Umbrella sampling in GROMACS

BioExcel Summer School 2018

Paul Bauer <paul.bauer@scilifelab.se>

SciLifeLab
KTH

29/05/2018

Sampling problems in Molecular Dynamics

figures/system.pdf

Covering of complete phase space is unrealistic

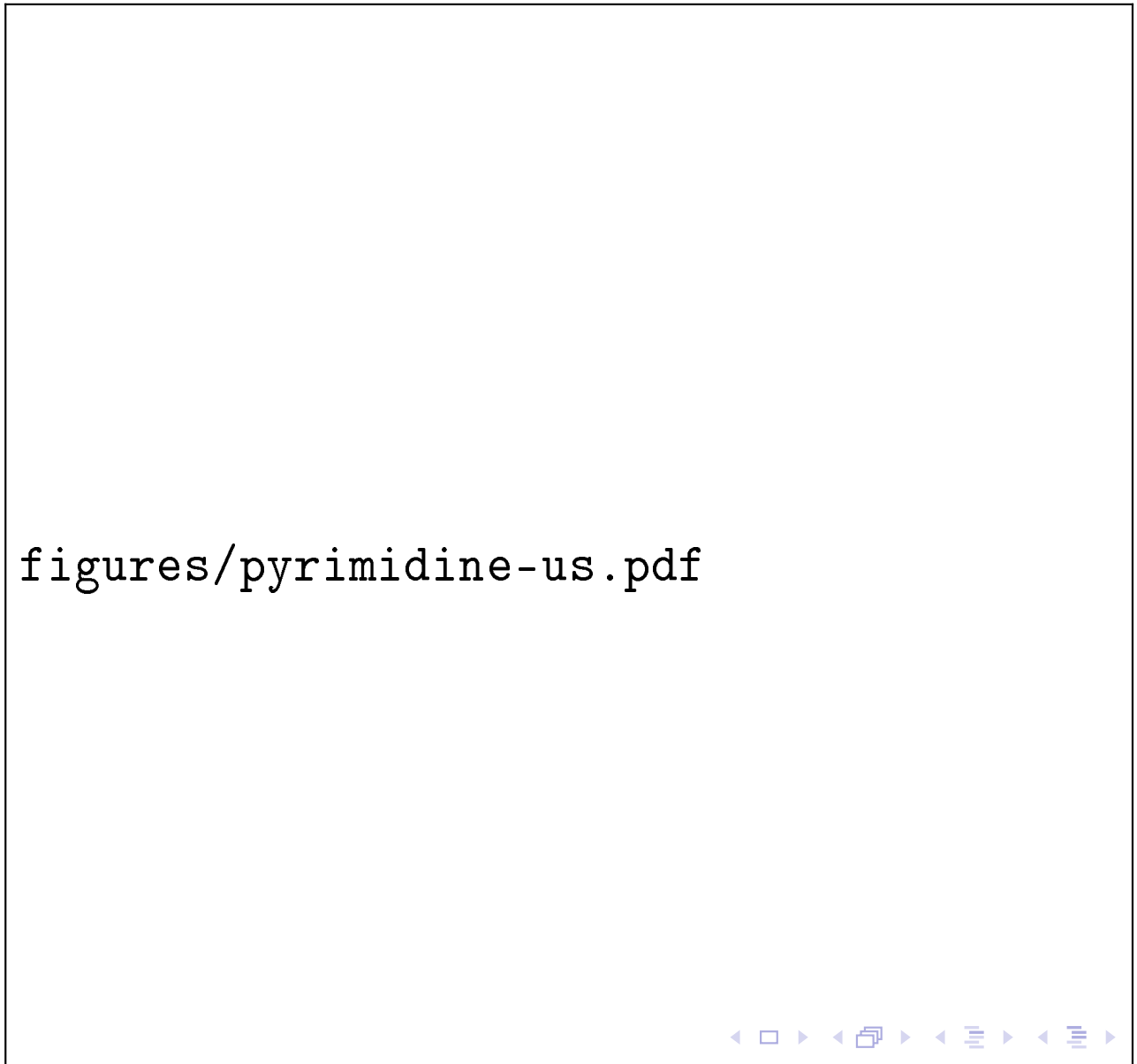
- Systems explore phase space according to boundary conditions given by the ensemble
- Larger differences in energy between states make crossing less likely
- Interesting changes often involve large changes in free energy and large free energy barriers
- System needs to be forced to either cross barriers or ignore them

Biasing potentials help crossing of free energy barriers

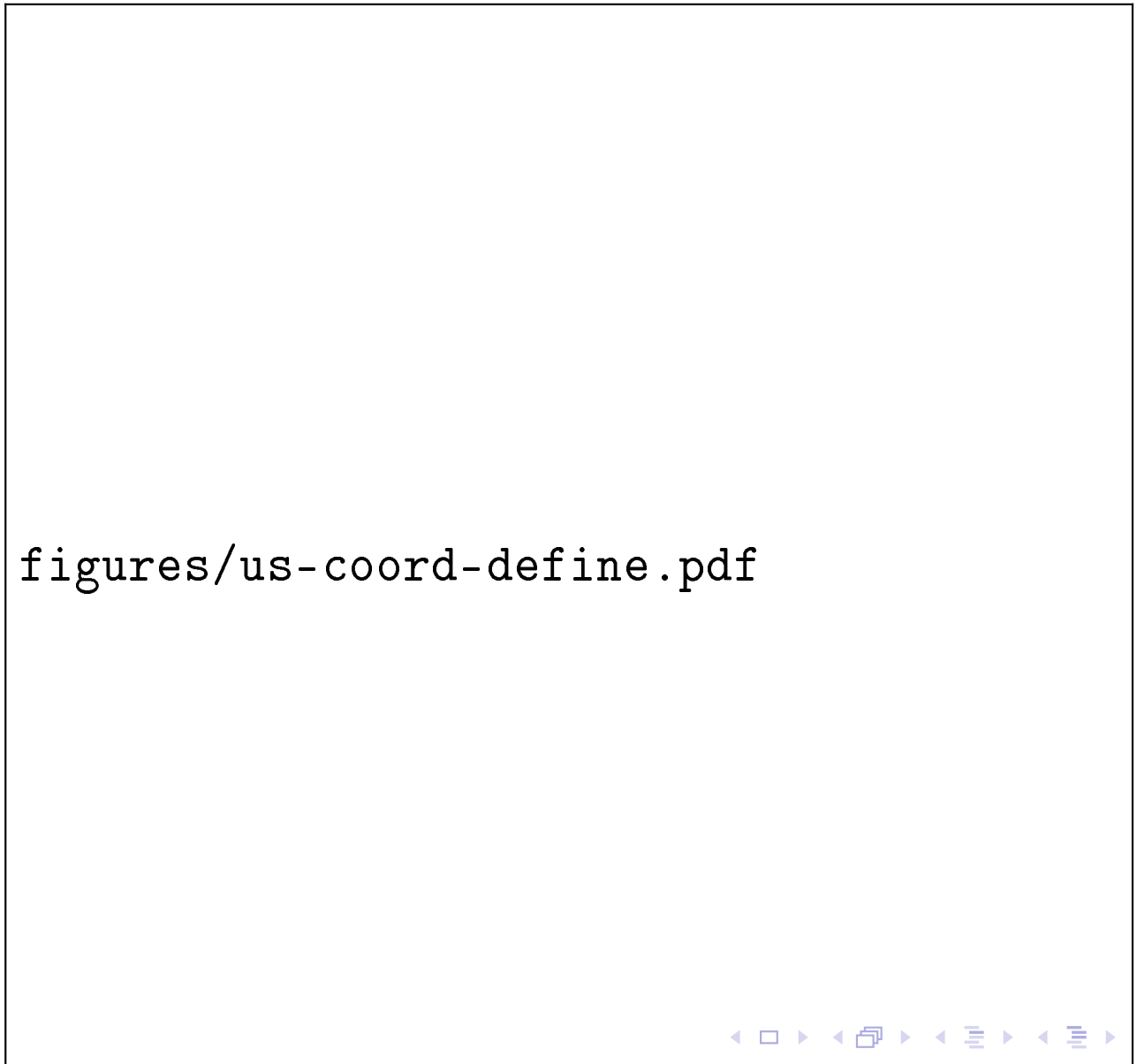
- Basic math behind use of Umbrella Sampling
-

(1.1)

Example visualization of US windows



Ways to define US window coordinates



Minimum requirements for converging simulation

- Each window needs to represent physical configuration
- Windows need to have sufficient sampling overlap
- No large changes between configurations that can't be sampled

How to generate initial configurations?

figures/us-generate-conf.pdf

Some application examples

- Calculation of dimerization energies
- Clearly defined change within simulated structure
- QM/MM for chemical reactions in enzyme systems
- What can you think of?

US basics and requirements

- All here based on reasonable recent version of GROMACS, tried on v2018
- Examples and later tutorials based on [link here]
- Need to understand pull code, index groups, (partially) WHAM
- Topology setup and definition not covered

Introduction to index groups

figures/index-groups.pdf

Index group basics

- List of atoms being grouped together according to some requirement
- Used to specify important parts of simulation for running/analysis/data processing
- Defaults generate by all GROMACS tools internally
- User defined groups available through either gmx make_ndx or gmx select

How to generate index files with make_ndx

figures/demo.pdf

Setting up the simulation

- Define groups that should be restrainted by US windows
- Define coordinates to be used for restraint
- Set window spacing or generate initial configurations
- Set US window restraint force constant

Different kinds of coordinates and geometries

figures/pull-coordinates.pdf

Basics of US histogram analysis

- All information here based on [link]
- Full disclosure: I'm not an expert in WHAM
- More intendet as opening point to look for further information

Math behind WHAM

Analysing GROMACS US simulations

- Needs at least run input file and pullx or pullf file for each window
- Will by default calculate both individual histograms and PMF
- Able to also obtain autocorrelation times and error estimates

Common pitfalls

- Insufficient sampling overlap
- Bad choice of reaction coordinate
- Bad choice of restraint parameters
- Insufficient equilibration

Thanks for your attention/patience

Questions? Comments? Suggestions?