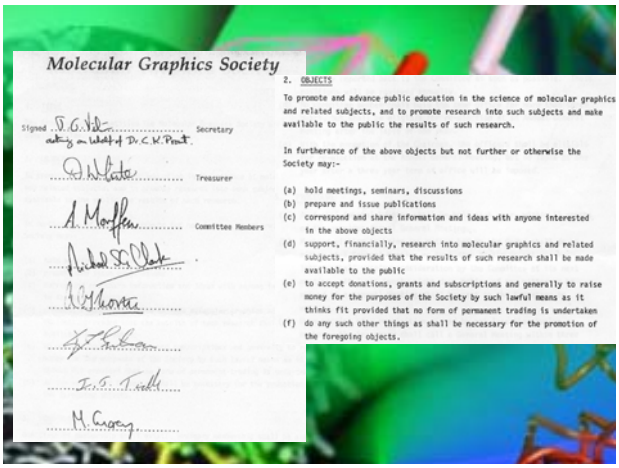
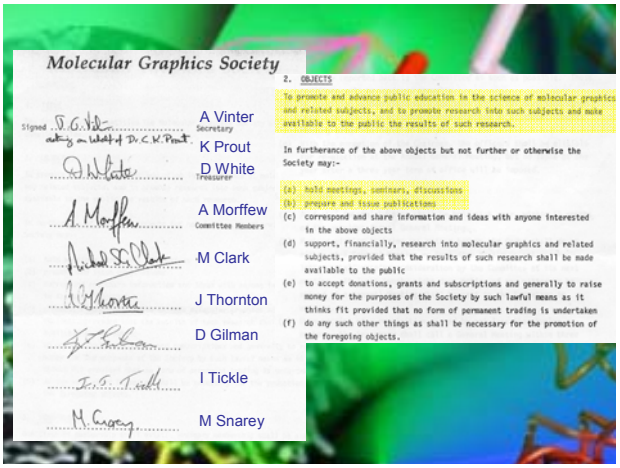
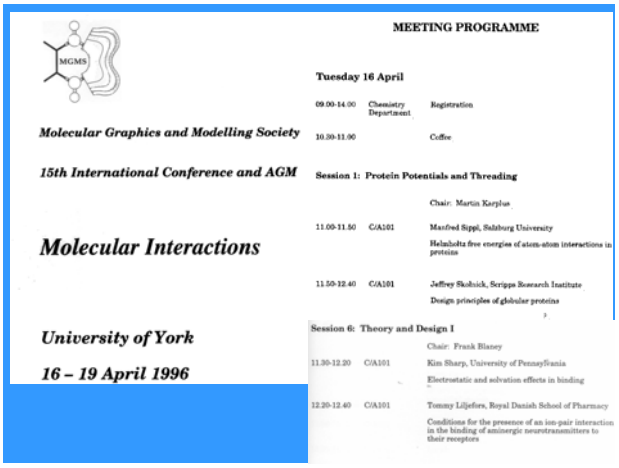
# Towards Improved Models of Carbohydrate Structure and Dynamics

Richard A. Bryce
*School of Pharmacy & Pharmaceutical Sciences*
*University of Manchester*

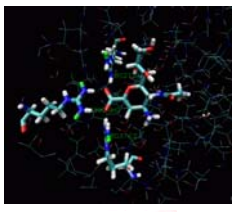


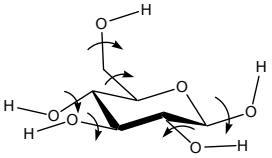

## Carbohydrates

- polar
- anomeric (stereoelectronic) effect
- flexible
- hypervalent
- biomolecular recognition



## Computational Challenge I

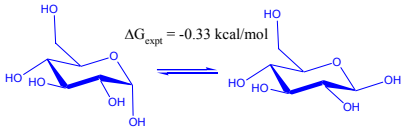
- floppy conformations (especially in solution)
  - adequate sampling difficult *eg*. one glucose anomer has $3^6 = 729$ conformations



## Computational Challenge II

- carbohydrate force fields
  - CHARMM
    - Rasmussen (1979), Brady (1988)
  - MM*x (x=2,3,4)*
    - Allinger (1976), Jeffrey (1980), Allinger (2003)
  - AMBER
    - Homans (1990), Merz (1992), Woods (1995), Kollman (1998), Woods (2000)
  - OPLS-AA
    - Jorgensen (1997), van Gunsteren (2000)

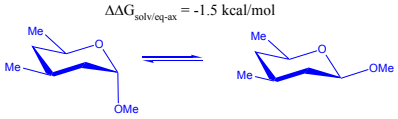
## Anomeric Equilibrium




- glucose
  - Karplus JACS 1991 (+0.31 kcal/mol)
- xylose
  - Karplus JACS 1996 (+0.15 kcal/mol)

## Effect of Solvent Model

- anomeric equilibrium of model solute (tetrahydropyran derivative)




- reference interaction site model (RISM)

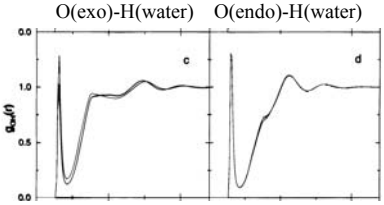
$$h_{\alpha\beta} = \sum_{\alpha'\gamma'} \omega_{\alpha'\alpha} c_{\alpha'\gamma'} \omega_{\gamma'\gamma} + \rho \sum_{\alpha'\gamma'} \omega_{\alpha\alpha'} c_{\alpha'\gamma'} h_{\gamma'\gamma}$$

## Differential Solvation Energies

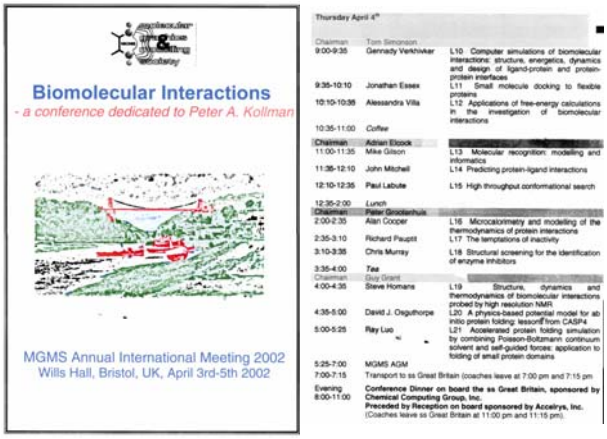
| Model | Method | $\Delta\Delta G_{eq-ax}$ |
|---|---|---|
| DMT | calorimetry | -1.5 |
| DMT | AM1/SM5.4 | -0.6 |
| DMT | PCM | -0.7 |
| MTP | OPLS/FEP | -2.1 |
| DMT | RISM/HNC | -1.7 |
| DMT | RISM/GF | -1.7 |

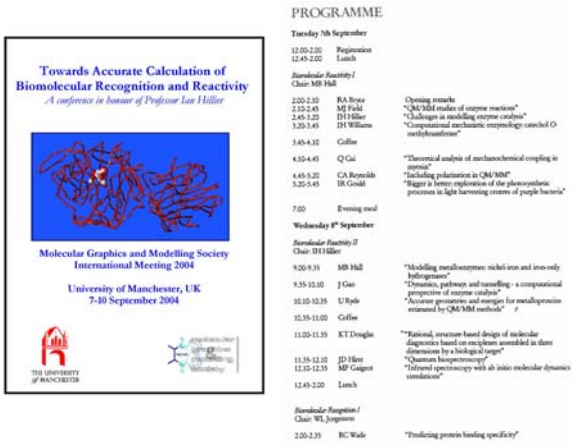
Differential Solvation Energies (kcal/mol)

## Differential Solvation Energies

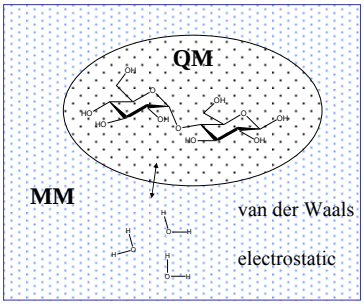



- differential solvation mainly at exocyclic O






# Hybrid QM/MM Methods

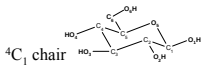


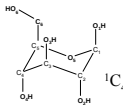

$$E_{tot} = E_{QM} + E_{QM/MM} + E_{MM}$$
$$= \langle \Psi_0 | \hat{H}_{QM}^{elec} - \sum_{k} \frac{Q_k}{r_{ik}} | \Psi_0 \rangle + E_{QM}^{nuc}$$
$$+ E_{QM/MM}^{nuc} + E_{QM/MM}^{vdW} + E_{QM/MM}^{bad} + E_{MM}$$

$$V_{\mu\nu}^{chg} = -\sum_{s} Q_s \langle \mu | r_{is}^{-1} | \nu \rangle$$
$$E_{QM/MM}^{elec} = 2 \sum_{\mu\nu} D_{\mu\nu} V_{\mu\nu}^{chg}$$

# Subtle Conformational Energetics

- Quantum mechanics:


$^4C_1$ chair


$^1C_4$ chair

| Conf | comp | MP2/ 6-31G* | MP2/cc-pVDZ | MP2/cc-pVTZ$^D$ | B3LYP/ 6-31G* | B3LYP/ 6-31+G* | HF/ 6-31G* |
|---|---|---|---|---|---|---|---|
| $^4C_1$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^4C_1$ | 0.27 | -0.45 | -0.50 | 0.07 | -0.70 | -0.12 | -0.15 |
| $^1C_4$ | 6.41 | 0.66 | -0.11 | 4.13 | 1.65 | 6.92 | 6.73 |
| $^1C_4$ | 6.99 | -0.60 | -0.81 | 3.65 | -0.60 | 4.93 | 6.76 |

Barrows *et al. Carbohydr. Res.* **1995**, *276*, 219

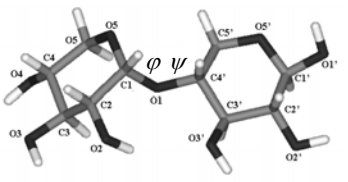
# Semi-empirical methods

| Conf | comp | PM3 |
|---|---|---|
| $^4C_1$ | 0.00 | 0.00 |
| $^4C_1$ | 0.27 | -1.50 |
| $^1C_4$ | 6.41 | -0.90 |
| $^1C_4$ | 6.99 | 2.30 |

- PM3 applied to carbohydrates:
  - Woods et al., *Chem. Comm.* 1991
  - Cramer & Truhlar, *JACS* 1993, 1994
  - PM3 also overstabilises $^1C_4$
- *cf.* condensed phase calculations using QM/MM approach:
  - treat carbohydrate solute *via* PM3 Hamiltonian
  - treat solvent using TIP3P force field

# Molecular Dynamics

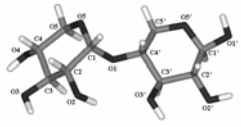
- PM3/TIP3P dynamics of dixylose

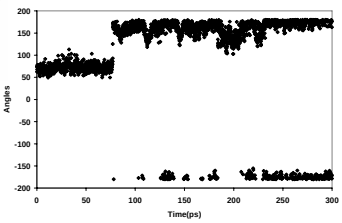



4-O-α-D-xylopyranosyl-α-D-xylopyranose

# Molecular Dynamics

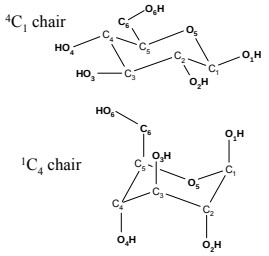
- PM3/TIP3P dynamics of dixylose





## $^1C_4$ Global Minimum

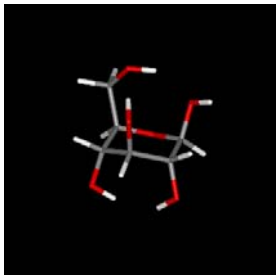
- pyranose ring conformation (Barrows et al., 1995)




| Conf | B3LYP | PM3 |
|---|---|---|
| $^4C_1$/cc/tg | 0.00 | 0.00 |
| $^1C_4$/cc/gg | 9.48 | 1.44 |
| $^1C_4$/cc/g$^-$g$^-$ | 8.77 | 0.82 |
| $^1C_4$/cl/g$^-$g | 8.69 | -2.43 |

## $^1C_4$/β/cl/g$^-$g

- underestimation of H-H repulsion



## *cf* B3LYP Structure

## Semi-empirical Theory

- minimal basis in field of nuclear cores
- neglect/parametrize integrals of Hamiltonian
  - Neglect of Diatomic Differential Overlap (NDDO) approximation

$$(\mu\nu \mid \lambda\sigma) = (\mu^A\nu^A \mid \lambda^B\sigma^B)$$

= 0 if μ,ν not both on A
= 0 if λ,σ not both on B
no 3 or 4 centre integrals

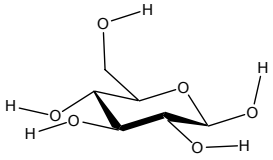
  - used in MNDO, AM1, PM3
  - developed to handle "large" systems
  - limited use, as difficult to identify systematic errors
  - poor treatment of weak interactions (H-bonds and dispersion)

## Semi-empirical Methods

- Specific Reaction Parameters (SRP)
  - Rossi & Truhlar, *Chem. Phys. Lett.* 1995
  - fitted parameters for H, C, Cl to structure and energetics of *ab initio* QM calculations
- fit to *ab initio* energies/geometries *cf.* a force field
  - fit H and O parameters
  - require training set

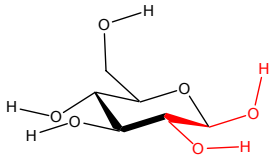
## SRP for Carbohydrates

- base on fitting to small representative molecules

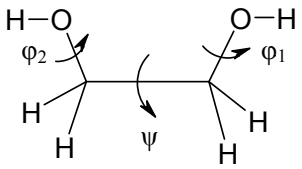


## SRP for Carbohydrates

- base on fitting to small representative molecules



## Ethanediol Conformations

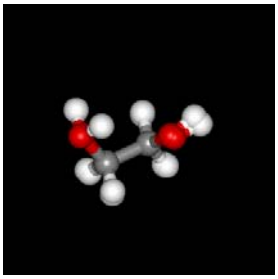



10 unique conformers

## Ethanediol Energetics

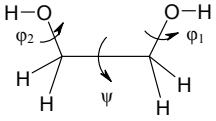
- intramolecular hydrogen bond (Cramer & Truhlar, 1994)

| Conf | MP2 | AM1 | PM3 |
|---|---|---|---|
| g⁻Gg⁻ | 1.19 | -0.98 | -0.78 |
| gGg | 3.15 | -0.98 | -0.82 |
| gGg⁻ | 0.30 | -0.88 | -1.39 |
| gTg | 3.00 | 0.39 | -0.43 |
| gTg⁻ | 2.81 | 0.20 | -0.67 |
| tGg | 3.81 | 2.56 | 1.76 |
| tGg⁻ | 0.00 | 0.00 | 0.00 |
| tGt | 3.48 | 3.19 | -0.78 |
| tTg | 2.92 | 1.55 | 1.07 |
| tTt | 2.87 | 2.98 | 2.56 |



## 1,2-Ethanediol Conformers

| Conf | MP2 | AM1 | PM3 | PM3CARB-1 |
|---|---|---|---|---|
| g⁻Gg⁻ | 1.19 | -0.98 | -0.78 | 1.45 |
| gGg | 3.15 | -0.98 | -0.82 | 1.75 |
| gGg⁻ | 0.30 | -0.88 | -1.39 | 1.10 |
| gTg | 3.00 | 0.39 | -0.43 | 2.87 |
| gTg⁻ | 2.81 | 0.20 | -0.67 | 2.50 |
| tGg | 3.81 | 2.56 | 1.76 | 3.63 |
| tGg⁻ | 0.00 | 0.00 | 0.00 | 0.00 |
| tGt | 3.48 | 3.19 | -0.78 | 2.56 |
| tTg | 2.92 | 1.55 | 1.07 | 2.40 |
| tTt | 2.87 | 2.98 | 2.56 | 2.23 |




•RMS error in energy = 0.57 kcal/mol (PM3 = 2.56 kcal/mol)

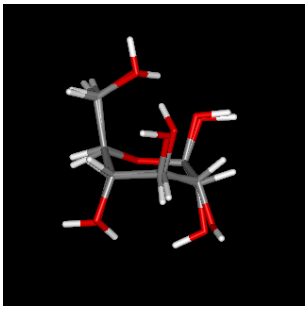
•RMS error in structure = 6.5 deg (PM3 = 11.6 deg)

## β-glucopyranose

- improved ranking of ring conformers

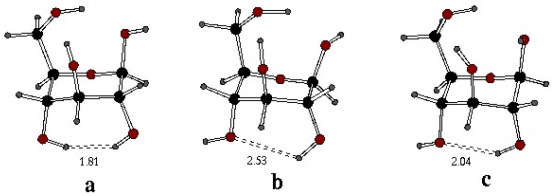
| Conf | B3LYP | PM3 | PM3CARB-1 |
|---|---|---|---|
| $^4C_1$/cc/tg | 0.00 | 0.00 | 0.00 |
| $^1C_4$/cc/gg | 9.48 | 1.44 | 6.52 |
| $^1C_4$/cc/g⁻g⁻ | 8.77 | 0.82 | 6.31 |
| $^1C_4$/cl/g⁻g | 8.69 | -2.43 | 4.73 |

## β-glucopyranose



## β-glucopyranose

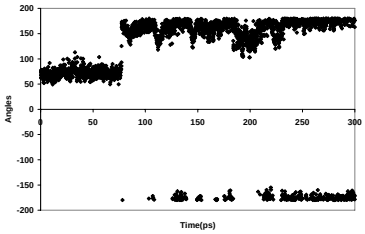
- improvement in structure




PM3 PM3CARB-1 HF/6-31G*

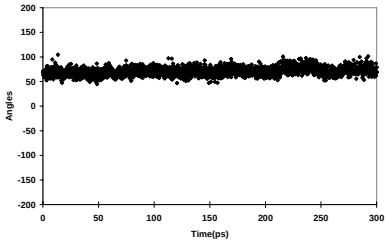
## QM/MM Molecular Dynamics

- PM3/TIP3P dynamics of dixylose



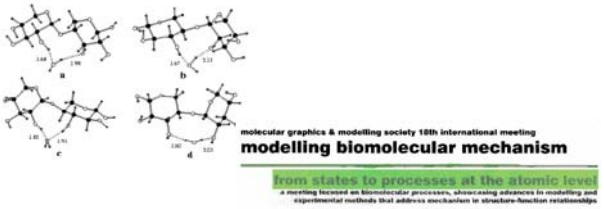

## QM/MM Molecular Dynamics

- PM3CARB-1/TIP3P dynamics of dixylose
- stable $^4C_1$



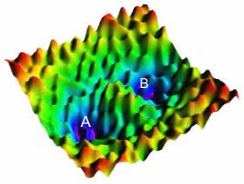

## Modelling carbohydrate dynamics

– oligosaccharides can adopt different conformations
– large conformational space
  – complex conformational transitions
  – sufficient to describe by a single TS, pathway?



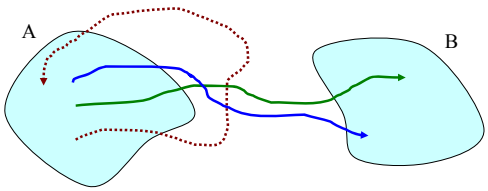

## Transition Path Sampling

- generates *ensemble* of transition pathways and transition states
  - developed by Chandler to study transition events in complex systems
  - models conformational or reactive transitions without assumption of specific reaction coordinates




## Generation of the TP Ensemble

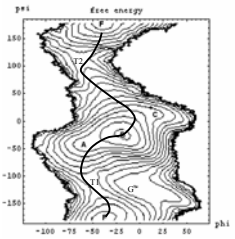
- apply Metropolis importance sampling to bias towards transitional trajectories



## TPS Moves

- Shooting
  - perform small modification of velocities of transitional configuration
  - then shoot forwards/backwards in time

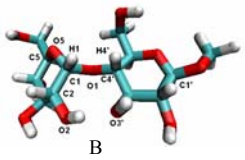

- Shifting
  - perform small translation of existing transitional trajectory in time
  - then complete trajectory by integrating forwards or backwards as necessary
  - in both move types, accept move if it connects desired minima




## Model Disaccharide: Methyl β-Maltoside

*in vacuo* $\varphi\psi$ potential of mean force

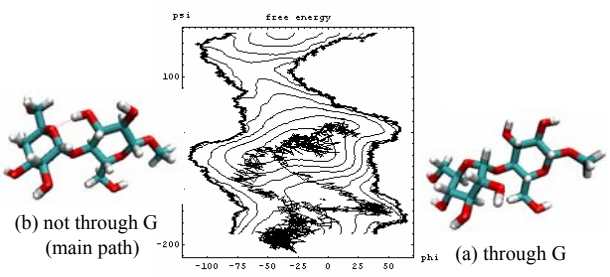
- consider transition between conformations B (-20º,-20º) and F (-40º,-180º)
- "G" plateau
- two MEPs for B→F
- ΔG barriers are 7 and 11 kcal/mol




B

## TPS of Negative $\psi$ Route

- two types of transition path detected

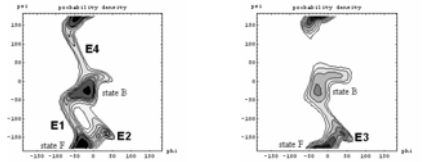



(b) not through G (main path)

(a) through G

## Importance of the O5···H-O3' Hydrogen Bond for Pathway

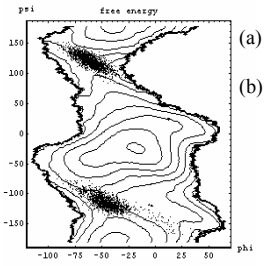
- computer experiments:
  - **E2**: constrain paths to travel through G
  - **E3**: switch off electrostatics of O5/H-O3' interaction




- none of **E2** had the hydrogen bond; all **E3** passed through G

## Transition State Ensembles

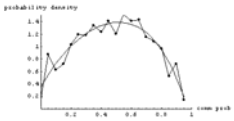
- committor probability = 0.5




Two ways to check $\varphi\psi$ as coordinates for main path:

(a) compare location of TS ensemble to PMF transition states on $\varphi\psi$ map

(b) calculate committor probability profile in locality around PMF transition states – is it parabolic?



## Rate of B to F transition

- From transition state theory and PMF
  - $k = 4 \times 10^7\ s^{-1}$
- From TPS
  - $k = 6 \times 10^6\ s^{-1}$
  - *ie.* one transition every 160 ns
  - rate calculation takes 25x effort of PMF
  - generation of a TP ensemble is 2.5x PMF

## Conclusions

- TPS detected an extra B to F route
- TPS and PMF confirms $\varphi\psi$ are reasonable coordinates for the B to F transition *in vacuo*
- Challenges
  - aqueous phase more difficult – diffusive
  - methods rely upon good potential energy functions for carbohydrates
- Future
  - potential to combine TPS and umbrella sampling for studying large-scale changes (*cf*. protein-DNA interactions, Schlick 2004)


## Acknowledgements


– Hoda Abdel-Aal, Abdul-Mueed Muslim, Richard Dimelow, Jonathan McNamara, Mark Vincent, Neil Burton, Ian Hillier

- Funding:
  – EPSRC, Colgate-Palmolive