic level. In general, molecular dynamics simulations of protein folding can cover only a portion (e.g., microseconds) of the estimated 10 to 100 milliseconds required for a small protein to fully fold (27). Such an approach is therefore not applicable to the study of protein adsorption, which is a process that typically takes seconds to hours to reach an energetic steady state (28). Thus, we feel that simplified mesoscopic chain-on-a-lattice models, such as those developed by Dill (29) and Socci and Onuchic (30), remain the most practical way of investigating at a fundamental level the dependence of chain conformational entropy on system properties.

### 2.2 Theory

### 2.2.1 Dynamic Monte Carlo simulations of HP-model polymers

The HP model, first proposed by Dill (29) and used extensively by others (see, for example, 31,32 ), assumes that the major contribution to the free energy of the native conformation of a protein is due to interactions between hydrophobic $(\mathrm{H})$ amino acids. A hydrophobic core in the folded structure tends to form that is shielded from the
surrounding solvent by polar ( P ) amino acids in the polypeptide sequence. In the HP model, the amino acid sequence of a peptide-like chain is represented mesoscopically as a binary string of H and P monomers. The HP model therefore ignores the fact that some amino acids cannot be unambiguously classified as being either hydrophobic or hydrophilic. Despite the simplicity of the HP model, folding processes for HP-model chains simulated by the dynamic Monte Carlo (dMC) method or the molecular dynamics method appear to show important similarities with the protein folding process (33), allowing a number of research groups to successfully evaluate new hypotheses for protein structure formation and driving forces for protein folding (30, 34, 35). For example, the HP model has been successfully used to understand better the formation of compact secondary structures and hydrophobic cores in proteins (36). It has also been used to specify energy pathways and intermediates of protein folding which in turn have led to a probable explanation for the Levinthal paradox (37). The success of the model stems in part from the fact that the simple HP representation of the peptide makes it possible to enumerate and consider all possible chain conformations and associated system energetics. Given their ability to document changes in chain conformation resulting from perturbations to system environment, dMC simulations of HP-model chains may also provide a useful means of evaluating putative models for protein adsorption, including those that postulate that conformational changes in the peptide during adsorption result in an increase in the entropy of the chain that drives the adsorption process and its apparent irreversibility.

Recently, dMC simulations have been used to describe adsorption of mesoscopic proteinlike chains at a liquid-liquid interface. Anderson et al. (38) simulated the adsorption behaviour at an oil-water interface of a uniquely folding linear copolymer in which intraand intermolecular interactions between chain segments were defined by the interaction energies of Miyazawa and Jernigan (39). Solvent-chain interactions were considered, as well as the mixing energies of the oil and water phases. They observed that due to the unfavorable energetics between the oil and water, the chain is overwhelmingly ( $>99 \%$ ) likely to adhere to the interface in an unfolded state. The transition to this denatured state requires overcoming significant energy barriers, but once attained, is apparently irreversible. A key result from this important study is that the conformational entropy of the denatured adsorbed chain is greater than that of the native-state protein in solution.

Zhdanov and Kasemo $(40,41)$ used dMC simulations to investigate the relative rates of denaturation of HP chains in the presence of a solid-liquid interface. Their results showed that at relatively high temperatures, unfolding of the chains follows an apparent first-order rate equation, similar to what was observed without a surface present. At lower temperatures however, the denaturation pathway at the surface differed in that metastable states were formed. More recently, Castells et al. (42) used a dMC approach to demonstrate that mesoscopic protein-like chains attach to a surface in an unfolded state, and that the degree of unfolding is dependent on the degree of attraction of the residues to the surface.

Here we apply dMC simulations to Dill's HP model in two dimensions to further investigate conformations of protein-like chains adsorbed at a solid-liquid interface and the contribution of changes in chain conformation to adsorption energetics and system entropy. Results from these model calculations are then interpreted in the context of real protein adsorption systems and previous experimental results of Norde et al. $(13,43)$ and others $(44,45)$ that suggest, albeit indirectly, that nonspecific protein adsorption to hydrophilic surfaces is often driven by an increase in the conformational entropy of the polypeptide.

### 2.2.2 The model system

The 2D protein-like HP chains we used are taken from the work of Dill et al. $(29,33)$ and contain $L$ residues, connected through $L-1$ vectors, all one lattice unit $a$ in length. The chain is placed on a Cartesian coordinate grid in a self-avoiding configuration so that at any instance, no more than one residue occupies a given lattice site. Each residue is therefore uniquely positioned on a coordinate point (lattice site) of the grid (i.e. the first residue lies on coordinates $\left(x_{1}, y_{1}\right)$, the second on $\left(x_{2}, y_{1}\right)$, etc. ) and all connecting vectors run parallel to either the $x$ or $y$-axis. Spaces unfilled by the chain are assumed to be solvent units. The grid size is defined in each simulation as grid height $=$ grid width $=$ 10x the fully stretched chain length, thereby creating a lattice of sufficient size to allow the HP chain to freely sample all conformational space. For adsorption studies, a potential is applied at selected walls to represent a model surface, while in all other cases, walls remain athermal with respect to interactions with the remaining components of the system. The model chain is initially placed away from the adsorbing interface and
allowed to undergo Verdier-Stockmeyer moves as shown in Figure 2.1. The model is ergodic, as the set of three allowable Verdier-Stockmeyer moves (crankshaft, flip, and turn) enable the chain to adopt all possible conformations within the lattice.

Our dMC simulations are based on the Metropolis algorithm where chain movements are allowed or disallowed depending on the change in energy of the system (46). The total energy of the lattice of N sites is defined as the sum of the interaction energies between all contacting elements of the system, excluding interactions between directly connecting chain residues. In all simulations reported here, four components are considered: hydrophobic $(\mathrm{H})$ and polar $(\mathrm{P})$ residues of the chain, the solvent $(\mathrm{S})$, and the active wall (W). Interaction energies between each of the components and residues are defined by an associated set of Flory energy parameters, $\chi_{i j}=\omega_{i j} / k T$, where $i$ and $j$ index the four components/residues in the system. The overall lattice (system) energy can therefore be calculated as follows for each sequence in a particular conformation and position in the grid,
$E=\sum_{i}^{N-1} \sum_{j>i}^{N} \omega_{i j} \delta\left(r_{i}-r_{j}\right)-\sum_{l}^{L-1} \omega_{l, l+1}$
The first term in Eq. [2.1] sums over all nearest-neighbor interactions in the lattice, where $r_{i}$ and $r_{j}$ are the coordinates of lattice sites $i$ and $j$, respectively, and $\delta$ is a function based on the positioning of the involved sites. For neighboring residues, $\delta(a)=1$, while for nonneighboring residues, $\delta\left(\left|r_{i}-r_{j}\right|>a\right)=0$. Interactions between connected chain residues are not considered. The second term in Eq. [2.1] therefore subtracts energies discounted due to connected residues within the chain of length $L$.

The dMC simulation algorithm used for both HP chain annealing studies in solution and all adsorption studies is based on randomly selecting a segment of the HP chain and evaluating all Verdier-Stockmeyer moves consistent with the position of the segment. Where multiple move options are allowed for a particular chosen segment, a single move is randomly selected and the new system energy is evaluated. If the new system energy is found to be equal or less than that of the original conformation, the move is accepted and a new cycle started. If the new conformation results in an increased system energy, then the difference in energy $\delta E$ is weighted using a Boltzmann relation,
$P=\exp (-\delta E / k T)$
and the calculated probability $P$ is compared to a random number, $n$, where $0 \leq n \leq 1$. Moves with a calculated probability higher than $n$ are accepted while moves of lower probability are rejected. Overall, the weighted method allows for both favorable and unfavorable moves to take place in the simulation. The frequency of acceptance of unfavorable moves, however, is significantly lower, but sufficient to allow the chain to escape local energy minima and sample all conformational space. The time coordinate of the simulations is presented in units of attempted moves.

### 2.2.3 Calculation of thermodynamic parameters

The on-lattice dMC simulations reported here represent a canonical ensemble. The energy of the lattice, $E$, is therefore equivalent to the internal energy, $U$, and the natural free energy of the system is the Helmholtz energy, $A$. The direct observable from each simulation is the energy histogram $p_{i}\left(E_{i}\right)$, generated through Monte Carlo sampling at
system temperature, $T$, from which one can calculate the ensemble-averaged internal energy of the system, $U$
$U \equiv<E>=\sum_{i} p_{i}\left(E_{i}\right) E_{i}$
where $p_{i}$ is the probability that the energy of the lattice is $E_{i}$.

The Helmholtz energy of the lattice, $A$, is calculated directly from the partition function, $Q$
$A \equiv<A>=-k T \ln Q$

In the simplest representation of the HP model, only contacts between adjacent hydrophobic segments are considered, with all other interactions assumed athermal. In this case, folding is solely driven by net favorable interactions between hydrophobic residues within the protein-like HP chain (i.e. $\omega_{H H} / k T<0$, with all other interaction energies set equal to zero). Thus, hydrophobic effects associated with a repulsive energy of interaction between water and other components are not considered here.

Each model sequence folds into a finite number of conformations, each having an energy value based on $h$, the number of formed hydrophobic contacts. For this simple model, the partition coefficient, $Q$, is given by
$Q=\sum_{h=0}^{h_{N}} \Omega(h) \exp \left(\frac{-h \omega_{H H}}{k T}\right)$
where $\Omega(h)$ is the density of states for which the number of hydrophobic contacts is $h$, $\omega_{H H}$ is the interaction energy between hydrophobic residues, $k$ is Boltzmann's constant, $T$
is the temperature, and $h_{N}$ is the number of hydrophobic contacts formed in the most stable conformation(s). Extension of Eq. [2.5] to include contributions from other segment $i$ - segment $j$ interactions is straightforward.

Finally, the system entropy, $S$, is given by the standard thermodynamic relation
$-S=\frac{A-U}{T}$
For the simple model described by Eq. [2.1], $S$ represents the chain entropy at infinite dilution.

### 2.3 Results

### 2.3.1 Dynamics and thermal unfolding behaviour of model sequences

Two HP model chains that fold into unique "native-state" conformations at low temperatures when driven by intramolecular interactions between hydrophobic residues (33) were selected for our simulation studies (Figure 2.2). The first model sequence, referred to here as sequence $I$, consists of 18 units and folds into a globular native-state (lowest energy) conformation defined by 9 pair-wise contacts between two hydrophobic residues not directly connected on the chain. It is predominantly hydrophobic, with an entirely hydrophobic interior and partially hydrophobic exterior. Sequence II is comprised of 20 units and folds into a globular native-state conformation having 8 hydrophobic contacts. In this lowest energy state, sequence II has an entirely polar exterior.

In the HP model, the temperature dependence of the ensemble-average chain conformation is captured by the Flory interaction parameter $\chi_{H H}\left(=\omega_{H H}(T) / k T\right)$ between hydrophobic residues of the chain. The conformational state of the chain at a given temperature can therefore be specified by the average number of contacts formed between these residues. A low number of hydrophobic contacts indicates an open "denatured" conformation, whereas a high number of contacts indicates a more compactly arranged chain, with the maximum hydrophobic contact number indicating the unique lowest energy conformation.

For model sequence I, Figure 2.3 shows dMC simulation results for the ensemble average number of intramolecular hydrophobic contacts formed as a function of the reduced temperature $T^{*}\left(=T / T_{m}\right)$. The simulation data for sequence I resemble a thermal denaturation curve for a single domain globular protein, showing a transition from a highly structured chain to a largely unstructured chain at a defined 'melting' temperature $T_{m}$. At very low temperatures (far below $T_{m}$ ), the protein-like chain forms an average number of hydrophobic HH contacts of 9 , indicating that the chain is essentially always found in its lowest energy native-state conformation when the temperature is sufficiently low (i.e., near 0 K ).

However, at temperatures $10^{\circ} \mathrm{C}$ to $80^{\circ} \mathrm{C}$ below $T_{m}$, corresponding to $\chi_{H H}$ values between -1 and -5 , the conformation of the chain in solution is not static. Conformational fluctuations of the folded chain in solution when $\chi_{H H}=-4$ are shown in Figure 2.4 in the form of a probability histogram for the system energy. The dynamic nature of the chain
conformation in solution is apparent, in that the chain is found in its unique lowestenergy conformation only a fraction of the time. The remainder of the time, the chain is in one of a relatively large number of conformations of slightly higher energy.

It is well known that protein denaturation results in a large increase in chain entropy that is compensated in part by an associated increase in the internal energy of the chain due to the loss of favorable intramolecular contacts between chain segments. These compensating effects are captured in the HP model. For sequence I, Table 2.1 lists the ensemble-average energy $\Delta_{N-D}<E>$ and entropy $T \Delta_{N-D}<S>$ of the denatured state relative to that of the native state at reduced temperatures $T^{*}$ of 0.44 and 0.71 assuming zero excess heat capacity for the denatured state. As expected, chain denaturation is energetically unfavorable (i.e. $\Delta_{N-D}<A>$ is greater than 0 ) at both temperatures due to an unfavorable increase in chain internal energy. Denaturation, however, is favored by the increase in chain entropy that would accompany the process at either temperature.

### 2.3.2 Chain adsorption to the surface

The presence of a sorbent surface was introduced into the simulations by setting a favorable interaction energy between a hydrophobic segment of the chain and an adjacent lattice segment of the simulation boundary (i.e., the wall). At the start of each simulation, the chain was positioned in close proximity to the wall to ensure that contact was made early in the simulation. In cases where adsorption to a planar surface was studied, the length of the active dimension of the grid was set to be greater than the fully-stretched length of the chain. Consequently, the chain could interact with the surface without
experiencing physical constraint from the opposite and two adjacent walls, which were all assigned to be energetically neutral. The approach of the chain to the wall was simulated by Monte Carlo moves and random diffusion, while movement along the wall following initial chain contact resulted solely from Monte Carlo moves.

Figure 2.5 shows the system energy over the initial $3 \times 10^{8}$ cycles of a dMC simulation of the adsorption of sequence I, initially in its folded lowest energy conformation in solution, to a planar hydrophobic surface. Relatively strong interaction energies are considered here by setting $\chi_{H H}=\chi_{H W}=-4$. Initially, the energy $E_{i}$ of the system is $\geq-$ $36 k T$, indicating that the chain is not in contact with the surface. After a relatively short number of cycles, the system energy decreases by $c a .-12 k T$ to a new energy minimum of $-48 k T$, consistent with favorable side-on adsorption of the native-state chain to the hydrophobic wall (sequence I in its lowest energy native-state conformation has a maximum of 3 exposed hydrophobic residues per solvent-exposed side). For the initial ca. $2.5 \times 107$ simulation cycles following chain adsorption, $E_{i} \geq-48 k T$ and the chain fluctuates among a large number of conformations with energy similar to that of the native-state conformation. Examples of these conformational states are shown in Figure 2.6.

The chain then adopts a new set of adsorbed conformations characterized by a system energy of $-52 k T$ and in a relatively small number of additional cycles begins to assume one or more conformations characterized by an $E_{i}$ equal to $-56 k T$, which represents the global energy minimum for the system. The probability of finding the system in this
lowest energy state remains high for the remainder of the simulation, but the dynamics of the system are such that the chain frequently adopts higher energy conformations, including conformations close to its native-state conformation on occasion.

Snapshots of conformations of the chain when the system is at its global energy minimum are shown in Figure 2.7. Unlike the chain in solution, which is characterized by a single lowest energy conformation, the adsorbed chain can assume at least 84 unique conformations on the surface to reach the global energy minimum of $-56 k T$. The chain conformations corresponding to this lowest energy state vary in dimensions and aspect ratio, but are generally flat, taking on the landscape view of a pancake or mound lying on a plate. Each such conformation differs noticeably from the chain in its native state.

A similar set of dMC simulations with $\chi_{H H}=\chi_{H W}=-4$ was completed for the adsorption of sequence II. As with sequence I, a severe perturbation of the chain conformation away from the native state is required to access the new ensemble-average system energy of $<E\rangle=c a .-40 k T$. Conformational degeneracy at the global energy minimum is again observed. In this case, 5 unique adsorbed chain conformations are observed at $-44 k T$, the global energy minimum of the system.

The total change in entropy requires solution of Eq. [2.6] for the adsorption process. It is, however, instructive to first evaluate the contribution to the overall change in chain entropy of the degeneracy in the conformation of the adsorbed chain at the global energy minimum for the system. This analysis, we believe, is analogous to popular models for
nonspecific protein adsorption which view the protein in solution as having a single native-state conformation, but capable of accessing multiple conformational states on the surface $(5,20,21,47)$. For sequence $I$, for example, the change in the entropy of the chain for the hypothetical case where it is restricted to its global energy minimum in the absence and presence of the sorbent surface is $+4.4 k T$, or an average of $+0.25 k T$ per chain residue. This per residue entropy change is similar to those estimated for nonspecific protein adsorption based on measurement of the total entropy change for adsorption and model estimates of the chain entropy. If it represented the true change in chain entropy, it would certainly provide a strong driving force for adsorption. However, as shown in the energy histograms provided in Figure 2.4, the HP chain in our model system accesses a range of energy states and a large number of associated conformations when the sorbent surface is both absent and present. Thus, as will be shown below, the actual change in chain entropy for the adsorption process can differ both in magnitude and in sign from the crude estimate provided above.

### 2.3.3 Thermodynamics of chain adsorption

For sequence I adsorbing to a planar sorbent, Table 2.2 reports changes in the ensembleaverage internal energy $\Delta_{a d s}<E>$, Helmholtz energy $\Delta_{a d s}<A>$, and chain entropy $\left.T \Delta_{a d s}<S\right\rangle$ under conditions where the sorbent surface is relatively hydrophilic $\left(\chi_{H W}=-1\right)$. When the stability of the native-state conformation of the chain is low $\left(\chi_{H H}=-1\right)$, a positive change in system entropy $T \Delta_{\text {ads }}<S>$ of $3.6 k T$ is found to make the dominant contribution to the overall Helmholtz energy change driving the adsorption process. In this case, intramolecular and intermolecular energies of interaction are symmetric and weak,
allowing the adsorbed chain to sample a very large number of conformations of similar energy on the sorbent surface.

An increase in chain entropy continues to make the dominant contribution to $\Delta_{a d s}<A>$ when the stability of the folded chain in solution is increased two-fold $\left(\chi_{H H}=-2\right)$. However, a further increase in the stability of the native-state conformation in solution results in an adsorption process characterized by either a relatively small increase ( $\chi_{H H}=$ -3 ) or a slight decrease $\left(\chi_{H H}=-4\right)$ in the chain entropy. For the $\chi_{H H}=-3$ system, $\Delta_{a d s}\langle E\rangle=c a .-3 k T$. Side-on adsorption of the chain in its native-state conformation would yield a $\Delta_{a d s}\langle E\rangle=-3 k T$, and indeed one of the observed lowest energy conformations of the adsorbed chain is the native state. However, as shown in Figure 2.8, five additional non-native conformations are observed at the global energy minimum. In each of these conformations, the adsorbed chain accesses the lowest energy state by reducing its total number of intramolecular contacts to form a larger number of new intermolecular contacts. To achieve a $\Delta_{a d s}<E>$ of $-3 k T$, 6 intermolecular HW contacts must form in place of a unit reduction in the total number of intramolecular HH contacts. This requirement severely limits the conformational space that is effectively available to the adsorbed chain. The adsorbed chain entropy is further limited by the impenetrable nature of the sorbent surface, which removes a degree of freedom for any chain segment in the lattice layer adjacent to it due to the fact that the segment is not free to step into the sorbent surface. In addition, like for the chain in solution, the interaction energy between the solvent and hydrophobic segments of the adsorbed chain is net unfavorable and therefore favors adsorbed state conformations that limit such contacts.

Finally, Table 2.3 reports thermodynamic values for adsorption of sequence I on a hydrophobic surface $\left(\chi_{H W}=-4\right)$ at conditions where the chain in solution is denatured and therefore fluctuates among a very large set of random coil conformations (achieved by setting $\chi_{H H}=-0.25$ ). As one would expect, the random-coil chain in solution loses conformational entropy upon adsorption due to the required localization of hydrophobic chain segments on the sorbent surface.

### 2.3.4 Influence of sorbent surface geometry

Although the effect is not easily studied through experiment, the topography (roughness, porosity, surface area to volume ratio) of the sorbent surface is thought to influence adsorption behaviour, including adsorption thermodynamics. The effect of a non-planar sorbent surface on chain adsorption was investigated by dMC simulations by activating two connected walls of the simulation cell, thereby allowing the chain to interact with a corner of the grid.

When $\chi_{H H}=\chi_{H W}=-4$, sequence I preferentially adsorbs to the corner of the grid in order to maximize solvent-free intermolecular contact area. As shown in Table 2.4, adsorption of sequence I to the grid corner results in a lower ensemble-average energy $\Delta_{a d s}<E>$ and a lower global energy minimum $E_{\text {min }}$ than observed for adsorption of the same sequence to a planar surface. When adsorbed to the corner, sequence I forms a maximum of 17 favorable hydrophobic contacts (sum of intramolecular and intermolecular contacts), 3 more than when adsorbed to the planar surface. This results in adsorbed chain conformations that are non-native, but which retain a globular structure similar in aspect
ratio to that of the native-state in solution. In comparison, sequence I adopts pancake or mound-like conformations to access the global energy minimum on a planar surface at otherwise identical conditions.

The preferential adsorption of the chain to the corner region of the sorbent surface is reflected in the more favorable $\Delta_{a d s}<A>$ for this process (Table 2.4). The change in $\Delta_{a d s}<A>$ is not as great as the change in $\Delta_{a d s}<E>$ due to the lower entropy of the chain adsorbed to the corner. The formation of a larger average number of hydrophobic contacts effectively restricts the conformational freedom of the chain adsorbed to the corner. This effect is reflected in the conformational degeneracy of the adsorbed chain at the global energy minimum. Sequence I forms a total of 15 unique lowest energy conformations when adsorbed at the hydrophobic corner, much less than the 84 conformations found for adsorption on the planar sorbent.

When $\chi_{H H}=\chi_{H W}=-4$, sequence II. also preferentially adsorbs to the corner, forming a total of 12 favorable contacts, or 1 more than formed on the planar surface. In this case, however, chain conformations at the global energy minimum are more pancake in nature and similar to those observed when sequence II adsorbs to the equivalent planar surface.

### 2.3.5 Dependence on the total hydrophobicity

The influence of the total hydrophobicity of the system was also investigated. Table 2.5 reports values $\Delta_{a d s}<E>, \Delta_{a d s}<A>$, and $T \Delta_{a d s}<S>$ for dMC simulations of adsorption of sequence I to a planar sorbent surface under conditions where the value of $\chi_{H H}=\chi_{H W}$ is
varied from -1 to -4 . In the absence of the sorbent, the distribution of chain energies narrows as $\omega_{H H}$ becomes more favorable, resulting in a higher probability for the chain to be in its lowest energy native-state conformation and an overall lower chain entropy.

Despite the lower conformational entropy of the chain in solution when $\chi_{H H}=-4$, adsorption of the chain to the hydrophobic planar surface results in a net decrease in chain entropy. Under these conditions, the chain adopts non-native conformations on the surface to achieve a total of two additional favorable contacts ( HH and HW ) relative to adsorption in its native-state conformation. The observed loss in conformational entropy of the chain during this transition is due to the additional constraints imposed by the impenetrable, inflexible sorbent surface and the connected nature of the chain, which make it possible to reach the lowest energy state (and fluctuate around it) through only a limited number of conformational trajectories.

### 2.4 Discussion

Despite the simplicity of the model, dynamic Monte Carlo simulations of model HP chains capture many of the phenomenological events that are frequently associated with nonspecific protein adsorption at a solid-liquid interface. For example, protein adsorption, particularly to hydrophobic sorbents, often results in experimentally observable changes in protein conformation (7, 13, 48). Our results provide a mesoscopic picture of this phenomenon. The dMC data show that the conformation of the adsorbed chain is constantly changing, so that the system fluctuates in the total number of favorable contacts (HH and HW) formed. As a result, the system energy also
fluctuates in time between energy states at or relatively near the global energy minimum. Therefore, adsorbed proteins are not frozen into a fixed, energetically most favorable conformation or energetically degenerate set of lowest energy conformations. Instead, as shown in the adsorbed-state energy probability distribution (Figure 2.4), the system is not always in its lowest-energy state due to the low conformational degeneracy of that state, and is often at a slighter higher energy $E_{i}$ because the conformational degeneracy of those energy states is orders of magnitude larger than that of the global energy minimum. Thus, the adsorbed HP chains exist as an ensemble of energies and conformations, a conclusion that has recently been proposed by others to describe the adsorbed protein state (49, 50).

Based on this concept, many researchers have argued that an increase in conformational entropy of the peptide chain contributes to the driving force for protein adsorption. Direct (isothermal titration calorimetry) and indirect (e.g., adsorption isotherms measured at several temperatures) measures of heats of protein adsorption indicate that the process is often endothermic, so that the dríving force for adsorption is provided by a positive change in system entropy $(13,51)$. However, whether this entropy change is due, at least in part, to an increase in the conformational entropy of the protein remains unclear. For adsorption to hydrophilic surfaces, indirect evidence suggests that the conformational entropy of the protein often increases upon adsorption (13). In contrast, adsorption to hydrophobic surfaces has recently been shown to result in an increase in the secondary structure of a protein, suggesting a loss in chain conformational entropy $(24,25,49,50$, 52).

Our dMC simulation results suggest that changes in the conformational entropy of a protein-like chain during adsorption to a solid-liquid interface can provide a substantial, even dominant driving force for adsorption under certain conditions. Consistent with inferences made from experiment, substantial increases in the conformational entropy of HP sequence I are observed when the stability of the native-state fold is low $\left(\chi_{H H}=-1\right)$ and the surface is relatively hydrophilic $\left(\chi_{H W}=-1\right)$. However, if the stability of the native state or the hydrophobicity of the sorbent is increased, any favorable contribution of chain conformational entropy to $\Delta_{a d s}<A>$ is quickly lost.

For instance, for adsorption to a more hydrophobic surface $\left(\chi_{H W}=-4\right)$, sequence I loses conformational entropy upon adsorption, a result that agrees with the experimental studies cited above. Careful comparison of the simulation data for the model HP chain in solution and adsorbed to the surface provides an understanding of why this is the case. Two phenomena are responsible. First, although the adsorbed chain does sample a large number of conformational states, these accessible conformations are those corresponding to system energies near the global energy minimum. To reach energies near this global energy minimum, the adsorbed chain must maintain a well-defined number of favorable intramolecular ( HH ) and intermolecular (HW) contacts with the rigid impenetrable sorbent surface, effectively restricting the number of accessible chain conformations. As a result, although the adsorbed chain does indeed adopt many conformations, these accessible conformations represent a relatively small fraction of all possible conformations of the linear chain. This point is illustrated in the dMC results for
adsorption of sequence I in its denatured state in solution to the same hydrophobic surface $\left(\chi_{H W}=-4\right)$. In this case, the chain in solution samples all conformational space with near equal probability, so the conformational entropy of the chain is a maximum. The conformational entropy change $T \Delta_{a d s}<S>$ for adsorption of this denatured chain is $12 k T$ (see Table 2.3 ), nearly identical to the loss in the conformational entropy of the chain in the denatured to native-state transition $\left(T \Delta_{D-N}<S>=-13 k T\right)$. Thus, relative to the denatured chain in solution, the adsorbed chain and the folded chain in solution have similar conformational entropy.

The second phenomemon of importance to understanding changes in chain entropy upon adsorption is the realization that the native-state chain in solution also fluctuates between a large number of conformational states of energy near that of the unique native-state conformation (see Figure 2.4). That is, while each HP chain may assume only one conformation to access its lowest energy state in solution, energy states very near the lowest energy are populated by a large number of conformations. As a result, the ensemble-average conformational entropy $\langle S\rangle_{\text {chain }}$ of the chain in solution is significantly larger than one would anticipate by treating the structure in solution as that corresponding to its unique lowest-energy solution state. For proteins, the possible existence of this phenomenon is supported by proton-exchange and other NMR relaxation measurements by Hwang et al. (53) and others (54) which reveal considerable mobility in the peptide backbone, and even more in the side chains of proteins in solution.

Finally, it is important to note that in their dMC simulations of mesoscopic protein-like chains adsorbing to a liquid-liquid interface, Anderson et al. (38) observed larger increases in the conformational entropy of the folded chain upon adsorption. The rigidity of the sorbent interface therefore appears to influence adsorption thermodynamics. In our system, the surface is rigid and impenetrable, so that the chain must adopt very specific conformations and points of contact to access low energy states. In the liquid-liquid system, the fluidity of the interface provides the adsorbed chain additional degrees of freedom, such that the chain can adopt a large number of conformations involving partial chain penetration into the sorbent (oil) phase. Nevertheless, Anderson et al. (38) found that the large positive changes in total entropy $\Delta_{a d s}<S>$ often observed in protein adsorption to oil-water interfaces are likely to be largely due to interfacial dehydration effects and not changes in chain conformational entropy.

### 2.4.1 Protein adsorption kinetics are linked to energetic barriers that frustrate conformational trajectories of the peptide chain

The dMC simulation algorithm is, in general, restricted to determination of equilibrium thermodynamic properties. However, several studies (see for example references 38, 40) have shown that the dMC approach can yield insights into dynamic and/or kinetic events if a sufficiently large number of configurational states are sampled at each step number. In this context, it is important to note that the trend in the energy trajectory shown in Figure 2.5 was observed for many independent simulations at the same conditions, indicating that the adsorption path is general in nature and not an artifact of the dMC simulation algorithm. That is, although the chain conformation at any given simulation
step is quite different in each independent simulation, the energy trajectory remains the same, showing $E_{i}$ values $\geq-48 k T$ for the initial $3 \times 10^{7}$ cycles following chain contact, followed by a relatively rapid drop in $E_{i}$ to $-56 k T$. In such cases, several groups have shown that dMC simulations can provide insights into the reaction pathway $(38,40-42)$ in addition to their classic use in determining equilibrium properties. Here, we use this approach to provide a few qualitative observations concerning the adsorption pathway for sequence I.

The energy $\langle E\rangle$ trajectory (a typical one is shown in Figure 2.5) for adsorption of sequence I shows that the folded chain fluctuates during the initial $1 \times 10^{7}$ to $1 \times 10^{8}$ simulation steps following contact with the sorbent surface between a large number of adsorbed conformations that are close to the native-state in structure and energy. All of these conformations are of energy $8 k T$ or higher than the global energy minimum of the system ( -56 kT ). Thus, the simulations identify a barrier to relaxation of the chain to those conformations associated with or close to the global energy minimum.

Analysis of our dMC data suggests that this barrier occurs because trajectories for conformational change are limited both by the compact segment density of the nativestate chain when it contacts the surface and by the additional constraints placed on chain conformational freedom imposed by favorable HW contacts formed during the initial phase of adsorption. These two constraints serve to frustrate the chain's attempts to access the global energy minimum through conformational change. As a result, any
attempted move by the chain generally leads to an increase in system energy, such that the adsorbed native state and sequences close to it sit in a local energy minimum.

Eventually, the chain randomly samples a conformation on the surface that allows a new segment move to be selected that results in a decrease in system energy. As shown in Figure 2.5 , the energy of the system then falls toward the global energy minimum through a series of previously inaccessible conformational trajectories. For 2D proteinlike HP chains such as sequence $I$, our dMC simulations therefore predict that a significant number of simulation steps are required for the protein to find and follow an adsorbed-state conformational trajectory that enables it to reach the global energy minimum. This result is qualitatively analogous to the typical time-course of a nonspecific protein adsorption event. Experimental observations of desorption of nonspecifically bound proteins during the initial stages (short times) of protein contact are numerous $(20,48)$, and several groups have proposed models for the kinetics of nonspecific protein adsorption that segregate the process into a fast reversible adsorption step, followed by a slow irreversible adsorption process (5). Protein exchange experiments by Balasubramanian et al. (55) and Bentaleb et al. (56) show that direct exchange between solution and surface-bound proteins occurs during short contact times with the sorbent and then rapidly diminishes with longer adsorption times.

In our simulations, once the system. first accesses its global energy minimum, the chain appears to be irreversibly adsorbed to the surface, at least with respect to the longest simulation interval we could achieve (a series of ten $1 \times 10^{10}$ step dMC simulations of the
adsorbed protein in which the starting conformation of the new simulation corresponded to the final conformation of the last). Therefore, a striking feature of our dMC results is the observation that a series of reversible actions (all dMC moves are intrinsically reversible) can result in effectively irreversible adsorption behaviour, at least with respect to the duration of the simulation. This same observation was made by Anderson et al.(38) in their dMC studies of the adsorption of protein-like chains to a liquid-liquid interface. More importantly, slow (or non-existent) desorption kinetics following longtime exposure of a protein to a solid surface have been reported in most experimental studies of nonspecific protein adsorption (1). For example, studies by Suttiprasit et al. (57) and Giacomelli et al. (28) show that conformational changes originating from adsorption occur slowly and can be detected for hours to days after an initial adsorption event.

### 2.5 Summary

There is now ample experimental evidence showing that protein conformational changes occur during the nonspecific adsorption of proteins to solid-liquid interfaces. However, it remains unclear under what conditions these structural changes contribute to adsorption energetics through a concomitant gain in chain conformational entropy. Dynamic Monte Carlo simulations of short protein-like HP chains on a solvent-filled two-dimensional lattice were used to show that adsorption to a solid-liquid interface of chains having a global energy minimum in solution characterized by a single native-state conformation results in a new lowest energy state populated by multiple adsorbed-chain conformations. While the HP chains used are not proteins, we have shown that their simulated adsorption
properties qualitatively mirror those of real proteins. Thus, dMC simulations of HP chains may provide a simple but useful means of understanding molecular aspects of nonspecific protein adsorption.

In simulations of low-stability HP chains adsorbing to relatively hydrophilic sorbents, the conformational entropy of the chain increases significantly, in part due to the conformational degeneracy of the global energy minimum. However, gains in the conformational entropy of the chain quickly disappear when either the stability of the native-state fold or the hydrophobicity of the sorbent is increased. Under these conditions, energetically accessible conformations of the adsorbed chain are severely restricted due both to the need to maximize the total number of favorable intramolecular (HH) and intermolecular bonds formed within the chain and with the surface, respectively, and to the physical restrictions imposed by the impermeable nature of the sorbent surface.

The adsorption trajectory of our model protein-like chains exhibit features consistent with experimental studies of nonspecific protein adsorption. Initially, the model chain adsorbs reversibly at energies well above the global energy minimum. The dMC simulations indicate that this is due to local energy barriers that frustrate attempts by the chain to access conformational trajectories leading to the global energy minimum. Eventually, however, a conformational trajectory is accessed that results in the chain becoming irreversibly adsorbed to the surface, at least within the duration of the simulation.

### 2.6 References

1. Haynes, C.A. and Norde, W., Globular protein at solid/liquid interfaces. Colloids Surfaces B, 2 (1994) 517-566.
2. Malmsten, M., Formation of adsorbed protein layers. J. Colloid Interface Sci., 207 (1998) 186-99.
3. Brash, J.L. and Horbett, T.A.,Reversibility and the mechanism of protein adsorption. In "Proteins at Interfaces II: Fundamentals and Applications" (T.A. Horbett and J.L. Brash, Eds.), p.1-23. American Chemical Society, Washington, DC, 1995.
4. Kondo, A., Urabe, T. and Yoshinaga, K., Adsorption activity and conformation of alpha-amylase on various ultrafine silica particles modified with polymer silane coupling agents. Colloids Surfaces A, 109 (1996) 129-36.
5. Krisdhasima, V., McGuire, J. and Sproull, R., Surface hydrophobic influences on $\beta$ lactoglobulin adsorption kinetics. J. Colloid Interface Sci., 154 (1992) 337-350.
6. Kull, T., Nylander, T., Tiberg, F. and Wahlgren, N.M., Effect of surface properties and added electrolyte on the structure of $\beta$-casein layers adsorbed at the solid/aqueous interface. Langmuir, 13 (1997) 5141-5147.
7. Norde, W. and Giacomelli, C.E., BSA structural changes during homomolecular exchange between the adsorbed and dissolved states. J. Biotechnol., 79 (2000) 259268.
8. Kondo, A. and Fukuda, H., Effects of adsorption conditions of kinetics of protein adsorption and conformational changes at ultrafine silica particles. J. Colloid Interface Sci., 198 (1998) 34-41.
9. Elwing, H., Protein adsorption and ellipsometry in biomaterial research. Biomaterials, 19 (1998) 397-406.
10. McNay, J.L. and Fernandez, E.J., How does a protein unfold on a reversed-phase liquid chromatography system? $J$ Chromatography $A, 849$ (1999) 135-148.
11. Buijs, J. and Hlady, V., Adsorption kinetics, conformation, and mobility of the growth hormone and lysozyme on solid surfaces, studied with TIRF. J. Colloid Interface Sci., 190 (1997) 171-181.
12. Buijs, J., Britt, D.W. and Hlady, V., Human growth hormone adsorption kinetics and conformation on self-assembled monolayers. Langmuir, 14 (1998) 335-341.
13. Haynes, C.A. and Norde, W., Structures and stabilities of adsorbed proteins. J. Colloid Interface Sci., 169 (1995) 313-328.
14. Zoungrana, T. and Norde, W., Thermal stability and enzymatic activity of alphachymotrypsin adsorbed on polystyrene surfaces. Colloids Surfaces B, 9 (1997) 157167.
15. Froberg, J.C., Arnebrant, T., McGuire, J. and Claesson, P.M., Effect of structural stability on the characteristics of adsorbed layers of T4 lysozyme. Langmuir, 14 (1998) 456-62.
16. Wang, J. and McGuire, J., Surface tension kinetics of the wild type and four synthetic stability mutants of T4 phage lysozyme at the air-water interface. J. Colloid Interface Sci., 185 (1997) 317-23.
17. Billsten, P., Wahlgren, M., Arnebrant, T., McGuire, J. and Elwing, H., Structural changes of T4 lysozyme upon adsorption to silica nanoparticles measured by circular dichroism. J. Colloid Interface Sci., 175 (1995) 77-82.
18. Billsten, P., Freskgard, P.-O., Carlsson, U., Jonsson, B.-H. and Elwing, H., Adsorption to silica nanoparticles of human carbonic anhydrase II and truncated forms induce a molten-globule-like structure. FEBS Letters, 402 (1997) 67-72.
19. Billsten, P., Carlsson, U., Jonsson, B.-H., Olofsson, G., Hook, F. and Elwing, H., Conformation of human carbonic anhydrase II variants adsorbed to silica nanoparticles. Langmuir, 15 (1999) 6395-6399.
20. Wahlgren, M., Arnebrant, T. and Lundstrom, I., The adsorption of lysozyme to hydrophilic silicon oxide surfaces: Comparison between experimental data and models for adsorption kinetics. J. Colloid Interface Sci., 175 (1995) 506-514.
21. Van Tassel, P.R., Guermouri, L., Ramsden, J.J., Tarjus, G., Viot, P. and Talbot, J., A particle-level model of irreversible protein adsorption with a postadsorption transition. J. Colloid Interface Sci., 207 (1998) 317-23.
22. Norde, W. and Lyklema, J., The adsorption of human plasma albumin and bovine pancreas ribonuclease at negatively changed polystyrene surfaces: I. Adsorption isotherms, effects of charge, ionic strength and temperature. J. Colloid Interface Sci., 66 (1978) 257-265.
23. Norde, W. and Lyklema, J., Why proteins prefer interfaces. J. Biomater. Sci. Polymer Edn., 2 (1991) 183-202.
24. Giacomelli, C.E. and Norde, W., Influence of hydrophobic teflon particles on the structure of amyloid beta-peptide. Biomacromolecules, 4 (2003) 1719-1726.
25. Sane, S.U., Cramer, S.M. and Przybycien, T.M., Protein structure perturbations on chromatographic surfaces. J. Chromatogr. A, 849 (1999) 149-159.
26. Ben-Tal, N., Honig, B., Bagdassarian, C.K. and Ben-Shaul, A., Association entropy in adsorption processes. Biophys. J., 79 (2000) 1180-1187.
27. Fersht, A.R., On the simulation of protein folding by short time scale molecular dynamics and distributed computing. PNAS, 99 (2002) 14122-14125.
28. Giacomelli, C.E., Esplandiu, M.J., Ortiz, P.I., Avena, M.J. and De Pauli, C.P., Ellipsometric study of bovine serum albumin adsorbed onto $\mathrm{Ti} / \mathrm{TiO}_{2}$ electrodes. $J$. Colloid Interface Sci., 218 (1999) 404-411.
29. Lau, K.F. and Dill, K.A., A lattice statistical mechanics model of the conformational and sequence spaces of proteins. Macromolecules, 22 (1989) 3986-3997.
30. Socci, N.D. and Onuchic, J.N., Folding kinetics of proteinlike heteropolymers. J. Chem. Phys., 101 (1994) 1519-1528.
31. Camacho, C.J. and Thirumalai, D., Kinetics and thermodynamics of folding in model proteins. PNAS, 90 (1993) 6369-6372.
32. Konig, R. and Dandekar, T., Solvent entropy-driven searching for protein modeling examined and tested in simplified models. Protein Eng., 14 (2001) 329-35.
33. Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D. and Chan, H.S., Principles of protein folding - A perspective from simple exact models. Protein Sci., 4 (1995) 561-602.
34. Dill, K.A., Dominant forces in protein folding. Biochemistry, 29 (1990) 7133-7155.
35. Schymkowitz, J.W.H., Rousseau, F. and Serrano, L., Surfing on protein folding energy landscapes. PNAS, 99 (2002) 15846-15848.
36. Chan, H.S. and Dill, K.A., Compact polymers. Macromolecules, 22 (1989) 4559-73.
37. Levinthal, C., Are there pathways for protein folding? J. Chim. Phys. PCB, 65 (1968) 44-45.
38. Anderson, R.E., Pande, V.S. and Radke, C.J., Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. J. Chem. Phys., 112 (2000) 9167-85.
39. Miyazawa, S. and Jernigan, 'R.L., Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. Macromolecules, 18 (1985) 534-552.
40. Zhdanov, V.P. and Kasemo, B., Simulations of denaturation of adsorbed proteins. Phys. Rev. E, 56 (1997) 2306-9.
41. Zhdanov, V.P. and Kasemo, B., Monte Carlo simulations of denaturation of adsorbed proteins. Proteins: Structure, Function, and Genetics, 30 (1998) 168-76.
42. Castells, V., Yang, S. and Van Tassel, P.R., Surface-induced conformational changes in lattice model proteins by Monte Carlo simulation. Phys. Rev. E, 65 (2002)
43. Norde, W., MacRitchie, F., Nowicka, G. and Lyklema, J., Protein adsorption at solidliquid interfaces: Reversibility and conformation aspects. J. Colloid Interface Sci., 112 (1986) 447-456.
44. Welzel, P.B., Investigation of adsorption-induced structural changes of proteins at solid/liquid interfaces by differential scanning calorimetry. Thermochim. Acta, $\mathbf{3 8 2}$ (2002) 175-188.
45. Jackler, G., Steitz, R. and Czeslik, C., Effect of temperature on the adsorption of lysozyme at the silica/water interface studied by optical and neutron reflectometry. Langmuir, 18 (2002) 6565-6570.
46. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., Equation of state calculations by fast computing machines. J. Chem. Phys., 21 (1953) 1087-92.
47. Norde, W., Adsorption of proteins from solution at the solid-liquid interface. $A d v$. Colloid Interface Sci., 25 (1986) 267-340.
48. Wertz, C.F. and Santore, M.M., Effect of surface hydrophobicity on adsorption and relaxation kinetics of albumin and fibrinogen: Single-species and competitive behaviour. Langmuir, 17 (2001) 3006-3016.
49. Przybycien, T.M. A phase diagram for the structural response of proteins to hydrophobic chromatography. (2003).
50. Fernandez, E.J. Toward modeling of hydrophobic interaction chromatography (hic): Multiple tools to connect thermodynamics, adsorption and conformation. (2003).
51. Omanovic, S. and Roscoe, S.G., Interfacial behaviour of beta-lactoglobulin at a stainless steel surface: An electrochemical impedance spectroscopy study. J. Colloid Interface Sci., 227 (2000) 452-460.
52. Zoungrana, T., Findenegg, G.H. and Norde, W., Structure, stability, and activity of adsorbed enzymes. J. Colloid Interface Sci., 190 (1997) 437-48.
53. Hwang, P.M., Choy, W.Y., Lo, E.I., Chen, L., Forman-Kay, J.D., Raetz, C.R.H., Prive, G.G., Bishop, R.E. and Kay, L.E., Solution structure and dynamics of the outer membrane enzyme PagP by NMR. PNAS, 99 (2002) 13560-13565.
54. Prompers, J.J. and Bruschweiler, R., Reorientational eigenmode dynamics: A combined MD/NMR relaxation analysis method for flexible parts in globular proteins. J. Am. Chem. Soc., 123 (2001) 7305-7313.
55. Balasubramanian, V., Grusin, N.K., Bucher, R.W., Turitto, V.T. and Slack, S.M., Residence-time dependent changes in fibrinogen adsorbed to polymeric biomaterials. J. Biomed. Mater. Res., 44 (1999) 253-260.
56. Bentaleb, A., Abele, A., Haikel, Y., Schaaf, P. and Voegel, J.C., FTIR-ATR and radiolabeling study of the adsorption of ribonuclease A onto hydrophilic surfaces: Correlation between the exchange rate and he interfacial denaturation. Langmuir, 14 (1998) 6493-6500.
57. Suttiprasit, P., Krisdhasima, V. and McGuire, J., The surface-activity of alphalactalbumin, beta-lactoglobulin, and bovine serum-albumin .1. surface-tension measurements with single-component and mixed-solutions. J. Colloid Interface Sci., 154 (1992) 316-326.

### 2.7 Tables

Table 2.1: Thermodynamic changes for the native to denatured state transition of sequence I at reduced temperatures $T^{*}=0.44$ and $T^{*}=0.71$.

| Reduced | Internal energy | Helmholtz energy | Entropy |
| :---: | :---: | :---: | :---: |
| temperature |  |  |  |
| $T^{*}$ | $\Delta_{N-D}<E>$. | $\Delta_{N-D}<A>$ |  |
|  |  |  |  |

Table 2.2: Thermodynamic changes for the adsorption of sequence I to a relatively hydrophilic surface $\left(\chi_{H W}=-1\right)$ as a function of chain stability $\left(\chi_{H H}\right)$.

| Energy ratio | Internal energy | Helmholtz energy | Entropy |
| :---: | :---: | :---: | :---: |
| $\chi_{H H} / \chi_{H W}$ | $\Delta_{\text {ads }}<E>$ |  |  |
|  |  |  |  |

Table 2.3: Thermodynamic changes for the adsorption of denatured sequence I ( $\chi_{H H}=-$ 0.25 ) on a hydrophobic surface $\left(\chi_{H W}=-4\right)$

| Internal energy | Helmholtz energy | Entropy |
| :---: | :---: | :---: |
| $\Delta_{a d s}<E>$ | $\Delta_{a d s}<A>$ | $T \Delta_{a d s}<S>$ |
| $(k T)$ | $(k T)$ | $(k T)$ |
| -53.9 | -41.8 | -12.1 |
|  |  |  |

Table 2.4: Thermodynamic changes for the adsorption of sequence I on surfaces of varying geometry with $\chi_{H H}=\chi_{H W}=-4$.


Table 2.5: Thermodynamic changes for the adsorption of sequence I to a planar surface when $\chi_{H H}=\chi_{H W}$.

| Interaction energy | Internal energy | Helmholtz energy | Entropy |
| :---: | :---: | :---: | :---: |
| $\chi_{H H}=\chi_{H W}$ | $\Delta_{a d s}<E>$ | $\Delta_{a d s}<A>$ | $T \Delta_{a d s}<S>$ |
| $(k T)$ | $(k T)$ | $(k T)$ | $(k T)$ |
| -1 | -2.1 | -5.7 | +3.6 |
| -2 | -11.3 |  |  |
| -3 | -17.7 |  |  |
|  |  |  |  |

### 2.8 Figures

Figure 2.1: Examples of Verdier-Stockmeyer moves used to manipulate the protein-like chains during simulations. Shown are a 2-bead crankshaft move, a 1-bead flip and an example of an end-bead turn.


Figure 2.2: Schematic diagrams of the two model chains: sequence I and sequence II. Filled circles represent hydrophobic $(\mathrm{H})$ units while open circles represent polar $(\mathrm{P})$ units.


Figure 2.3: Simulation results demonstrating the conformational dependence of sequence I on temperature. Shown is the ensemble-averaged number of intramolecular hydrophobic contacts as a function of reduced temperature, $T^{*}$. The line drawn indicates the trend of the data.


Figure 2.4: System energy probability histograms for sequence I: filled grey bars - chain in solution when $\chi_{H H}=-4\left(T^{*}=0.71\right)$; filled black bars - chain adsorbed on a relatively hydrophobic surface $\left(\chi_{H W}=-4\right)$ when $\chi_{H H}=-4$. The error bars refer to the standard deviation of 5 runs.


Figure 2.5: Energy trajectory for the adsorption of sequence I to a planar hydrophobic surface when $\chi_{H H}=\chi_{H W}=-4$.


Figure 2.6: Representative conformational states of sequence I shortly after adsorbing to a planar hydrophobic surface when $\chi_{H H}=\chi_{H W}=-4$. All conformations correspond to system energies $E_{i} \geq-48 k T$.


Figure 2.7: Representative examples of lowest energy states of sequence I adsorbed to a hydrophobic surface when $\chi_{H H}=\chi_{H W}=-4$. At the global minimum, $E_{i}=-56 k T$.


Figure 2.8: The 6 lowest energy states for sequence I when $\chi_{H H}=-3$ and $\chi_{H W}=-1$. The structure in the centre is the adsorbed chain in its native state.


## 3 Mesoscopic dynamic Monte Carlo simulations of the adsorption of protein-like HP chains within laterally constricted spaces*

### 3.1 Introduction

The process and consequences of protein adsorption to solid-liquid interfaces have received significant attention over the past half century, due in part to the importance of control of protein adsorption to the design and performance of biomedical implants (e.g., artificial hip and knee joints, contact lenses, vascular grafts), chromatography columns, food processing equipment, etc., and to the inherent complexity of nonspecific protein adsorption that effectively limits our ability to understand and model the process using classic adsorption theories. Much of what we know about nonspecific protein adsorption has been derived from experimental studies, beginning with the now classic works of Vroman on blood proteins $(1,2)$, and the early seminal contributions of Norde and Lyklema (3-8). Consistent with the chemical heterogeneity of proteins, experiments have shown that intra- and intermolecular Coulombic, hydration, hydrogen-bonding, and shortrange van der Waals forces can each contribute, either favorably or unfavorably, to the kinetics and energetics of protein adsorption (9). Intermolecular contacts between the sorbent and the adsorbed protein are often formed at the expense of intramolecular contacts that are stable in the native-state conformation of the protein. As a result, changes in protein conformation are often observed during adsorption $(9,10)$.

[^1]Conformational changes in proteins during adsorption have been shown to depend on a number of factors, including protein and sorbent surface hydrophobicity (11, 12), and the thermodynamic stability of the native-state conformation of the protein (13, 14). Percent occupation of total binding sites on the sorbent surface also affects the extent of conformational change, with the magnitude of the change typically being largest at low surface coverage ( $10,15,16$ ). For example, Norde and Favier (17) used circular dichroism spectra of proteins displaced from the surfaces of various sorbents to show that proteins adsorbed at high surface coverage retain more secondary structure than do proteins adsorbed at lower surface concentrations. Norde et al. (18) also used differential scanning calorimetry (DSC) to show that denaturation temperatures of adsorbed proteins tend to decrease with decreasing surface coverage. These DSC studies also show that denaturation enthalpies (per mole of protein) are often smaller for adsorbed proteins and tend toward zero as the surface coverage decreases. Enzyme activity assays $(16,18)$, NMR and Raman spectroscopy (19, 20), total internal reflectance fluorescence (21, 22), atomic force microscopy (AFM) $(23,24)$, and ellipsometry $(25,26)$ are but some of the many other experimental techniques that have confirmed changes in protein conformation upon adsorption and the dependence of these changes on the adsorption conditions.

One consequence of increased surface coverage is increased excluded volume effects between proximally adsorbed protein molecules that prohibit one protein molecule from accessing the volume occupied by that of a second. As a result, at high surface coverage, the peptide chain of an adsorbed protein molecule cannot easily 'spread out' to sample low-energy extended conformations. Ellipsometry $(26,27)$ data support this model by
showing that the thickness of adsorbed protein layers tends to increase with increasing surface coverage. However, little is known about how changes in accessible sorbent surface area and restrictions in the volumetric space above it affect the energy and entropy of an adsorbed protein macromolecule.

Recently, we have shown that mesoscopic 2D dynamic Monte Carlo simulations of the adsorption to a planar surface of a uniquely folding linear copolymer comprised of a specified sequence of hydrophobic $(\mathrm{H})$ and polar $(\mathrm{P})$ segments (i.e., the HP chain model of Dill (28)) share many properties characteristic of nonspecific protein adsorption (29). Although they are based on an idealized representation of protein-like chains, the simulations predict energy and chain entropy changes consistent with those observed in experiments of protein adsorption and therefore add to our general knowledge of the adsorption process by providing molecular-level insights and the ability to precisely quantify contributions of the various adsorption sub-processes (e.g., conformational changes in the chain and the associated change in chain entropy, etc.). Here, mesoscopic dynamic Monte Carlo simulations of simple protein-like HP chains are used to explore the dependence of adsorbed-chain conformations and energetics on the amount of sorbent surface area and proximal volumetric space available to the native-state sequence for adsorption. In the dynamic Monte Carlo method, the conformational trajectory of the adsorbing chain molecule is determined by the change in system energy that results from a change in chain conformation and any associated change in the number of intra- and intermolecular contacts. The method therefore allows one to calculate changes in chain conformational space and entropy, as well as changes in chain and system energy and
their dependence on adsorption conditions, including the amount of sorbent surface area and associated volume available to the adsorbing chain. Results from our simulations are compared with experimental data for the adsorption of hen egg-white lysozyme to silica with the aim of qualitatively connecting observations drawn from the simple mesoscopic simulations to relevant macroscopic data for a real protein adsorption process.

### 3.2 Dynamic Monte Carlo simulations of chain adsorption

Dynamic Monte Carlo (dMC) simulations have been used by our group (29) and others (30-32) to study adsorption of simple protein-like chains to planar liquid-liquid (oilwater) (31) and solid-liquid interfaces. The details of the simulation method and its application to the adsorption of flexible copolymer chains can be found in those references. However, application of the dMC technique to analysis of the dependence of adsorbed-state conformations and energetics on the available sorbent surface area and interfacial volume, the purpose of this study, required modification of the model and our previously reported dMC simulation code. A brief description of the basic dMC simulation method and necessary modifications made for this study are therefore presented below.

Chain movements are carried out by standard Verdier-Stockmeyer moves on a 2D lattice where the axial distance between adjacent lattice sites is $x$ (or $y$ ) $=a$. The model is ergodic, as the set of Verdier-Stockmeyer moves allow the chain to adopt all possible conformations within the lattice. Chain dynamics and associated system energetics are determined using the Metropolis algorithm, which uses Boltzmann-weighted statistics to
determine allowed moves. The energy, $E_{i}$, of the lattice is defined by the sum of pairwise interaction energies between all non-connecting neighboring units of the system:

$$
\begin{equation*}
E_{i}=\sum_{i}^{N-1} \sum_{j>i}^{N} \omega_{i j} \delta\left(r_{i}-r_{j}\right)-\sum_{l}^{L-1} \omega_{l, l+1} \tag{3.1}
\end{equation*}
$$

where $N$ represents the total number of lattice sites in the grid, $\omega_{i j}$ is the interaction energy between the components occupying lattice sites $i$ and $j, r_{i}$ and $r_{j}$ are the respective coordinates of lattice sites $i$ and $j$, and $\delta$ is a delta function based on the relative positioning of the two interacting species. For neighboring residues, $\delta(a)=1$, while for non-neighboring residues, $\delta\left(\left|r_{i}-r_{j}\right|>a\right)=0$. The second term on the right-hand side of Eq. [3.1] subtracts all contributions from pair-wise energies between all directly connected residues on the polymer chain. The interaction energy $\omega_{i j}$ is related to the Flory interaction parameter $\chi_{i j}(T)$ by $\chi_{i j}=\omega_{i j} / k T$. Finally, $l$ indexes the chain bond number and therefore counts from 1 to $L-1$, where $L$ is the number of segments in the chain.

In the dMC simulations reported here, each lattice site is occupied by one of five different components: a hydrophobic $(\mathrm{H})$ chain segment, a polar ( P ) chain segment, solvent $(\mathrm{S})$, a sorbent wall (W) segment, or an athermal wall (A) segment. Reflective boundary conditions are used at each wall. The positions of wall segments are fixed in the lattice and solvent and chain segments are not allowed to occupy or pass through wall segments. The Cartesian lattice is allowed to vary in width in one lattice unit increments, with the
top and bottom surfaces of the lattice always presenting a planar sorbent wall ( S ) and the two side surfaces serving as athermal walls (A) which act only to limit the sorbent surface area and volumetric space available to the chain for adsorption. Changes in the available surface area and associated volumetric space above it for an individual proteinlike chain to adsorb are therefore achieved in the dMC simulations by moving the parallel side walls of the lattice closer or further apart.

All simulations involve the adsorption of individual protein-like chains to the planar sorbent surface at the bottom of the lattice. At the start of each run, the chain is placed directly on the surface and centered equidistantly from the parallel athermal side walls (A) which are impenetrable to the chain. As a result, the sorbent surface and the two side walls act to confine the chain by restricting chain movement in three of four directions. In cases when the initial chain dimensions match the grid width, contact with the adjacent walls at the start of the simulation is permitted. The starting chain conformation is dependent on the width of the sorbent surface. In most simulations, the chain is initially in its lowest energy "native state" conformation; however, simulations in which the chain is initially in a random denatured conformation were also performed to ensure that the simulation results are independent of initial chain conformation. Non-native chain conformations were also used to initiate dMC simulations in which the distance of separation between the athermal side walls of the lattice is less than the width of the chain in its native state.

Adsorption thermodynamics were calculated from probability histograms generated from Monte Carlo sampling according to standard statistical mechanical methods. Details of the sampling method and calculation procedures can be found in Liu and Haynes (29). All simulation runs were repeated at least 5 times and averaged to obtain thermodynamic values.

The HP model of Dill and coworkers $(28,33)$ was used to specify chain composition in the simulations. In the HP model, a linear chain is comprised of two types of segments: hydrophobic $(\mathrm{H})$ and polar $(\mathrm{P})$. The sequence of H and P segments is specified such that the chain adopts a single lowest-energy conformation in the solvent. As a result, the chain exhibits protein-like properties. At very low temperatures, the chain is almost exclusively in its lowest-energy "native state" conformation. As the temperature is raised, higher entropy conformations are increasingly observed. Analogous to real proteins, the model chain also experiences a melting transition over a limited temperature range, such that below $T_{m}$ (the melting temperature) the chain is most often observed in a compact folded conformation, while above $T_{m}$ the chain conformation is that of a random coil (denatured state). In making this connection, we are not implying that the HP chain captures all physical and thermodynamic properties of real proteins. It clearly does not. The value of the model lies in the fact that the conformation, entropy and energy of the HP chain are sensitive to solution environment in a manner that is qualitatively similar to the observed behavior of globular proteins (see (29) and references therein for a more detailed discussion of the protein-like properties of the HP chain model). However, unlike for a complex peptide chain, the simplicity of the mesoscopic HP chain allows one
to consider all possible conformations and energetics accessible to the chain in a given environment. It therefore provides a simple but useful model for exploring the connection between conformational states of adsorbed chains and adsorption thermodynamics that is relevant to nonspecific protein adsorption.

The chain is placed on a Cartesian coordinate lattice in a self-avoiding configuration and all lattice sites not occupied by the chain are by definition occupied by solvent. In all cases, the sorbent surface at the bottom of the lattice is specified to be hydrophobic, and therefore interacts favorably with the H-residues of the chain. Simulations are run for a sufficient number of steps to allow for sampling of the entire conformational space (usually ca. $10^{9}$ to $10^{10}$ cycles).

The chain used in the simulations (shown in Figure 3.1) is one used in our previous study (29). The sequence, an 18 -mer, is predominantly hydrophobic and folds into a lowestenergy (native) state having 9 intramolecular HH bonds.

### 3.3 Materials and Methods

### 3.3.1 Reagents

Hen egg-white lysozyme (HEWL) and $\beta$ - N -acetylglucosaminidase (NAHase), extracted from jack bean, were purchased from Sigma Chemicals (St. Louis, MO) and used without further purification. The substrate, p-nitrophenyl penta- N -acetyl $-\beta$-chito-pentaoside
(pNP-C5) was purchased from Seikagaku Co. (Japan). All buffer and background salts were purchased from Fisher Scientific (Nepean, Canada).

All water used in the experiments was distilled and filtered through a Sybron/Barnstead NANOpure II system.

### 3.3.2 Measurement of Adsorption Isotherms

Nonporous microcrystalline silica particles (Sigma, St. Louis, MO) with a size distribution of $0.5 \mu \mathrm{~m}$ to $10 \mu \mathrm{~m}$ (over $80 \%$ of the particles between 1 to 5 microns) were used as the sorbent. The surface of the silica particles used is hydrophilic and has a point of zero charge of $c a . \mathrm{pH} 3$. The silica was washed overnight in a stirring solution of $0.7 \%$ sodium persulfate (BDH Ltd., Poole, England) in concentrated sulphuric acid (Fisher Scientific, Nepean, ON). Once cleaned, the mixture was transferred to pyrex glass tubes, spun down, and the acid discarded. The silica was then repeatedly rinsed with water to remove all residual acid and placed overnight in a Precision vacuum oven at $100{ }^{\circ} \mathrm{C}$. Cleaned and dried silica was stored in a dessicator at room temperature until ready for use.

Stock solutions of HEWL in 50 mM sodium phosphate/sodium citrate buffer (1:1 parts), pH 6.0 were made prior to each experiment. Concentrations of lysozyme were measured by spectral absorption using a Carey 1E UV/VIS Spectrometer. The extinction coefficient used was $24.826 \mathrm{~mL} \mathrm{mg}^{-1} \mathrm{~cm}^{-1}(10 \mathrm{mg} / \mathrm{mL}, \lambda=280 \mathrm{~nm})$. A stock suspension of silica in sodium phosphate/sodium citrate buffer was also prepared. The mixture was
sonicated for 5 minutes. The sonication step effectively disperses the particles in solution with no measurable change in specific surface area. The specific surface area of the silica particles, $A_{s}$, was determined by multipoint BET measurements using nitrogen gas and a Quantisorb BET apparatus (Quantachrome Corporation). The specific surface area, $A_{s}$, was found to be $5.6 \pm 0.45 \mathrm{~m}^{2} / \mathrm{g}$ silica.

Isotherms for adsorption of HEWL' on silica particles were measured by the depletion method. Adsorption experiments were carried out in 1.5 mL polypropylene microcentrifuge tubes. To each tube, specific amounts of buffer, HEWL and silica stock solutions were added in proper proportions to achieve a predetermined total protein concentration and sorbent surface area. The total volume of solution in each tube was 1 mL . Samples were left to turn end-over-end at room temperature $\left(22^{\circ} \mathrm{C}\right)$ for at least 12 hours, giving the system sufficient time to reach steady state. Samples were then centrifuged at 14000 rpm for 1 minute, and the supernatant recovered, filtered to remove any residual silica, and analyzed by absorbance at 280 nm to determine the free protein concentration. A total HEWL mass balance was then used to determine the corresponding concentration of adsorbed HEWL.

The centrifuged HEWL-loaded sorbent was recovered and then washed three times with a five-fold excess volume of 50 mM sodium phosphate/sodium citrate buffer (1:1 parts), pH 6.0. The protein bearing sorbent was then resuspended in the same buffer to a total volume of 0.5 mL and allowed to equilibrate for 24 to 48 hours. The solution phase was
then assayed for desorbed protein by adsorbance at 280 nm . In all cases, adsorption of HEWL to the silica particles was found to be irreversible.

### 3.3.3 Lysozyme activity measurements

HEWL activity, either in solution or adsorbed to silica particles was determined by measuring the rate of hydrolysis of the soluble substrate pNP-C5 according to a modified version of the assay originally described by Nanjo et al. (34). HEWL-catalyzed hydrolysis of pNP-C5 yields short chitooligosaccahrides that are susceptible to further hydrolysis by NAHase to release p-nitrophenol (pNP), a colorimetric compound that can be detected at 400 nm .

HEWL was adsorbed to silica at either $20 \%$ or full surface saturation according to the method used to determine the adsorption isotherm. Steady-state adsorbed HEWL samples (see above) were centrifuged at 14000 rpm for 1 minute and the supernatant discarded. Fresh buffer was added and the silica pellet resuspended. The centrifugation, decanting and resuspension steps were then repeated 5 to 7 times in order to eliminate protein that might still be in solution. Following the final rinse, an appropriate amount of buffer was added to adjust the adsorbed protein concentration to 0.0035 mM . To each lysozyme/silica/buffer mixture, appropriate volumes of NAHase and pNP-C5 stock solution were added. The resulting concentrations of the enzyme (NAHase) and substrate were $3.8 \times 10^{-5} \mathrm{mM}$ and 0.2 mM , respectively. Samples were left to rotate end-over-end at room temperature for a given incubation period (either 20,40 or' 60 minutes), after which, the solutions were centrifuged and the supernatant removed.

Membrane syringe filters ( $0.2 \mu \mathrm{~m}$ Gelman PVDF membrane, Pall Corporation, Ann Arbor, MI) were used to remove any residual silica particulates from the solutions. The absorbance of each sample was then measured at 400 nm to quantify released pNP . Readings from the spectophotometer were adjusted to compensate for any losses of the product during the filtration step.

Measurements were also made of the corresponding activity of the buffer solution and HEWL in buffer solutions containing no silica. In these samples, absorbance was measured continuously by absorption spectrophotometry at 400 nm .

### 3.3.4 Isothermal Titration Calorimetry

An aqueous titrate solution containing $3.5 \mathrm{mg} \mathrm{mL}^{-1} \mathrm{HEWL}$ was prepared in 50 mM KCl , adjusted to pH 7.0 through appropriate addition of HCl or KOH . Clean dry silica was also prepared as a $50-\mathrm{mM} \mathrm{KCl}$ titrand solution adjusted to pH 7.0 . Isothermal titration calorimetry (ITC) experiments were conducted in a Calorimetry Sciences Corp. Model 4200 Isothermal Titration Calorimeter. The titrand, reference, and titrate solutions were thoroughly degassed prior to loading. Once the system reached thermal equilibrium (ca. 2 hrs .), $25 \quad 10-\mu \mathrm{L}$ aliquots of protein solution were sequentially injected into the sample cell, which contained 1 mL of well-mixed, thermally equilibrated silica suspension ( 0.005 $\mathrm{g} / \mathrm{mL}$ ), with a thermally equilibrated silica suspension of identical volume and composition serving as reference. The time between injections was set at 2400 s to allow a return to baseline signal after each thermal peak. The titrations were carried out so that
the surface coverage reached $95 \%$ of the adsorption plateau after 10 to 15 injections of titrate solution. Samples were continuously mixed with an internal Rushton turbine-type blade rotating at 100 rpm , which was sufficient to fully suspend the silica and eliminate mass-transfer effects that might broaden the thermal peak.

### 3.4 Results and discussion

### 3.4.1 dMC simulations of HP chain adsorption within confined spaces

Figure 3.2 reports the energy $E_{i}$ trajectory for the first $1.5 \times 10^{5}$ steps of a typical dMC simulation of the adsorption of the HP chain to a planar sorbent surface (bottom wall of the lattice) when the length of the sorbent surface is large (i.e. lattice length is greater than the fully extended chain length). The chain is initially in its lowest-energy folded state in solution. Prior to chain contact with the sorbent surface $E_{i} \geq-9 k T$. Following chain contact, $E_{i}$ rapidly decreases to $c a .-40 k T$, an energy level corresponding to adsorption of the chain to the sorbent surface in an ensemble of relatively compact nonnative conformations. In this simulation, the stability of the native-state conformation of the chain is low $\left(\chi_{H H}=-1\right)$ and formation of intermolecular contacts between a hydrophobic (H) residue of the chain and a sorbent surface site (W) is energetically highly favorable since the wall is hydrophobic $\left(\chi_{H W}=-4\right)$. A relatively small number of simulation steps are therefore required to allow the chain to adopt new more surface associated conformations that further reduce $E_{i}$ toward the global energy minimum $E_{\text {min }}$ of the system, $-56 k T$. The probability of finding the system at or near the global energy minimum then remains high for the remainder of the simulation. The conformation of
the chain is dynamic however, such that the chain frequently adopts higher energy conformations, including conformations close to the native-state conformation. Thus, even under conditions where the native state of the chain is only marginally stable and the hydrophobic segments of the chain have a strong preference for the interface, the adsorbed chain will adopt the native-state conformation, albeit very infrequently. While the results in Figure 3.2 represent a single dMC trajectory, they are consistent with results obtained from a large set of independent dMC trajectories for the same system, indicating that the reported results are not unique to a single simulated adsorption trajectory but are reflective of the general dynamic properties of the chain at equilibrium.

Unlike in solution, where the lowest-energy state is occupied by a single conformation (the native state), the global energy minimum for the adsorbed HP chain can include a rather large number of distinct conformations. Figure 3.3 plots the conformational degeneracy of the global energy minimum, $E_{m i n}$, for the adsorbed HP chain as a function of the width, $x_{S}$, of the sorbent surface available for chain contact. When $x_{S}$ is equal to or greater than the length of the fully extended chain (18 lattice units), 84 unique conformations of the adsorbed HP chain are observed at $E_{\text {min }}$. However, as $x_{S}$ is decreased, the conformational degeneracy of $E_{\text {min }}$ falls rapidly, such that when $x_{S}$ equals 6 lattice units (i.e., 2 lattice units wider than the smallest dimension of the native state of the HP chain), $E_{\text {min }}$ is occupied by only two distinct conformations. At all sorbent widths, the adsorbed chain also adopts a large number of higher energy conformations, as suggested in Figure 3.2. However, our simulation results indicate that the more than one order of magnitude loss in conformational degeneracy observed at $E_{\text {min }}$ is consistent with
that observed at all other accessible energy levels, indicating significantly lower chain entropy when adsorption occurs within a confined volume.

The energies accessible to the adsorbed sequence are shown for two different $x_{S}$ values in Figure 3.4 in the form of probability histograms, which report the probability at any given step number of the system being at energy $E_{i}$, and the chain in a conformation unique to that energy. Reducing the sorbent surface area (and the volumetric space directly above it) reduces the number of accessible energies in addition to reducing the conformational degeneracy of the chain at each $E_{i}$. However, the native-state conformation of the chain remains accessible at all $x_{S} \geq 4$.

Figure 3.5 reports the probability of finding the adsorbed HP chain in its native-state conformation at any step number following system equilibration. When the native-state conformation of the HP chain is relatively stable in solution $\left(\chi_{H H}=-4\right)$, reducing $x_{S}$ enhances the probability of finding the adsorbed chain in its native-state conformation, irrespective of the hydrophobicity of the sorbent surface. Confinement eliminates many expanded conformations of the chain (see above) while still allowing the chain to access its native-state conformation. In addition, side-on adsorption of the chain in its nativestate conformation becomes one of the lowest energy conformations when $x_{S}=4$. As a result, restricting the volumetric space available to the adsorbed chain significantly increases the stability of its native state conformation. Previously, Zhou and Dill (35) developed an elegant theory to show that confining a model protein obeying two-state unfolding thermodynamics to a small inert space increases the stability of the native state
by as much as $15 \mathrm{kcal} / \mathrm{mol}$. For our adsorbed HP chain, this same stabilizing effect is somewhat weaker due to the fact that one of the walls of the confining space is no longer inert, but instead serves as a sorbent surface offering a favorable energy of interaction with hydrophobic segments of the chain. When $\chi_{H H}=-4$ and $\chi_{H W}=-1$, chain confinement on three sides (by setting $x_{S}=4$ ) increases the probability of finding the adsorbed HP chain in its native state conformation by over an order of magnitude relative to that observed in solution. This corresponds to an adsorption process that results in a net increase in the stability of the native state of $\Delta \Delta A=-2 \mathrm{kcal} / \mathrm{mol}$ despite the energetic driving force to denature the chain provided by the favorable energy of interaction between hydrophobic segments of the chain and the sorbent surface.

Hyperstabilization of the native-state conformation in the adsorbed chain relative to the same chain in solution is no longer observed when the energy of interaction between the sorbent surface and hydrophobic chain becomes more favorable $\left(\chi_{H W}=-4\right)$. In this case, the probability of finding the HP chain in its native-state conformation is reduced by nearly three orders of magnitude following adsorption to the sorbent surface when $x_{S}=$ 18. A decrease in $x_{S}$ again results in a dramatically higher probability that the adsorbed chain will adopt its native-state conformation, such that the probability of finding the adsorbed chain in its native-state conformation becomes equal to (but not greater than) that observed for the chain free in solution.

The ability to stabilize native-state conformations of adsorbed proteins through volumetric confinement has been observed experimentally. Eggers and Valentine (36)
showed that the melting temperature of $\alpha$-lactalbumin could be increased by as much as $32^{\circ} \mathrm{C}$ by confinement in the pores of silica glass.

An alternate means of restricting the volumetric space available to an adsorbed protein is to increase the concentration of protein on the sorbent surface to near monolayer coverage. Excluded volume forces with neighboring adsorbed protein macromolecules then effectively eliminate adsorbed state conformations requiring chain extension in lateral directions. Figure 3.6 reports, in the form of initial reaction rates, the molar activity of HEWL in solution and when adsorbed to nonporous particulate silica at two different surface concentrations. All HEWL in the solution phase was removed by washing prior to analyzing the activity of the adsorbed protein. In all three experiments, the total mass of HEWL in the reaction cell was held constant. Based on activity being a sensitive measure of protein conformation, the results show that the native-state conformational stability of adsorbed HEWL increases with increasing surface coverage. At 20\% monolayer coverage, adsorbed HEWL catalyzes the hydrolysis of pNP-C5 at an initial rate approximately $6 \%$ of that observed for an equivalent loading of HEWL in solution. At monolayer coverage, the initial rate of pNP-C5 hydrolysis increases to $34 \%$ of that observed in solution, indicating, as predicted in our simulations for simple proteinlike HP chains, that the stability of the native-state of an adsorbed protein can be increased by reducing the lateral volume (i.e. along the sorbent surface) available to the adsorbed protein to adopt extended conformations. The generality of our model prediction is supported by a number of previous studies $(10,15-26)$ on the adsorption of a
number of globular proteins to different sorbents which show that structural changes in adsorbed proteins tend to decrease with increasing surface coverage.

When qualitatively linked to the experimental results of Eggers and Valentine (36) and those reported in Figure 3.6, our dMC simulations on protein-like HP chains suggest that any physically meaningful isotherm model for globular protein adsorption to solid-liquid interfaces must account for the dependence of the adsorbed-protein partition function on surface concentration. Confinement reduces both the number of realistically accessible energies, $E_{i}$, and, to an even greater degree, the conformational degeneracy of each accessible $E_{i}$. As a result, when an adsorbed protein becomes confined, its partition function, which represents the weighted sum of all possible chain conformations, decreases substantially.

Our simulation results also suggest strategies for improving the performance of technologies and processes based on protein adsorption. For example, operation of selective adsorptive chromatography columns under shock-wave, high-feedconcentration conditions may serve to reduce irreversible denaturation and inactivation of desired protein products during their purification, particularly in the case of proteins of relatively low native-state stability in solution. Similarly, entrapment of purified proteins through adsorption into volumetrically defined matrices may serve as a powerful formulation tool for long-term storage of protein products in their functionally active form.

Recent experimental data indicate that nonspecific protein adsorption to hydrophobic sorbents can result in an observed increase in secondary structures (20). Our HP model is not capable of defining chain secondary structure at a level where one can observe distinct changes in either $\alpha$-helix or $\beta$-sheet content. Instead, the model provides a measure of segment packing densities. Our simulations show that under certain adsorption conditions, particularly at higher surface coverages, the average segment packing density within the adsorbed chain is similar to or greater than that in solution. Thus, although denaturation of the chain occurs, it does not necessarily lead to a more open average chain conformation on the sorbent surface. Instead it leads to an ensemble of non-native, high segment density conformers of low energy and a concomitant reduction in the conformational entropy of the chain.

### 3.4.2 Influence of lateral confinement on the thermodynamics of HP chain adsorption

As reported above, our dMC simulations of simple protein-like HP chains agree with the previous work of Zhou and Dill by predicting that adsorption of chains possessing highly stable native-state conformations in solution to a restricted volume for which the confining walls are either inert or offer only a very weak attraction for the chain results in stabilization of the native-state conformation by shifting the equilibrium away from the denatured state. For the HP chain under adsorption conditions where $\chi_{H H}<-4$ and $-1<$ $\chi_{H W} \leq 0$, the ensemble averaged energy of adsorption $\Delta U / k T$ approaches the energy difference for side-on adsorption of the native-state as $x_{S}$ nears 4 , the edge-length of the chain in its native state. Thus, the chain has little tendency to change conformation
during adsorption and adsorption isotherm models that ignore conformational changes in the chain are adequate.

However, our dMC simulation results also suggest that adsorption thermodynamics will quickly become intimately linked with perturbations in chain conformation when the stability of the native state is reduced, the attraction for the sorbent surface is increased, or average sorbent surface area available per protein macromolecule is increased. Here, we report adsorption thermodynamics for these more interesting cases.

For a weakly stable $\left(\chi_{H H}=-1\right)$ native state of the HP chain adsorbing to a weakly attractive sorbent surface $\left(\chi_{H W}=-1\right)$, ensemble-averaged thermodynamic properties calculated from our dMC simulation data indicate that binding is both energetically $\left(\Delta U / k T \equiv \Delta<E>/ k T=-2.1 \pm 0.1\right.$ when $\left.x_{S}=18\right)$ and entropically $(\Delta S / k \equiv \Delta\langle S>/ k=3.7 \pm 0.1)$ favored, particularly when $x_{S}$ is large (Figure 3.7). The favorable entropy of adsorption arises in this case because the adsorbed chain accesses a significantly larger number of conformations and associated energies $E_{i}$ than does the native-state chain in solution.

Energy indexing of all chain conformations shows that laterally-stretched conformations of the adsorbed chain preferentially populate energy states ( $E_{i}$ values) near the global energy minimum $E_{\text {min }}$ of the system. When $x_{S}$ is reduced sufficiently, these stretched conformations are no longer accessible and the probability of finding the system in the associated low-energy state is either substantially decreased or zero. As a result, $\Delta U / k T$ is a function of surface coverage. The entropic contribution to adsorption $\Delta S / k$ also
decreases with decreasing $x_{S}$ due to the inherent loss in conformational degrees of freedom that accompanies a reduction in free volume.

Evidence of a change in adsorption energy with increasing surface coverage is provided in Table 3.1, which shows the dependence of the molar enthalpy of adsorption $\left(\Delta H_{\text {ads }}\right)$, measured by isothermal titration calorimetry, for binding of HEWL to silica as a function of percent surface coverage. At low surface coverage, $\Delta H_{a d s}$ (per mol of adsorbed protein) is large and endothermic, indicative of an adsorption process which results in significant changes in protein conformation and the disruption of a large number of favorable intramolecular contacts. At high surface coverage, $\Delta H_{\text {ads }}$ becomes exothermic, in qualitative agreement with our simulation results which predict that adsorption thermodynamics (and the ensemble of conformations accessible to the adsorbed protein) will vary with percent surface coverage.

Adsorption thermodynamics of the HP chain and the influence of confinement are quite sensitive to the stability of the native state of the chain in solution, and the energy of attraction between hydrophobic segments of the chain and the sorbent surface. In the system reported in Figure 3.7, $\Delta U / k T$ and $\Delta S / k$ are both favorable and their sensitivities to available sorbent surface area both influence the dependence of the overall driving force for adsorption $\Delta A / k T$ on $x_{s}$. In contrast, Figure 3.8 reports adsorption thermodynamics regressed from dMC simulation data for the HP chain when $\chi_{H H}=\chi_{H W}=-4$. In this case, the formation of intramolecular HH and intermolecular HW contacts are both highly favorable and the dominant contribution to $\Delta A / k T$ is provided by the energy change
$\Delta U / k T$. As a result, adsorbed chain conformations that maximize the total sum of HH and HW contacts are favored and the probability of finding the system at or near $E_{\min }$ is high. Since the sorbent surface is rigid and impenetrable, the chain must adopt very specific conformations and points of contact to access these low energy states. At $x_{S}=$ 18 , the simulation results in Figure 3.8 therefore show the entropy of adsorption $\Delta S / k$ to be unfavorable; that is, the conformational entropy of the HP chain is lower in the adsorbed state than in solution. Although the conformation of the adsorbed chain is far from its native state, the chain entropy is lower than in solution due to the conformational restrictions imposed by the need to maximize energetically favorable contacts.

When $x_{S}$ is reduced, $\Delta S / k$ increases linearly, as opposed to the highly nonlinear dependence observed when both $\chi_{H H}$ and $\chi_{H W}$ are set to -1 . Our dMC results indicate that this linear trend arises because, although it does decrease as $E_{i}$ moves sufficiently far away from $E_{\text {min }}$, the weighted contribution of laterally elongated conformations of the HP chain at each $E_{i}$ remains fairly constant for energies near $E_{\text {min }}$. As the accessible energy landscape for the system is effectively restricted to these low $E_{i}, \Delta S / k$ exhibits a linear dependence on $x_{s}$.

Finally, Figure 3.9 reports ensemble-averaged adsorption thermodynamics for the HP chain when the attraction of the sorbent surface for hydrophobic $(\mathrm{H})$ segments of the HP chain is stronger than that between two H segments $\left(\chi_{H W}=-2\right.$ and $\left.\chi_{H H}=-1\right)$. In this asymmetric system, the adsorption energy, $\Delta U / k T$, favors breakage of $n \mathrm{HH}$ contacts to form $0.5 n+1$ or a greater number of HW contacts. At equilibrium, the average number
of HW contacts therefore increases relative to that observed in the two symmetric adsorption systems described above, and the weighted contribution of laterally stretched conformations of the chain also increases at all $E_{i}$ since these conformations favor formation of HW contacts. As a result, $\Delta U / k T$ and $\Delta A / k T$ depend more strongly on $x_{S}$, increasing with decreasing sorbent surface area at all $x_{S}$ less than 18 , the length of the fully stretched chain.

### 3.5 Summary

Dynamic Monte Carlo simulations of a protein-like HP chain were used to investigate the influence of lateral chain confinement on adsorption thermodynamics and adsorbed chain conformational space. Adsorption results in a net loss (generally substantial) of nativestate conformation when the volume available for adsorption, defined by the volumetric space proximal to and directly above the available sorbent surface area per adsorbing chain, is large compared to the fully stretched chain length. Confinement of the adsorbed chain is shown to dramatically stabilize the native-state conformation due to selective removal of denatured (particularly elongated) chain conformations from the ensemble of accessible states. For the case where the native-state conformation of the chain is relatively stable in solution and its energy of attraction to the sorbent surface is relatively weak, adsorption to an expansive sorbent surface results in destabilization of the native-state. However, increasing lateral confinement of the adsorbed chain leads to hyperstabilization of the native-state conformation of the chain in the adsorbed state relative to free in solution. This stabilization effect correlates with a loss in the
conformational degeneracy of favorable low energy states that results in an increased probability of finding the adsorbed chain in its native-state conformation.

Lateral confinement of the adsorbed chain has a more complicated effect on overall adsorption thermodynamics, with $\Delta U / k T$ and $\Delta S / k$ typically showing significantly different dependences on $x_{s}$. When the sorbent surface is relatively hydrophilic, $\Delta S / k$ becomes progressively more unfavorable as $x_{S}$ is decreased, consistent with the decreased free volume and the associated hyperstabilization of the unique native-state conformation. For adsorption to hydrophobic surfaces, however, $\Delta S / k$ is thermodynamically unfavorable at all $x_{S}$, indicating that the number of conformations accessible to the chain in the adsorbed state is largely dictated by the energetic penalty associated with conformations that lead to a net reduction in the total number of favorable intermolecular HW contacts.

### 3.6 References

1. Vroman, L., Blood. Natural History Press, Garden City, NY, 1967.
2. Vroman, L. and Adams, A.L., Adsorption of proteins out of plasma and solutions in narrow spaces. J. Colloid Interface Sci., 111 (1986) 391-402.
3. Norde, W. and Lyklema, J., The adsorption of human plasma albumin and bovine pancreas ribonuclease at negatively changed polystyrene surfaces: I. Adsorption isotherms, effects of charge, ionic strength and temperature. J. Colloid Interface Sci., 66 (1978) 257-265.
4. Norde, W. and Lyklema, J., The adsorption of human plasma albumin and bovine pancreas ribonuclease at negatively changed polystyrene surfaces: II. Hydrogen ion titrations. J. Colloid Interface Sci., 66 (1978) 266-276.
5. Norde, W. and Lyklema, J., The adsorption of human plasma albumin and bovine pancreas ribonuclease at negatively changed polystyrene surfaces: III. Electrophoresis. J. Colloid Interface Sci., 66 (1978) 277-284.
6. Norde, W. and Lyklema, J., The adsorption of human plasma albumin and bovine pancreas ribonuclease at negatively changed polystyrene surfaces: IV. The charge distribution in the adsorbed state. J. Colloid Interface Sci., 66 (1978) 285-294.
7. Norde, W. and Lyklema, J., The adsorption of human plasma albumin and bovine pancreas ribonuclease at negatively changed polystyrene surfaces: V. Microcalorimetry. J. Colloid Interface Sci., 66 (1978) 295-302.
8. Norde, W. and Lyklema, J., Why proteins prefer interfaces. J. Biomater. Sci. Polymer Edn., 2 (1991) 183-202.
9. Haynes, C.A. and Norde, W., Globular protein at solid/liquid interfaces. Colloids Surfaces B, 2 (1994) 517-566.
10. Malmsten, M., Formation of adsorbed protein layers. J. Colloid Interface Sci., 207 (1998) 186-99.
11. Wertz, C.F. and Santore, M.M., Effect of surface hydrophobicity on adsorption and relaxation kinetics of albumin and fibrinogen: Single-species and competitive behaviour. Langmuir, 17 (2001) 3006-3016.
12. Vermeer, A.W.P., Giacomelli, C.E. and Norde, W., Adsorption of IgG onto hydrophobic teflon. Differences between the Fab and Fc domains. Biochim. Biophys. Acta, 1526 (2001) 61-69.
13. Tian, M., Lee, W.-K., Bothwell, M.K. and McGuire, J., Structural stability effects on adsorption of bacteriophage T4 lysozyme to colloidal silica. J. Colloid Interface Sci., 200 (1998) 146-154.
14. Billsten, P., Carlsson, U., Jonsson, B.-H., Olofsson, G., Hook, F. and Elwing, H., Conformation of human carbonic anhydrase II variants adsorbed to silica nanoparticles. Langmuir, 15 (1999) 6395-6399.
15. Galisteo, F. and Norde, W., Adsorption of lysoyzme and alpha-lactalbumin on poly(styrenesulphonate) latices 2. Proton titrations. Colloid. Surface. B, 4 (1995) 389-400.
16. Kondo, A., Urabe, T. and Yoshinaga, K., Adsorption activity and conformation of alpha-amylase on various ultrafine silica particles modified with polymer silane coupling agents. Colloids Surfaces A, 109 (1996) 129-36.
17. Norde, W. and Favier, J.P., Structure of adsorbed and desorbed proteins. Colloids Surfaces A, 64 (1992) 87-93.
18. Zoungrana, T. and Norde, W., Thermal stability and enzymatic activity of alphachymotrypsin adsorbed on polystyrene surfaces. Colloids Surfaces B, 9 (1997) 157167.
19. Sane, S.U., Cramer, S.M. and Przybycien, T.M., Protein structure perturbations on chromatographic surfaces. J. Chromatogr. A, 849 (1999) 149-159.
20. McNay, J.L. and Fernandez, E.J., Protein unfolding during reversed-phase chromatography: I. Effect of surface properties and duration of adsorption. Biotech. Bioeng., 76 (2001) 225-32.
21. Robeson, J.L. and Tilton, R.D.; Spontaneous reconfiguration of adsorbed lysozyme layers observed by total internal reflection fluorescence with a pH -sensitive fluorophore. Langmuir, 12 (1996) 6104-6113.
22. Buijs, J. and Hlady, V., Adsorption kinetics, conformation, and mobility of the growth hormone and lysozyme on solid surfaces, studied with TIRF. J. Colloid Interface Sci., 190 (1997) 171-181.
23. Cullen, D.C. and Lowe, C.R., AFM studies of protein adsorption .1. Time-resolved protein adsorption to highly oriented Pyrolytic-Graphite. J. Colloid Interface Sci., 166 (1994) 102-108.
24. Sagvolden, G., Giaever, I. and Feder, J., Characteristic protein adhesion forces on glass and polystyrene substrates by atomic force microscopy. Langmuir, 14 (1998) 5984-5987.
25. Elwing, H., Protein adsorption and ellipsometry in biomaterial research. Biomaterials, 19 (1998) 397-406.
26. Giacomelli, C.E., Esplandiu, M.J., Ortiz, P.I., Avena, M.J. and De Pauli, C.P., Ellipsometric study of bovine serum albumin adsorbed onto $\mathrm{Ti} / \mathrm{TiO}_{2}$ electrodes. $J$. Colloid Interface Sci., 218 (1999) 404-411.
27. Malmsten, M. and Lassen, B.,Ellipsometry studies of protein adsorption at hydrophobic surfaces. In "Proteins at Interfaces II: Fundamentals and Applications" (T.A. Horbett and J.L. Brash, Eds.), p.228-238. American Chemical Society, Washington, DC, 1995.
28. Lau, K.F. and Dill, K.A., A lattice statistical mechanics model of the conformational and sequence spaces of proteins. Macromolecules, 22 (1989) 3986-3997.
29. Liu, S.M. and Haynes, C.A., Mesoscopic analysis of conformational and entropic contributions to nonspecific adsorption of HP copolymer chains using dynamic Monte Carlo simulations. J. Colloid Interface Sci., In press (2004)
30. Zhdanov, V.P. and Kasemo, B., Simulations of denaturation of adsorbed proteins. Phys. Rev. E, 56 (1997) 2306-9.
31. Anderson, R.E., Pande, V.S. and Radke, C.J., Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. J. Chem. Phys., 112 (2000) 9167-85.
32. Castells, V., Yang, S. and Van Tassel, P.R., Surface-induced conformational changes in lattice model proteins by Monte Carlo simulation. Phys. Rev. E, 65 (2002)
33. Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D. and Chan, H.S., Principles of protein folding - A perspective from simple exact models. Protein Sci., 4 (1995) 561-602.
34. Nanjo, F., Sakai, K. and Usui, T., p-nitrophenyl penta-N-acetyl-beta-chitopentaoside as a novel synthetic substrate for the colorimetric assay of lysozyme. J. Biochemistry, 104 (1988) 255-258.
35. Zhou, H.-X. and Dill, K.A., Stabilization of proteins in confined spaces. Biochemistry, 40 (2001) 11289-93.
36. Eggers, D.K. and Valentine, J.S., Molecular confinement influences protein structure and enhances thermal protein stability. Protein Sci., 10 (2001) 250-261.

### 3.7 Tables

Table 3.1: Molar enthalpy change $\Delta H_{a d s}$ as a function of percent sorbent surface coverage for the adsorption of HEWL to particulate silica in $50-\mathrm{mM} \mathrm{KCl}(\mathrm{pH} 7)$ at $37^{\circ} \mathrm{C}$. $\Delta H_{a d s}$ is expressed on a per mole of HEWL adsorbed basis.

| $100 \times \Gamma / \Gamma^{\max }$ | $\begin{gathered} \Delta H_{a d s} \\ \left(\mathrm{kcal} \mathrm{~mol}^{-1}\right) \end{gathered}$ |
| :---: | :---: |
| 4.5 | $8.1 \pm 6$ |
| 9.2 | $8.0 \pm 5$ |
| 18.2 | $7.7 \pm 6$ |
| 32.1 | $4.8 \pm 4$ |
| 63.8 | $0.2 \pm 4$ |
| 95.0 | $-3.4 \pm 5$ |

### 3.8 Figures

Figure 3.1: Schematic diagram of the model HP chain. Filled circles represent hydrophobic ( H ) chain segments while open circles represent polar ( P ) chain segments.


Figure 3.2: Energy trajectory for adsorption of the HP chain to a planar sorbent surface when $\chi_{H H}=-1$ and $\chi_{H W}=-4$. All other segment-segment interaction energies set equal to zero.


Figure 3.3: Dependence of the degeneracy of the lowest energy state (i.e., the total number of unique chain conformations) on accessible sorbent surface area and volumetric space above it. The $x$-axis indicates the number of lattice sites on the sorbent surface available for binding. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-$ 1.


Figure 3.4: Energy probability histograms for the HP chain adsorbed to sorbent surfaces of two different widths: etched grey bars - width of sorbent surface, $x_{S}$, is 4 lattice units; solid black bars $-x_{S}$ is 18 lattice units. The error bars refer to the standard deviation of 5 runs. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-1$.


Figure 3.5: Ratio of the probability of the HP chain being in its native-state conformation when adsorbed, $P_{N S}{ }^{\text {adsorbed }}$, to that in solution, $P_{N S}{ }^{\text {solution }}$, as a function of the width of the available adsorption site, $x_{s}$. Non-zero segment-segment interaction energies are $\chi_{H H}=$ $\chi_{H W}=-4$ (open squares), and $\chi_{H H}=-4, \chi_{H W}=-1$ (open circles). The lines drawn indicate the trends of the data.


Figure 3.6: HEWL-catalyzed pNP-C5 hydrolysis kinetics when HEWL is dissolved in aqueous solution ( $\mathrm{pH} 6,22{ }^{\circ} \mathrm{C}$ ) and when HEWL is nonspecifically adsorbed to particulate silica at different levels of surface coverage. In all experiments, the amount of HEWL and the initial concentration of pNP-C5 are held constant at $50 \mu \mathrm{~g}$ and 0.20 mM , respectively. Shown are measurements taken for HEWL in solution with no surface (squares), HEWL adsorbed on silica at monolayer coverage (circles), and HEWL adsorbed at $20 \%$ monolayer coverage (triangles).


Figure 3.7: Thermodynamics of adsorption of the HP chain to a planar sorbent surface as a function of accessible sorbent surface area and volumetric space above it. The $x$ - axis indicates the number of lattice sites on the sorbent surface available for binding. Nonzero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-1$. Trends indicated are $\Delta U / k T$ (solid line), $\Delta A / k$ (dashed line) and $\Delta S / k$ (dash-dotted line). To avoid excessive clutter, simulation points are shown for $\triangle U / k T$ only. Error bars are calculated from the standard deviation of 5 runs.


Figure 3.8: Thermodynamics of adsorption of the HP chain to a planar sorbent surface as a function of accessible sorbent surface area and volumetric space above it. Non-zero segment-segment interaction energies are $\chi_{H H}=\chi_{H W}=-4$. Symbology is the same as in Figure 3.7.


Figure 3.9: Thermodynamics of adsorption of the HP chain to a planar sorbent surface as a function of accessible sorbent surface area and volumetric space above it. Non-zero segment-segment interaction energies are $\chi_{H H}=-1, \chi_{H W}=-2$. Symbology is the same as in Figure 3.7.


## 4 Energy landscapes for adsorption of protein-like HP chains as a function of native-state stability*

### 4.1 Introduction

More than a decade ago, Arai and Norde (1) published their landmark paper describing the dependence of globular-protein adsorption on the thermodynamic stability of the protein in solution at adsorption temperature and pH . Their experimental data indicate that although changes in protein conformation are likely to occur in all adsorption events, the perturbation of a protein's conformation away from the native state on adsorption tends to increase with decreasing thermodynamic stability of the native state in solution. This effect correlated well with their adsorption isotherm data, which revealed a general tendency for less stable proteins to adsorb with higher affinity, particularly in systems where the sorbent surface is weakly attractive. To specifically capture the contribution of native-state thermal stability, Arai and Norde coined the term "soft" and "hard" proteins, with the degree of softness reflecting the susceptibility to denaturing changes in protein conformation upon adsorption. While this concept is perhaps a bit too simplistic, it has remained a part of protein adsorption dogma, in part because it provides a concise and qualitatively useful measure of the importance of protein conformation and stability to the overall adsorption process.

[^2]More recent studies by Malmsten and others $(2,3)$ have contributed further to our understanding of the connection between structural stability and adsorption thermodynamics. For example, Haynes and Norde (4) used differential scanning microcalorimetry to compare denaturation enthalpies of adsorbed hen egg-white lysozyme and bovine milk $\alpha$-lactalbumin. Their results revealed that the thermodynamically less stable protein, $\alpha$-lactalbumin, adsorbed more strongly and denatured more extensively on hematite, a weakly attractive hydrophilic surface.

These findings are supported by a number of adsorption studies involving variations in solution conditions which indicate that proteins generally adsorb with higher affinity under conditions where the temperature $(5,6)$ or solution $\mathrm{pH}(7)$ render the protein less stable. As well, the addition of an ion known to bind specifically and thereby stabilize (through mass action effects) the native state of a protein has been shown to reduce the net force of adhesion between a protein and a sorbent surface $(8,9)$.

Although comparisons of different proteins over a range of adsorption conditions are useful, interpretation of the results is limited by the fact that factors other than protein stability are altered as well. Perhaps a more specific approach to understanding the connection between protein structural stability and adsorption thermodynamics is to study the adsorption of a family of site-directed variants of a protein designed to alter the native-state stability of that protein. McGuire and coworkers investigated the adsorption of wild-type T4 bacteriophage lysozyme and a series of single-site and multi-site mutants of lysozyme to silica and mica surfaces (10-14). Their circular dichroism data indicate
that proteins of lower native-state stability lose larger amounts of $\alpha$-helix content during adsorption (10, 12, 13). Interferometric surface-force measurements on the same adsorption system reinforce this conclusion by showing that the tertiary structures of the lower stability mutants are more severely compromised during adsorption to the negatively charged mica surface, resulting in significantly stronger forces of adhesion (12). Surface-force measurements on T4 lysozyme variants have also been used to show that lower stability variants displace proteins of higher stability during adsorption, indicating a larger force of adhesion for low stability proteins (14).

Billsten et al. $(15,16)$ observed similar trends for the adsorption of site-directed variants of human carbonic anhydrase II to silica. Using a combination of circular dichroism, fluorescence spectroscopy and differential scanning calorimetry, they showed that adsorption resulted in much larger changes in the conformation of the less stable variants.

Precisely how protein adhesion forces are strengthened by a decrease in native-state stability remains unclear, in part because the contributions of the various reaction subprocesses (e.g., changes in chain entropy, dehydration effects, etc.) cannot be precisely defined. Previously, we have shown that two-dimensional (2D) dynamic Monte Carlo (dMC) simulations of uniquely folding copolymers composed of a linear sequence of hydrophobic $(\mathrm{H})$ and polar $(\mathrm{P})$ segments (i.e., the HP chain model of Dill (17)) allow one to unambiguously compute changes in system entropy and energy accompanying nonspecific adsorption of the chain (18). Although they are based on an idealized protein-like chain, simulation results for this simple model add to our general
knowledge of the adsorption process by providing molecular-level insights and the ability to enumerate all conformations and energies available to the chain in solution and adsorbed to the sorbent surface. This is something that cannot at present be done with real proteins.

Here, we use dMC simulations to improve our general understanding of the thermodynamics of nonspecific adsorption of a protein-like HP chain and its dependence on the stability of the native-state conformation of the adsorbing chain. Of particular concern to us is how the stability of the native-state conformation limits changes in chain conformation and chain conformational entropy upon adsorption. Calorimetric studies of nonspecific protein adsorption reveal that the process often results in a net increase in system entropy (4). However, the source of this entropy gain and its dependence on chain stability remain unclear, as a number of subprocesses, including dehydration of ordered water molecules and an increase in chain conformational entropy, could be at least in part responsible $(2,19)$.

### 4.2 Protein-like HP chain and dMC simulation algorithm

Simulation results reported here are for two linear copolymers (Figure 4.1), sequence I and II, both having specific sequences of hydrophobic $(\mathrm{H})$ and polar $(\mathrm{P})$ residues. The basic chain architectures are drawn from the HP chain model of Dill and coworkers (17, 20), and the sequences are designed so that the chains fold at their global energy minimum into unique compact conformations (hereafter referred to as their native states, respectively) (18). HP chains like that shown in Figure 4.1 have been studied extensively
by Dill (17, 20) and many others (see for instance, 21-23) using Monte Carlo and molecular dynamics simulation techniques. Despite their obvious simplicity, the folding dynamics and solution thermodynamics of these model chains have significantly enriched our fundamental understanding of protein folding (20-23) and other macromolecular association events (24, 25).

The details of the algorithm and fundamental equations used in our dynamic Monte Carlo simulations of HP chain adsorption to a planar solid-liquid interface are described in a previous paper (18). We therefore restrict ourselves to a brief description of the simulation method and its specific application to the interrogation of the influence of the chain's native-state stability on adsorption thermodynamics. A single protein-like HP chain is placed on a Cartesian 2D lattice of sufficient size to allow all possible chain conformations and to prevent the chain from interacting with more than one lattice boundary (which serves as the sorbent surface) during the simulation. The four walls of the lattice, including the one selected as the sorbent surface, are impenetrable and reflective boundary conditions are employed. The system energy is defined by Florytype interaction energies between adjacent components within the lattice and its boundaries. These components include hydrophobic $(\mathrm{H})$ and polar ( P ) chain segments, the sorbent surface (W) and solvent (S).

In all simulations reported, folding of the HP chain into its native state in solution is exclusively driven by the value of $\chi_{H H}$, the Flory parameter between two H residues not directly connected within the chain. Increasingly negative values of $\chi_{H H}$ shift equilibrium
toward stabilization of the native-state conformation. Solvent interactions are assumed to be athermal (i.e., $\chi_{H S}=0$ ) and hydration effects in the model are therefore represented only indirectly, since the favourable HH interaction (and HW interaction when the sorbent wall is active) implicitly makes the HS interaction (and the SW interaction) net unfavourable.

Unless otherwise stated, all dMC simulations were run for $1 \times 10^{9}$ or more cycles, with system equilibration typically observed within the first $5 \times 10^{8}$ cycles. Samples were taken every 5000 cycles, giving $1 \times 10^{5}$ or more data points for each run. Each model condition was simulated 5 or more times and average values are reported. The change in system energy upon chain adsorption is computed from the resulting energy distribution functions for two well-defined conditions: the chain is initially placed within the lattice in its lowest energy conformation and all boundaries are athermal (initial state), and a previously athermal wall is made attractive to H -segments of the chain (final state). The free energy change $\Delta_{a d s} A$ and entropy change $\Delta_{a d s} S$ for this process are then computed using standard thermodynamic integration algorithms described by Allen and Tildesley (26) and previously used by Socci and Onuchic (22).

### 4.3 Results and discussion

### 4.3.1 HP chain adsorption thermodynamics

### 4.3.1.1 Adsorption of sequence I

Table 4.1 reports the calculated change in Helmholtz free energy, $\Delta_{a d s} A$, internal energy, $\Delta_{a d s} U$, and entropy, $\Delta_{a d s} S$, resulting from adsorption of sequence I to planar sorbent surfaces with different degrees of attraction for H segments of the chain. Consistent with the soft and hard protein model originally proposed by Arai and Norde based on their experimental studies of nonspecific protein adsorption (1), the mesoscopic adsorption thermodynamics for our simple protein-like HP chain reveal a strong correlation between adsorption affinity and thermal stability of the native-state conformation of the chain in solution. The value of $\Delta_{\text {ads }} A$, which provides a measure of the overall affinity of the chain for the surface, moves toward positive values with increasing stability of the nativestate conformation, irrespective of the hydrophobicity of the sorbent surface. For adsorption on the weakly attractive sorbent $\left(\chi_{H W}=-1\right), \Delta_{a d s} A$ increases from $-5.67 k T$ to 3.77 kT , and therefore becomes thermodynamically less favourable, with a change in $\chi_{H H}$ from -1 to -4 . On the higher affinity surface $\left(\chi_{H W}=-4\right)$, the change in $\Delta_{\text {ads }} A$ is even more pronounced, with the affinity reduced by nearly half when $\chi_{H H}$ is changed from -1 to -4 .

In contrast to the more complex process of nonspecific protein adsorption, simple thermodynamic integration algorithms (26) can be used to compute $\Delta_{a d s} A$ for the adsorption of our simple HP chain, allowing one to identify the molecular basis for the
observed dependence of both $\Delta_{a d s} A$ and $\Delta_{a d s} S$ on the thermal stability of the native-state conformation. In our model, the overall change in system entropy, $\Delta_{\text {ads }} S$, reflects both the change in chain conformational entropy upon adsorption and the increase in the total degrees of freedom that accompanies introduction of the sorbent surface into the system. Figure 4.2 reports both $\Delta_{a d s} A / k T$ and $\Delta_{a d s} S / k$ as a function of $\Delta_{D-N} A / k T$, the Helmholtz energy difference between the native and fully denatured (zero energy) conformations of sequence I for the case where $\chi_{H W}=-1$. The dMC data reveal a transition in both $\Delta_{a d s} A / k T$ and $\Delta_{a d s} S / k$ centered near $\Delta_{D-N} A / k T=-12$. At low absolute values of $\Delta_{D-N} A / k T$ (i.e. low native-state stability), the chain in solution trades weakly favourable intramolecular HH interactions for conformational entropy. As a result, the chain in solution forms an average of only ca. 4 HH contacts, significantly less than the 9 HH contacts that define its lowest-energy (native state) conformation. Along with hydrophobic residues on the surface of the native state, additional unpaired hydrophobic residues are therefore present, and each may form an intermolecular HW contact to reduce the energy of the system.

The system energy may also be lowered by reducing the average number of HH contacts to form a set of lower energy HW contacts, and this concept has been applied frequently in the protein adsorption literature. In particular, many have correctly argued that disruption of specific intramolecular interactions to form more favourable intermolecular contacts can perturb the average chain conformation away from the native state towards a larger density of (denatured) conformational states on the sorbent surface $(4,27)$. In such cases, adsorption may be expected to result in a net increase in entropy due to a net
decrease in HH contacts. Our dMC simulations identify certain adsorption conditions where a net decrease in intramolecular (HH) contacts is observed, but rarely does this decrease lead to an increase in chain entropy. Moreover, as shown in Table 4.2, a net decrease in HH contacts often does not occur upon adsorption of sequence I, even in cases where an increase in entropy is observed. For instance, adsorption of HP chain sequence I when $\chi_{H H}=\chi_{H W}=-1$ results in relatively little change in the average number of HH contacts within the chain, but favourable changes in both $\Delta_{a d s} U / k T$ and $\Delta_{a d s} S / k$ are observed. The resulting decrease in $\Delta_{a d 5} A / k T$ that drives adsorption is not due to an increase in chain entropy resulting from disruption of specific intramolecular contacts. Rather, it results from the energy decrease and entropy increase generated by the large density of chain conformations that allow unpaired hydrophobic residues (including solvent-exposed residues in the native state structure) to form contacts with the sorbent.

In contrast, at high absolute values of $\Delta_{D-N} A / k T$, such as when $\chi_{H H}=-4$, sequence I in solution forms an average of 8 HH contacts and is often observed in its lowest energy conformation. Conformations of the chain in solution which expose additional (relative to those exposed on the surface of the native state) unpaired H residues to the sorbent surface are therefore greatly reduced. Moreover, for adsorption of this more stable chain to the $\chi_{H W}=-1$ surface, breakage of an intramolecular HH contact is energetically unfavourable unless a significantly greater number of intermolecular HW contacts can be formed as a result. Thus, as shown in Table 4.2, we again observe no net reduction in HH contacts upon chain adsorption. Conformations the adsorbed chain can adopt to lower the system energy are therefore limited by the general need to form specific
intermolecular contacts with the sorbent while retaining highly favourable intramolecular HH contacts. When the sorbent surface is weakly attractive $\left(\chi_{H W}=-1\right)$, adsorption then results in a small loss in entropy $\left(\Delta_{a d s} S<0\right)$ and is driven entirely by the decrease in internal energy that accompanies chain adhesion.

In the transition region, the system (both in the presence and absence of the sorbent surface) responds to $\chi_{H H}$ taking on more negative values by trading access to high-energy/high-entropy states to increase the number of energetically favourable HH contacts. The derivatives
$\frac{\partial \Delta_{a d s} U / k T}{\partial \chi_{H H}}$ and $\frac{\partial \Delta_{a d S} S / k}{\partial \chi_{H H}}$.
carry the same sign; both are positive and thereby confer on the adsorption process a weak form of energy-entropy compensation.

A further examination of the statistical mechanical definition of $S$ and $U$ in terms of the partition function shows that they depend in the same qualitative way on the distribution of the system among different energy levels. If the system is closed and the perturbation adiabatic (e.g., a change in system temperature), the entropy must decrease as the mean energy of the system decreases since no new degrees of freedom have been added into the system (28). In our system, however, the perturbation is isothermal and involves the introduction of the sorbent surface. As a result, the total degrees of freedom are
significantly increased and positive values of $\Delta_{a d s} S$ can be observed despite a decrease in the mean energy.

### 4.3.1.2 Adsorption of sequence II

Sequence II differs significantly from sequence I. While sequence I is overall quite hydrophobic with an amphipolar surface in its native state, sequence II is more hydrophilic and symmetric in nature, displaying a completely hydrophilic surface and hydrophobic core in its native state. Nevertheless, adsorption thermodynamics for sequence II are qualitatively similar to those reported for sequence I (Tables 4.3 and 4.4). For example, adsorption affinity again weakens with increasing thermodynamic stability of the native-state conformation, irrespective of the attractiveness of the sorbent surface (Table 4.3). This effect is sufficiently strong to disfavor adsorption of sequence II when the native-state conformation is highly stable $\left(\chi_{H H}=-4\right)$ and the attraction of the sorbent for H segments of the chain is weak $\left(\chi_{H W}=-1\right)$. In this case, non-native chain conformations that allow contact between hydrophobic residues of the chain and the sorbent surface are disfavoured relative to the native-state conformation.

Adsorption of sequence II is entropically favored under several conditions (Table 4.3). For example, when $\chi_{H H}=\chi_{H W}=-4, \Delta_{a d s} S$ is positive, due in large part to the associated net decrease in HH contacts (Table 4.4) that allows the chain to access a significantly larger density of conformational states on the sorbent surface. This mechanism for increasing the system entropy is well known, primarily through the work of Norde and
coworkers $[1,2,4]$. However, as with sequence I, adsorption of sequence II may lead to a favourable $\Delta_{\text {adS }} S$ through other mechanisms. When $\chi_{H H}=\chi_{H W}=-1, \Delta_{a d S} S / k=+5.62$ and adsorption is driven purely by entropy (i.e., $\Delta_{a d s} U / k T \approx 0$ ), despite a small to insignificant increase in the average number of HH contacts. The origin of this favourable $\Delta_{a d s} S$ is less obvious, but can be understood through more careful analysis of the energy landscapes for the system when the sorbent wall is first athermal (initial state) and then attractive (final adsorbed state).

### 4.3.2 Thermally-averaged energy landscape analysis of HP chain adsorption

Energy landscape analysis has become a central tool in understanding the folding of HP and other protein-like chains $(20,29,30)$, allowing one, for instance, to visualize the ensemble of parallel pathways a chain may follow to fold into its native-state conformation. For relatively simple protein-like chains possessing a global energy minimum occupied by a single chain conformation, the energy landscape for the chain in solution has been shown to resemble a funnel, with the lowest energy state, the native state, occupied with a large Boltzmann weight at temperatures well below the native-todenatured state transition temperature, $T_{m}$, but still high enough that the chain folding kinetics are not limited by the inability of the chain to escape from conformations that represent local energy minima.

In its simplest thermally-averaged form, energy landscape analysis of chain folding yields a symmetric funnel whose shape is specified by the density of conformational states of the chain as a function of the system energy. Due to its symmetry, the energy landscape can be displayed by plotting the contour of the funnel wall without loss of information content. Figure 4.3 compares the thermally-averaged energy landscape for HP chain sequence I in solution to that for the same chain adsorbed to the sorbent surface when $\chi_{H H}=\chi_{H W}=-1$. Comparing the volumes of the two funnels shows that although the total number of possible chain conformations is the same in the two systems, introduction of the sorbent surface increases the total density of unique states to $c a .4 .6$ times that for the same chain in solution. This increase in total degrees of freedom is observed because, while a given chain conformation has a single energy in solution, it can be found at different system energies in the presence of a sorbent surface depending on its position and orientation relative to the surface. In a 2 D lattice with a planar sorbent surface, a given chain conformation can reside either off the surface with an energy identical to that in the sorbent-free system (initial state), or adsorbed to the surface in one of four possible orientations. Thus, when the sorbent surface is present, any given chain conformation can appear in at most five different energy levels, with the average for all chain conformations being 4.6 different energies due to conformational symmetry effects. The adsorbed-state system $(a d s)$ thereby has a higher or equal density of states $\left(\Omega_{i}\right)$ at every energy level $E_{i}$. This is the dominant reason why, under certain conditions, $\Delta_{\text {ads }} S$ can be positive in an adsorption process that lowers the mean energy of the system.

Because $\Omega_{i, a d s}\left(N, V, E_{i}\right) \geq \Omega_{i, \text { free }}\left(N, V, E_{i}\right)$ at all $E_{i}$,
$Q_{a d s}-Q_{f r e e}=\sum_{E_{1}}\left(\Omega_{i, a d s}\left(N, V, E_{i}\right)-\dot{\Omega}_{i, f r e e}\left(N, V, E_{i}\right)\right)^{-E_{i} / k T}>0$
where $Q_{i}$ is the partition function of state $i$ and the sum covers all energy levels $i$ available to the adsorbed state. The sign and magnitude of $\Delta_{a d s} S$ is then determined by the fundamental statistical thermodynamic relation,

$$
\begin{equation*}
\Delta_{\text {ads }} S / k=\Delta_{\text {ads }} U / k T+\ln \left(\frac{Q_{\text {ads }}}{Q_{\text {free }}}\right) \tag{4.2}
\end{equation*}
$$

which states that a positive $\Delta_{\text {ads }} S$ can be observed in an isothermal adsorption process in which the mean energy of the system ( $\cup_{a d s} U$ ) is reduced when $Q_{\text {ads }}$ is sufficiently larger than $Q_{\text {free }}$. Because $Q_{\text {ads }} / Q_{\text {free }}$ is attenuated by the natural logarithm, $Q_{\text {ads }}$ must in general be substantially larger than $Q_{\text {free }}$ for an adsorption process to provide a net increase in entropy when $\Delta_{a d s} U<0$.

Our MC simulations show that

$$
\begin{equation*}
\sum_{E_{i}} \Omega_{i, a d t} \cong 4.6 \sum_{E_{i}} \Omega_{i, \text { free }} \tag{4.3}
\end{equation*}
$$

for adsorption of the HP chain to the planar sorbent surface. Eq. [4.3] further constrains Eq. [4.1] such that the value of $Q_{\text {ads }}-Q_{\text {free }}$ can only be increased by altering how the
excess density of states in the adsorbed system is distributed among the energy levels. Taken together with the requirement that $\Omega_{i, a d s}\left(N, V, E_{i}\right) \geq \Omega_{i, \text { free }}\left(N, V, E_{i}\right)$ at all $E_{i}$, it is then easy to prove that $Q_{a d s}-Q_{\text {free }}$ is maximized when the 3.6 -fold new states in the adsorbed system are preferentially distributed among the lowest energy levels for that system; that is, at adsorbed-state energies with the largest Boltzmann weightings. Thus, widening the energy landscape funnel for the adsorbed state system at the lowest energy levels (i.e., the funnel tip) increases $\Delta_{a d s} S$ toward more positive, thermodynamically favourable values.

To fix ideas, consider again the thermally-averaged energy landscapes for sequence I shown in Figure 4.3 for the case where $\chi_{H H}=\chi_{H W}=-1$. In solution, the chain must trade considerable conformational entropy to find its lowest-energy state, which is occupied by a single chain conformation, the native-state conformation. The thermally-averaged energy landscape for this process is therefore a "closed-tip" funnel that intersects the abscissa at $E_{i} / k T=-9$, the minimum energy level for the chain in solution. In contrast, our MC simulation data show that sequence I may adopt any one of 84 unique conformations at its lowest energy level $E_{\text {min }}$ in the adsorbed state. The degeneracy (number of unique conformational states) at $E_{\min }$ and energy levels just above $E_{\min }$ is significant and a positive $\Delta_{\text {ads }} S$ is observed.

Table 4.5 reports $\Omega_{\text {min,ads }}\left(E_{\min }\right)$ for sequence I as a function of the ratio $\chi_{H H} / \chi_{H W}$. There is a coarseness to the data because of the short length of our HP chain. However, our results show that the degeneracy of the lowest energy state is maximized when $\chi_{H H} / \chi_{H W}=$ 1 , irrespective of the value of $\chi_{H H}$ and thus, the stability of the native state. Likewise, on
any given sorbent surface the second term on the right side of Eq. [4.2] is maximized when the two contact energies are equal.

Our results further show that $Q_{a d s}-Q_{\text {free }}$ decreases quickly as the value of $\chi_{H H} \chi_{H W}$ diverges away from unity in either direction because the lowest energy levels for the adsorbed chain become significantly less populated with unique conformational states. At $\chi_{H H} / \chi_{H W}=4$, which includes the case of adsorption of a stable HP chain $\left(\chi_{H H}=-4\right)$ on the weakly attractive surface $\left(\chi_{H W}=-1\right)$, only one chain conformation is observed at $E_{\min }$. The adsorbed-state energy landscape is therefore a closed-tip funnel, and a negative $\Delta_{\text {ads }} S$ is observed due to a dramatic reduction in $Q_{a d s}$. Our results therefore suggest that the distinct differences noted by Arai and Norde (1) in the adsorption behaviour of proteins having relatively stable ("hard") versus relatively unstable ("soft") native-state conformations may be due to the manner in which the two systems distribute the excess states generated by the introduction of the sorbent surface.

### 4.3.3 The deformation entropy for HP chain adsorption

In response to a change in solvent quality, a linear polymer chain in solution will adopt new conformations in order to decrease repulsive and increase attractive contacts within the chain and between the chain and solvent. As a result, the total number of accessible conformations of the polymer molecule will change and will reach a maximum under socalled theta solvent conditions (31), where the chain assumes an "undeformed" randomflight configuration with overall dimensions solely determined by the bond lengths and
angles within the chain. In its theta solvent, the chain will therefore have maximum conformational entropy $S_{\text {conf }}=S_{\text {conf }}^{o}$. If the quality of the solvent is made poorer, net repulsive interactions with the solvent will cause the polymer chain to collapse. Conversely, the volume of the polymer chain will expand in response to net attractive interactions with a good solvent. In either case, $S_{\text {conf }}$ will decrease as a result of the lower number of accessible chain conformations. Flory called this entropy loss the chain deformation entropy, $\Delta_{\text {def }} S$, where
$\Delta_{d e f} S=S_{c o n f}-S_{c o n f}^{o}$

Analogous with the elastic properties of a linear chain in solution, our dMC simulations show that the conformational entropy of an adsorbed HP chain is a strong function of the quality of the sorbent surface, such that $S_{\text {conf }}$ shows a maximum at $\chi_{H H} / \chi_{H W}=1$. Like in solution, when $\chi_{H H} / \chi_{H W}>1$ the sorbent quality is relatively poor and the chain will collapse on itself to increase the density of more favourable HH contacts. When $0<$ $\chi_{H H} / \chi_{H W}<1$, the chain will expand along the sorbent surface to maximize contact area. Our objective is to correlate $\Delta_{d e f} S$ with a measurable property of an adsorbed chain. One such property is the average end-to-end distance of the chain at the sorbent surface, which is proportional to the average diameter of the adsorbed chain and can therefore be estimated at monolayer coverage with knowledge of the available sorbent surface area and an appropriate adsorption isotherm model such as the random sequential adsorption model $(32,33)$.

Flory has shown for a linear chain on a one-dimensional lattice where one end of the chain is fixed at the origin that the position of the remaining chain end is described by a Gaussian-type distribution, $W(x)$, of the form

$$
\begin{equation*}
W(x)=\frac{\beta}{\sqrt{\pi}} \exp \left(-\frac{1}{2} \beta^{2} x^{2}\right) \tag{4.5}
\end{equation*}
$$

where $W(x)$ is the probability of finding the chain end between $x$ and $x+d x$. To simplify our analysis in the case of a multidimensional lattice, we invoke the approximation that the effect of the deformation on $\Delta_{\text {def }} S$ in the $y$ dimension is equal to that in the $x$ dimension, so that

$$
\begin{equation*}
W(x, y) \approx(W(x))^{2}=\frac{\beta^{2}}{\pi} \exp \left(-\beta^{2} x^{2}\right) \tag{4.6}
\end{equation*}
$$

where

$$
\begin{equation*}
\beta=\frac{1}{\sqrt{2 n \overline{l_{x}^{2}}}} \tag{4.7}
\end{equation*}
$$

$n$ is the number of bonds within the chain and $\overline{l_{x}^{2}}$ is the average square of the projection of each bond vector on the $x$ axis. The average end-to-end distance of the adsorbed chain is then $n \sqrt{\overline{l_{x}^{2}}}$. We seek to use this result to develop a simple analytical model for $\Delta_{\text {def }} S\left(\overline{l_{x}}\right)$ that captures the dependence of $\Delta_{a d s} S$ on $\chi_{H H} / \chi_{H W}$ observed in our dMC
simulations. Let $W_{o}(x, y)$, given by Eq. [4.6], represent the distribution of chain end-to-end-distances at the $\chi_{H H} / \chi_{H W}$ value where $S_{\text {conf }}=S_{\text {con } f}^{o}$ and $\bar{l}_{x}=\bar{l}_{x o}$. A change in $\chi_{H H} / \chi_{H W}$ will then shift the average length of each bond vector from $\bar{l}_{x o}$ to a new value $\bar{l}_{x}$, giving a new distribution of end-to-end distances

$$
\begin{equation*}
W(x, y)=\frac{\left(\beta^{\prime}\right)^{2}}{\pi} \exp \left(-\left(\beta^{\prime}\right)^{2} x^{2}\right)=\frac{\beta^{2}}{\alpha_{x}^{2} \pi} \exp \left(-\beta^{2} \frac{x^{2}}{\alpha_{x}^{2}}\right) \tag{4.8}
\end{equation*}
$$

where

$$
\begin{equation*}
\beta=\beta \frac{\overline{l_{x o}}}{\overline{l_{x}}}=\frac{\beta}{\alpha_{x}} \tag{4.9}
\end{equation*}
$$

The entropy change associated with the deformation in adsorbed chain structure can be estimated from the change in the total density of states through the fundamental relation

$$
\begin{equation*}
\Delta_{d e f} S=S_{c o n f}-S_{c o n f}^{o}=k \ln \left(\frac{\Omega}{\Omega_{o}}\right) \tag{4.10}
\end{equation*}
$$

where,

$$
\begin{equation*}
\Omega=\frac{N!}{\Pi_{i} n_{i}} \Pi_{i} n_{i} W_{i} \text { and } \Omega_{o}=\frac{N!}{\Pi_{i} n_{o i}} \Pi_{i} n_{o i} W_{o i} \tag{4.11}
\end{equation*}
$$

In Eq. [4.11], $\Omega_{0}$ is the density of states at maximum conformational entropy, $N$ is the total number of end-to-end distances sampled and $n_{i}$ is the number within that sample
with end-to-end distances in the $x$ dimension between $x_{i}$ and $x_{i}+d x$. Substitution of Eq. [4.11] into Eq. [4.10], followed by application of the summation relations [ $\sum_{i} n_{i}=\sum_{i} n_{o i}=N$ and $\left.\sum W_{i}=\sum_{i} W_{o i}=1\right]$ then gives our desired result after some algebraic manipulations

$$
\begin{equation*}
\Delta_{d e f} S / k=a\left[2 \ln \alpha_{x}-\left(\alpha_{x}^{2}-1\right)\right]=q\left[2 \ln \left(\frac{\overline{l_{x}}}{\overline{l_{x o}}}\right)-\left(\left(\frac{\overline{l_{x}}}{\overline{l_{x o}}}\right)^{2}-1\right)\right] \tag{4.12}
\end{equation*}
$$

where $a$ is the dimensionality of the system ( $a=2$ in our lattice). Use of Eq. [4.12] requires values for $\alpha_{x}$, which can be determined from our dMC simulation data. Figure 4.4 reports the dependence of $\bar{l}_{x}$ on $\chi_{H H} / \chi_{H W}$ for our HP chain. As expected, $\bar{l}_{x}$ increases with increasing quality of the sorbent surface (i.e., with decreasing $\chi_{H H} / \chi_{H W}$ ). The asymptote observed at high values of $\chi_{H H} / \chi_{H W}$ corresponds to $\bar{l}_{x}$ for native-state conformation of the chain.

Based on the data in Figure 4.4, Figure 4.5 plots values of $\Delta_{\text {def }} S$ predicted using Eq. [4.12] as a function of $\bar{l}_{x}$ for sequence I. Despite the rather crude assumptions embodied in Eq. [4.6], our simple model (Eq. [4.12]) shows qualitative agreement with dMC simulation data. $\Delta_{\text {def }} S$ is predicted to be a maximum at $\chi_{H H} / \chi_{H W}=1$, where $\alpha_{x}$ is also equal to unity. When $\chi_{H H} / \chi_{H W}<1$, the adsorbed chain is observed to spread on the sorbent surface $\left(\alpha_{x}>1\right)$ and a decrease in $\Delta_{d e f} S$ and therefore $\Delta_{a d s} S$ is predicted as $\alpha_{x}$ increases. A decrease in entropy is also predicted and observed when $\chi_{H H} / \chi_{H W}>1$ due
to a collapse of the adsorbed chain onto itself that leads to $\alpha_{x}$ values less than unity. Here, however, the dependence of $\Delta_{d e f} S$ on $\alpha_{x}$ is predicted to be stronger, indicating that the dependence of the conformational entropy on chain segment density is enhanced at high segment densities. As a result, small deformations of the adsorbed chain can lead to relatively large changes in entropy.

As shown in Figure 4.5, dMC simulation results for the adsorption of sequence I show a stronger dependence of $\Delta_{d e f} S$ on $\bar{l}_{x}$ than predicted by Eq. [4.12]. This is due to additional limitations to chain expansion or compression in the $y$ dimension arising from the impenetrable nature of the sorbent that are not accounted for in our simple model. As a result, the contribution to $\Delta_{\text {def }} f$ of deformations in the dimension normal to the sorbent surface are somewhat higher than predicted. To empirically include this effect, we have treated $a$ in Eq. [4.12] as an adjustable parameter. With $a=2.8$, Eq. [4.12] shows good agreement with our dMC simulation data under conditions in which the sorbent surface is relatively attractive and chain expansion is observed. However, the model under-predicts the loss in entropy due to chain collapse since it does not account for the unique property of a protein-like chain to energetically favour a single conformation upon collapse.

### 4.4 Summary

In this paper, we have used dynamic Monte Carlo simulations to explore the relationship between the native-state stability of protein-like chains and the thermodynamics of adsorption of the chains onto a solid-liquid interface. Our results provide molecular
insights that help explain the well-known differences in the adsorption behaviour of proteins of low and high native-state stability. Increases in entropy have been observed in protein adsorption to a solid-liquid interface, particularly when the native-state stability of the protein is low. Such increases are most often attributed to a combination of solvent dehydration effects and conformational changes in the protein upon adsorption that increase chain entropy through a net loss in intramolecular interactions stabilizing the native state of the protein. Our dMC simulations directly probe the latter effect for simple protein-like chains and show that a net loss in intramolecular HH contacts leading to an increase in chain conformational entropy can be observed under certain conditions. However, the effect is far from general. Instead, positive $\Delta_{a d s} S$ are more directly related to the ability of the system to preferentially distribute new states generated by the sorbent surface into adsorbed-state energy levels with the largest Boltzmann weightings. This situation is favoured when the average intramolecular contact energy $\left(\chi_{H H}\right)$ equals the average intermolecular contact energy $\left(\chi_{H W}\right)$, and maxima in $\Delta_{a d s} S$ are always observed at this condition. This result therefore provides a possible new explanation for why positive and negative values of $\Delta_{a d s} S$ are observed for proteins of low and high native-state stability, respectively, adsorbing on a weakly attractive (e.g., hydrophilic) sorbent.

Finally, a simple analytical model based on Flory's theory of polymer elasticity was derived and used to correlate changes in adsorption entropy with a measurable physical parameter, the average diameter of an adsorbed chain. Analogous to the behaviour of a polymer in solution, the model predicts that the conformational entropy of an adsorbed chain will be a maximum when $\chi_{H H} / \chi_{H W}=1$, which can be loosely thought of as the theta
condition for the sorbent. $\Delta_{a d s} S$ is predicted to decrease when the sorbent becomes more attractive due to expansion of the chain along the sorbent surface to maximum more favourable HW contacts. When the sorbent becomes less attractive, an even stronger correlation with $\Delta_{\text {ads }} S$ is predicted (and observed in our dMC results) due to collapse of the chain into a compact structure on the surface to maximize more favourable HH contacts.

### 4.5 References

1. Arai, T. and Norde, W., The behavior of some model proteins at solid-liquid interfaces 1. Adsorption from single protein solutions. Colloid. Surface., 51 (1990) 1-15.
2. Haynes, C.A. and Norde, W., Globular protein at solid/liquid interfaces. Colloids Surfaces B, 2 (1994) 517-566.
3. Malmsten, M., Formation of adsorbed protein layers. J. Colloid Interface Sci., 207 (1998) 186-99.
4. Haynes, C.A. and Norde, W., Structures and stabilities of adsorbed proteins. J. Colloid Interface Sci., 169 (1995) 313-3.28.
5. Kondo, A. and Fukuda, H., Effects of adsorption conditions of kinetics of protein adsorption and conformational changes at ultrafine silica particles. J. Colloid Interface Sci., 198 (1998) 34-41.
6. Jackler, G., Steitz, R. and Czeslik, C., Effect of temperature on the adsorption of lysozyme at the silica/water interface studied by optical and neutron reflectometry. Langmuir, 18 (2002) 6565-6570.
7. Buijs, J., Norde, W. and Lichtenbelt, J.W.T., Changes in the secondary structure of adsorbed IgG and F(ab' $)_{2}$ studied by FTIR spectroscopy. Langmuir, 12 (1996) 160513.
8. Norde, W. and Anusiem, A.C.I., Adsorption, desorption and re-adsorption of proteins on solid surfaces. Colloid. Surface., 66 (1992) 73-80.
9. Larsericsdotter, H., Oscarsson, S. and Buijs, J., Thermodynamics analysis of proteins adsorbed on silica particles: Electrostatic effects. J. Colloid Interface Sci., 237 (2001) 98-103.
10. Billsten, P., Wahlgren, M., Arnebrant, T., McGuire, J. and Elwing, H., Structural changes of T4 lysozyme upon adsorption to silica nanoparticles measured by circular dichroism. J. Colloid Interface Sci., 175 (1995) 77-82.
11. McGuire, J., Wahlgren, M.C. and Arnebrant, T., Structural stability effects on the adsorption and dodecyltrimethylammonium bromide-mediated elutability of bacteriophage-T4 lysozyme at silica surfaces. J. Colloid Interface Sci., 170 (1995) 182-192.
12. Froberg, J.C., Arnebrant, T., McGuire, J. and Claesson, P.M., Effect of structural stability on the characteristics of adsorbed layers of T4 lysozyme. Langmuir, 14 (1998) 456-62.
13. Tian, M., Lee, W.-K., Bothwell, M.K. and McGuire, J., Structural stability effects on adsorption of bacteriophage T4 lysozyme to colloidal silica. J. Colloid Interface Sci., 200 (1998) 146-154.
14. Lee, W.K., McGuire, J. and Bothwell, M.K., Competitive adsorption of bacteriophage T4 lysozyme stability variants at hydrophilic glass surfaces. J. Colloid Interface Sci., 269 (2004) 251-254.
15. Billsten, P., Freskgard, P.-O., Carlsson, U., Jonsson, B.-H. and Elwing, H., Adsorption to silica nanoparticles of human carbonic anhydrase II and truncated forms induce a molten-globule-like structure. FEBS Letters, 402 (1997) 67-72.
16. Billsten, P., Carlsson, U., Jonsson, B.-H., Olofsson, G., Hook, F. and Elwing, H., Conformation of human carbonic anhydrase II variants adsorbed to silica nanoparticles. Langmuir, 15 (1999) 6395-6399.
17. Lau, K.F. and Dill, K.A., A lattice statistical mechanics model of the conformational and sequence spaces of proteins. Macromolecules, 22 (1989) 3986-3997.
18. Liu, S.M. and Haynes, C.A., Mesoscopic analysis of conformational and entropic contributions to nonspecific adsorption of HP copolymer chains using dynamic Monte Carlo simulations. J. Colloid Interface Sci., In press (2004).
19. Norde, W., Energy and entropy of protein adsorption. J. Disper. Sci. Technol., 13 (1992) 363-377.
20. Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D. and Chan, H.S., Principles of protein folding - A perspective from simple exact models. Protein Sci., 4 (1995) 561-602.
21. Camacho, C.J. and Thirumalai, D., Kinetics and thermodynamics of folding in model proteins. PNAS, 90 (1993) 6369-6372.
22. Socci, N.D. and Onuchic, J.N., Kinetic and thermodynamic analysis of proteinlike heteropolymers: Monte Carlo histogram technique. J. Chem. Phys., 103 (1995) 4732-44.
23. Konig, R. and Dandekar, T., Solvent entropy-driven searching for protein modeling examined and tested in simplified models. Protein Eng., 14 (2001) 329-35.
24. Gupta, P., Hall, C.K. and Voegler, A.C., Effect of denaturant and protein concentrations upon protein refolding and aggregation: A simple lattice model. Protein Sci., 7 (1998) 2642-2652.
25. Nguyen, H.D. and Hall, C.K., Effect of rate of chemical or thermal renaturation on refolding and aggregation of a simple lattice protein. Biotechnology and Bioengineering, 80 (2002) 823-834.
26. Allen, M.P. and Tildesley, D.J., Computer simulations of liquids. Oxford University Press, New York, NY, 1987.
27. Norde, W. and Lyklema, J., Why proteins prefer interfaces. J. Biomater. Sci. Polymer Edn., 2 (1991) 183-202.
28. Hill, T.L., An introduction to statistical thermodynamics. Dover Books, New York, NY, 1986.
29. Dill, K.A. and Chan, H.S., From Levinthal to pathways to funnels. Nature Structural Biol., 4 (1997) 10-19.
30. Anderson, R.E., Pande, V.S. and Radke, C.J., Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. J. Chem. Phys., 112 (2000) 9167-85.
31. Flory, P.J., Principles of polymer chemistry. Cornell University Press, Ithaca, NY, 1953.
32. Widom, B., Random sequential addition of hard spheres to a volume. J. Chem. Phys., 44 (1966) 3888.
33. Tarjus, G., Schaaf, P. and Talbot, J., Generalized random sequential adsorption. J. Chem. Phys., 93 (1990) 8352-8360.

### 4.6 Tables

Table 4.1: Adsorption thermodynamics for HP chain sequence I as a function of nativestate stability and sorbent surface affinity.

| $\chi_{H W}$ | $\chi_{H H}$ | $\Delta_{\text {ads }} A / k T$ | $\Delta_{\text {ads }} U / k T$ | $\Delta_{\text {ads }} S / k$ |
| :---: | :---: | :---: | :---: | :---: |
| -1 | -1 | -5.67 | -2.07 | 3.60 |
| -4 | -1 | -3.77 | -4.13 | -0.36 |
| -41.87. | -49.61 | -7.75 |  |  |
|  |  |  |  |  |

Table 4.2: The average number of favorable intramolecular and intermolecular contacts formed by sequence I in solution and when adsorbed onto a planar surface.

| $\chi_{H W}$ | $\chi_{\text {HH }}$ | Average number of HH contacts |  | Average number of HH + HW contacts |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Solution state | Adsorbed state | Adsorbed state |
| -1 | -1 | 4.3 | 4.4 | 6.3 |
|  | -4 | 7.7 | 7.8 | 12.0 |
| -4 | -1 | 4.3 | 0.3 | 13.5 |
|  | -4 | 7.7 | 5.8 | 13.7 |

Table 4.3: Adsorption thermodynamics for sequence II as a function of native-state stability and sorbent surface affinity.

| $\chi_{H W}$ | $\chi_{H H}$ | $\Delta_{\text {ads }} A / k T$ | $\Delta_{\text {ads }} U / k T$ | $\Delta_{\text {ads }} S / k$ |
| :---: | :---: | :---: | :---: | :---: |
| -1 | -1 | -5.79 | -0.17 | +5.62 |
| -4 | -1 | -0.03 | -0.19 | -0.22 |
|  | -18.93 | -30.88 | -11.95 |  |
|  |  |  |  |  |

Table 4.4: The average number of favorable intramolecular and intermolecular contacts formed by sequence II in solution and when adsorbed onto a planar surface.


Table 4.5: The density of unique conformational states for the adsorbed chain at the global energy minimum energy state as a function of $\chi_{H H} / \chi_{H W}$.

| $\chi_{H H^{\prime}} / \chi_{\text {HW }}$ | $\Omega_{\text {min,ads }}$ |
| :---: | :---: |
| 0.25 | 4 |
| 0.50 | 4 |
| 1.0 | 84 |
| 2.0 | 5 |
| 3.0 | 6 |
| 4.0 | 1 |

### 4.7 Figures

Figure 4.1: Schematic of HP chain sequences used in these simulations: sequence I and sequence II. Hydrophobic (H) segments are filled and polar segments (P) units are unfilled.


Figure 4.2: Helmholtz energy ( $\Delta_{\text {ads }} A / k T$, squares) and entropy ( $\Delta_{\text {ads }} S / k$, circles) of adsorption for the HP chain sequence I adsorbing on a weakly attractive surface ( $\chi_{H W}=-$ 1) as a function of $\Delta_{D-N} A / k T$, the stability of the native-state of the chain relative to its fully denatured state. Lines drawn indicate data trends. Error is within $11 \%$ and $22 \%$ for values of $\Delta_{a d s} A / k T$ and $\Delta_{a d s} S / k$, respectively.


Figure 4.3: Thermally-averaged energy landscapes for the HP chain sequence I in solution (squares) and adsorbed (circles) to a weakly attractive sorbent surface $\left(\chi_{H W}=-1\right)$ under conditions where the stability of the native-state conformation is low $\left(\chi_{H H}=-1\right)$. Lines drawn indicate data trends.


Figure 4.4: The average length, $\overline{l_{x}}$, in the $x$ dimension of each bond vector in HP chain sequence I as a function of $\chi_{H H} / \chi_{H W}$, the ratio of the average intramolecular to intermolecular contact energies. The dotted line represents the average length of each bond vector in the $x$ dimension when the chain is in its native-state conformation. The line drawn indicates the data trend.


Figure 4.5: Comparison of $\Delta_{\text {def }} f / k$ values calculated with Eq. [4.12] to dMC simulation data values over a range of chain deformations, $\overline{l_{x}} / \overline{l_{x 0}}$ : dashed line $a=2$, solid line $a=$ 2.8. The dMC data (open squares) shown correspond to adsorption conditions where $\chi_{H H} / \chi_{H W}=0.25\left(\overline{l_{x}} / \overline{l_{x o}}=2.2\right), 1\left(\overline{l_{x}} / \overline{l_{x o}}=1\right)$, and $4\left(\overline{l_{x}} / \overline{l_{x o}}=0.85\right)$, respectively.


## 5 Conclusion

The objective of designing surfaces to control protein adsorption is certainly not new and, as noted in Chapter 1, a great deal of both experimental and theoretical work has been conducted over the past half century in an effort to understand and control protein adsorption. While much has been learned, very little is yet understood regarding how to actually control protein adsorption behaviour, and even less is understood regarding the submolecular events involved in protein adsorption processes; these interactions must be understood before protein adsorption can be predicted and controlled. New approaches to understand protein adsorption behaviour are thus needed. One of the most powerful techniques to study complex molecular behaviour today is computational chemistry. Very significant advancements have been made in this field over the past decade to improve both the size of the systems that can be modelled and the accuracy of simulations. These ever increasing capabilities have enormous potential for helping us to understand protein adsorption at a submolecular level and to provide a path toward the goal of proactively designing biomaterial surfaces to control biological response.

In general, the initial protein adsorption process must be governed by a balance of the intermolecular interactions between the residues presented by the protein's surface with the functional groups presented by the sorbent surface as a function of separation distance (i.e., residue-surface interactions) and the intramolecular interactions between the residues within the protein itself (i.e., residue-residue interactions). While numerous computational chemistry studies have been conducted in the area of protein folding to understand the energetics of residue-residue interactions, very little is currently
understood regarding the energetics of protein residue-surface interactions. If these interactions can be quantified, then theoretically they should be able to be combined with an understanding of intramolecular residue-residue interactions to provide a thermodynamic basis for the prediction of protein-surface adsorption behaviour. This concept is similar to the approach used in numerous other biomolecular simulations, such as those used to predict ligand-protein, protein-protein, protein folding and RNA folding interactions (see for example, 1-3) In each of these approaches, contributions of the enthalpy, entropy, and/or free energy associated with specific functional group interactions, and an overall accounting of these interactions are used to predict the free energy of binding and structural organization of the system.

This thesis is among the first attempts to use advanced computational chemistry, in particular, dynamic Monte Carlo simulations of a simple coarse-grained protein-like chain, to gain insights into the underlying molecular physics of the adsorption to solidliquid interfaces of chain molecules that preferentially adopt specific compact low-energy conformations in solution. Globular proteins are the most obvious example of such chain molecules, and it is hoped that results from the model developed in this thesis have improved our understanding of the complexities of protein adsorption.

The results of this work give a unique perspective on the mechanisms driving adsorption of protein-like chains and the factors that influence them. Because they allow the direct connection of adsorption thermodynamics to adsorbed-chain conformational space, they also present a richer view of the process that establishes some unique features of the
adsorption of a protein-like chain as compared to the well-known behaviour of a randomcoil polymer adsorbing to a solid-liquid interface. Due to anchoring effects associated with multiple points of contact between the sorbent and the chain, unstructured polymer adsorption necessarily results in a decrease in the conformational entropy of the polymer, a fact confirmed by both experiment (4) and theoretical considerations (5). Nonspecific protein adsorption, however, is often an endothermic process, so that an increase in entropy must drive the adsorption process. While dehydration effects almost certainly contribute to the entropy increase, our results confirm that chain conformational entropy can also increase during adsorption. This is due in part to the restricted number of low energy conformations accessible to the HP chain in solution. The results of this thesis reveal that an increase in chain conformational entropy arises through an ability of the system to preferentially distribute new states generated by the sorbent surface into adsorbed-state energy levels with the largest Boltzmann weightings. If the new ground state energy level is highly degenerate, an increase in chain conformational entropy will generally be observed. The simulation results reported here indicate that a highly degenerate ground-state energy is favoured when the average intramolecular contact energy $\left(\chi_{H H}\right)$ equals or is near the average intermolecular contact energy ( $\chi_{H W}$ ). Similarly, results from the simulations carried in this work provide insights into the affect of a number of important systems variables (e.g., sorbent hydrophobicity, protein-like chain sequence, sorbent geometry and macromolecular confinement) on the conformational freedom of the adsorbed chain and adsorption thermodynamics.

The results of this work therefore indicate that a simple mesoscopic (coarse-grain) model can be useful in helping to further understand adsorption phenomenon unique to proteins and solid/liquid systems. The simulations have successfully demonstrated that adsorption behaviour and thermodynamic properties of HP model chains on a simulated surface can resemble those of real protein adsorption systems and therefore provide a useful simple model for testing and understanding certain fundamental concepts related to nonspecific protein adsorption.

This is not to say that our model is without weaknesses. It is clearly a highly simplified view of the adsorption of proteins and protein-like chain molecules. As discussed in several places in this thesis, the approach developed here has several clear limitations that must be kept in mind in order to properly interpret and appreciate the results from the molecular simulations. First of all, due to computational limitations, relatively short protein-like HP chains were used and were modeled in only two dimensions. Secondly, more complex events, such as protein-protein lateral interactions and entropic effects emanating from the solvent have not been considered. It also must be recognized, however, that one must walk before learning to run, and this work must therefore be recognized as a humble beginning toward understanding a very complex problem and not the final analysis.

### 5.1 References

1. Zhou, Y.Q. and Karplus, M., Folding of a model three-helix bundle protein: A thermodynamic and kinetic analysis. J. Mol. Biol., 293 (1999) 917-951.
2. Nakajima, N., Higo, J., Kidera; A. and Nakamura, H., Free energy landscapes of peptides by enhanced conformational sampling. J. Mol. Biol., 296 (2000) 197-216.
3. Kellogg, G.E., Ligand docking and scoring: New techniques and applications in drug discovery. Med. Chem. Res., 9 (1999) 439-442.
4. Denoyel, R., Durand, G., Lafuma, F. and Audebert, R., Adsorption of cationic polyelectrolytes onto montmorillonite and silica - microcalorimetric study of their conformation. J. Colloid Interface Sci., 139 (1990) 281-290.
5. Cohen-Stuart, M.A., Fleer, G.J. and Bijsterbosch, B.H., The adsorption of polyvinyl pyrrolidone onto silica. 1. Adsorbed amount. J. Colloid Interface Sci., 90 (1982) 310-320.

## Appendix

## Description of dynamic Monte Carlo Simulations and related program code

The investigations reported in this thesis describe the behaviour of a protein-like HP chain within and adsorbed to the surface boundary of a two-dimensional lattice. The movements of the chain are carried out using dynamic Monte Carlo algorithms. The purpose of using this method is not necessarily to mimic the exact movements of a protein in solution or on the surface, but to be able to sample the system within a reasonable period of time in order to compute all energetic and conformational states of the system.

Computer programs for the dynamic Monte Carlo simulations were developed specifically for this work. Below are brief explanations of the major program elements and an example of a program used.

## The lattice grid

The program shown here is written for a 2D simulation only, although simulations on 3D lattices were also performed. The lattice space used in the simulations is defined as a Cartesian coordinate grid having only positive $x$ and $y$ coordinates ranging from (and including) the lowest value at 0 to a highest value of GRIDLIM. The values of GRIDLIM are given in the starting input file and may differ in each dimension. The
lattice is therefore bounded by $x=0$ and $\operatorname{GRIDLIM}(x)$ (left and right boundaries, respectively) and $y=0$ and $\operatorname{GRIDLIM}(y)$ (bottom and top boundaries, respectively). Lattice site centers are assumed to be a distance, $a$, apart. All boundaries are reflective.

The input parameter, WALL_SWITCH, indicates whether a boundary is assigned a potential. A lattice boundary designated as a sorbent surface is referred to as an active wall (W), whose interaction energies with the other simulation components are defined. Non-active walls are assumed to be athermal. Four possibilities are allowed by the WALL_SWITCH function: horizontal active walls (choice 1), vertical active walls (choice 2), one wall active (choice 3 ), and all walls active (choice 4).

## The HP chain

The specific sequence of the protein-like HP chain is entered into the input file by the user. The chain length, $L$, is specified, as are the sequence of H and P residues. The chain conformation is entered in the input file as a series of $L$ coordinates (i.e. $\left(x_{1}, y_{1}\right),\left(x_{2}\right.$, $\left.\left.y_{2}\right) \ldots\left(x_{L}, y_{L}\right)\right)$. The chain is then placed in the program with each residue occupying a unique coordinate position within the allotted grid space. Throughout the simulation, the chain's position and configuration are defined by the coordinates of its first residue (START_PT) and a series of unit-length vectors (VECTOR). In this array, the structure of the chain is described by $L-1$ vectors (i.e. $\left(\right.$ vector $_{x 1}$, vector $\left._{y 1}\right)$, (vector ${ }_{x 2}$, vector $_{y 2}$ )... (vector ${ }_{x L-I}$, vector $\left._{y L-1}\right)$ ), each running parallel to either the $x$ or $y$ axis. Vector subscript
numbers refer to the chain residue the vector originates from (i.e. vector $r_{1}$ refers to the vector originating from $\left(x_{1}, y_{1}\right)$ and ends at $\left.\left(x_{2}, y_{2}\right)\right)$.

## The algorithm

The program runs as follows. Initial parameters are read into the main program using READ_DATA. The initial chain position and conformation are scanned to ensure that they are workable within the given simulation conditions. Chain characteristics are also analyzed at this point. Calculations of the chain energy and conformational likeness to a reference structure (usually the native-state lowest-energy structure) are made. Whether the chain conformation should be stored or not is also considered. In certain simulations, only the lowest energy conformations are stored. In other simulations, conformations at other or all energy levels are collected.

At this point, the main loop of the main subroutine begins. Upon entering the loop, all sampling frequencies are checked to determine whether information should be written to output files. The frequency of the WHIRLING function is also evaluated to determine whether the chain should be translated or rotated at this time.

The program then enters the move algorithm section. A chain residue is randomly selected. Depending on the position of the residue in relation to the overall chain conformation, a successful move may or may not occur. A successful move is one that results in a conformational change without chain overlap and without exceeding the
boundaries of the grid space (after having been once reflected off the walls). A successful move also has to pass the energy criteria given by the Metropolis algorithm.

For successful moves, conformation and energy parameters are calculated for the newly formed conformation and a transition protocol then takes place to replace old simulation parameters with new ones. Finally, just prior to returning to the beginning of the loop for the next attempted move, a scan is made to decide which data for the new conformation should be saved in the storage files.

If a move is unsuccessful, the program returns to the start of the loop, with only the simulation step number having changed throughout the process.

## Ending the program

The program is designed to terminate itself. This can happen in a few ways.

The program can end at its last designated cycle.
For instance, if the number of times to run the main subroutine, CHANGE_LIM, is 2, and the number of cycles in each main subroutine, LIMIT, is $1 \times 10^{9}$ cycles, then the program will end in its $2 \times 10^{9}$-th cycle, as long as the assigned storage capacities are not exceeded.

The program can end when its storage files are full.

For instance, if the size of the storage files, NSTOR, is 50 , the simulation will end when the 50th unique conformation having energy lower than the defined maximum, E_MAX, is found.

## Program files

Running of the simulations involves 3 types of files - the main program, the input file, and output files.

Main program <*.for>

The dynamic Monte Carlo simulation programs for these investigations were programmed in Fortran 77. All program coding is original work by the author, except for the random number generators, RANLUX and GASDEV. Coded by F. James in 1993, RANLUX is a well-known subroutine easily found on shareware sites (e.g. http://tonic.physics.sunysb.edu/docs/num meth.html). GASDEV is taken from "Numerical Recipes in Fortran77" and is also found on numerous web sources (e.g. $\underline{\text { http://lib-www.lanl.gov/numerical/index.html). }}$

## Input file <*.dat>

The input file holds initial simulation parameters for the program. Information such as the chain sequence, grid parameters, interaction energies and sampling frequency are specified here. Also included is designation of the type of output generated by the program. The name of the input file is important in that it becomes the name of the output file. For example, a program run using the input file <test 2 .dat $>$ will have the output files <test2.out>. Also important is that the spacing and format of the input file remain the same during the editing procedure. All numbers in the input file are rightjustified.

## Output files <*.out>, <Screen>, etc.

The number of output files varies with each program.

As some simulations run for long periods of time, it is often necessary to have multiple outputs so that files remain reasonably sized. Having more than one output file also allows for parallel sampling, and consequently, avoidance of having to store large amounts of information in a single file during the simulation.

In all programs, the main output file is $<^{*}$.out $>$, where "*" refers to the title of the input file. <*.out> generally presents detailed sampling data collected at the frequency indicated by NWRITE. There is also a screen output where data sampled at the frequency of NSIGN is written to the screen.

## Editing, compiling and running the program

For most part, the simulation programs used for this work were edited, compiled and run on LINUX systems supported by the Biotechnology Laboratory (www.biotech.ubc.ca) and the Institute of Applied Mathematics (www.iam.ubc.ca), both at the University of British Columbia. A portion of the work was conducted on Windows run PC's, using either Microsoft Fortran PowerStation or the shareware program, GNU g77 (found at http://www.cs.yorku.ca/Courses/1540/ftn.htm).

Instructions for running a program on a LINUX server will be given here.

## Running a single program

To compile the program,
Type "f77 <program.for> -o <commandname>-O3"
where,
f77: Calls the Fortran77 compiler.
<program.for> or <program.f>: The name of the program.
-o <commandname>: The program is compiled to become a command, and the name of the command can be referred to here.
-O3: Optimization command. Some brief trials were run, and it was found that level 3 optimization resulted in the most efficient program.

To run the program in background mode,
Type "nice <commandname> <inputfile.dat>> <screenfile> \&"
where,
nice: A command used so that the computationally-intensive program doesn't overtake the entire server.
<commandname>: Running the command.
<inputfile.dat>: The input file.
$><$ screenfile $>$ : Channelling the screen output to a storage file.
$\&:$ Sets the program to run in the background.

## Running a collection of programs in series

Compile the program.
Set up a command file listing the programs to be run on separate lines. Type in an open editor window:
<commandname1> <inputfile1.dat>> <screenfile1>
<commandname2> <inputfile2.dat>> <screenfile2>
(etc.)
Save the window as <commandfilename>.
Type "chmod $+x<$ commandfilename $>$ " to activate the command file.
Type "nice <commandfilename $>$ \&".

## Example program for dynamic Monte Carlo simulation, CONTACT9

CONTACT9 is a program designed to frequently sample a chain adsorbed onto an active surface. Simulations using this program begin with the chain sitting on the surface. CONTACT9 carries out the dynamic Monte Carlo moves, recording unique chain conformations into file $<^{*}$.out> while regularly sampling energy and structural data into <*.gph>.

## Main program, <contact9.for>

    PROGRAM CONTACT9
    * *************************

* 
* General program description:
* This program simulates the adsorption of a protein-like chain to a
* surface in a 2D Cartesian coordinate lattice. Monte Carlo
* dynamics are used to manoeuver the chain into random configurations.
* The internal energy of the chain is calculated and weighted using
* the method outlined by Metropolis.
* 
* The chain is of length L , and is composed of a connected
* series of hydrophobic (H) and hydrophilic ( P ) units. It is
* simulated to be in a space whose dimensions are independent of one
* another (GRIDLIM(X) vs GRIDLIM(Y)). The remainder of the lattice is
* filled with solvent units. Spatial boundaries are reflective.
* 
* Sorbent surfaces are simulated to be at specific boundary planes.
$\star$

* 
* CONTACT9 program description
* This program is a modification of DEGEN10, a program designed to store
* conformations with energies less than the specified maximum. The program
* continually runs without restarting. Acceptable conformations are
* first evaluated by comparison of previously stored conformations and
* their isomers.
$\star$
* The program can run with a fully active 4-sided surface or active * parallel $2-s i d e d$ surface. It is possible to have differing sizes
* of the $x$ and $y$ axes. No diffusion mechanism exists in this
* simulation; however, a switch to control the WHIRLING (transposition
* and rotation) function is added. When set on automatic, the whirling
* mechanism shuts off when the chain contacts an active surface.
* 
* Energy for hydrophobic-solvent and solvent-wall interactions can be
* expressed in terms of CHI $=$ KAPPA $/ T-P S I$.
* 
* The CONTACT version of the simulation counts the number of
* original native contacts in lowest energy conformations. First,
* the non-neighbouring contacts of the native state are catalogued.
* The non-neighbouring contacts of the lowest conformations found
* by the simulation are then compared against the original contacts
* and the numbers are compared.
* The CONTACT9 version is designed to sample a large number of data
* points for a simulation. In it, the energy and contact information
* are recorded (but not the conformation coordinates themselves to
* minimize the size of data files). The data can then be analyzed
* using the Monte Carlo histogram technique to calculate the
* energy of adsorption. The native state conformation has been added
* to the read file in order to calculate native state contacts.
* There is also an option in this version to record radius of gyration
* information. The calculation can be made at request using the switch,
* RADIUS_SWITCH.
* 
* 
* 


$\star$

```
* List of subroutines:
*********************
CENTRE_OF_MASS = Determines the centre of mass coordinates for a
    configuration. Coordinates used as a point of
    rotation for movement of the chain in WHIRLING.
* CENTRE_OF_MASS_REAL = Determines the centre of mass values for a
    configuration. Results are real numbers, used for
    radius of gyration calculations
* COORD MAKER = caculates series of coordinates describing
    positions and conformation of chain
    = transfers information from active part of the
    program into the storage files assigned to record
    unique lowest energy conformations
* FILTER = screens conformations which are tagged as
    having the specified criteria (e.g. having energy
    lower than the given maximum threshold) as being
    unique. Comparison of the conformation and its
    reflective and rotational versions to all previously
    stored structures are carried out.
* EIND_SITUATION = the position of a specifically chosen bead relative
    to connecting beads is assessed for the purpose of
    determining a possible future move.
* INCONSISTANCY_CHECK = scans the initial chain conformation given at the
* start of the simulation for errors, such as diagonal
* vectors or skipped coordinates.
* ISWITCH = switches two given integers. Used to reflect
    coordinates or vectors in the x=y line.
* MOVE_2D = using vector data, calculates new chain
    coordinates after a selected move has taken place.
*
* NEIGHBOUR_CATALOG_SINGLE = compares the intramolecular contacts of the
*
* *
* *
*
*
* NEIGHBOUR_COUNT
*
*
* NEIGHBOUR_ID
*
*
* OVERLAP_CHECK = scans a proposed chain configuration for possible
*
*
* PROTEIN_ADSORTPION_2 = the main process subroutine. A single run of this
*
*
*
*
* *
*
*
*
* RANDOM_REAL = converts a random number determined in RANLUX
*
*
* RANLUX
*
* READ_DATA
    working conformation with the conformation given
    in the input file for comparison (usually the native
    state conformation).
* FILE
*
*
*
*
*
*
*
*
*
*
*
*
*
* ISWITCH = switches two given integers. Used to reflect
*
*
MOVE 2D
    = counts the number of contacts made for a
    specific conformation.
    = identifies the series of intramolecular contacts
*
    overlap.
PROTEIN_ADSORTPION_2
    subroutine cannot exceed 2**32 cycles, and therefore
    had to be looped within a larger program for the
    possibility of longer simulations.
    = converts a random number determined in RANLUX
    from RVEC to integer value between two specified
    limits
    from RVEC to real value between two specified
    limits
        = random number generator, gives an array of
        random numbers between 0 and 1
    = reads initial input data from<*.DAT>.
```


PARAMETER ( $M A X D=2, \operatorname{MAXC}=100$, MAXSTOR $=100000$, MAXCONTACTS $=1000$ )
INTEGER NDIM, LIMIT, NWRITE, NSTOR, NWHIRL_START, NWHIRL,

+ NTRANS, L, COORD (MAXD,MAXC), STOR,
+ NCHANGE, CHANGE_LIM, NFRAME, NSTARTREC,
+ XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW,
+ NHOURS, NMIN, NSEC, WHIRL_SWITCH, RADIUS_SWITCH, WALL_SWITCH,
+ MAP_SWITCH, PAST_STORTAG, NNEIGHBOUR_ORIGIN,
+ MAP_FLAG,NNATIVE_ORIGIN,
+ START ORIGIN(MAXD),
+ PAST_START (MAXD), START_PT (MAXD), GRIDLIM(MAXD),
+ PAST_VECTOR (MAXD, MAXC), VECTOR (MAXD, MAXC),
+ COORD_ORIGIN (MAXD, MAXC), VECTOR_ORIGIN (MAXD, MAXC),
+ LOW_STEP(MAXSTOR), LOW_XHH (MAXSTOR),
+ LOW_XHP (MAXSTOR), LOW_XPP (MAXSTOR), LOW_XHS (MAXSTOR),
+ LOW_XPS (MAXSTOR), LOW_XHW(MAXSTOR), LOW_XPW(MAXSTOR),
+ LOW XSW (MAXSTOR),
+ LOW_START (MAXD, MAXSTOR), LOW_VECTOR (MAXD,MAXC, MAXSTOR),
+ LOW_CHANGE (MAXSTOR), NEIGHBOUR_ORIGIN(2,MAXCONTACTS)
INTEGER*2 TIME
REAL GRIDX, GRIDY,

+CHI _SW, KAPPA_HS, PSI_HS, KAPPA_SW, PSI_SW,
+ TEMP, E_MAX, TIME_COUNT, ETIME, TARRAY( $\overline{2})$, LOW_E(MAXSTOR)
CHARACTER*1 TYPE (MAXC)
CHARACTER*40 ARG, F_Name

```
    COMMON / CHI / CHI_HH, CHI_PP, CHI_HP, CHI_HS,
    + CHI_PS, CHI_HW, CHI_PW, CHI_SW,
```



```
    COMMON / GRAPH_LIM / L, NDIM, NVECTOR
    COMMON / TEMP / TEMP
    COMMON / WHIRL / NWHIRL, NTRANS
    COMMON / LATTICE / GRIDLIM
    COMMON / TYPE / TYPE
    COMMON / NEIGHBOUR / NEIGHBOUR ORIGIN, NNEIGHBOUR ORIGIN,
    +
                                    NNATIVE_ORIGIN
* Open files
*************
    IF ( IArgC()..LT. 1 ) THEN
        WRITE(*,*) ' SPECIFY INPUT FILE PLEASE '
                STOP
    END IF
    Call GetArg(1, F_Name)
    IF ( Access( F_NAME ,'r') .NE. 0 ) THEN
        WRITE(*,*) '-FILE IS NOT ACCESSIBLE!!'
        STOP
    END IF
    inc = index(F_Name,'.') -1
    IF ( INC .lt. 0 ) THEN
        inc = index(E_Name,' ') - 1
    END IF
    OPEN (UNIT=10, FILE=F_NAME, STATUS='OLD', ERR=1000,
    + IOSTAT=IERROR)
        OPEN (UNIT=20, FILE=F_Name(1:INC)//('.Out'),ERR=1000,
    + IOSTAT=IERROR)
    OPEN (UNIT=42, FILE=F_Name(1:INC)//('.gph'),ERR=1000,
    + IOSTAT=IERROR)
        TIME_COUNT = ETIME(TARRAY) .
* Initialize parameters and external loop
*****************************************
    NCHANGE =0
        NTOTALSTEP = 0
        MAP_FLAG = 0
    CALL READ_DATA (NDIM, CHANGE_LIM, LIMIT, NWRITE, NSTOR, GRIDX,
    + GRI\overline{DY, NWHIRL_START, NTRRANS, NSIGN, NFRAME,}
    + NSTARTREC, MAP_SWITCH,
    + START_PT, WHIRI_SWITCH, RADIUS_SWITCH, WALL_SWITCH,
    + CHI_HH, CHI_PP, CHI_HP, CHI_HS, CHI_PS, CHI_HW, CHI_PW,
    + CHI_SW, KAP\overline{PA_HS, PS}\mp@subsup{\overline{S}}{-}{-}HS, K\overline{P}PPA_SW,
        TEMP, E_MAX, \overline{L}, COORD, TYPE,START_ORIGIN, COORD_ORIGIN)
        NVECTOR = L - 1
    GRIDLIM(1) = INT(GRIDX * REAL(L))
    GRIDLIM(2) = INT(GRIDY * REAL(L))
    CALL VECTOR_MAP (COORD_ORIGIN, MAP_SWITCH, VECTOR_ORIGIN,
    + START_ORIGIN}, MAP_FLAG\overline{)
    IF (MAP FLAG.EQ.1) THEN
        WRI\overline{TE}(20,*)'THE GRID IS TOO SMALL FOR (NATIVE) CONFORMATION'
        WRITE(*,*) 'THE GRID IS TOO SMALL FOR (NATIVE) CONFORMATION '
        STOP
    ENDIF
        CALL NEIGHBOUR COUNT (VECTOR, START PT, XHH, XPP, XHP,
    + XHS, XPS, XHW, XPW, XSW, WALL_SWITCH
        WRITE(20,*) '***********************************************'
```

```
    WRITE(20,*) 1**********************************************'
    WRITE(20,*) 'DATA FILE: ', F_NAME
    WRITE(20,7) 'GRID DIMENSIONS X:', GRIDLIM(1),
    + 'Y:',GRIDLIM(2)
    WRITE(42,*) '********************************************'
    WRITE(42,*) '************************************************'
    WRITE(42,*) 'DATA FILE: ', F_NAME
    WRITE(42,7) 'GRID DIMENSIONS X:', GRIDLIM(1),
    + 'Y:',GRIDLIM(2)
    FORMAT (1X,A21,I5,8X,A2,I5)
    CALL WRITE DATA (NDIM, CHANGE_LIM, LIMIT, NWRITE, NSTOR, GRIDX,
    G GRIDY, NWHIRL_START, NTRANS, NSIGN, NFRAME,
        NSTARTREC, MAP_SWITCH,START_PT, WHIRL_SWITCH, WALL_SWITCH,
        CHI_HH, CHI_PP, CHI_HP, CHI_HS, CHI_PS, CHI_HW, CHI_PW,
        CHI_SW, KAP\overline{PA_HS, PSI__HS, KAPPPA_SW,}\mp@subsup{}{~}{-}PSI_SW,
        TEMP, E_MAX, L, COORD, TYPE,START_ORIGIN, COORD_ORIGIN)
    WRITE (20,*) '**************************************************'
    WRITE(20,*) '***********************************************'
    WRITE(42,*) '************************************************'
    WRITE(42,*) '**********************************************'
    DO 20 DIM = 1, NDIM
                PAST_START (DIM) = 0
    CONTINUE
    DO }25\mathrm{ VNUM = 1, NVECTOR
                DO 23 DIM = 1, NDIM
                    PAST_VECTOR(DIM,VNUM) = 0
                CONTINUE
    CONTINUE
    DO 40 STOR = 1, NSTOR
        LOW_STEP(STOR) = 0
        LOW_XHH (STOR) = 0
        LOW_XPP(STOR) = 0
        LOW_XHP(STOR) = 0
        LOW_XHS (STOR) = 0
        LOW_XPS(STOR) = 0
        LOW_XHW (STOR) = 0
        LOW_XPW(STOR) = 0
        LOW_-XSW(STOR) = 0
        LOW E(STOR) = E MAX + 0.0001
        DO }\overline{3}0\mathrm{ DIM = 1, N
                LOW_START(DIM,STOR) = 0
        CONTINUE
        DO 35 VNUM = 1, NVECTOR
        DO }34\mathrm{ DIM = 1, NDIM
                LOW_VECTOR(DIM,VNUM, NSTOR) = 0
                CONTINUE
        CONTINUE
        LOW_CHANGE (STOR) = 0
    CONTINÜE
    CALL VECTOR_MAP (COORD, MAP_SWITCH, VECTOR, START_PT, MAP_FLAG)
    IF (MAP_FLAG.EQ.1) THEN
        WRITE(20,*)'THE GRID IS TOO SMALL FOR THE CHAIN CONFORMATION'
        WRITE(*,*) 'THE GRID IS TOO SMALL FOR THE CHAIN CONFORMATION'
        STOP
    ENDIF
* Set diffusion on or off automatically according to switch
*****************************************************************
IF (WHIRL_SWITCH.EQ.1) THEN
            IF (XHW.GT.0) THEN
                                    NWHIRL = LIMIT
            ELSE
            NWHIRL = NWHIRL_START
        ENDIF
    ELSE
```

```
        NWHIRL = NWHIRL_START
    ENDIF
* Calculate contact information for original conformation
    CALL NEIGHBOUR_ID (START_ORIGIN, VECTOR_ORIGIN, NEIGHBOUR_ORIGIN,
    + NNEIGHBOUR_ORIGIN)
        DO 45 HOOD = 1, NNEIGHBOUR_ORIGIN
            IF (NEIGHBOUR_ORIGIN(2,HOOD).GT.0) THEN
                        NNATIVE_ORIGIN = NNATIVE_ORIGIN + 1
            ENDIF
45
    CONTINUE
* Start of external loop
**************************
50 CONTINUE
* Ending program
*****************
    IF ((NCHANGE.GE.CHANGE_LIM).OR.(PAST_STORTAG.EQ.NSTOR)) THEN
    WRITE(*,*) 'PROGRAM IS COMPLETE'
    WRITE (20,*)'PROGRAM IS COMPLETE'
    TIME_COUNT = ETIME (TARRAY) - TIME_COUNT
    NHOU\overline{RS}= INT (TIME COUNT/3600).
    NMIN = INT (TIME_COUNT - (NHOURS*3600)) / 60
    NSEC = INT (TIME COUNT) - (NHOURS*3600) - (NMIN*60)
    WRITE (20,60) 'PRÖGRAM RUN-TIME: ',NHOURS,' HOURS ',
    +
                    NMIN,' MINUTES ', NSEC,' SECONDS '
    WRITE(42,60) 'PROGRAM RUN-TIME: ',NHOURS,' HOURS ',
    +
    FORMAT (A17,I4, A7,I2, A9, I2, A9)
    CLOSE (UNIT=42)
    CLOSE (UNIT=20)
    CLOSE (UNIT=10)
    STOP
    ENDIF
* External loop counter
************************
    NCHANGE = NCHANGE + 1
* Call simulation
*****************
CALL PROTEIN ADSORPTION 2 (NCHANGE, PAST_START, PAST_VECTOR,
+ PAST_STORTAG, LOW_STEP, LOW_XHH, LOW_XHP, LOW_XPP; •
+ LOW_XHS, LOW_XPS, LOW_XHW, LOW_XPW, प्LOW_XSW, LOW_E,
+ LOW_-START, LOW_VECTOR, LOW_CHANGE)
* End external loop
*******************
GOTO 50
*Error statements:
******************
1000 WRITE(*,*) '****TROUBLE OPENING FILE***'
WRITE(*,*) 'IOSTAT IS', IERROR
WRITE(*,*)
END
```

```
*********************************************************************
\star\start******t************************************************************
*SUBROUTINES:
```

```
********************************************************************
```

********************************************************************
SUBROUTINE PROTEIN ADSORPTION 2 (NCHANGE, PAST START, PAST VECTOR,
+ PAST_STORTAG, LOW_STEP, LOW_XHH, LOW_XHP, LOW_XPP,
+ LOW XHS, LOW XPS, LOW XHW, LOW XPW, LOW_XSW, LOW_E, LOW_START,
+ LOW_VECTOR, LOW_CHANGE
*

* This is the main subroutine that runs the simulation. Given the data
* read from <*.dat>, it establishes the lattice and its boundaries,
* the interaction energies and chain conformation and position. This
* subroutine governs chain movements, sampling to the output files and
* screen, and determination of which conformations should be stored.
* 

**************************************************************************
*

* Parameter list:
*****************

```
PARAMETER ( \(\operatorname{MAXD}=2, \operatorname{MAXC}=100, \operatorname{MAXSTOR}=100000\), MAXCONTACTS=1000)
INTEGER L, POSITION, CONFIG, DIM, VNUM,
\(+\quad\) NDIM, NVECTOR, NCHANGE, RAND BEAD,
\(+\quad\) NSTEP, NSUCCESS, LIMIT, CHANḠE_LIM,
\(+\quad \mathrm{XHH}, \mathrm{XPP}, \mathrm{XHP}, \mathrm{XHS}, \mathrm{XPS}, \mathrm{XHW}, \bar{X} P W, ~ X S W\),
+ MOVE_FLAG, OVER_FLAG, MAP_FLAG, INCON_FLAG,
+ WHIRL_FLAG, NTOTALSTEP, NSTARTREC,
+ NEW_XHH, NEW_XPP, NEW XHP, NEW_XHS, NEW_XPS, NEW_XHW,
    NEW_XPW, NEW_XSW, NWHIRL, NTRANS, STOR, NWHIRL_START,
    NWRITE, STORTAG, NSTOR, NNEIGHBOUR_ORIGIN, NNATIVE_ORIGIN,
    NFRAME, WHIRL_SWITCH, RADIUS_SWITCH, WAIL_SWITCH, PAST_STORTAG,
            START PT (MAXD), NEW_START (MAXD), START_ORIGIN (MAXD),
            GRIDLIM (MAXD), COORD (MAXD, MAXC), COORD_ORIGIN (MAXD, MAXC),
            VECTOR (MAXD, MAXC), NEW VECTOR (MAXD, MAXC \(), ~ T Y P E \_N U M(M A X C), ~\)
            LOW XHH (MAXSTOR), LOW XPP (MAXSTOR), LOW_XHP (MAXXSTOR),
            LOW_XHS (MAXSTOR), LOW_XPS (MAXSTOR), LOW_XHW (MAXSTOR),
            LOW XPW (MAXSTOR), LOW XSW (MAXSTOR), LOW_STEP (MAXSTOR),
            LOW_START (MAXD, MAXSTOR), LOW_VECTOR (MAXD, MAXC, MAXSTOR),
    LOW_CHANGE (MAXSTOR), NEIGHBOUR_ORIGIN (2, MAXCONTACTS),
    \(+\quad\) PAST_START (MAXD), PAST_VECTOR (MAXD, MAXC),
\(+\quad\) NNATĪVE CONTACTS, NNONNATIVE_CONTACTS, NWALL_CONTACTS
REAL GRIDX, GRIDY, TEMP,
\(+\quad \mathrm{CHI}\) + \(\mathrm{HH}, \mathrm{CHI}\) PP, CHI _HP, CHI _HS, CHI _PS, \(\mathrm{CHI} \_\mathrm{HW}, \mathrm{CHI} \_\mathrm{PW}\),
\(+\mathrm{CHI}_{-} \mathrm{SW}, \mathrm{KAP} \mathrm{\overline{PA}} \mathrm{HS}, \mathrm{PSI} \mathrm{HS}, \mathrm{KAPPA} \bar{S}^{-} \mathrm{SW},{ }^{-} \mathrm{PSI}\) _SW,
\(+\quad E N G\), NEW_E, D-E, P_CAL̄C, P_RAND,
+ LOW_E (MAXXSTOR \(\overline{)}\), NATIVE_FRAC , E_MAX, SUCCESS,
\(+\quad\) R_G_AVG, R_G_DEFORM

\section*{CHARACTER*1 TYPE (MAXC)}
```

COMMON / CHI / CHI_HH, CHI_PP, CHI_HP, CHI_HS,

```
\(+\mathrm{CHI}_{-}^{-} \mathrm{PS}, \mathrm{CHI}_{-}^{-} \mathrm{HW}, \mathrm{CHI}_{-}^{-} \mathrm{PW}, \mathrm{CHI}_{-}^{-} \mathrm{SW}\),
\(+\quad\) KAPPA_HS, PSI_HS, KĀPPA_SW, PSI_SW
COMMON / GRAPH LIM / L, NDIM, NVECTOR
COMMON / TEMP / TEMP
COMMON / WHIRL / NWHIRL, NTRANS
COMMON / LATTICE / GRIDLIM
COMMON / TYPE / TYPE
COMMON / NEIGHBOUR / NEIGHBOUR_ORIGIN, NNEIGHBOUR_ORIGIN,
+ NNATIVE_ORIGIN
```

* Read data
***********

```

CALL READ DATA (NDIM, CHANGE LIM, LIMIT, NWRITE, NSTOR, GRIDX, + GRIDY, NWHIRL_START, NTRĀNS, NSIGN, NFRAME,
```

    + NSTARTREC, MAP SWITCH,
    + START_PT, WHIRL_SWITCH, RADIUS_SWITCH, WALL_SWITCH,
+ CHI_H啇, CHI_PP, CHI_HP, CHI_HS},\mp@subsup{\textrm{CHI}}{-}{\prime}PS, CHI_HW, CHI_PW
+ CHI_SW, KAPP\overline{_}_HS, PSI__HS, KA\overline{PPA_SW, \overline{PSI_SW,}}\mathbf{+},\mp@code{M}
+ TEM\overline{P}, E_MAX, \overline{L}, COORD, TYPE,START_ORIGIN},\mp@subsup{\overline{N}}{~}{\prime
* Initialize internal loop parameters:
****************************************
NVECTOR = L - 1
NSTEP = 0
NSUCCESS = 0
INCON FLAG = 0
OVER_\overline{FLAG = 0}
MAP \overline{FLAG = 0}
WHIT
STORTAG = 0
NTOTALSTEP = (NCHANGE - 1) * LIMIT
* Creating vector map:
**********************
CALL VECTOR_MAP (COORD, MAP_SWITCH, VECTOR, START_PT, MAP_FLAG)
IF (MAP_FLAG.EQ.1) THEN
WRITE(20,*)'THE GRID IS TOO SMALL FOR THE CHAIN CONFORMATION'
WRITE(*,*) 'THE GRID IS TOO SMALL FOR THE CHAIN CONFORMATION'
STOP
ENDIF
IE (NCHANGE.GT.1) THEN
STORTAG = PAST_STORTAG
DO 50 DIM = 1, }\mp@subsup{}{}{-}\textrm{NDIM
START_PT(DIM) = PAST_START(DIM)
CONTINUE
DO }60\mathrm{ VNUM = 1, NVECTOR
DO 55 DIM = 1, NDIM
VECTOR(DIM,VNUM) = PAST_VECTOR(DIM,VNUM)
5 5
60
CONTINUE
ENDIF
    * Checking for problems in original conformation:
*************************************************
CALL INCONSISTANCY_CHECK (VECTOR, INCON_FLAG)
IF (INCON FLAG.EQ.1) THEN
WRITE (20,*) 'THE ORIGINAL CONFORMATION IS INCONSISTENT'
WRITE(*,*) 'THE ORIGINAL 'CONFORMATION IS INCONSISTENT'
STOP
ENDIF
CALL OVERLAP_CHECK (VECTOR, START_PT, OVER_FLAG)
IF (OVER_FLAG.EQ.1) THEN
WRITE(20,*) 'THE ORIGINAL CONFORMATION HAS OVERLAP'
WRITE(*,*) 'THE ORIGINAL CONFORMATION HAS OVERLAP'
STOP
ENDIF
* Count contacts and calculate energy of original configuration
***************************************************************
CALL NEIGHBOUR COUNT (VECTOR, START_PT, XHH, XPP, XHP, XHS,
    + XPS, XHW, XPW, XSW, WALL_SWIT\overline{C}H)

```
```

    ENG = F_ENERGY (XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW)
    CALL NEIGHBOUR_CATALOG_SINGLE (START_PT,VECTOR,
    + NATIVE_FRAC, NNATIVE_CONTACTS, NNONNATIVE_CONTACTS,
    NWALL_CONTACTS)
    IF (ENG.LE.E MAX) THEN
                CAL\overline{L}}\mathrm{ FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP,
    + XHS, XPS, XHW, XPW, XSW, START_PT, VECTOR, NCHANGE,
    + LOW_STEP, LOW_E, LOW_XHH, LOW_XPP, LOW_XHP, LOW_XHS,
    + LOW_XPS, LOW_\overline{XHW, LOW_XPW, LOW_XSW,}
    LOW_START, LOW_VECTOR, LOW_CHANGE)
    ENDIF
    * Set bead type to numerical value
*******************t***************
DO 70 BEAD = 1, L
IF (TYPE (BEAD).EQ.'P')' THEN
TYPE_NUM(BEAD) = 0
ELSE
TYPE_NUM(BEAD) = 1
ENDIF
CONTINUE
* Start of internal loop
**************************
100 CONTINUE
* Ending internal loop at at given limit and storing final results
IF ((NSTEP.EQ.LIMIT).OR.(STORTAG.EQ.NSTOR)) THEN
SUCCESS = 100 * (REAL (NSUCCESS))/(REAL (NSTEP))
PAST_STORTAG = STORTAG
DO 123 DIM = 1, NDIM
PAST_START(DIM) = START_PT(DIM)
1 2 3
CONTINUE
DO }125\mathrm{ VNUM = 1, NVECTOR
DO 124 DIM = 1, NDIM
PAST VECTOR(DIM,VNUM) = VECTOR(DIM,VNUM)
124
125
CONTINUE
CONTINUE
IF ((STORTAG.EQ.NSTOR).OR.(NCHANGE.EQ.CHANGE_LIM)) THEN
WRITE (20,*)'LIMIT REACHED FOR SIMULATION \#',NCHANGE
WRITE(20,*) '%AGE OF SUCCESSFUL MOVES IS ', SUCCESS
WRITE (20,*)
WRITE(20,*) l******* THE LOWEST ENERGY CONFORMATIONS ******''
WRITE (20,*) 'MAXIMUM ENERGY RECORDED = ', E_MAX
WRITE (20,*)

```
* *Output option \#1: short listing of unique conformations

                                    DO 215 STOR = 1, STORTAG - 1
                                    WRITE \((20,206)\) 'LEC', STOR, 'SIM',
    \(+\quad\) LOW_CHANGE(STOR),
\(+\quad\) '\#', LOW_STEP(STOR), 'E', LOW_E(STOR)
    \(+\quad\) LOW_CHANGE(STOR),
\(+\quad\) '\#', LOW_STEP(STOR), 'E', LOW_E(STOR)
206
215
    FORMAT ( \(2 \mathrm{X}, \overline{\mathrm{A}} 3, \mathrm{I} 7,2 \mathrm{X}, \mathrm{A} 3, \mathrm{I} 4,2 \mathrm{X}, \mathrm{A} 1, \overline{\mathrm{I}} 10,2 \mathrm{X}, \mathrm{A} 1, \mathrm{F9} .3\) )
        CONTINUE
* \(\quad\) Output option \(\# 2:\) detailed listing of unique conformations
```

* DO 215 STOR = 1, STORTAG - 1
* WRITE (20, *)
*206
*     + 
*     + 
*     + 
*     + 
*     + 

*208
*
*212
*
*

*     + 
* 210
* 

*215
*

* *Output end
* ***********************************************************
ENDIF

```
```

WRITE (20,*)

```
WRITE (20,*)
WRITE (20,*)
WRITE (20,*)
RETURN
ENDIF
```

```
* Write and sampling data
```

* Write and sampling data
*******************************
IF ((MOD(NSTEP,NFRAME).EQ.0).AND. (NTOTALSTEP.GT.NSTARTREC)) THEN
IF (RADIUS_SWITCH.EQ.O) THEN
WRITE(42,220) NCHANGE, NSTEP, ENG, NATIVE_FRAC,
    + NNATIVE_CONTACTS, 'NNONNATIVE_CONTACTS,
    + NWALL_CONTACTS
F20 FORMAT (I }\overline{4},I10,F9.3,F9.3,I4,I4,I4
ELSE
CALL R_G_CALC(VECTOR, START_PT, R_G_AVG,
R G-DEFORM)
WRITE (\overline{42,222) NCHANGE, NSTEP, ENG, NATIVE_FRAC,}
NNATIVE_CONTACTS, NNONNATIVE_CONTACTS,
NWALL_CONTACTS, R_G_AVG, R_G_DEFORM
222
FORMAT (\overline{I}4,I10, F7.2, F7. 2,I4,I\overline{4},\overline{I}4,F9.3,F9.3)
ENDIF
ENDIF
* Data sampling to screen
***************************
IF (MOD (NSTEP, NSIGN).EQ.O) THEN WRITE(*,*) 'SIM', NCHANGE, ' STEP\#',NSTEP,
+ ENDIF $\quad$ ENERGY ', ENG

```
```

* Data sampling to output file

```
* Data sampling to output file
********************************
********************************
IF (MOD (NSTEP, NWRITE).EQ.0) THEN
```

    WRITE(20, 235) 'HH',XHH, 'HP', XHP,'PP',XPP, 'HS',
    ```
```

                        XHS, 'PS', XPS, 'HW', XHW, 'PW', XPW
    FORMAT (2X, 7 (A2, I3, 4X))
    WRITE(20,*) ' ENERGY ', ENG
    WRITE(20,*) ' POSITION', START PT(1),
        START_PT(2)
    WRITE(20,*)}\mp@subsup{}{}{-
    DO 240 VNUM = 1, NVECTOR
                            WRITE(20,*) VECTOR(1,VNUM), VECTOR(2,VNUM)
    CONTINUE
    WRITE(20,*) '*******************************************'
    WRITE(20,*)
    ENDIF
    *Counter and flag reset for internal loop
NSTEP = NSTEP + 1
NTOTALSTEP = NTOTALSTEP + 1
MOVE FLAG = 0
OVER_FLAG = 0

* Translate or rotate polymer
*****************************
IF (MOD(NSTEP,NWHIRL).EQ.0) THEN
CALL WHIRLING(VECTOR, START_PT, NEW_VECTOR, NEW_START,
WHIRL_FLAG)
IF (WHIRL_\overline{FLAG.EQ.1) THEN}
GOTO 100
ENDIF
CALL OVERLAP CHECK (NEW_VECTOR, NEW_START, OVER FLAG)
IF (OVER_FLAG.EQ.1) THEN
GOTO 100
ENDIF
CALL NEIGHBOUR_COUNT (NEW_VECTOR, NEW_START,
NEW_XHH, NEW_XPP, NEW_XHP, NEW_XHS, NEW_XPS,
+
NEW_E = F_ENERGY(NEW_XHH, NEW_XPP, NEW_XHP, NEW_XHS,
NEW-}\mp@subsup{}{}{-}\mathrm{ XPS, NEW XHW, NEW XPWW, NEW X-
D_E = NEW_E - ENG
IF (D_E.LE. (0.0)) THEN
CALL REFRESH (VECTOR, START_PT,
XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW, ENG,
NEW VECTOR, NEW START,
NEW_XHH, NEW_XPP, NEW_XHP, NEW_XHS, NEW_XPS,
NEW_XHW, NEW_XPW, NEW-XSW, NEW - E)
CALL NEIGHBOUR_CATALOG_SINGLE (START_PT,VECTOR,
NATIVE_FRAC,
NNATIVE_CONTACTS, NNONNATIVE_CONTACTS,
NWALL CONTACTS)
NSUCCESS = NSUCCESS + 1
ELSE
P CALC = F PROB(D E)
CALLL RANDOM REAL (\overline{P}}\mathrm{ RAND,0.0,1.0)
IF (P_RAND.LE.P_CALC) THEN
CALL REFRESH (VECTOR, START_PT,
XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW,
ENG, NEW_VECTOR, NEW_START,
NEW XHH, NEW XPP, NEW XHP, NEW XHS,
NEW_XPS, NEW-XHW, NEW - XPW, NEW XSW, NEW E)
CALL NEIGHBOUR_CATALOG_SINGLE (START_PT,
VECTOR, NATIVE_FRAC,
NNATIVE_CONTACTS, NNONNATIVE CONTACTS,
NWALL_CO-NTACTS)
NSUCCESS = NSUCCESS + 1
ELSE

```
```

                    GOTO 100
            ENDIF
        ENDIF
        IF (ENG.LE.E_MAX) THEN
            CALL FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP,
        XHS, XPS, XHW, XPW, XSW, START PT, VECTOR, NCHANGE,
        LOW_STEP, LOW_E, LOW XHH, LOW_XPP, LOW_XHP, LOW_XHS,
        LOW_XPS, LOW_XHW, LOW_XPW, LOW_XSW,
        LOW_START, LŌW_VECTOR_, LOW_CHANGGE)
        ENDIF
        GOTO 100
    ENDIF
    * Randomly choose a bead, find its position and situation
CALL RANDOM_INT (RAND_BEAD, 1, L)
CALL FIND_SITUATION (RAND_BEAD,. VECTOR, POSITION, CONFIG)
* Calculating a potential move
********************************
CALL MOVE_2D (RAND_BEAD, VECTOR, START_PT, POSITION, CONFIG,
    + NEW_VECTOR, NEW_START, MOVE_FLAG)
IF (MOVE_FLAG.EQ.1) THEN
The following statement (IF-THEN loop) is only necessary when recording
the rate of occurance of low energy conformations; take out when
scanning for unique conformations because it is redundant.
**************************************
IF (ENG.LE.E_MAX) THEN
CALL FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP,
XHS, XPS, XHW, XPW, XSW, START_PT, VECTOR, NCHANGE,
LOW_STEP, LOW_E, LOW_XHH, LOW_XPP, LOW_XHP, LOW_XHS,
LOW_XPS, LOW_\overline{XHW, LOW_XPW, LOW_XSW,}
LOW_START, LOWW_VECTOR, LOW_CHANGE)
ENDIF
GOTO 100
ENDIF
* Check potential move for overlap, reset new configuration
**************************************************************
CALL OVERLAP_CHECK(NEW_VECTOR, NEW_START, OVER_FLAG)
IF (OVER_FLAG.EQ.1) THEN
The following statement (IF-THEN loop) is only necessary when recording
the rate of occurrance of low energy conformations ; take it out when
scanning for unique conformations because it is redundant.
**************************************
IF (ENG.LE.E_MAX) THEN
CALL FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP,
XHS, XPS, XHW, XPW, XSW, START_PT, VECTOR, NCHANGE,
LOW_STEP, LOW_E, LOW_XHH, LOW_XPP, LOW_XHP, LOW_XHS,
LOW_XPS, LOW X XHW, LOW_XPW, LOW_XSW,
LOW_START, LOW_VECTOR, LOW_CHANGE)
ENDIF
GOTO 100
ENDIF

```
```

* Identify contacts and calculate energy of new conformation
************************************************************
CALL NEIGHBOUR_COUNT (NEW VECTOR, NEW START,
    + NEW_XHH, NEW_XPP, NEW_XHP, NEW_XHS, NEW_XPS, NEW_XHW, NEW_XPW,
    + NEW_XSW, WALL_SWITCH)
NEW_E = F_ENERGY(NEW_XHH, NEW_XPP, NEW_XHP, NEW XHS, NEW_XPS,
    +         - NEW_XHW, NEW_X\overline{PW, NEW_X\overline{SW)}}\mathbf{-}W\mathrm{ ( }
D_E = NEW_E - ENG
IF (D E.LE.(0.0)) THEN
CALL REFRESH (VECTOR, START_PT,
XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW, ENG,
NEW_VECTOR, NEW_START,
NEW_XHH, NEW_XPP, NEW_XHP, NEW_XHS, NEW_XPS, NEW_XHW,
NEW_XPW, NEW XSW, NEW_E)
CALL NEIGHBOUR_CATÄLOG_SINGIE (START_PT,VECTOR, NATIVE_FRAC,
NNATIVE_CONTACTS, NNONNATIVE_CONTACTS, NWALL_CONTACTS)
NSUCCESS =- NSUCCESS + 1
ELSE
P_CALC = F_PROB(D_E)
CALLL RANDOM REAL(\overline{P}}\mathrm{ RAND, 0.0,1.0)
IF (P_RAND.İE.P_CAILC) THEN
CALL REFRESH (VECTOR, START_PT,
XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW, ENG,
NEW VECTOR, NEW START,
NEW_XHH, NEW_XPP, NEW_XHP, NEW_XHS, NEW_XPS,
NEW XHW, NEW XPW, NEW XSW, NEW E)
CALL NEIGGHOUR_C\ATALOG_SİNGLE (START_PT,VECTOR,
NATIVE FRAC,
NNATIVE_CONTACTS, NNONNATIVE_CONTACTS,
NWALL CONTACTS)
NSUCCESS = NSUCCESS + 1
ELSE
GOTO 100
ENDIF
ENDIF
* Store lowest energy configuration
***********************************
IF (ENG.LE.E_MAX) THEN
CALL FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP,
XHS, XPS, XHW, XPW, XSW, START PT, VECTOR, NCHANGE,
LOW_STEP, LOW_E, LOW_XHH, LOW_XPP, LOW_XHP, LOW_XHS,
LOW XPS, LOW XHW, LOW XPW, LOW XSW,
LOW_START, LOW_VECTOR, LOW_CHANGE)
ENDIF
* End loop
**********
GOTO 100
END
*********************************************************************
SUBROUTINE CENTRE_OF_MASS (VECTOR, START_PT, CENTRE)
*********************************************************************
* Calculates the centre of mass. Takes the position and structure
* of the chain and determines the coordinates closest to the centre
* of mass for the chain.
* Main parameters:
* VECTOR: the vector set describing the chain conformation
* START PT: the coordinates of the first bead establishing the
chain position

```
```

* CENTRE: the coordinates closest to the chain centre of mass
* 

************************************************************************

* Parameter list
****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, NDIM, NVECTOR, VNUM, DIM, FACTOR,
    + START_PT(MAXD), TCENTRE(MAXD), CENTRE(MAXD),
    + VECTOR (MAXD,MAXC)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Initialize parameters
***********************
DO 20 DIM = 1, NDIM
TCENTRE(DIM) = 0
CENTRE(DIM) = 0
20 CONTINUE
* Calculate centre of mass
*****************************
DO }60\mathrm{ VNUM = 1, NVECTOR
FACTOR = L - VNUM
DO 40 DIM = 1, NDIM
TCENTRE(DIM) = TCENTRE(DIM) + FACTOR * VECTOR(DIM,VNUM)
CONTINUE
60 CONTINUE
DO }80\mathrm{ DIM = 1, NDIM
CENTRE(DIM) = NINT(REAL(TCENTRE(DIM))/REAL(L))
CENTRE(DIM) = CENTRE(DIM) + START_PT(DIM)
CONTINUE
RETURN
END
SUBROUTINE CENTRE_OF_MASS_REAL (VECTOR, START_PT, CENTRE_REAL)

```

```

* 
* Calculates the centre of mass. Takes the position and structure
* of the chain and determines the exact location of the centre
* of mass for the chain in real numbers.
* 
* Main parameters:
* VECTOR: the vector set describing the chain conformation
* START_PT: the coordinates of the first bead establishing the
* START_Pr chain position
* CENTRE: the chain centre of mass (real numbers)
**************************************************************************
* Parameter list
*****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, NDIM, NVECTOR, VNUM, DIM, FACTOR,
+ START_PT(MAXD), TCENTRE (MAXD),
+ VECTOR (MAXD,MAXC)
REAL TAB (MAXD) , CENTRE_REAL (MAXD)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Initialize parameters
************************

```
```

    DO 20 DIM = 1, NDIM
        TCENTRE(DIM) = 0.0
        CENTRE_REAL(DIM) = 0.0
    CONTINUE
    * Calculate exact centre of mass
**********************************
DO 60 VNUM = 1, NVECTOR
FACTOR = L - VNUM
DO 40 DIM = 1, NDIM
TCENTRE(DIM) = TCENTRE (DIM) + FACTOR * VECTOR(DIM,VNUM)
CONTINUE
4 0
6 0
CONTINUE
DO }80\mathrm{ DIM = 1, NDIM
CENTRE REAL(DIM) = REAL(TCENTRE(DIM))/REAL(L)
CENTRE_REAL(DIM) = CENTRE_REAL(DIM) + REAL(START_PT (DIM))
CONTINUE
RETURN
END
***************************************************************************
SUBROUTINE COORD_MAKER (VECTOR, START_PT, COORD_OUT)
***********************\overline{x}**********************\overline{*}**********\overline{*}****************
* 
* Converts starting point and vector data into lattice coordinates.
* 
* Main parameters:
* VECTOR: the vector set describing the chain conformation
* START PT: the coordinates of the first bead of the chain
* COORD_OUT: chain coordinates on the grid
*********脑***************************************************************
* Parameter list
*****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, DIM, NDIM, NVECTOR, BEAD, $+\quad$ START_PT (MAXD), VECTOR (MAXD, MAXC), COORD_OUT (MAXD, MAXC)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Calculate coordinates
***********************
DO 20 DIM $=1$, NDIM COORD_OUT (DIM, 1 ) = START_PT (DIM)
CONTINUE
DO $60 \mathrm{BEAD}=2$, L DO 40 DIM $=1$, NDIM COORD_OUT (DIM, BEAD) $=$ COORD_OUT $(D I M, B E A D-1)+$ $+\quad$ VECTOR(DIM, BEAD-1) CONTINUE
RETURN
END

```
```

SUBROUTINE FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP, $+\quad X H S, X P S, ~ X H W, ~ X P W, ~ X S W, ~ S T A R T \_P T, ~ V E C T O R, ~ N C H A N G E, ~$ + LOW STEP, LOW_E, LOW_XHH, LOW XPP, LOW_XHP, LOW_XHS, $+\quad$ LOW_XPS, LOW_X $H W$, LOW_XPW, LOW_XSW, + LOW START, LOW VECTOR, LOW CHANGE)

```
```

***************\overline{\#}**********\overline{*}***********\overline{\star}********************************

```
***************\overline{#}**********\overline{*}***********\overline{\star}********************************
*
```

```
* Files information into the storage files.
* Main parameters:
* STORTAG: variable used to identify the next unfilled storage position
* NSTEP: cycle number
* ENG: energy of the active conformation
* XHH, etc. number of contacts between the two specified components
    for the active conformation (e.g. XHH is the number of
    HH contacts)
    START_PT: the coordinates of the first bead of the active chain
    VECTOR: the vector set describing the active chain conformation
    NCHANGE: continuous repeat number
* LOW_*, etc:storage array for the given variable (e.g. LOW_XHH is the
    storage array for XHH values)
* Parameter list
****************
    PARAMETER (MAXD=2, MAXC=100, MAXSTOR=100000)
    INTEGER STORTAG, NSTEP, XHH, XPP, XHP, XHS,
    + XPS, XHW, XPW, XSW,
    + L, DIM, VNUM, NDIM, NVECTOR, NCHANGE,
    + START_PT (MAXD), VECTOR(MAXD,MAXC),
    + LOW_XHH (MAXSTOR), LOW_XPP(MAXSTOR), LOW_XHP(MAXSTOR),
    + LOW_XHS (MAXSTOR), LOW_XPS (MAXSTOR), LOW_XHW (MAXSTOR),
    LOW_XPW (MAXSTOR), LOW_XSW(MAXSTOR), LOW_STEP (MAXSTOR),
    LOW_START (MAXD, MAXSTOR), LOW_VECTOR (MAXD,MAXC,MAXSTOR),
    + LOW_CHANGE (MAXSTOR)
    REAL ENG, LOW_E(MAXSTOR)
    COMMON / GRAPH_LIM / L, NDIM, NVECTOR
    STORTAG = STORTAG + 1
    LOW_STEP(STORTAG) = NSTEP
    LOW E(STORTAG) = ENG
    LOW_XHH(STORTAG) = XHH
    LOW_XPP(STORTAG) = XPP
    LOW XHP(STORTAG) = XHP
    LOW_XHS (STORTAG) = XHS
    LOW_XPS(STORTAG) = XPS
    LOW_XHW (STORTAG) = XHW
    LOW XPW(STORTAG) = XPW
    LOW_XSW(STORTAG) = XSW
    DO 10 DIM = 1, NDIM
            LOW_START(DIM,STORTAG) = START_PT(DIM)
        CONTINUE
        DO 40 VNUM = 1, NVECTOR
            DO 20 DIM = 1, NDIM
                                    LOW_VECTOR(DIM,VNUM,STORTAG) = VECTOR(DIM,VNUM)
                CONTINUE
20 CON
            LOW_CHANGE(STORTAG) = NCHANGE
            RETURN
        END
    *********************************************************************
        SUBROUTINE FILTER (STORTAG, NSTEP, ENG, XHH,
        + XPP, XHP, XHS, XPS, XHW, XPW, XSW, START_PT, VECTOR,
        + NCHANGE, LOW STEP, LOW E, LOW XHH, LOW XPP, LOW XHP,
        + LOW_XHS, LOW_XPS, LOW_XHW, LOW_XPW, LOW_XSW, LOW_START,
        + - LOW VECTOR, LOW_CHANGE)
    *
    * Screens conformations for those which are original or isomers of
```

```
* past recorded conformations with energies lower than the given
* maximum.
* Files information into the storage files.
* Main parameters:
* STORTAG: variable used to identify the next unfilled storage position
* NSTEP: cycle number
* ENG: energy of the active conformation
* XHH, etc. number of contacts between the two specified components
                    for the active conformation (e.g. XHH is the number of
                    HH contacts)
* START_PT: the coordinates of the first bead of the active chain
* VECTOR: the vector set describing the active chain conformation
* NCHANGE: continuous repeat number
* LOW_*, etc:storage array for the given variable (e.g. LOW_XHH is the
* storage array for XHH values)
* Parameter list
****************
        PARAMETER (MAXD=2, MAXC=100, MAXSTOR=100000)
        INTEGER STORTAG, NSTEP, XHH, XPP, XHP, XHS,
        + XPS, XHW, XPW, XSW,
        + STOR, L, DIM, VNUM, NDIM, NVECTOR, NCHANGE,
        + START_PT(MAXD), VECTOR(MAXD,MAXC),
        + LOW_XHH(MAXSTOR), LOW XPP(MAXSTOR), LOW_XHP(MAXSTOR),
        LOW_XHS (MAXSTOR), LOW_XPS (MAXSTOR), LOW_XHW(MAXSTOR),
        LOW_XPW (MAXSTOR), LOW_XSW(MAXSTOR), LOW_STEP(MAXSTOR),
        LOW_START (MAXD,MAXSTOR), LOW_VECTOR(MAXD,MAXC,MAXSTOR),
    LOW_CHANGE (MAXSTOR)
        REAL ENG, LOW_E(MAXSTOR)
        COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Initialize parameter
***********************
    COUNT = 0
* Loop for checking previous stored conformation and their isomers
************************************************************************
    DO 700 STOR = 1, STORTAG
* Comparison with previously stored conformation
**************************************************
            DO 60 VNUM = 1, NVECTOR
                DO 40 DIM = 1, NDIM
                IF (VECTOR (DIM, VNUM).NE.
                                    LOW_VECTOR(DIM,VNUM,STOR)) THEN
                                    GOTO }8
                                    ENDIF
                CONTINUE
40
60 CONTINUE
    GOTO 750
80
    CONTINUE
```

* Comparison with 90 rotation of previous conformation
*************************************************************)
DO 100 VNUM $=1$, NVECTOR
IF ((-1*VECTOR(2,VNUM)).NE.LOW_VECTOR(1,VNUM,STOR))THEN
GOTO 120
ELSEIF (VECTOR(1,VNUM).NE.LOW_VECTOR(2,VNUM,STOR)) THEN
GOTO 120
ENDIF

```
100
    CONTINUE
    GOTO 750
120
    CONTINUE
* Comparison with 270 rotation of previous conformation
************************************************************
    DO }150\mathrm{ VNUM = 1, NVECTOR
    IF (VECTOR(2,VNUM) .NE.LOW_VECTOR(1,VNUM,STOR)) THEN
                                    GOTO 170
            EISEIF ((-1*VECTOR(1,VNUM)).NE.
            LOW_VECTOR(2,VNUM,STOR)) THEN
                GOTO 170
                    ENDIF
                CONTINUE
                GOTO 750
                CONTINUE
170
* Comparison with }180\mathrm{ rotation of previous conformation
************************************************************
    DO 210 VNUM = 1, NVECTOR
                                    DO 200 DIM = 1, NDIM
                                    IF((-1*VECTOR (DIM, VNUM)) .NE.
                                    LOW_VECTOR(DIM,VNUM,STOR)) THEN
                                    GOTO 220
                                    ENDIF
200
    CONTINUE
210
    CONTINUE
    GOTO 750
    CONTINUE
* Comparison with reflection of previous conformation in x=0
******************************************************************
DO 250 VNUM \(=1\), NVECTOR
IF ((-1*VECTOR(1,VNUM)).NE.LOW_VECTOR(1,VNUM,STOR)) THEN GOTO 270
ELSEIF (VECTOR (2,VNUM) .NE.LOW_VECTOR (2,VNUM, STOR) ) THEN GOTO 270
ENDIF
CONTINUE
GOTO 750
270
CONTINUE
* Comparison with reflection of previous conformation in \(y=0\)
```



```
DO 300 VNUM \(=1\), NVECTOR
IF (VECTOR(1,VNUM).NE.LOW_VECTOR(1,VNUM,STOR)) THEN GOTO 320
ELSEIF ((-1*VECTOR (2,VNUM)) .NE.LOW_VECTOR (2,VNUM, STOR))
THEN
GOTO 320
ENDIF
GOTO 750
CONTINUE
```

```
300 CONTINUE
* Comparison with reflection of previous conformation in \(x=y\)
```



```
DO 350 VNUM \(=1\), NVECTOR
IF (VECTOR (1,VNUM) . NE.LOW_VECTOR ( \(2, \operatorname{VNUM}, \operatorname{STOR}\) )) THEN GOTO 370
ELSEIF ((VECTOR(2,VNUM)).NE.LOW_VECTOR(1,VNUM,STOR))
THEN
GOTO 370
ENDIF
CONTINUE
GOTO 750
```

```
* Comparison with reflection of previous conformation in x=-y
*******************************************************************
DO }400\mathrm{ VNUM = 1, NVECTOR
                                    IF((-1*VECTOR(1,VNUM)) .NE.LOW VECTOR (2,VNUM,STOR) )THEN
                                    GOTO 420
                                    ELSEIF((-1*VECTOR(2,VNUM)).NE.LOW_VECTOR(1,VNUM,STOR))
                                    THEN
                                    GOTO 420
                                    ENDIF
                    CONTINUE
                    GOTO 750
                    CONTINUE
4 2 0
CONTINUE
    STORTAG = STORTAG + 1
    CALL FILER (STORTAG, NSTEP, ENG, XHH, XPP, XHP,
    + XHS, XPS, XHW, XPW, XSW, START_PT, VECTOR, NCHANGE,
    + LOW_STEP, LOW_E, LOW_XHH, LOW_XPP, LOW_XHP, LOW_XHS,
    + LOW_XPS, LOW_\overline{XHW, LOW_XPW, LOW_XSW, LOW_START,}
    + LOW_VECTOR, LOW_CHANGE) .
750 CONTINUE
    RETURN
    END
*************************************************************************
    SUBROUTINE FIND_SITUATION (RAND_BEAD, VECTOR, POSITION, CONFIG)
****************************************\overline{*}**********************************
*
* Determine general position and situation of a particular bead:
* Position labels (POSITION):
* 1 = first bead
* 2 = second to L-2 bead
* 3 = second last bead
* 4 = last bead
* Configuration labels (CONFIG):
* 1 = first bead at right angle
* 2 = first bead in straight line
* 3 = middle bead positioned for 3 pt flip
* 4 = middle bead positioned for 4 pt flip
* (includes vector previous and 2 vectors after)
* 5 = middle bead in straight line
* 6 = last bead at right angle
* 7 = last bead in straight line
*
* Main parameters:
* RAND_BEAD: the chain residue currently being evaluted
* VECTOR: the vector set describing the active chain conformation
* POSITION: label of where the residue is relative to the rest of the chain
* CONFIG: configurational situation of the residue relative to neighbouring
* residues
*
**********************************************************************
* Parameter list
************
PARAMETER (MAXD=2, MAXC=100)
    INTEGER L, DIM, RAND_BEAD, POSITION, CONFIG, N_VECTOR,
    + DEPTH, F DOT, NDIM, PROD FLAG,
    + V_LINE1(\overline{3}), V_LINE2(3), \overline{FLAG (2), OPPOS (2),}
    + D\overline{OT_ARRAY (3), - V_ARRAY (2,3), NVECTOR,}
```

$+\operatorname{VECTOR}$ (MAXD, MAXC)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR

* Initialize parameters
***********************
DO 60 DEPTH $=1,3$ DO 40 DIM $=1$, NDIM $\mathrm{V} \operatorname{ARRAY}(\mathrm{DIM}, \mathrm{DEPTH})=0$
Continue DOT_ARRAY(DEPTH) $=0$
continue
DO 80 DIM $=1$, NDIM $\mathrm{V}_{-}$LINE1 (DIM) $=0$ $\mathrm{V}_{-}^{-}$LINE2 $(\mathrm{DIM})=0$
CONTINUE
* Define POSITION and CONFIG

IF (RAND_BEAD.EQ.1) THEN POSITION = 1 N VECTOR $=2$
DO 140 DEPTH $=1,2$
DO 120 DIM $=1$, NDIM
V_ARRAY(DIM, DERTH) $=$ VECTOR(DIM, DEPTH) CONTIÑE
continue
ELSE IF (RAND_BEAD.EQ.(L-1)) THEN POSITION $=3$ N_VECTOR $=2$
D $\overline{-} 220$ DIM $=1$, NDIM
V_ARRAY(DIM, 1) $=\operatorname{VECTOR}(D I M, L-2)$
V_ARRAY (DIM, 2 ) $=\operatorname{VECTOR}(D I M, L-1)$
continue
ELSE IF (RAND BEAD.EQ.L) THEN . POSITION $=4$ N VECTOR $=2$
DO 240 DIM $=1$, NDIM
V_ARRAY(DIM, 1) $=\operatorname{VECTOR}(D I M, L-2)$
V_ARRAY(DIM, 2) $=\operatorname{VECTOR}(D I M, L-1)$

ELSE
POSITION $=2$
N_VECTOR $=3$
DO 260 DIM = 1, NDIM
V_ARRAY(DIM, 1) $=$ VECTOR(DIM, RAND_BEAD-1)
$\mathrm{V}_{-}^{-}$ARRAY (DIM, 2) $=$VECTOR (DIM, RAND_BEAD)
$\mathrm{V}_{-}^{-} \operatorname{ARRAY}(\mathrm{DIM}, 3)=\operatorname{VECTOR}\left(\mathrm{DIM}, \mathrm{RAND}_{-}^{-} \mathrm{BEAD}+1\right)$
continue
END IF
DO 290 DEPTH $=1,3$ DO 280 DIM $=1$, NDIM

V_LINE1(DIM) = V_ARRAY (DIM, DEPTH+1)
$\mathrm{V}_{-}^{-}$LINE2 (DIM) $=\mathrm{V}_{-}^{-}$ARRAY (DIM, DEPTH)
CONTINUE DOT_ARRAY (DEPTH) = F_DOT (V_LINE1, V_LINE2)
CONTINUE
IF (POSITION.EQ.1) THEN
IF (DOT_ARRAY(1).EQ.0) THEN
CŌNFIG = 1
ELSE
CONFIG $=2$

```
            END IF
            ELSE IF (POSITION.EQ.2) THEN
            IF (DOT_ARRAY(1).EQ.1) THEN
                        CONFIG = 5
            ELSE
            DO 350 DIM = 1, NDIM
                                    FLAG(DIM) = 0
                                    OPPOS(DIM) = -1* V ARRAY(DIM,1)
                                    IF (OPPOS(DIM).EQ.V_ARRAY(DIM,3)) THEN
                                    FLAG(DIM) = 1-
                                    END IF
350
        CONTINUE
        PROD FLAG = 1
        DO 3\overline{60 DIM = 1, NDIM}
                PROD FLAG = PROD FLAG * FLAG(DIM)
        CONTINUE
        IF (PROD_FLAG.EQ.1) THEN
                        CONFIG = 4
                ELSE
                        CONFIG = 3
            END IF
    END IF
    ELSEIF (POSITION.EQ.3) THEN
        IF (DOT_ARRAY (1).EQ.1) THEN
            CONFIG = 5
        ELSE
            CONFIG = 3
    END IF
    ELSE
        IF (DOT_ARRAY(1).EQ.0) THEN
            CONFIG = 6
    ELSE
        CONFIG = 7
    END IF
ENDIF
RETURN
END
******************************************************************************
    SUBROUTINE INCONSISTANCY CHECK (VECTOR, INCON FLAG)
******************************\overline{\star}t*********************\overline{x}******************
*
* Scans the initial chain conformation given at the start of the
* simulation for errors, such as diagonal vectors or skipped
* coordinates.
* Main parameters:
* VECTOR: the vector set describing the active chain conformation
* ICON_FLAG: flag signalling whether chain is acceptable or not
    (=0, chain is fine, continue program)
    (=1, chain in inconsistent, need to end program)
*
***************************************************************************
* Parameter list
****************
    PARAMETER (MAXD=2, MAXC=100)
    INTEGER L, NDIM, NVECTOR, DIM, VNUM,
    + INCON_ELAG, VECTOR(MAXD,MAXC)
    COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Checking for skipped coordinate
*********************************
```

```
    DO }320\mathrm{ VNUM = 1, NVECTOR
            DO 300 DIM = 1, NDIM
                IF (ABS(VECTOR(DIM,VNUM)).GE.2) THEN
                WRITE(*,*) 'BAD ORIGINAL CONFORMATION'
                WRITE(*,*) 'ONE VECTOR IS GREATER THAN 2'
                        INCON_FLAG = 1
                        GOTO }\overline{3}3
                            ENDIF
    CONTINUE
300
* Checking for vector not following lattice
**********************************************
    IF (ABS (VECTOR(2,VNUM)).EQ.
    + (ABS(VECTOR(1,VNUM)))) THEN
                                    WRITE(*,*) 'BAD ORIGINAL CONFORMATION'
                                    WRITE(*,*) 'ILLEGAL DIAGONAL STEP TAKEN'
                                    INCON FLAG = 1
                                    GOTO \overline{3}30
                ENDIF
320
        CONTINUE
330 CONTINUE
        RETURN
        END
            SUBROUTINE ISWITCH (A, B)
************************************************************************
*
* Switches integers A and B
*
****************************************************************************
            INTEGER A, B, C
            C=A
            A=B
            B=C
            RETURN
            END
*************************************************************************
            SUBROUTINE MOVE 2D (RAND BEAD, VECTOR, START PT, POSITION,
    + CONEIG, NEW_VECTOR, NEW_START, MOVE_FLAGG)
***********************
*
* Carries out a potential move for a given bead. The WALL-BOUNCE function
* is used here so that the resulting conformation is within the grid.
*
* POS CONFG CODE MOVE
* 1 1 1 straighten
* 1 1 1 1 2 180
* 1 1 2 % 1 % 90 counter-clockwise (90
* 1 2 2 90 clockwise (270*)
* 2 3 3 NA 
* 2 4 NA 4 pt flip
* 3 3 NA 3 pt flip
* 4 6 1 straighten
* 4 6 2 180
* 4 4 7 1 90 counter-clockwise (90')
* 4 4 8 2 2 90 clockwise (270
* NA 5 NA no move
*
* Main parameters:
* RAND BEAD: the chain residue currently being evaluted
* VECTOR: the vector set describing the active chain conformation
* START PT: the coordinates of the first residue of the active chain
* POSITION: label describing where the residue is relative to the rest of
```

```
* the chain
* CONFIG: configurational situation of the residue relative to neighbouring
    residues
* NEW VECTOR:the vector set describing the newly calculated conformation
* NEW_START: the coordinates of the newly calculated conformation
* MOVE_FLAG: flag signalling whether the move is plausible
    (0 = move is accepted)
    (1 = move is unacceptable)
*******************************************************************************
* Parameter list
*****************
    PARAMETER (MAXD=2, MAXC=100)
    INTEGER L, DIM, NDIM, NVECTOR, RAND_BEAD, POSITION, CONFIG,
    + MOVE_FLAG, BOUNCE_FLAG, VNUM, BEAD, RAND_MOVE,
    + STAR\overline{T}PT (MAXD), C\overline{HECK_PT (MAXD), NEW_START (MAXD),}
    + GRIDLIM(MAXD),
    + VECTOR (MAXD,MAXC), NEW VECTOR (MAXD,MAXC)
        COMMON / GRAPH_LIM / L, NDIM, NVECTOR
        COMMON / LATTICE / GRIDLIM
* Initialize parameters
************************
    MOVE_FLAG = 0
    RAND MOVE = 0
    BOUNCE_FLAG = 0
    DO }20\mathrm{ VNUM = 1, NVECTOR
        DO 10 DIM = 1, NDIM
                NEW_VECTOR(DIM,VNUM) = VECTOR(DIM,VNUM)
            CONTINUE
10
    CONTINUE
    DO 25 DIM = 1, NDIM
    NEW_START(DIM) = START_PT(DIM)
    CONTINUE
* No move for middle bead in straight line
**********************************************
    IF (CONFIG.EQ.5) THEN
        MOVE FLAG = 1
        GOTO 300
    ENDIF
* Moves for bead 1
*******************
    IF (POSITION.EQ.1) THEN
        CALL RANDOM INT (RAND MOVE, 1, 2)
        IF (CONFIG.EQ.1) THEN
                IF (RAND MOVE.EQ.1) THEN
                                    DO }30\mathrm{ DIM = 1, NDIM
                            NEW VECTOR(DIM,1) = VECTOR(DIM,2)
                            NEW_START(DIM) = START PT(DIM) +
                                    VECTOR(DIM,1) - VECTOR(DIM,2)
```

$+$

```
CONTINUE ELSE
```

```
                                    DO 40 DIM = 1, NDIM
```

                                    DO 40 DIM = 1, NDIM
                        NEW VECTOR(DIM,1) = -1 * VECTOR(DIM,1)
                        NEW VECTOR(DIM,1) = -1 * VECTOR(DIM,1)
                        NEW_START(DIM) = START_PT(DIM) +
                        NEW_START(DIM) = START_PT(DIM) +
                                    2 * VECTOR(DIM,1)
                                    CONTINUE
                ENDIF
    ELSE
                IF (RAND_MOVE.EQ.1) THEN
    ```
```

    NEW_VECTOR(1,1) = -1 * VECTOR (2,1)
    NEW_VECTOR(2,1) = VECTOR(1,1)
    ELSE
    NEW_VECTOR(1,1) = VECTOR(2,1)
    NEW_VECTOR (2,1) = -1 * VECTOR(1,1)
    ENDIF
    DO 60 DIM = 1, NDIM
    NEW_START(DIM) = START PT(DIM) +
        VECTOR(DIM,1) - NEW VECTOR(DIM,1)
    CONTINUE
    ENDIF
    * Moves for beads 2 to L-2
**************************
ELSEIF (POSITION.EQ.2) THEN
IF (CONFIG.EQ.3) THEN
DO 80 DIM = 1, NDIM
NEW_VECTOR(DIM, RAND BEAD-1) =
VECTOR(DIM, RAND BEAD)
NEW_VECTOR(DIM, RAND_BEAD) =
VECTOR(DIM, RAND_BEAD-1)
CONTINUE
ELSE
DO 100 DIM = 1, NDIM
NEW VECTOR(DIM, RAND BEAD-1) =
-1 * VECTOR(DIMM, RAND_BEAD-1)
NEW_VECTOR(DIM, RAND BEAD+1) =
-1 * VECTOR(DIM, RAND_BEAD+1)
CONTINUE
ENDIF
* Moves for beads L-1
ELSEIF (POSITION.EQ.3) THEN
DO 120 DIM = 1, NDIM
NEW_VECTOR(DIM, RAND_BEAD-1) =
VECTOR(DIM, RAND BEAD)
NEW_VECTOR(DIM, RAND_BEEAD) =
+ NEW_VECTOR(DIM, RAND_BEAD)
120 CONTINUE
* Moves for bead L
or bead
ELSE
CALL RANDOM_INT (RAND_MOVE, 1, 2)
IE (CONFIG.EQ.6) THEN
IF (RAND_MOVE.EQ.1) THEN
DO 140 DIM = 1, NDIM
NEW_VECTOR(DIM,L-1) = VECTOR(DIM,L-2)
140 CONTINUE
ELSE
DO 160 DIM = 1, NDIM
NEW_VECTOR(DIM, L-1) = -1
* VECTOR(DIM, L-1)
1 6 0
CONTINUE
ENDIF
ELSE
IF (RAND_MOVE.EQ.1) THEN
NEW VECTOR (1,L-1) = -1 * VECTOR (2,L-1)
NEW_VECTOR(2,L-1) = VECTOR(1,L-1)
ELSE
NEW_VECTOR(1,L-1) = VECTOR(2,L-1)
NEW VECTOR(2,L-1) = -1 * VECTOR(1,L-1)
ENDIF
ENDIF
ENDIF

```
* Check whether newly formed conformation is acceptable
```

DO 200 DIM = 1, NDIM
CHECK_PT(DIM) = NEW_START(DIM)
IF ((CHECK PT(DIM).GT.GRIDLIM(DIM)).OR.
(CHEC\overline{K_PT(DIM).LT.0)) THEN}
CALL WALL BOUNCE (NEW VECTOR, NEW START,
BOUNTCE FLAG)
IF (BOUNCE FLAG.EQ.1) THEN
MOVE_\overline{FLAG = 1}
GOTO 300
ENDIF
ENDIF
200 CONTINUE
DO 260 BEAD = 2, L
DO 240 DIM = 1, NDIM
CHECK_PT(DIM) = CHECK PT(DIM) +
NEW_VECTOR(DIM, BEAD-1)
IF ((CHECK PT(DIM).GT.GRIDLIM(DIM)).OR
(CHE\overline{CK_PT(DIM).LT.0)) THEN}
CALL WALL_BOUNCE (NEW_VECTOR, NEW_START,
BOUNCE. FLAG)
IF (BOUNCE FLAG.EQ.1) THEN
MOVE_FLAG = 1
GOTO 300
ENDIF
ENDIF
240 CONTINUE
260 CONTINUE
300 CONTINUE
RETURN
END
*********************************************************************
SUBROUTINE NEIGHBOUR_COUNT (VECTOR, START_PT, XHH, XPP, XHP,
+ XHS, XPS, XHW, X}PW, XSW, WALL SWITCH)
****************************************"******************************

* Counts the contacts existing within a conformation.
* NEARBY array
* north (0, 1)
* south (0,-1)
* east (1, 0)
west (-1,0)
* Main parameters:
* VECTOR: the vector set describing the active chain conformation
* START_PT: the coordinates of the first bead of the active chain
* XHH, etc. number of contacts between the two specified components
for the active conformation (e.g. XHH is the number of
HH contacts)
WALL_SWITCH:flag indicating the walls active in the simulation
(1 = vertical walls on, x=0, GRIDLIM)
(2 = horizontal walls on, y=0, GRIDLIM)
(4 = all walls active)
**********************************************************************
* Parameter list
******************
PARAMETER ( $M A X D=2, \operatorname{MAXC}=100, \operatorname{MAXS}=3$ )
INTEGER L, NDIM, NVECTOR, BEAD, DIM, QUAD, NSIDES, SIDE,
$+\quad \mathrm{XHH}, \mathrm{XPP}, \mathrm{XHP}, \mathrm{XHS}, \mathrm{XPS}, \mathrm{XHW}, \mathrm{XPW}, \mathrm{XSW}$,
+ PROD_FLAG, SCAN_BEAD, WALL_SWITCH, WALL_AREA,
$+\quad$ START_PT (MAXD), CHECK_PT (MAXD), ELAG (MAXD), SIDE_PT (MAXD),
$+\quad$ PREVIOUS (MAXD), GRIDLIM(MAXD),

```
\(+\quad\) NEARBY (MAXD, 4), SIDES (MAXD, MAXS), COORD_MAP (MAXD, MAXC),
\(+\quad\) VECTOR (MAXD, MAXC)
CHARACTER*1 TYPE (MAXC)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
COMMON / LATTICE / GRIDLIM
COMMON / TYPE / TYPE
DATA NEARBY/0, \(1,0,-1,1,0,-1,0 /\)
```

* Initialize parameters

```
    \(X H H=0\)
    \(X P P=0\)
    \(X H P=0\)
    \(X H S=0\)
    \(X P S=0\)
    \(X H W=0\)
    \(X P W=0\)
    \(\mathrm{XSW}=0\)
    CALL COORD_MAKER(VECTOR,START_ET,COORD_MAP)
* Establishing non-connecting neighbouring sites for selected bead

    DO \(600 \mathrm{BEAD}=1\), L
        DO 70 DIM \(=1\), NDIM
                CHECK_PT (DIM) = COORD_MAP (DIM, BEAD)
        CONTINUE
        IF (BEAD.EQ.1) THEN
                NSIDES \(=3\)
                SIDE \(=1\)
                DO 120 QUAD \(=1,4\)
                    IF ( (NEARBY (1, QUAD) .NE.VECTOR (1, 1))
    \(+\quad . \operatorname{OR} \cdot(\operatorname{NEARBY}(2, \operatorname{QUAD}) \cdot \operatorname{NE} \cdot \operatorname{VECTOR}(2,1)))\) THEN
                        DO 80 DIM \(=1\), NDIM
                                    SIDES (DIM, SIDE) \(=\) NEARBY (DIM, QUAD)
                                    CONTINUE
                                    SIDE \(=\) SIDE + 1
                                    ENDIF
                CONTINUE
        ELSEIF (BEAD.EQ.L) THEN
                NSIDES \(=3\)
                SIDE \(=1\)
                DO 220 QUAD = 1, 4
                        DO 140 DIM \(=1\), NDIM
                        PREVIOUS (DIM) \(=-1\) * VECTOR(DIM, L-1)
                        CONTINUE
                            IF ((NEARBY (1, QUAD) .NE. PREVIOUS (1)). OR.
```

                        (NEARBY(2,QUAD) .EQ.VECTOR(2,BEAD)))) THEN
                                    CONTINUE
                                    ELSE
                                    DO 280 DIM = 1, NDIM
                                    SIDES (DIM,SIDE) = NEARBY (DIM, QUAD)
                                    CONTINUE
                                    SIDE = SIDE + 1
                    ENDIF
                            CONTINUE
    ENDIF
    * Establishing number of intramolecular and wall contacts
DO 500 SIDE = 1, NSIDES
DO 340 DIM = 1, NDIM
SIDE_PT(DIM) = CHECK_PT(DIM) + SIDES(DIM,SIDE)
CONTINUE
IF (WALL_SWITCH.EQ.1) THEN
IF ((SIDE PT(1).GT.GRIDLIM(1)).OR.
(SIDE_PT(1).LT.0)) THEN
IF (TYPE(BEAD).EQ.'H') THEN
XHW = XHW + 1
ELSE
XPW = XPW + 1
ENDIF
GOTO 500
ENDIF
ELSEIF (WALL_SWITCH.EQ.2) THEN
IF ((SIDE PT(2).GT.GRIDLIM(2)).OR.
(SIDE_PT(2).LT.0)) THEN
IF (TYPE(BEAD).EQ.'H') THEN
XHW = XHW + 1
ELSE
XPW = XPW + 1
ENDIF
GOTO 500
ENDIF
ELSE
DO 360 DIM = 1, NDIM
IF ((SIDE PT (DIM).GT.GRIDLIM(DIM)).OR.
CONTINUE
ENDIF
DO 400 SCAN_BEAD = 1, L
PROD_FLAG = 1
DO 380 DIM = 1, NDIM
IF (SIDE_PT(DIM).EQ.
COORD MA\overline{P}(DIM,SCAN BEAD)) THEN
FLAG(DIM) = 1
. ELSE
FLAG(DIM) = 0
ENDIF
PROD_FLAG = PROD_ELAG * FLAG(DIM)
CONTINUE
IF (PROD_ELAG.EQ.1) THEN
IF ((TYPE (BEAD).EQ.'H').AND.
(TYPE(SCAN BEAD).EQ.'H')) THEN
XHH = XHH }+
GOTO 500
ELSEIF ((TYPE (BEAD).EQ.'P').AND.

```
```

                        (TYPE(SCAN_BEAD).EQ.'P')) THEN
                        XPP = XPP + 1
                                    GOTO 500
                                    ELSE
                                    XHP = XHP + 1
                                    GOTO 500
                                    ENDIF
                    ENDIF
    400
CONTINUE
IF (TYPE(BEAD).EQ.'H') THEN
XHS = XHS + 1
ELSE
XPS = XPS + 1
ENDIF
5 0 0 ~ C O N T I N U E ~
600 CONTINUE
XHH = XHH / 2
XPP = XPP / 2
XHP = XHP / 2

* Determine number of solvent-wall contacts
********************************************
IE (WALL SWITCH.EQ.4) THEN
WAL\vec{L_AREA = (2*GRIDLIM(1)) + (2*GRIDLIM(2))}
ELSEIF (WALLL SWITCH.EQ.1) THEN
WALL_AREA = 2*GRIDLIM(2)
ELSE
WALL_AREA = 2*GRIDLIM(1)
ENDIF
XSW = WALL_AREA - XHW - XPW
RETURN
END
*********************************************************************
SUBROUTINE NEIGHBOUR_CATALOG_SINGLE (START_PT,VECTOR, NATIVE_FRAC,
    + NNATIVE CONTACTS, NNONNATIVE CONTACTS, NWALL CONTACTS)
**************************************\overline{****************\overline{*}}\boldsymbol{*}+************
* Compares the intramolecular contacts of the working conformation with
* the conformation given in the input file for comparison (usually the
* native state conformation).
* 
* Main parameters:
* START_PT: the coordinates of the first bead of the active
* . chain
* VECTOR: the vector set describing the active chain
* conformation
* NATIVE_FRAC: fraction of contacts matching those of comparison
* case over all in comparison
* NNATIVE_CONTACTS: number of contacts matching those in given
*     - comparison case
* NNONNATIVE_CONTACTS: number of contacts in given conformation which
*     - do not match those of comparison
* NWALL_CONTACTS: number of contacts of the active conformation
* with the grid boundary (either active or
* non-active).
* 

*************************************************************************

* Parameter list
*****************
PARAMETER (MAXD=2, MAXC=100, MAXSTOR=100000, MAXCONTACTS=1000)
INTEGER L, NDIM, NVECTOR,

```
```

    + NNEIGHBOUR ORIGIN, HOOD,
    + NNEIGHBOUR, NNATIVE ORIGIN,
    + NNATIVE_CONTACTS, NNONNATIVE_CONTACTS, NWALL_CONTACTS,
    + START_PT (MAXD), VECTOR (MAXD, MAXC),
    + NEIGHBOUR_ORIGIN(2,MAXCONTACTS), NEIGHBOUR(2,MAXCONTACTS)
        REAL NATIVE_FRAC
    COMMON / GRAPH LIM / L, NDIM, NVECTOR
    COMMON / NEIGHBOUR / NEIGHBOUR_ORIGIN, NNEIGHBOUR_ORIGIN,
    +                   NNATIVE_ORIGIN
    * Initialize parameters
***********************
NNATIVE CONTACTS = 0
NNONNAT\overline{IVE_CONTACTS =0}00
NWALL CONTACTS = 0
NNEIGHBOUR = 1
CALL NEIGHBOUR_ID (START_PT, VECTOR,
    + NEIGHBOUR, NNEIGHBOUR)
* Determine number of matching, non-matching and wall contacts
**************************************************************
DO 100 HOOD = 1, NNEIGHBOUR .
DO 90 COMPARE = 1, NNEIGHBOUR_ORIGIN
IF (NEIGHBOUR(2,HOOD).LTT.0) THEN
NWALL CONTACTS =
    + NWALL_CONTACTS + 1
GOTO 100
ENDIF
IF ((NEIGHBOUR (1, HOOD).EQ.
    + NEIGHBOUR_ORIGIN(1,COMPARE)).AND.
    + (NEIGHBOUR(2,HOOD).EQ.
    + NEIGHBOUR_ORIGIN(2,COMPARE))) THEN
NNATIVE_CONTACTS = NNATIVE_CONTACTS + 1
GOTO 10\overline{O}
ENDIF
90 CONTINUE
NNONNATIVE_CONTACTS = NNONNATIVE_CONTACTS + 1
CONTINUE
NATIVE_FRAC = REAL(NNATIVE_CONTACTS) /
        + REAL(NNATIVE_ORIGIN)
RETURN
END
*****************************************************************************
SUBROUTINE NEIGHBOUR_ID (START_PT, VECTOR, NEIGHBOUR,
        + NNEIGHBOUR)
**
    * Identify the intramolecular contacts within a conformation
    * and contacts made with the grid boundary.
    * 
    * Contact codes for a 2D system (NEIGHBOUR array)
* y-axis wall ( }x=0\mathrm{ ) : -10
* (x = gridlim_x): -11
* x-axis wall ( }\textrm{y}=0\mathrm{ ) : -20
* (y = gridlim_y): -21
* 
* Main parameters:
* START_PT: the coordinates of the first bead of the active chain
* VECTOR: the vector set describing the active chain conformation
* NEIGHBOUR: array list of intramolecular and boundary contacts of a
* conformation
* NNEIGHBOUR:total number of contacts for the conformation, size of
NEIGHBOUR array

```
*

* Parameter list:

PARAMETER ( \(M A X D=2, ~ M A X C=100, ~ M A X S=3, ~ M A X C O N T A C T S=1000) ~\)
INTEGER L, NDIM, BEAD, DIM, SCAN_BEAD, PROD_FLAG,
\(+\quad\) NSIDES, SIDE, QUAD, NNEIGHBOUR,
\(+\quad\) PREVIOUS (MAXD), FLAG (MAXD), SIDE_PT (MAXD), GRIDLIM (MAXD),
\(+\quad\) START_PT (MAXD),
\(+\quad \operatorname{VECTOR}\) (MAXD, MAXC), COORD_MAP (MAXD, MAXC), NEARBY (MAXD, 4),
+ SIDES (MAXD,MAXS), NEIGHBOUR (2,MAXCONTACTS)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR COMMON / LATTICE / GRIDLIM

DATA NEARBY/0, \(1,0,-1,1,0,-1,0 /\)
* Initialize parameters
***********************

CALL COORD_MAKER(VECTOR,START_PT,COORD_MAP)
NNEIGHBOUR \(=0\)
```

* Establish non-connecting neighbouring sites for selected bead

```
*
    DO \(600 \mathrm{BEAD}=1, \mathrm{~L}-1\)
        DO 60 SIDE \(=1,3\)
                            DO 50 DIM= 1, NDIM
                        \(\operatorname{SIDES}(D I M, S I D E)=0\)
                    CONTINUE
        CONTINUE
        IF (BEAD.EQ.1) THEN
            NSIDES = 3
            SIDE = 1
            DO 120 QUAD \(=1,4\)
                IF ((NEARBY(1, QUAD).NE.VECTOR (1, 1))
        \(+\quad\).OR. (NEARBY \((2, \operatorname{QUAD}) . \operatorname{NE} . \operatorname{VECTOR}(2,1))) \mathrm{THEN}\)
                                    DO 80 DIM \(=1\), NDIM
                                    SIDES (DIM, SIDE) = NEARBY (DIM, QUAD)
                                    CONTINUE
                                    SIDE = SIDE + 1
                                    ENDIF
            CONTINUE
        ELSEIF (BEAD.EQ.L) THEN
                        NSIDES \(=3\)
                        SIDE = 1
                        DO 220 QUAD \(=1,4\)
                                DO 140 DIM \(=1\), NDIM
                                    PREVIOUS \((D I M)=-1 * V E C T O R(D I M, L-1)\)
                                    CONTINUE
                IF ((NEARBY(1, QUAD).NE. PREVIOUS(1)).OR.
                                    (NEARBY(2,QUAD).NE.PREVIOUS(2))) THEN
                                    DO \(180 \mathrm{DIM}=1\), NDIM
                                    SIDES (DIM, SIDE) = NEARBY (DIM, QUAD)
                                    CONTINUE
                                    SIDE \(=\) SIDE + 1
                                    ENDIF
                            CONTINUE
        ELSE
            NSIDES \(=2\)
            SIDE \(=1\)
            DO 320 QUAD \(=1,4\)
                DO 240 DIM \(=1\), NDIM
                    PREVIOUS (DIM) \(=-1\) * VECTOR(DIM, BEAD-1)
                CONTINUE
```

        IF (((NEARBY(1,QUAD).EQ.PREVIOUS (1)).AND.
    (NEARBY (2,QUAD) .EQ.PREVIOUS (2))).OR.
    ((NEARBY(1,QUAD) .EQ.VECTOR(1, BEAD)) .AND.
    (NEARBY(2,QUAD).EQ.VECTOR(2,BEAD)))) THEN
    CONTINUE
    ELSE
    DO 280 DIM = 1, NDIM
        SIDES (DIM,SIDE) = NEARBY(DIM, QUAD)
    CONTINUE
    SIDE = SIDE + 1
    ENDIF
320
CONTINUE
ENDIF

* Establish position of bead through evaluation of neighbouring sites
DO 500 SIDE = 1, NSIDES
340
CONTINUE
* Check for wall-chain contacts
*******************************

| DO 350 DIM $=1$, NDIM |  |
| :---: | :---: |
|  | NNEIGHBOUR $=$ NNEIGHBOUR + 1 |
|  | NEIGHBOUR ( 1, NNEIGHBOUR) = BEAD |
|  | NEIGHBOUR ( $2, \mathrm{NNEIGHBOUR} \mathrm{)}=-10 *$ DIM |
|  | GOTO 500 |
| ELSEIF | (SIDE_PT(DIM).GT.GRIDLIM(DIM)) THEN |
|  | NNEIGHBOUR $=$ NNEIGHBOUR +1 |
|  | NEIGHBOUR (1,NNEIGHBOUR) = BEAD |
|  | NEIGHBOUR (2,NNEIGHBOUR) $=(-10 *$ DIM $)-1$ |
|  | GOTO 500 |
| ENDIF |  |
| CONTINUE |  |

DO 450 SCAN_BEAD = BEAD + 1. L
PROD FLAG = 1
DO 4\overline{2}O DIM = 1, NDIM
IF (SIDE_PT(DIM).EQ.COORD_MAP(DIM,SCAN_BEAD)) THEN
FLAG(DIM) = 1
ELSE
FLAG(DIM) = 0
ENDIF
PROD_FLAG = PROD_FLAG * FLAG(DIM)
CONTINUE
IF (PROD_FLAG.EQ.1) THEN
NNEIGHBOUR = NNEIGHBOUR + 1
NEIGHBOUR (1,NNEIGHBOUR) = BEAD
NEIGHBOUR(2,NNEIGHBOUR) = SCAN_BEAD
GOTO 500
ENDIF
450 CONTINUE
500 CONTINUE
600 CONTINUE
RETURN
END
SUBROUTINE OVERLAP_CHECK (VECTOR, START_PT, OVER_FLAG)
*************************汭********************\overline{*}*********\overline{*}****************

* Checking for overlap in the conformation of the chain. The
* evaluation of whether or not a potential move is acceptable.

```
```

* 
* Main parameters:
* VECTOR: the vector set describing the active chain conformation
* START_PT: the coordinates of the first bead of the active chain
* OVER_FLAG: flag indicating whether overlap of beads occurs
(0 = initial value of flag, no overlap detected)
(1 = overlap detected, present conformation not possible)
* 

****************************************************************************

* Parameter list
*****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER NDIM, NVECTOR, DIM, BEAD, PREV_BEAD,
    + PROD_FLAG, L, OVER_FLAG,
    + CHECK_PT (MAXD), FLAG (MAXD), START_PT (MAXD),
    + PREV_\overline{PT (MAXD,MAXC), VECTOR (MAXD,M\overline{AXC)}}\mathbf{+}\mathrm{ (M)}
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Initialize parameters
***********************
DO }80\mathrm{ DIM = 1, NDIM
CHECK_PT(DIM) = START_PT(DIM)
80
CONTINUE
* Identify possible overlap
****************************
DO 190 BEAD = 2, L
DO 110 DIM = 1, NDIM
CHECK_PT(DIM) = CHECK_PT(DIM) + VECTOR(DIM,BEAD-1)
PREV_\overline{PT}(DIM,BEAD) = CHECK_PT(DIM)
CONTINUE
DO 140 PREV_BEAD = BEAD-1, 1, -1
PROD \overline{FLAG = 1}
DO 1\overline{3}0 DIM = 1, NDIM
FLAG(DIM) = 0
PREV PT(DIM, PREV BEAD) =
PREV_PT(DIM,PREV_BEAD+1)
- VECTOR(DIM,PREV BEAD)
IF (PREV_PT(DIM,PREV_BEADD).EQ.CHECK_PT (DIM))
THEN
FLAG(DIM) = 1
ENDIF
PROD_FLAG = PROD_FLAG * FLAG(DIM)
130
CONTINUE
IF (PROD_FLAG.EQ.1) THEN
OVER FLAG = 1
RETURN
ENDIF
140 CONTINUE
190 CONTINUE
RETURN
END
SUBROUTINE RANDOM_INT (R_INT, LBOUND, UBOUND)
************************************************************************
* 
* Choosing random integer between and including upper and lower bounds.
* Converts a random number between 0 and 1 to the required integer.
* 
* Main parameters:

```
```

* R INT: random integer chosen
* L\overline{BOUND: lower boundary for integer}
* UBOUND: upper boundary for integer
* 
* Parameter list
****************
INTEGER DIFF, LBOUND, UBOUND, R_INT
REAL RVEC(1)
* Convert random number to integer within range
*************************************************
DIFF = UBOUND - LBOUND
CALL RANLUX(RVEC,1)
R_INT = INT( RVEC(1) *(REAL(DIFF) + 0.9999999))
    +         + LBOUND
RETURN
END
SUBROUTINE RANDOM_REAL (R_REAL, LREAL, UREAL)

```

```

* 
* Choosing random real number between and including the upper and lower
* bounds. Converts a random number between 0 and 1 to the real number
* needed.
* 
* Main parameters:
* R_REAL: random real number chosen
* LREAL: lower boundary for real number
* UREAL: upper boundary for real number
* 

****************************************************************************

* Parameter list
*****************
REAL R_REAL, DIFF, LREAL, UREAL, RVEC(1)
DIFF = UREAL - LREAL
* Convert random number to be within required range
****************************************************)
CALL RANLUX(RVEC,1)
R_REAL = (RVEC (1) * DIFE) + LREAL
RETURN
END
**************************************************************************
SUBROUTINE RANLUX(RVEC,LENV)
*************************************************************************
* Random number generator, giving an array of LENV random numbers between
* zero and one.
* 

C Subtract-and-borrow random number generator proposed by
C Marsaglia and Zaman, implemented by F. James with the name
C RCARRY in 1991, and later improved by Martin Luescher
in 1993 to produce "Luxury Pseudorandom Numbers".
Fortran }77\mathrm{ coded by F. James, }199
references:
C M. Luscher, Computer Physics Communications 79 (1994) }10

```
```

F. James, Computer Physics Communications 79 (1994) 111
LUXURY LEVELS.
level 0 (
and Zaman, very long period, but fails many tests.
level 1 ( }\textrm{p}=48\mathrm{ ): considerable improvement in quality over level 0,
now passes the gap test, but still fails spectral test.
level 2 ( }\textrm{p}=97\mathrm{ ): passes all known tests,, but theoretically still
defective.
level 3 ( }\textrm{p}=223\mathrm{ ): DEFAULT VALUE. Any theoretically possible
correlations have very small chance of being observed.
level 4 (p=389): highest possible luxury, all 24 bits chaotic.
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
Calling sequences for RANLUX: ++
CALL RANLUX (RVEC, LEN) returns a vector RVEC of LEN . ++
32-bit random floating point numbers between ++
zero (not included) and one (also not incl.). ++
CALL RLUXGO(LUX,INT,K1,K2) initializes the generator from ++
one 32-bit integer INT and sets Luxury Level LUX ++
which is integer between zero and MAXLEV, or if ++
LUX .GT. 24, it sets p=LUX directly. K1 and K2 ++
should be set to zero unless restarting at a break++
point given by output of RLUXAT (see RLUXAT). ++
CALL RLUXAT(LUX,INT,K1,K2) gets the values of four integers++
which can be used to restart the RANLUX generator ++
at the current point by calling RLUXGO. K1 and K2++
specify how many numbers were generated since the ++
initialization with LUX and INT. The restarting ++
skips over K1+K2*E9 numbers, so it can be long.++
A more efficient but less convenient way of restarting is by: ++
CALL RLUXIN(ISVEC) restarts the generator from vector ++
ISVEC of 25 32-bit integers (see RLUXUT) ++
CALL RLUXUT(ISVEC) outputs the current values of the 25 ++
32-bit integer seeds, to be used for restarting ++
ISVEC must be dimensioned 25 in the calling program ++
+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
DIMENSION RVEC(LENV)
DIMENSION SEEDS(24), ISEEDS(24), ISDEXT(25)
PARAMETER (MAXLEV=4, LXDFLT=3)
DIMENSION NDSKIP(0:MAXLEV)
DIMENSION NEXT(24)
PARAMETER (TWOP12=4096., IGIGA=1000000000,JSDFLT=314159265)
PARAMETER (ITWO24=2**24, ICONS=2147483563)
SAVE NOTYET, I24, J24, CARRY, SEEDS, TWOM24, TWOM12, LUXLEV
SAVE NSKIP, NDSKIP, IN24, NEXT, KOUNT, MKOUNT, INSEED
INTEGER LUXLEV
INTEGER*2 TIME
LOGICAL NOTYET
DATA NOTYET, LUXLEV, IN24, KOUNT, MKOUNT /.TRUE., LXDFLT, 0,0,0/
DATA I24,J24,CARRY/24,10,0.1
C default
C Luxury Level 0 1 2 * ** 4
DATA NDSKIP/0, 24, 73, 199, 365/
Corresponds to p=24 48 97 223
C time factor 1 2 % 3 6 6 10 on slow workstation
1 1.5 2 3 5 5 on fast mainframe
NOTYET is .TRUE. if no initialization has been performed yet.
Default Initialization by Multiplicative Congruential
IF (NOTYET) THEN
NOTYET = .FALSE.
JSEED = INT(TIME())
INSEED = JSEED
WRITE(6,'(A,I12)') ' RANLUX DEFAULT INITIALIZATION: ',JSEED
LUXLEV = LXDFLT
NSKIP = NDSKIP(LUXLEV)
LP = NSKIP + 24
IN24 = 0

```
```

        KOUNT = 0
        MKOUNT = 0
            WRITE(6,'(A,I2,A,I4)') ' RANLUX DEFAULT LUXURY LEVEL = ',
                        LUXLEV,' p =',LP
            TWOM24 = 1
        DO 25 I= 1, 24
            TWOM24 = TWOM24 * 0.5
        K = JSEED/53668
        JSEED = 40014*(JSEED-K*53668) -K*12211
        IF (JSEED .LT. 0) JSEED = JSEED+ICONS
        ISEEDS(I) = MOD(JSEED,ITWO24)
        CONTINUE
        TWOM12 = TWOM24 * 4096.
        DO 50 I= 1,24
        SEEDS(I) = REAL(ISEEDS(I))*TWOM24
        NEXT(I) = I-1
    50 CONTINUE
        NEXT(1) = 24
        I24 = 24
        J24 = 10
        CARRY = 0.
        IF (SEEDS(24) .EQ. 0.) CARRY = TWOM24
        ENDIF
        The Generator proper: "Subtract-with-borrow",
        as proposed by Marsaglia and Zaman,
        Florida State University, March, 1989
    DO 100 IVEC= 1, LENV
    UNI = SEEDS(J24) - SEEDS(I24) - CARRY
    IF (UNI .LT. O.) THEN
            UNI = UNI + 1.0
            CARRY = TWOM24
        ELSE
            CARRY = 0.
        ENDIF
        SEEDS(I24) = UNI
        I24 = NEXT(I24)
        J24 = NEXT(J24)
        RVEC(IVEC) = UNI
    C small numbers (with less than 12 "significant" bits) are "padded".
IF (UNI .IT. TWOM12) THEN
RVEC(IVEC) = RVEC(IVEC) + TWOM24*SEEDS(J24)
and zero is forbidden in case someone takes a logarithm
IF (RVEC(IVEC) .EQ. O.) RVEC(IVEC) = TWOM24*TWOM24
ENDIF
Skipping to luxury. As proposed by Martin Luscher.
IN24 = IN24 + 1
IF (IN24 .EQ. 24) THEN
IN24 = 0
KOUNT = KOUNT + NSKIP
DO }90\mathrm{ ISK= 1, NSKIP
UNI = SEEDS(J24) - SEEDS(I24) - CARRY
IF (UNI .LT. O.) THEN
UNI = UNI + 1.0
CARRY = TWOM24
ELSE
CARRY = 0.
ENDIF
SEEDS(I24) = UNI
I24 = NEXT(I24)
J24 = NEXT(J24)
CONTINUE
ENDIF
100 CONTINUE
KOUNT = KOUNT + LENV
IF (KOUNT .GE. IGIGA) THEN
MKOUNT = MKOUNT + 1
KOUNT = KOUNT - IGIGA
ENDIF
RETURN

```
                Entry to input and float integer seeds from previous run
    ENTRY RLUXIN(ISDEXT)
    * IF block added by Phillip Helbig after correpondence with James
        IF (NOTYET) THEN
            WRITE(6,'(A)') ' PROPER RESULTS ONLY WITH INITIALISATION FROM
        \(\$ 25\) INTEGERS OBTAINED WITH RLUXUT'
            NOTYET = .EALSE.
        ENDIF
        TWOM24 \(=1\).
        DO \(195 \mathrm{I}=1,24\)
        \(\operatorname{NEXT}(I)=I-1\)
    195 TWOM24 \(=\) TWOM24 * 0.5
        \(\operatorname{NEXT}(1)=24\)
        TWOM12 = TWOM24 * 4096.
    WRITE(6,'(A)') ' FULL INITIALIZATION OF RANLUX WITH 25 INTEGERS:'
    WRITE(6,'(5X,5I12)') ISDEXT
    DO \(200 \mathrm{I}=1,24\)
    \(\operatorname{SEEDS}(I)=\operatorname{REAL}(\operatorname{ISDEXT}(I)) * T W O M 24\)
    200 CONTINUE
    CARRY \(=0\).
    IF (ISDEXT (25) .LT. 0) CARRY = TWOM24
    ISD \(=\) IABS (ISDEXT (25) )
    I24 = MOD (ISD, 100)
    ISD \(=\) ISD/100
    \(\mathrm{J} 24=\operatorname{MOD}(I S D, 100)\)
    ISD \(=\) ISD/100
    IN24 \(=\operatorname{MOD}(I S D, 100)\)
    ISD \(=\) ISD/ 100
    LUXLEV = ISD
        IF (LUXLEV .LE. MAXLEV) THEN
            NSKIP \(=\) NDSKIP (LUXLEV)
            WRITE ( \(\left.6, '(A, I 2)^{\prime}\right)\) ' RANLUX LUXURY LEVEL SET BY RLUXIN TO: ',
    \(+\)
                                    LUXLEV
            ELSE IF (LUXLEV .GE. 24) THEN
            NSKIP \(=\) LUXLEV -24
            WRITE (6,'(A, I5)') ' RANLUX P-VALUE SET BY RLUXIN TO:',LUXLEV
            ELSE
                NSKIP \(=\) NDSKIP (MAXLEV)
                WRITE ( \(\left.6,(\mathrm{~A}, \mathrm{I} 5)^{\prime}\right)\) ' RANLUX ILLEGAL LUXURY RLUXIN: ', LUXLEV
                LUXLEV \(=\) MAXLEV
            ENDIF
        INSEED \(=-1\)
        RETURN
C
C
                    Entry to ouput seeds as integers
    ENTRY RLUXUT (ISDEXT)
    DO \(300 \mathrm{I}=1,24\)
        ISDEXT(I) \(=\) INT(SEEDS (I)*TWOR12*TWOP12)
    300 CONTINUE
    ISDEXT \((25)=I 24+100 * \mathrm{~J} 24+10000 * I N 24+1000000 * \operatorname{LUXLEV}\)
    IF (CARRY .GT. 0.) ISDEXT (25) \(=-\operatorname{ISDEXT}(25)\)
    RETURN
C
                    Entry to output the "convenient" restart point
    ENTRY RLUXAT (LOUT, INOUT, K1, K2) •
    LOUT \(=\) LUXLEV
    INOUT \(=\) INSEED
    K1 = KOUNT
    K2 \(=\) MKOUNT
    RETURN
C
C Entry to initialize from one or three integers
    ENTRY RLUXGO(LUX,INS,K1, K2)
        IF (LUX .LT. 0) THEN
            LUXLEV = LXDFLT
        ELSE IF (LUX . LE. MAXLEV) THEN
            LUXLEV = LUX
        ELSE IF (LUX .LT. 24 .OR. LUX .GT. 2000) THEN
            LUXLEV \(=\) MAXLEV
            WRITE ( \(\left.6, '(A, I 7)^{\prime}\right)\) ' RANLUX ILLEGAL LUXURY RLUXGO: ', LUX
```

        ELSE
            LUXLEV = LUX
            DO 310 ILX= 0, MAXLEV
                    IF (LUX .EQ. NDSKIP(ILX)+24) LUXLEV = ILX
                    CONTINUE
        ENDIF
    IF (LUXLEV .LE. MAXLEV) THEN
        NSKIP = NDSKIP (LUXLEV)
        WRITE(6,'(A,I2,A,I4)') ' RANLUX LUXURY LEVEL SET BY RLUXGO :',
    +
    ELSE
        NSKIP = LUXLEV - 24
        WRITE (6,'(A,I5)') ' RANLUX P-VALUE SET BY RLUXGO TO:',LUXLEV
    ENDIF
    IN24 = 0
    IF (INS .LT. 0) WRITE (6,'(A)')
    + ' Illegal initialization by RLUXGO, negative input seed'
    IF (INS .GT. O) THEN
        JSEED = INS
        WRITE(6,'(A,3I12)') ' RANLUX INITIALIZED BY RLUXGO FROM SEEDS',
    +
            JSEED, K1,K2
    ELSE
        JSEED = JSDFLT
        WRITE(6,'(A)')' RANLUX INITIALIZED BY RLUXGO FROM DEFAULT SEED'
    ENDIF
    INSEED = JSEED
    NOTYET = . FALSE.
    TWOM24 = 1.
        DO 325 I= 1, 24
            TWOM24 = TWOM24 * 0.5
        K = JSEED/53668
        JSEED = 40014*(JSEED-K*53668) -K*12211
        IF (JSEED .LT. 0) JSEED = JSEED+ICONS
        ISEEDS(I) = MOD(JSEED,ITWO24)
        CONTINUE
    TWOM12 = TWOM24 * 4096.
        DO 350 I= 1,24
        SEEDS(I) = REAL(ISEEDS(I))*TWOM24
        NEXT(I) = I-1
    350 CONTINUE
    NEXT(1) = 24
    I24 = 24
    J24 = 10
    CARRY = 0.
    IF (SEEDS(24) .EQ. O.) CARRY = 'TWOM24
        If restarting at a break point, skip K1 + IGIGA*K2
        Note that this is the number of numbers delivered to
        the user PLUS the number skipped (if luxury .GT. 0).
    KOUNT = K1
    MKOUNT = K2
    IF (K1+K2 .NE. 0) THEN
        DO 500 IOUTER= 1, K2+1
            INNER = IGIGA
            IF (IOUTER . EQ. K2+1) INNER = K1
            DO 450 ISK= 1, INNER
                    UNI = SEEDS(J24) - SEEDS(I24) - CARRY
                IF (UNI .LT. 0.) THEN
                    UNI = UNI + 1.0
                    CARRY = TWOM24
                ELSE
                    CARRY = 0.
                ENDIF
                SEEDS(I24) = UNI
                I24 = NEXT(I24)
                    J24 = NEXT(J24)
                CONTINUE
    450
    500 CONTINUE
    C Get the right value of IN24 by direct calculation
IN24 = MOD(KOUNT, NSKIP+24)
IF (MKOUNT .GT. 0) THEN
IZIP = MOD(IGIGA, NSKIP+24)

```
```

                    IZIP2 = MKOUNT*IZIP + IN24
                    IN24 = MOD(IZIP2, NSKIP+24)
            ENDIF
    C Now IN24 had better be between zero and 23 inclusive
IF (IN24 .GT. 23) THEN
WRITE (6,'(A/A,3I11,A,I5)')
+ ' Error in RESTARTING with RLUXGO:',' The values', INS,
+ K1, K2, ' cannot occur at luxury level', LUXLEV
IN24 = 0
ENDIF
ENDIF
RETURN
END
SUBROUTINE READ DATA (NDIM, CHANGE LIM, LIMIT, NWRITE, NSTOR,
+ GRIDX, GRIDY, NWHIRL START, NTRANS, NSIGN, NFRAME,
+ NSTARTREC, MAP SWITCH, START PT, WHIRL_SWITCH,
+ RADIUS_SWITCH, WALL_SWITCH, \overline{CHI_HH, CHI__PP, CHI_HP,}
+ CHI HS,- CHI PS, CHI-HW, CHI PW, CHI_SW, KAPPA_HS, PSI_HS,
+ KAPPA SW, PSI_SW, TEMP, E_MAXX, L, COOORD, TYPE,
+ START_ORIGIN,COORD_ORIGIN)
******************\overline{*}************\overline{*}******************************************
*

* Read start data from file <*.dat>
* 
* (LINE refers to blank in <*.dat> file.)
* 

```
PARAMETER (MAXD=2, MAXC=100)
    INTEGER NDIM, CHANGE_LIM, LIMIT, NWRITE, NSTOR, L, BEAD,
+ NWHIRL_START, NTRANS, NSIGN, NFRAME, MAP_SWITCH,
+ WHIRL SWITCH, RADIUS_SWITCH, WALL_SWITCH, NSTARTREC,
\(+\quad\) START_PT (MAXD), START_ORIGIN (MAXD),
+ COORD (MAXD, MAXC) , COORD_ORIGIN (MAXD, MAXC)
    REAL GRIDX, GRIDY,
\(+\quad \mathrm{CHI}\) + \(\mathrm{HH}, \mathrm{CHI} \mathrm{PP}, \mathrm{CHI}\) HP, CHI _HS, CHI _PS, CHI _HW, CHI_PW,
\(+\quad \mathrm{CHI}\) SW, KAPPA_HS, PSI_HS, KAPPA_SW, -PSI_SW, TEMP, E_MAX
    CHARACTER*1 LINE
    CHARACTER*1 TYPE (MAXC)
    REWIND (UNIT=10)
READ (10,'(A1)') LINE
READ (10,'(A1)') LINE
READ (10,'(A1)') LINE
READ (10,'(I15)') NDIM
READ (10,'(I15)') CHANGE_LIM
READ (10,'(I15)') LIMIT
READ (10,'(I15)') NWRITE
READ (10,'(I15)') NSTOR
READ (10,'(F15.2)') GRIDX
READ (10,' (F15.2)') GRIDY
READ (10,'(I15)') NWHIRL_START'
READ (10,'(I15)') NTRANS
READ (10,'(I15)') NSIGN
READ (10,' (I15)') NERAME
READ (10,'(I15)') NSTARTREC
READ (10,'(A1)') LINE
READ (10,'(I15)') MAP_SWITCH
READ (10,' (A1)') LINE
READ (10,'(I15)') START_PT(1)
READ ( \(\left.10,{ }^{\prime}(115)^{\prime}\right)\) START_PT (2)
READ (10,'(I15)') WHIRL_SWITCH
READ (10,'(A1)') LINE
READ (10,'(A1)') LINE
READ (10,'(A1)') LINE
```

    READ (10,'(I15)') RADIUS_SWITCH
    READ (10,'(A1)') LINE
    READ (10,'(A1)') LINE
    READ (10,'(I15)') WALL_SWITCH
    READ (10,'(A1)') LINE
    READ (10,'(A1)') LINE
    READ (10,'(A1)') LINE
    READ (10,'(A1)') LINE
    READ (10,40) CHI HH, CHI PP, CHI_HP, CHI_HS, CHI_PS, CHI_HW,
    ```

```

    + TEMP, E MAX
    40 FORMAT (13(F15.2,/),F15.2)
READ (10,'(I15)') L
READ (10,'(A1)') LINE
READ(10,'(A1)') LINE
DO 60 BEAD = 1, L
READ (10,80) COORD (1, BEAD), COORD (2, BEAD), TYPE (BEAD)
60 CONTINUE
80 FORMAT (4X,2I4,3X,A1)
READ (10,'(A1)') LINE
READ (10,'(A1)') LINE
READ (10,'(A1)') LINE
READ (10,'(I15)') START_ORIGIN(1)
READ (10,'(I15)') START_ORIGIN(2)
DO 120 BEAD = 1, L
READ (10,140) COORD_ORIGIN (1, BEAD),COORD_ORIGIN(2,BEAD)
120 CONTINUE
140 FORMAT (4X,2I4)
RETURN
END
*******************************************************************************)
SUBROUTINE REFRESH (VECTOR, STȦRT_PT,
XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW, ENG,
NEW VECTOR, NEW START,
NEW_XHH, NEW_XP\overline{P}, NEW_XHP, NEW_XHS, NEW_XPS, NEW_XHW,
NEW XPW, NEW XSW, NEW E)
**********************\star*********\overline{*}*********\overline{*}*********************************
*

* Replaces working variables with newly-calculated variables
* Main parameters:
* VECTOR: the vector set describing the active chain conformation
* START_PT: the coordinates of the first bead of the active chain
* XHH, etc. number of contacts between the two specified components
* for the active conformation (e.g. XHH is the number of
* HH contacts)
* ENG energy value for active chain
* NEW_VECTOR: the vector set describing the new conformation
* NEW_START: the coordinates of the first bead of the new
* conformation
* NEW_XHH, etc.: number of contacts between the two specified components
* for the new conformation (e.g. XHH is the
* number of HH contacts)
* NEW_E: energy value for new conformation
*******\overline{*****************************************************************}+\mp@code{***}
* Parameter list
****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, DIM, VNUM, NDIM, NVECTOR,
    + XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW,
    + START_PT (MAXD), VECTOR(MAXD, MAXC),
    + NEW XH}H, NEW_XPP, NEW_XHP, NEW_XHS, NEW_XPS, NEW_XHW

```

```

    REAL ENG, NEW_E
    ```

COMMON / GRAPH_LIM / L, NDIM, NVECTOR
```

* Replace variables
*********************
DO 40 VNUM = 1, NVECTOR
DO 20 DIM = 1, NDIM
VECTOR(DIM,VNUM) = NEW_VECTOR(DIM,VNUM)
20 CONTINUE
40 CONTINUE
DO 60 DIM = 1, NDIM
START_PT(DIM) = NEW_START (DIM)
CONTINUE
XHH = NEW_XHH
XPP = NEW XPP
XHP = NEW_XHP
XHS = NEW_XHS
XPS = NEW_XPS
XHW = NEW_XHW
XPW = NEW_XPW
XSW = NEW - XSW
ENG = NEW_E
RETURN
END
********************************************************************************
SUBROUTINE R_G_CALC (VECTOR, START_PT, R_G_AVG, R_G_DEFORM)

```

```

* 
* Calculates the radius of gyration parameters for the specified
* chain conformation.
* Main parameters:
* VECTOR: the vector set describing the active chain conformation
* START PT: the coordinates of the first bead of the active chain
* R_G_AVG: averaged radius of gyration of the conformation
* R_G_DEFORM: deformation ratio, R_G(x-dim)/R_G(y-dim)
* 

*****************************************************************************

* Parameter list
******************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, DIM, VNUM, NDIM, NVECTOR, BEAD,
+ START_PT (MAXD), COORD (MAXD,MAXC), VECTOR (MAXD, MAXC)
REAL R_G_AVG, R_G_DEFORM, CENTRE_REAL (MAXD), R_G(MAXD), SUM(MAXD)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Initialize parameters
***********************
R_G_AVG = 0.0
DO 10 DIM = 1, NDIM
R_G(DIM) = 0.0
10
CONTINUE
* Calculate R_G_AVG and R_G_DEFORM
************************************
CALL COORD_MAKER(VECTOR,START_PT,COORD)
CALL CENTRE_OF_MASS_REAL (VECTOR,START_PT,CENTRE_REAL)
DO 40 BEAD = 1, L
DO 20 DIM = 1, NDIM

```
```

                R_G(DIM) = R_G(DIM) +((CENTRE_REAL (DIM) -COORD(DIM,BEAD))**2)
                CONTIÑUE
    CONTINUE
    DO }60\mathrm{ DIM = 1, NDIM
        R_G(DIM) = R_G(DIM) / L
        R_G_AVG = R_\overline{G}_AVG + R_G(DIM)
        CONTINUE
    R_G_AVG = R_G_AVG**0.5
    R_G_DEFORM =
    RETURN
    END
    SUBROUTINE VECTOR_MAP (COORD, MAP_SWITCH, VECTOR, START_PT,
    + MAP_FLAG)
    ***********\overline{\#}****************************************************************

* Creates vector map from original coordinates, and situates the centre
* bead in the centre of the grid if requested.
* Main parameters:
* COORD: coordinates of the chain (read from <*.DAT>)
* MAP_SWITCH: flag indicating whether conformation should be placed
in middle of grid, or if location is indicated
(0 = switch off, exact location is given)
(1 = switch on, automatically place chain in grid centre)
* VECTOR: the vector set describing the active chain conformation
* START_PT: the coordinates of the first
* MAP_FLAG: indicates chain fit in grid
* (1 = chain extends outside of grid, need to restart
simulation)
* 

***************************************************************************

* Parameter list
*****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, NVECTOR, NDIM, VNUM, DIM, MAP_FLAG,
+ MID_L, MAP_SWITCH,
    + STA\overline{RT PT (MAXD), MID_GRID (MAXD), GRIDLIM(MAXD),}
    + COORD(MAXD, MAXC), VECTOR(MAXD,MAXC), COORD_MAP (MAXD,MAXC)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
COMMON / LATTICEE / GRIDLIM
* Create vector map
*******************
DO }90\mathrm{ VNUM = 1, NVECTOR
DO 50 DIM = 1, NDIM
VECTOR(DIM,VNUM) = COORD(DIM,VNUM+1) -
+ COORD(DIM,VNUM)
5 0 ~ C O N T I N U E ~
90 CONTINUE
* Locate centre of chain and grid
MID_L = INT((L+1)/2)
DO \overline{100 DIM = 1, NDIM}
MID_GRID(DIM) = INT((GRIDLIM(DIM)+1)/2)
100 CONTINUE
* Place chain on grid and check for inconsistancies
*****************************************************

```
```

            IF (MAP SWITCH.EQ.1) THEN
            DO- 115 DIM = 1, NDIM
                        START_PT(DIM) = MID_GRID(DIM)
                        CONTINUE
            DO 140 VNUM = MID L-1, 1, -1
                DO 120 DIM = 1, NDIM
                    START_PT(DIM)=START_PT(DIM)-VECTOR(DIM,VNUM)
                    IF ((START PT(DIM).LT.0).OR.
                        (START_PT(DIM).GT.GRIDLIM(DIM))) THEN
                        MAP_FLAGG = 1
                        GOTO 200
                            ENDIF
                CONTINUE
            CONTINUE
    ENDIF
DO 150 DIM = 1, NDIM
COORD_MAP(DIM,1) = START_PT(DIM)
CONTINUE
DO 170 BEAD = 2, L
DO 160 DIM = 1, NDIM
COORD_MAP(DIM,BEAD) = COORD_MAP(DIM,BEAD-1) +
VECTOR(DIM, BEAD-1)
IF ((COORD_MAP(DIM,BEAD).IT.0).OR.

+ (COORD_MAP(DIM, BEAD).GT.GRIDLIM(DIM))) THEN
MAP_FL\overline{A}G=1
GOTO 200
ENDIF
CONTINUE
continue
200 CONTINUE
RETURN
END
SUBROUTINE WALL_BOUNCE (VECTOR, START_PT, BOUNCE_FLAG)
*********************\overline{*}***********************秋***************************
* Checks for coordinates outside of boundaries and reflects outlying
* points back into the grid. The trespassed boundary is the
* boundary surface.
* Main parameters:
* VECTOR: the vector set describing the active chain conformation
* START PT: the coordinates of the first bead of the active chain
* BOUNCE_FLAG: indicates the state of the newly-reflected conformation
* (0 = all coordinates of the new conformation are within
the lattice space)
(1 = chain extends outside of grid, even after reflections,
the proposed conformation is unacceptable)
* 

**************************************************************************

* Parameter list
*****************
PARAMETER (MAXD = 2, MAXC = 100)
INTEGER L, NDIM, DIM, NVECTOR, BEAD, VNUM, BOUNCE FLAG,
    + GRIDLIM(MAXD), START_PT(MAXD),
    + COORD (MAXD,MAXC), NEW_COORD (MAXD, MAXC),
    + VECTOR (MAXD,MAXC)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
COMMON / LATTICE / GRIDLIM
* Initialize parameters
************************

```
```

    BOUNCE_FLAG = 0
    * Check starting point and reflect into grid if necessary
DO 20 DIM = 1, NDIM
COORD(DIM,1) = START_PT(DIM)
IF (COORD(DIM,1).GT.GRIDLIM(DIM)) THEN
NEW_COORD(DIM,1) = (-1 * COORD(DIM,1)) +
(2 * GRIDLIM(DIM))
ELSEIF (COORD(DIM,1).LT.0) THEN
NEW COORD(DIM,1) = (-1 * COORD(DIM,1))
ELSE
NEW_COORD(DIM,1) = COORD(DIM,1)
ENDIF
IF ((NEW_COORD(DIM,1).GT.GRIDLIM(DIM)).OR.
(NEW COORD(DIM,1).LT.0)) THEN
BOUNCE_FLAG = 1
GOTO 100
ENDIF
START PT(DIM) = NEW COORD(DIM,1)
2 0
* Check remaining coordinates and reflect into grid if necessary
*********************************************************************
DO 60 BEAD = 2, L
DO 40 DIM = 1, NDIM
COORD(DIM, BEAD) = COORD(DIM,BEAD-1)
+ VECTOR(DIM,BEAD-1)
IF (COORD(DIM, BEAD).GT.GRIDLIM(DIM)) THEN
NEW_COORD(DIM,BEAD) = (-1 * COORD(DIM,BEAD))
+ (2 * GRIDLIM(DIM))
ELSEIF (COORD(DIM,BEAD).LT.0) THEN
NEW_COORD(DIM,BEAD) = (-1 * COORD(DIM, BEAD))
ELSE
NEW_COORD(DIM,BEAD) = COORD(DIM,BEAD)
ENDIF
IF ((NEW_COORD(DIM, BEAD).GT.GRIDLIM(DIM)).OR.
(NEW COORD(DIM, BEAD).LT.O)) THEN
BOUNCE_ELAG = 1
GOTO 100
ENDIF
CONTINUE
60 CONTINUE
* Re-establish vector map from new coordinates
**********************************************
DO }90\mathrm{ VNUM = 1, NVECTOR
DO }80\mathrm{ DIM = 1, NDIM
VECTOR(DIM,VNUM) = NEW_COORD(DIM,VNUM+1) -
+
CONTINUE
CONTINUE
100 CONTINUE
RETURN
END
|**t************************************************************************
SUBROUTINE WHIRIING (VECTOR, START PT, NEW_VECTOR, NEW START
+ WHIRL_FLAG)
* 
* Randomly rotates or translates the polymer at a given frequency.
* The length of translation is given initially in <*.dat>
* 
* Possible maneouvres:

```
```

* CODE MANOEUVER
* rotate 90* (counter-clockwise)
* rotate 180
* rotate 270
* 4 translate north
* translate south
* translate east
* translate west
* Main parameters:
VECTOR: the vector set describing the active chain conformation
START PT: the coordinates of the first bead of the active chain
NEW_VECTOR: newly-calculated vector map
* NEW START: newly-calculated starting point
* WHIRL_FLAG: indicates the state of the newly-reflected conformation
(0 = all coordinates of the new conformation are within grid
boundaries)
(1 = chain extends outside of grid after reflections,
the proposed conformation is unacceptable)
* 
* Parameter list
****************
PARAMETER (MAXD=2, MAXC=100)
INTEGER L, NDIM, NVECTOR, DIM, WHIRL, BEAD, BOUNCE_FLAG, VNUM,
    + WHIRL_FLAG, GRIDLIM(MAXD), NWHIRL, NTRANS,
    + START_PT(MAXD), NEW_START(MAXD), DIFF(MAXD), CENTRE (MAXD),
    + ROT_DIFF(MAXD), CHECK_PT(MAXD),
    + VECTOR (MAXD, MAXC), NEWW_VECTOR (MAXD, MAXC)
COMMON / GRAPH LIM / L, NDIM, NVECTOR
COMMON / WHIRL / NWHIRL, NTRANS
COMMON / LATTICE / GRIDLIM
* Initialize parameters
************************
CALL CENTRE_OF_MASS (VECTOR, START_PT, CENTRE)
DO 20 DIM = 1, NDIM
NEW_START(DIM) = START PT (DIM)

```

```

        ROT_DIFF(DIM) =0
    CONTINUE
    DO 30 VNUM = 1, NVECTOR
        DO 25 DIM = 1, NDIM
                NEW_VECTOR(DIM,VNUM) = VECTOR(DIM,VNUM)
            CONTINUE
    25 CON

* Randomly chose move
**********************
CALL RANDOM_INT (WHIRL, 1, 7)
* Calculate new position of conformation
******************************************
IF (WHIRL.EQ.1) THEN
ROT_DIFF(1) = -1 * DIFF(2)
ROT_DIFF(2)= DIFF(1)
DO }\overline{40}\textrm{DIM}=1, NDI
NEW_START(DIM) = START_PT(DIM) + ROT_DIFF(DIM)
CONTINUE
DO }60\mathrm{ VNUM = 1, NVECTOR
NEW VECTOR (1,VNUM) = -1 * VECTOR (2,VNUM)
NEW_-VECTOR(2,VNUM) = VECTOR(1,VNUM)

```
```

60 CONTINUE
ELSEIF (WHIRL.EQ.2) THEN
DO 120 DIM = 1, NDIM
NEW_START(DIM) = CENTRE(DIM) - DIFF(DIM)
CONTINUE
DO }160\mathrm{ VNUM = 1, NVECTOR
DO 140 DIM = 1, NDIM
NEW_VECTOR(DIM,VNUM) = -1 * VECTOR(DIM,VNUM)
CONTINUE
CONTINUE
ELSEIF (WHIRL.EQ.3) THEN
ROT_DIFF(1) = DIFF(2)
ROT DIFF(2) = -1 * DIFF(1)
DO 220 DIM = 1, NDIM
NEW_START(DIM) = START_PT(DIM) + ROT_DIFE(DIM)
CONTINUE
DO 230 VNUM = 1, NVECTOR
NEW VECTOR(1,VNUM) = VECTOR(2,VNUM)
NEW_VECTOR(2,VNUM) =-1 * VECTOR(1,VNUM)
CONTINUE
ELSEIF (WHIRL.EQ.4) THEN
NEW START(2) = START PT(2) + NTRANS
ELSEIF (WHIIRL.EQ.5) THEN
NEW START(2) = START_PT(2) - NTRANS
ELSEIF (WHIIRL.EQ.6) THEN
NEW_START(1) = START_PT(1) + NTRANS
ELSE
NEW_START(1) = START_PT(1) - NTRANS
ENDIF

* Check whether newly-calculated conformation is within grid boundaries
DO 235 DIM = 1, NDIM
CHECK_PT(DIM) = NEW_START (DIM)
IF ((CHECK_PT(DIM).GT.GRIDLIM(DIM)).OR.
(CHEC̄K PT(DIM).LT.O)) THEN
CALL WĀLL_BOUNCE (NEW_VECTOR, NEW_START,
BOUNCE FLAG)
IF (BOUNCE_FLAG.EQ..1) THEN
WHIRL FLAG = 1
GOTO }\overline{300
ENDIF
ENDIF
235 CONTINUE
DO 260 BEAD = 2, L
DO 240 DIM = 1, NDIM
CHECK PT(DIM) = CHECK PT(DIM) +
NEW_VECTOR(DIM, \overline{BEAD-1)}
IF ((CHECK PT(DIM).GT.GRIDLIM(DIM)).OR.
(CHECK_\overline{PT}(DIM).IT.O)) THEN
CALL WALL BOUNCE (NEW VECTOR, NEW START,
BOUNCE_FLAG)
IF (BOUNCE FLAG.EQ.1) THEN
WHIRL_FLAG = 1
GOTO }\overline{3}0
ENDIF
ENDIF
CONTINUE
CONTINUE
CONTINUE
RETURN
END
SUBROUTINE WRITE_DATA (NDIM, CHANGE_LIM, LIMIT, NWRITE, NSTOR,
    + GRIDX, GRIDY, NWHIRL_START, NTRANS, NSIGN, NFRAME,
    + NSTARTREC, MAP_SWITCH, START_PT,

```
```

+ WHIRL_SWITCH, WALL_SWITCH, CHI_HH, CHI_PP, CHI_HP,
+ 
+ 
+ 

CHI H\overline{S}, CHI PS, CHI-HW, CHI_PW, - CHI_SW, - KAPPA_H\overline{S}, PSI_HS,
KAPPA_SW, PSI_SW, TEMP, E_MAXX, L, COOORD, TYPE,
START-ORIGIN,\overline{COORD_ORIGIN)}
*

* Write starting data from file <*.dat> to output file
* 

PARAMETER ( $\operatorname{MAXD}=2, \operatorname{MAXC}=100)$
INTEGER NDIM, CHANGE LIM, LIMIT, NWRITE, NSTOR, L, BEAD,

+ NWHIRL_START, NTRANS, NSIGN, NFRAME, MAP_SWITCH,
+ WHIRL_SWITCH, WALL_SWITCH, NSTARTREC,
+ START_PT (MAXD), START_ORIGIN(MAXD),
+ COORD (MAXD,MAXC), COORD_ORIGIN (MAXD,MAXC)
REAL GRIDX, GRIDY,
+ CHI_HH, CHI_PP, CHI_HP, CHI_HS, CHI_PS, CHI_HW, CHI_PW,
+ CHI_SW, KAP\overline{PA_HS, PSII_HS, KAPPPA_SW,-PSI_SW, -TEMP, E_MAX}
CHARACTER*1 TYPE (MAXC)
WRITE (20,'(A15,I15)') 'NDIM', NDIM
WRITE (20,'(A15,I15)') 'CHANGE_LIM', CHANGE_LIM
WRITE (20,'(A15,I15)') 'LIMIT', LIMIT
WRITE (20,'(A15,I15)') 'NWRITE', NWRITE
WRITE (20,'(A15,I15)') 'NSTOR',. NSTOR
WRITE (20,'(A15,F15.2)') 'GRIDX', GRIDX
WRITE (20,'(A15,F15.2)') 'GRIDY', GRIDY
WRITE (20,'(A15,I15)') 'NWHIRL_START', NWHIRL_START
WRITE (20,'(A15,I15)') 'NTRANS', NTRANS
WRITE (20,'(A15,I15)') 'NSIGN', NSIGN
WRITE (20,'(A15,I15)') 'NFRAME', NFRAME
WRITE (20,'(A15,I15)') 'NSTARTREC', NSTARTREC
WRITE (20,'(A15,I15)') 'MAP_SWITCH', MAP_SWITCH
WRITE (20,'(A15,I15)') 'START_PT(1)', ST\overline{ART_PT(1)}
WRITE (20,'(A15,I15)') 'START-PT(2)', START_PT(2)
WRITE (20,'(A15,I15)') 'WHIRL_SWITCH', WHIRL_SWITCH
WRITE (20,'(A15,I15)') 'WALL_S̄WITCH', WALL_SWITCH
WRITE (20,'(A15,F15.2)') 'CHI_HH', CHI_HH
WRITE (20,'(A15,F15.2)') 'CHI_PP', CHI_PP
WRITE (20,'(A15,F15.2)') 'CHI_HP', CHI_HP
WRITE (20,'(A15,F15.2)') 'CHI_HS', CHI_HS
WRITE (20,'(A15,F15.2)') 'CHI_PS', CHI_PS
WRITE (20,'(A15, F15.2)') 'CHI_HW', CHI_HW
WRITE (20,'(A15,F15.2)') 'CHI_PW', CHI_PW
WRITE (20,'(A15,F15.2)') 'CHI_SW', CHI_SW
WRITE (20,'(A15,F15.2)') 'KAPPA_HS', KAPPA_HS
WRITE (20,'(A15,F15.2)') 'PSI_HS'', PSI HS
WRITE (20,'(A15,F15.2)') 'KAPPA_SW', KAPPA_SW
WRITE (20,'(A15,F15.2)') 'PSI,SW', PSI_SW
WRITE (20,'(A15,F15.2)') 'TEMP', TEMP
WRITE (20,'(A15,F15.2)') 'E_MAX', E_MAX
WRITE (20,'(A15,I15)') 'L', L
DO }60\mathrm{ BEAD = 1, L
WRITE (20,80) COORD (1,BEAD), COORD (2,BEAD), TYPE(BEAD)
60
CONTINUE
FORMAT (4X,2I4,3X,A1)
WRITE (20,'(A30)') 'COMPARISON CONFORMATION'
WRITE (20,'(A15,I15)') 'START_ORIGIN(1)', START_ORIGIN(1)
WRITE (20,'(A15,I15)') 'START_ORIGIN(2)', START_ORIGIN(2)
DO 120 BEAD = 1, L
WRITE (20,140) COORD_ORIGIN(1,BEAD),COORD_ORIGIN (2,BEAD)
120 CONTINUE
140 FORMAT (4X,2I4)
WRITE (42,'(A15,I15)') 'NDIM', NDIM
WRITE (42,'(A15,I15)') 'CHANGE_LIM', CHANGE_LIM
WRITE (42,'(A15,I15)') 'LIMIT', LIMIT

```
```

    WRITE (42,'(A15,I15)') 'NWRITE', NWRITE
    WRITE (42,'(A15,I15)') 'NSTOR', NSTOR
    WRITE (42,'(A15,F15.2)') 'GRIDX', GRIDX
    WRITE (42,'(A15,F15.2)') 'GRIDY', GRIDY
    WRITE (42,'(A15,I15)') 'NWHIRL_START', NWHIRL_START
    WRITE (42,'(A15,I15)') 'NTRANS'', NTRANS
    WRITE (42,'(A15,I15)') 'NSIGN', NSIGN
    WRITE (42,'(A15,I15)') 'NFRAME', NFRAME
    WRITE (42,'(A15,I15)') 'NSTARTREC', NSTARTREC
    WRITE (42,'(A15,I15)') 'MAP_SWITCH', MAP_SWITCH
    WRITE (42,'(A15,I15)') 'START_PT(1)', START_PT(1)
    WRITE (42,'(A15,I15)') 'START_PT(2)', START_PT(2)
    WRITE (42,'(A15,I15)') 'WHIRL SWITCH', WHIRL SWITCH
    WRITE (42,'(A15,I15)') 'WALL_SWITCH', WALL_SWITCH
    WRITE (42,'(A15,F15.2)') 'CHI_HH', CHI_HH
    WRITE (42,'(A15,F15.2)') 'CHI_PP', CHI_PP
    WRITE (42,'(A15,F15.2)') 'CHI_HP', CHI_HP
    WRITE (42,'(A15,F15.2)') 'CHI.HS', CHI_HS
    WRITE (42,'(A15,F15.2)') 'CHI_PS', CHI_PS
    WRITE (42,'(A15,F15.2)') 'CHI_HW', CHI_HW
    WRITE (42,'(A15,F15.2)') 'CHI_PW', CHI_PW
    WRITE (42,'(A15,F15.2)') 'CHI_SW', CHI_SW
    WRITE (42,'(A15,F15.2)') 'KAP的_HS', KA\overline{APPA_HS}
    WRITE (42,'(A15,F15.2)') 'PSI_HS', PSI_HS
    WRITE (42,'(A15,F15.2)') 'KAPPA_SW', KAPPA_SW
    WRITE (42,'(A15,F15.2)') 'PSI_SW'', PSI_SW
    WRITE (42,'(A15,F15.2)') 'TEMP', TEMP
    WRITE (42,'(A15,F15.2)') 'E_MAX', E_MAX
    WRITE (42,'(A15,I15)') 'L', L
    DO 160 BEAD = 1, L
            WRITE (42,180) COORD (1, BEAD), COORD (2,BEAD), TYPE(BEAD)
    CONTINUE
    FORMAT (4X,2I4,3X,A1)
    WRITE (42,'(A30)') 'COMPARISON CONFORMATION'
    WRITE (42,'(A15,I15)') 'START_ORIGIN(1)', START_ORIGIN(1)
    WRITE (42,'(A15,I15)') 'START_ORIGIN(2)', START_ORIGIN(2)
    DO 220 BEAD = 1, L
        WRITE (42,240) COORD_ORIGIN (1, BEAD), COORD_ORIGIN (2,BEAD)
    CONTINUE
    FORMAT (4X,2I4)
    RETURN
    END
    *****************************************************************************

* FUNCTIONS:
***************************************************************************
INTEGER FUNCTION F DOT(VECTOR_A, VECTOR_B)
*************************\overline{*}**********\overline{*}**********\overline{*}************************
* 
* Dot product function
* 

**************************************************************************

* Parameter list
****************
PARAMETER (MAXD = 2)
INTEGER NDIM, DIM, NVECTOR,
    + VECTOR_A(MAXD), VECTOR_B (MAXD)
COMMON / GRAPH_LIM / L, NDIM, NVECTOR
* Initialize parameter
***********************
F_DOT = 0
* Calculate dot product

```
```

***********************
DO 20 DIM = 1, NDIM
F_DOT = F_DOT + VECTOR_A(DIM)*VECTOR_B(DIM)
CONTINUE
RETURN
END
*************************************************************************
REAL FUNCTION F ENERGY(XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW)
*********************\overline{\star}******************************************************)
*

* Calculates chain energy
* 

************************************************************************

* Parameter list
****************
INTEGER XHH, XPP, XHP, XHS, XPS, XHW, XPW, XSW
REAL CHI_HH, CHI_PP, CHI_HP, CHI_HS,
+ - CHI_PS, C\overline{HI_HW, CHI_PW, CHI_SW,}

```

```

        COMMON / CHI / CHI_HH, CHI_PP, CHI_HP, CHI_HS,
        + CHI_PS, CHI_HW, CHI_PW, CHI_SW,
        + KAP\overline{PA_HS, PSI_HS, KAPPPA_SW, - PSI_SW}
        COMMON / TEMP / TEMP
    * Option: If HS and SW interaction energies are functions of
* temperature, then the next two lines should be activated
************************************************************************
* CHI_HS = KAPPA_HS / TEMP + PSI_HS
* CHI_SW = KAPPA_SW / TEMP + PSI_SW
* Calculate energy
*******************
F_ENERGY = (CHI_HH*XHH) + (CHI_PP*XPP) + (CHI_HP*XHP) +
+-}(\textrm{CHI}HS*X\overline{HS})+(CHI_PS*X\overline{PS})+(CHI_HW*XMW) +
    + (CHI_PW*XPW) + (CHI_SW*XSW)
RETURN
END
*************************************************************************
REAL FUNCTION F_PROB(D_E)
*********************\overline{*}******\overline{*}*******************************************
* 
* Calculates probability according to Boltzmann relation
* 

***********************************************************************

* Parameter list
*****************
REAL D_E, K_bolz, TEMP
INTRINSIC EXP
COMMON /TEMP/ TEMP
* Boltzmann constant,k_bolz, in units of J/K
K_bolz = 1.381E-23
* Calculate probability
*************************
F_PROB = EXP(-1.0*D_E)
RETURN
END

```

\section*{Input data, <test.dat>}

The input file, <test.dat>, is shown below. The parameters were chosen to give a short demonstration of the possible outputs from CONTACT9. Explanations of input parameters are first given.
[NDIM]
The dimension number. For example, CONTACT9 is a 2 D simulation, and therefore \(\mathrm{NDIM}=2\).

\section*{[CHANGE_LIM]}

The maximum number of times the central subroutine, PROTEIN_ADSORPTION_2 is to be run. The computer has a counting limit of \(2^{32}\), and therefore it is necessary to run the main subroutine within an added outer loop to exceed this constraint for very long simulations. CHANGE_LIM defines the maximum number of times the subroutine is to be carried out in series.
[LIMIT]
The maximum number of cycles to be run in PROTEIN_ADSORPTION_2. It cannot exceed \(2^{32}\).
[NWRITE]

Frequency of sampling to output file <*.out>. Sampling occurs at multiples of the number listed. For example, when NWRITE \(=1000 \dot{0}\), sampling occurs at 0,10000 , 20000, etc. Information provided by this function includes the energy and configurational profiles of the chain.
[NSTOR]
Maximum size of storage array of unique low energy conformations.
[GRIDX]
The size of the grid in the \(x\)-dimension as multiple of chain length. For example, in a simulation of an 18 -residue chain when GRIDX \(=0.50\), the grid is 9 units wide. Since the chain occupies the coordinate sites (and not the spaces in between them), a lattice 9 units wide accommodates 10 chain residues.

\section*{[GRIDY]}

The size of the grid in the \(y\)-dimension as multiple of chain length (not necessarily equal to GRIDX).
[NWHIRL_START]
Frequency of a rotation or translation move. The type of move is chosen randomly.

\section*{[NTRANS]}

Length of each translation move.

Frequency of screen output. This function serves to inform the user on the progress of the simulation.

\section*{[NFRAME]}

Frequency of output to \(<^{*} . g p h>\) file. Information provided by this function is restricted to single line reports of chain energy and its resemblance to a comparison conformation. This function is designed to accommodate frequent sampling of long simulations.

\section*{[NSTARTREC]}

The number of cycles to pass before sampling takes place. This function is used when an equilibrium period is needed.

\section*{[MAP_SWITCH]}

The switch controlling the automatic chain placement function at the start of the simulation.
\(0=\) off, the chain is placed according to the coordinates given.
1 = auto, the chain is automatically placed in the middle of the lattice at the start of the simulation.
[START_PT] (starting \(x\)-coordinate)
[START_PT] (starting \(y\)-coordinate)

Coordinates of the first chain residue at the start of the simulation. Only valid if MAP_SWITCH \(=1\).
[WHIRL_SWITCH]
The switch controlling automatic suspension of the WHIRLING mechanism.
\(0=\) off, the WHIRLING mechanism continues at the frequency indicated by NWHIRL_START

1 = auto, the WHIRLING mechanism shuts off when contacts between the chain and an active wall are detected.
[RADIUS_SWITCH]
The switch controlling automatic calculation and presentation of radius of gyration data.
Data using this function are generated at the frequency indicated by NFRAME.
\(0=\mathrm{off}\), the report does not include \(R_{g}\) data
1 = on, the calculations includes \(R_{g}\) data
[WALL_SWITCH]
The switch controlling which boundaries become active surfaces.
4 = all sides are active walls
\(1=\) vertical walls on (i.e. surfaces at \(x=0\) and GRID_LIM \((y)\) )
\(2=\) horizontal walls on (i.e. surfaces at \(y=0\) and GRID_LIM \((x)\) )
[CHI_HH]

Interaction energy between HH chain residues.
[CHI_PP]
Interaction energy between PP chain residues.
[CHI_HP]
Interaction energy between HP chain residues.
[CHI_HS]
Interaction energy between H residues and solvent units.
[CHI_PS]
Interaction energy between P residues and solvent.
[CHI_HW]
Interaction energy between H residues and an active wall unit.
[CHI_PW]
Interaction energy between \(P\) residues and an active wall unit.
[CHI_SW]
Interaction energy between solvent and an active wall unit.
[KAPPA HS]
Internal energy portion of the temperature-dependent interaction energy between \(H\) residue and solvent units. There is a choice in the F_ENERGY function to set the HS interaction energy to be temperature dependent. (The program code must be accessed to control this function.) In this case, the interaction energy, CHI_HS, is divided into internal energy, KAPPA_HS, and entropy (PSI_HS) components using the relation: CHI_HS \(=(\) KAPPA_HS \(/\) TEMP \()+\) PSI_HS
[PSI_HS]
The entropy portion of the temperature-dependent interaction energy between H -residue and solvent units. (See KAPPA_HS for further explanation.)
[KAPPA_SW]
The internal energy portion of the temperature-dependent interaction energy between solvent and active wall units. There is a choice in the F_ENERGY function to set the SW interaction energy to be temperature dependent. (The program code must be accessed to control this function.) In this case, the interaction energy, CHI_SW, is divided into internal energy, KAPPA_SW, and entropy (PSI_SW) components using the relation: CHI_SW \(=(\) KAPPA_SW \(/\) TEMP \()+\) PSI_SW
[PSI_SW]
The entropy portion of the temperature-dependent interaction energy between solvent and active wall units. (See KAPPA_SW for further explanation.)
[TEMP]
System temperature.
[E_MAX]
Maximum energy limit when searching for low energy conformations. In searching for low energy structures, the selected conformations have to have calculated energies equivalent or lower than E_MAX.

\section*{[L]}

Chain length.
[COORD] and [TYPE]
Parameters describing the chain sequence. The sequence is entered into the input file as a series of coordinates and letters describing the residue type. The location of the coordinates with respect to the simulation space is not relevant because the program automatically places the chain either automatically in the centre of the grid or according to the position specified by START_PT. The format for the input is as follows:

Columns 1-4: indicator of residue number
Columns 5-8: \(x\)-coordinate (justify right)
Columns 9-12: \(y\)-coordinate (justify right)
Column 16: residue type, designated as H or P

The number of lines allotted for the sequence data is exactly the number of residues in the chain, and therefore no empty lines can be inserted between data sets in <*.dat>.
[START_ORIGIN] \(x\)-coordinate
[START_ORIGIN] \(y\)-coordinate
Coordinates for the first residue of the comparison conformation. A second chain conformation is entered into the data file to serve as a comparison for conformations generated by the simulation.

\section*{[COORD_ORIGIN]}

Sequence of the comparison chain. The input format here resembles that of the active chain sequence without the column indicating the residue type.

Columns 1-4: indicator of residue number

\section*{Columns 5-8: \(x\)-coordinate (justify right)}

Columns 9-12: \(y\)-coordinate (justify right)
```

Input data for 2D ADSORPTION PROGRAM - CONTACT9 EDITION
2 [NDIM] number of dimensions
1 [CHANGE LIM] number of simulations
5000000 [LIMIT] number of cycles for each simulation
50000 [NWRITE] frequency of conformations printed to output file
50 [NSTOR] size of storage array for lowest energy confns
1.00 [GRIDX] grid x-dimension (as multiple of chain length)
5.00 [GRIDY] grid y-dimension (as multiple of chain length)
10000 [NWHIRL_START] frequency of rotation or tranlation
5 [NTRANS] length of translation
20000 [NSIGN] frequency of output to screen
500 [NFRAME] frequency of output to movie or other output file
O [NSTARTREC] steps to pass before output recorded
Diffusion parameters
O [MAP_SWITCH] automatic placement of chain in centre
(0 = off, 1 = auto)
10 [START_PT] starting x-coordinate

```


\section*{Output files}
<Screen>

The output is a series of short statements printed to the screen (and in this case channelled to file \(<\) Screen \(>\) ) at the frequency specified by NSIGN.
"SIM" refers to which run of the main subroutine is currently being carried out. The number of times the subroutine is carried out is specified by CHANGE_LIM.
"STEP" refers to the cycle in the simulation (called NSTEP in the program).
"ENERGY" refers to the energy of the chain.

<test.out>

The first lines in <test.out> list simulation parameters given by the input file. Below this, reports of the chain's position and conformational state are given, sampled at the
frequency specified by NWRITE. Also included is a list of the number of contacts made at the sampling point.
"SIMULATION \#" refers to which run of the main subroutine is currently being carried out.
"STEP \#" refers to the cycle in the simulation.
"HH", "PP", etc. refer to the number of contacts made between the components indicated.
"ENERGY" refers to the energy of the chain.
"POSITION" refers to the coordinates of the first residue of the chain.
"VECTORS" refers to the series of vectors describing the chain conformation.

The final portion of \(<^{*}\).out \(>\) gives information regarding the low energy conformations stored by the program.
"\%AGE OF SUCCESSFUL MOVES" refers to the percentage of successful moves over the total number of cycles. One simulation cycle is equivalent to a single attempted move. The Monte Carlo algorithm specifies that only energetically favourable moves and a small weighted number of unfavourable moves are allowed. As a consequence, relatively few attempted moves are actually carried out. The ratio of these successful moves is calculated in the simulation.
"MAXIMUM ENERGY RECORDED" refers to the upper energy limit for the list of low energy conformations. It is the value specified by E_MAX.
"LOW ENERGY CONFORMATION \#" refers to the order in which the configurations were detected in the simulation.
"SIMULATION \#" refers to which run of the main subroutine is currently being carried out.
"STEP COUNT" refers to the cycle in the simulation.
"HH", "PP", etc. refer to the number of contacts made between the components indicated.
"POSITION" refers to the coordinates of the first residue of the chain.
"ENERGY" refers to the energy of the chain.
"VECTORS" refers to the series of vectors describing the chain conformation.

The last output line gives the run-time of the simulation.

```

| 1 | 1 | P |
| ---: | ---: | ---: |
| 2 | 1 | P |
| 2 | 0 | H |
| 3 | 0 | P |
| 3 | -1 | H |
| 2 | -1 | H |
| 2 | -2 | H |
| 3 | -2 | H |
| 3 | -3 | H |
| 2 | -3 | H |
| 1 | -3 | H |
| 1 | -2 | H |
| 0 | -2 | H |
| 0 | -1 | H |
| 1 | -1 | H |
| 1 | 0 | H |

COMPARISON CONFORMATION
START ORIGIN(1) 10
START_ORIGIN(2)
0}0
1 1
2
0
0
-1
-1
-2
-2
-3
-3
-3
-2
0
0
1
************************************************
************************************************

```

```

    ENERGY -48.
    POSITION 10 3
    VECTORS
    O1
    10
    10
    0-1
    10
    0 -1
    -1 0
    0-1
    10
    0
    -1 0
    -1 0
    0 1
    -1 0
    01
    10
    O 1
    SIMULATION \# 1 STEP \# 50000
HH
ENERGY -48.
POSITION }9
VECTORS
01
10
10

```
```

1
0 -1
-1 0
0-1
1 0
0 -1
-1 0
-1 0
-1 0
0 1
10
01
-1 0
*******************************************
(etc.)
SIMULATION \#
ENERGY -52.
POSITION 15 0
VECTORS
O 1
0 1
O 1
-1 0
-1 0
0-1
10
0-1
0-1
-1 0
-1 0
-1 0
-1 0
0 1
10
10
10
*******************************************
SIMULATION \#
ENERGY -52.
POSITION 15 0
VECTORS
01
01
0 1
-1 0
-1 0
0
1 0
0-1
0-1
-1 0
-1 0
-1 0
-1 0
01
10
10
10
*****************************************

```

LIMIT REACHED FOR SIMULATION \# 1 \%AGE OF SUCCESSFUL MOVES IS 5.97279978
****** THE LOWEST ENERGY CONFORMATIONS ****** MAXIMUM ENERGY RECORDED \(=-52\).
```

    LOW ENERGY CONFORMATION # 1
    SIMULATION # 1 STEP COUNT # 4048833
    HH 7 HP 2 Pr PP 0
    POSITION 150 ENERGY -52.
    VECTORS
    O 1
O 1
O 1
-1 0
-1 0
0-1
1 0
0 -1
0-1
-1 0
-1 0
-1 0
-1 0
01
10
1 0
10

```

(etc.)
```

LOW ENERGY CONFORMATION \# 9
SIMULATION \# 1 STEP COUNT \# 4874088
HH
POSITION 15 0 ENERGY -52.
VECTORS
0 1
0 1
0 1
-1 0
-1 0
0-1
10
0-1
0 -1
-1 0

```
```

-1 0
-1 0
-1 0
0 1
10
1 0
10
PROGRAM IS COMPLETE
PROGRAM RUN-TIME: 0 HOURS 0 MINUTES 33 SECONDS
<test.gph>

```

The first part of <test.gph> is the list of simulation parameters.

The second half of the file lists results sampled at the frequency specified by NFRAME, but starting only after the value defined by NSTARTREC. The simulation output shown here includes energy, conformation and radius of gyration data.

Columns 1 to 4: Simulation \#
Columns 5 to 14: Step \#
Columns 15 to 21: Energy
Columns 22 to 28: Fraction matching comparison configuration. The intramolecular contacts within the chain at the time of sampling are matched to those given for the comparison conformation. The fraction given is the number of similar contacts found in the sampled chain over the total number of intramolecular contacts of the comparison configuration.

Columns 29 to 32: Number of intramolecular contacts in the sampled chain matching those in the comparison configuration.

Columns 33 to 35: Number of intramolecular contacts in the sampled chain not matching those in the comparison configuration.

Columns 36 to 39: Number of contact made between the sampled chain and a boundary site (either active or non-active).

Columns 40 to 48: The calculated radius of gyration for the sampled chain.
Columns 49 to 57: The deformation ratio of the sampled chain. This value is dimensional ratio of the \(x\) and \(y\) components of the radius of gyration when calculating the averaged value.

```


[^0]:    * A version of this chapter is published in the Journal of Colloid and Interface Science. [Reference: Liu, S.M. and Haynes, C.A., Mesoscopic analysis of conformational and entropic contributions to nonspecific adsorption of HP copolymer chains using dynamic Monte Carlo simulations. J. Colloid Interface Sci. 275 (2004) 458-469.]

[^1]:    * A version of this chapter is currently in press in the Journal of Colloid and Interface Science.

[^2]:    * A version of this chapter is currently in press in the Journal of Colloid and Interface Science.

