Implementing Molecular Hydrophobicity Potential Measurment for the Analysis of Dynamic Biomolecular Interactions

Peleg Bar Sapir[1]

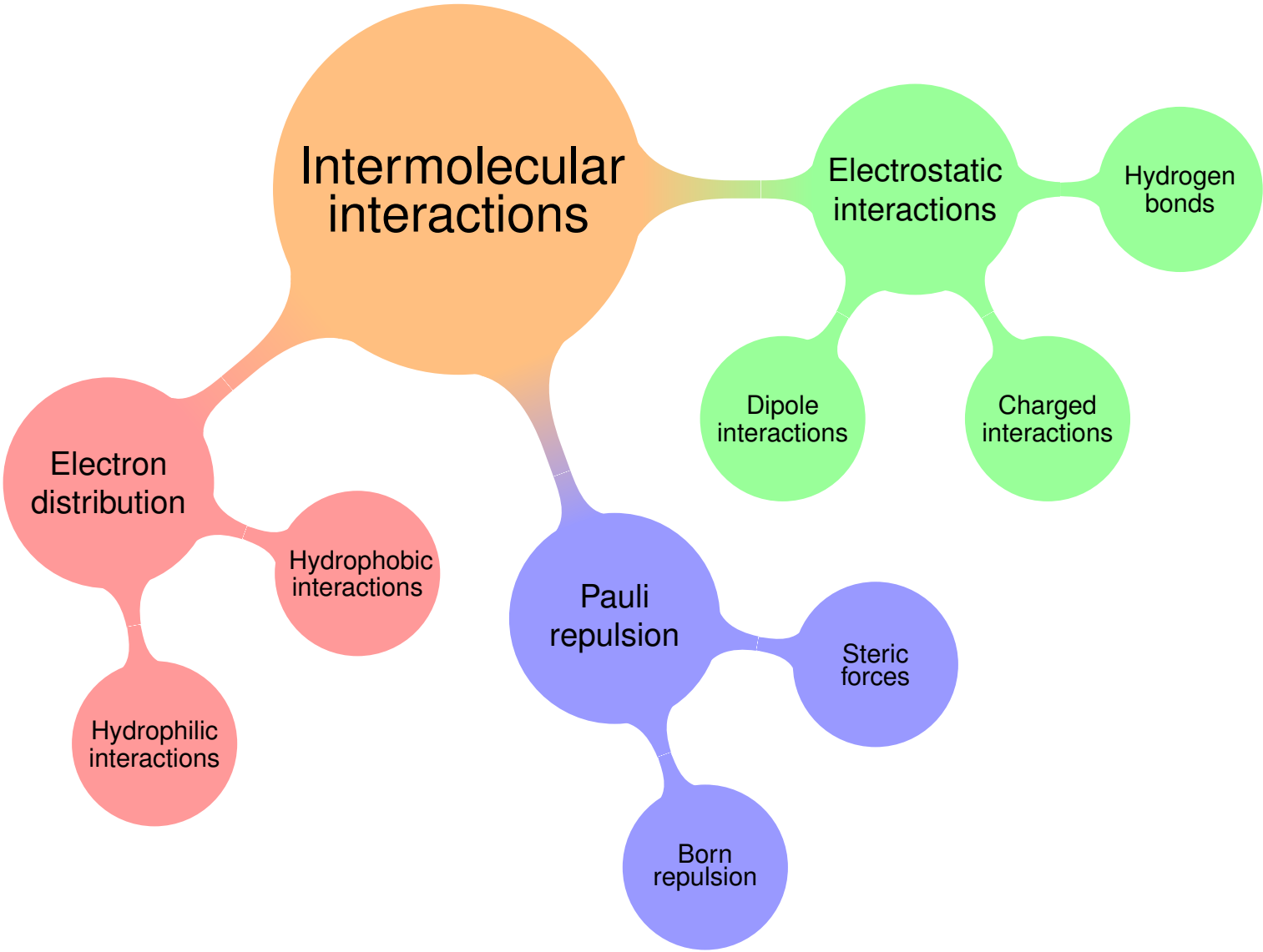
Under supervision of
Prof. Maria Andrea Mroginski[2]

[1]Freie Universität Berlin
[2]Techniche Universität Berlin

February 19, 2018

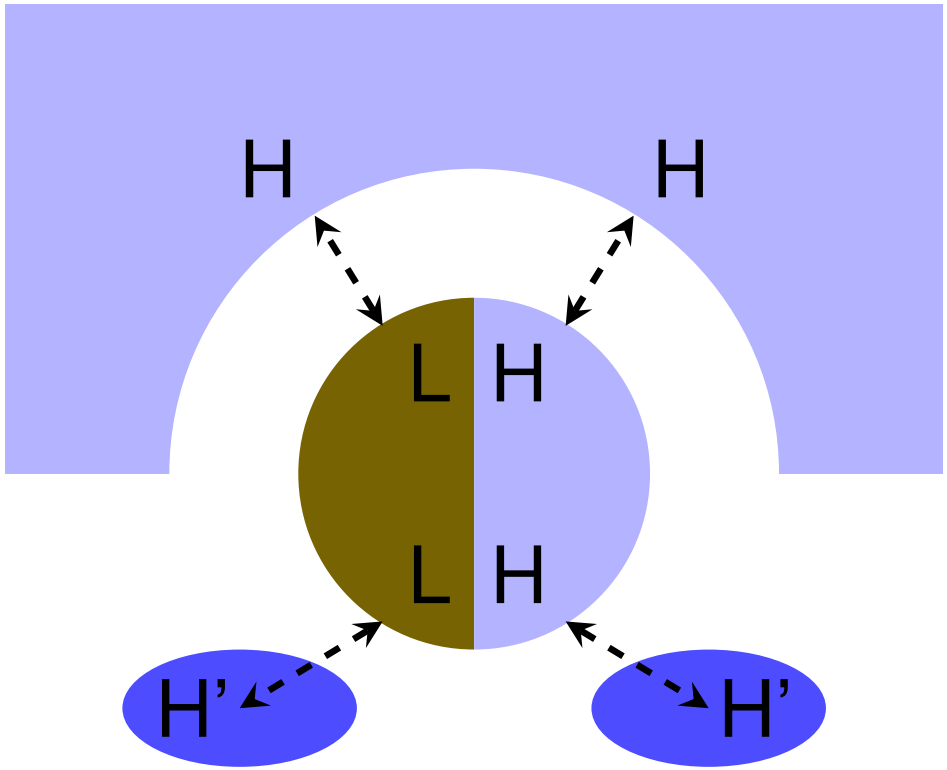
Outline

Hydrophobicity and log P



Hydrophilic/Hydrophobic Interactions

In short: like interacts best with like[1].



[1] Figure based on [1]

Partition Coefficient

Definition

The ratio of concentrations of a compound in a mixture of two immiscible phases at equilibrium

Partition Coefficient

Definition

The ratio of concentrations of a compound in a mixture of two immiscible phases at equilibrium

- Commonly used: water and octanol

Partition Coefficient

Definition

The ratio of concentrations of a compound in a mixture of two immiscible phases at equilibrium

- Commonly used: water and octanol
- Can be measured at an ionized or unionized state

Partition Coefficient

Definition

The ratio of concentrations of a compound in a mixture of two immiscible phases at equilibrium

- Commonly used: water and octanol
- Can be measured at an ionized or unionized state
- $\log P_{\text{octanol/water}} = \log \left(\frac{[\text{solute}]_{\text{water}}}{[\text{solute}]_{\text{octanol}}} \right)$

Partition Coefficient

Definition

The ratio of concentrations of a compound in a mixture of two immiscible phases at equilibrium

- Commonly used: water and octanol
- Can be measured at an ionized or unionized state
- $\log P_{\text{octanol/water}} = \log \left(\frac{[\text{solute}]_{\text{water}}}{[\text{solute}]_{\text{octanol}}} \right)$
- Hydrophobicity increases with the (common) $\log P$

What is Molecular Hydrophobicity Potential (MHP)?

- By measuring the $\log P$ of many (ca. 30,000) compounds[3]

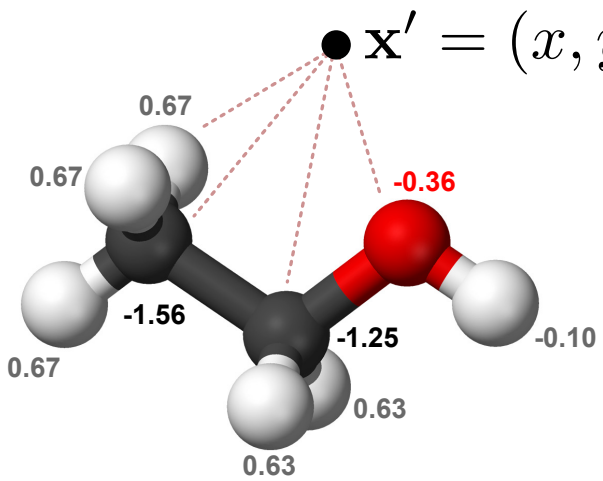
What is Molecular Hydrophobicity Potential (MHP)?

- By measuring the $\log P$ of many (ca. 30,000) compounds[3]
- Combining these values with a distance-depended decay function, a potential can be constructed.

What is Molecular Hydrophobicity Potential (MHP)?

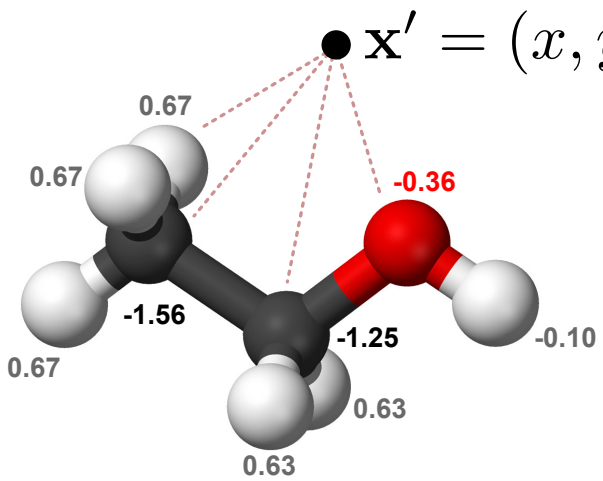
- By measuring the $\log P$ of many (ca. 30,000) compounds[3]
- Combining these values with a distance-depended decay function, a potential can be constructed.
- This potential predicts the local $\log P$ behaviour of fragments of a molecule.

The MHP Formula



$$\text{MHP}\left(\mathbf{x}'\right) = \sum_{i=1}^{k} \left[f_i \cdot D\left(\mathbf{x} - \mathbf{x}'_i\right) \right]$$

The MHP Formula

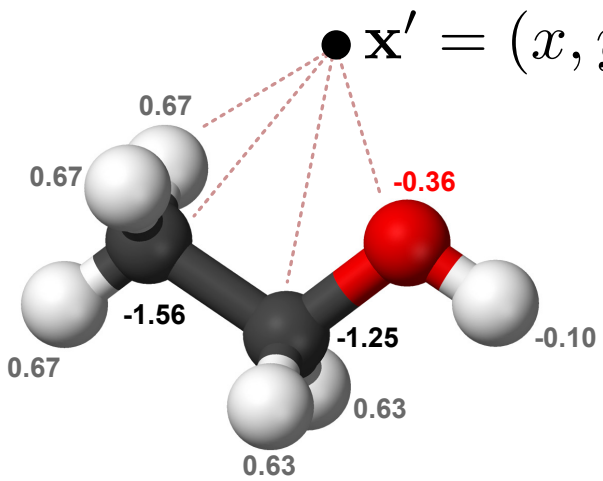


$$\text{MHP}\left(\mathbf{x}'\right) = \sum_{i=1}^{k} \left[f_i \cdot D\left(\mathbf{x} - \mathbf{x}'_i\right) \right]$$

$$\sum_{i=1}^{k}$$

Summing over all atoms

The MHP Formula

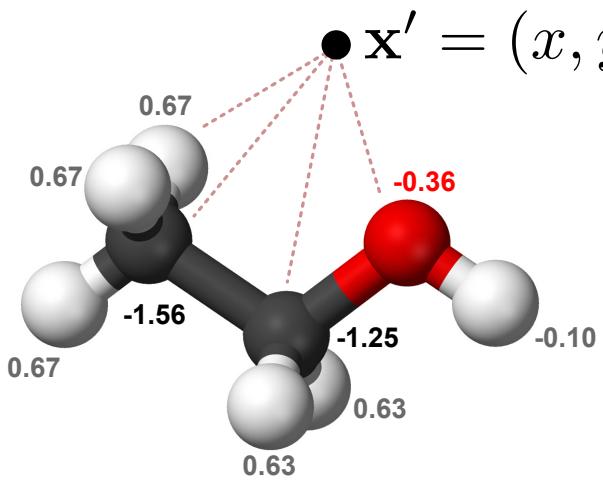


$$\text{MHP}\left(\mathbf{x}'\right) = \sum_{i=1}^{k} \left[f_i \cdot D\left(\mathbf{x} - \mathbf{x}'_i\right) \right]$$

$\sum_{i=1}^{k} f_i$

Summing over all atoms

The MHP Formula



$$\text{MHP}\left(\mathbf{x}'\right) = \sum_{i=1}^{k} \left[f_i \cdot D\left(\mathbf{x} - \mathbf{x}'_i\right) \right]$$

k

$D\left(\mathbf{x} - \mathbf{x}'_i\right)$

$i=1$

Summing over all atoms

Distance function

Force Constants - Carbon

Carbon atom contribution to hydrophobicity[2]

Type	Description	f_i value
	Carbon in:	
1	CH_3R	-1.5603
3	CHR_3	-0.6681
7	CH_2X_2	-1.0305
13	RCX_3	0.7894
17	$=CR_2$	0.0383
24	$R--CH--R$	-0.3251
25	$R--CR--R$	0.1492
26	$R--CX--R$	0.1539

[2]Source: Ghose et al, J. Phys. Chem. A 1998, 102, 3762-3772

Force Constants - Hydrogen

Hydrogen atom contribution to hydrophobicity[3]

Type	Description	f_i value
	Hydrogen attached to:	
46	C_{sp^3}, no X in α	0.7341
47	C_{sp^2}	0.6301
50	Heteroatom X	-0.1036
52	C_{sp^3}, 1 X in α	0.6666
54	C_{sp^3}, 3 X in α	0.6338

[3] Source: Ghose et al, J. Phys. Chem. A 1998, 102, 3762-3772

Force Constants - Oxygen

Oxygen atom contribution to hydrophobicity[4]

Type	Description	f_i value
	Oxygen in:	
56	Alcohol	-0.3567
57	Phenol, enol, carboxyl OH	-0.0127
58	Ketone	-0.0233
61	Nitro, N-oxides	1.0520
62	O^-	-0.7941

[4] Source: Ghose et al, J. Phys. Chem. A 1998, 102, 3762-3772

Force Constants - Various

Various atom contribution to hydrophobicity[5]

Type	Description	f_i value
66	N in Primary amine	-0.5427
67	N in Secondary amine	-0.3168
81	F attached to $\mathrm{C_{sp^3}}$	0.4797
106	S in $\mathrm{R{-}SH}$ (thiol)	1.0520
119	P in $\mathrm{PR_3}$ (phosphine)	-0.7941

[5] Source: Ghose et al, J. Phys. Chem. A 1998, 102, 3762-3772

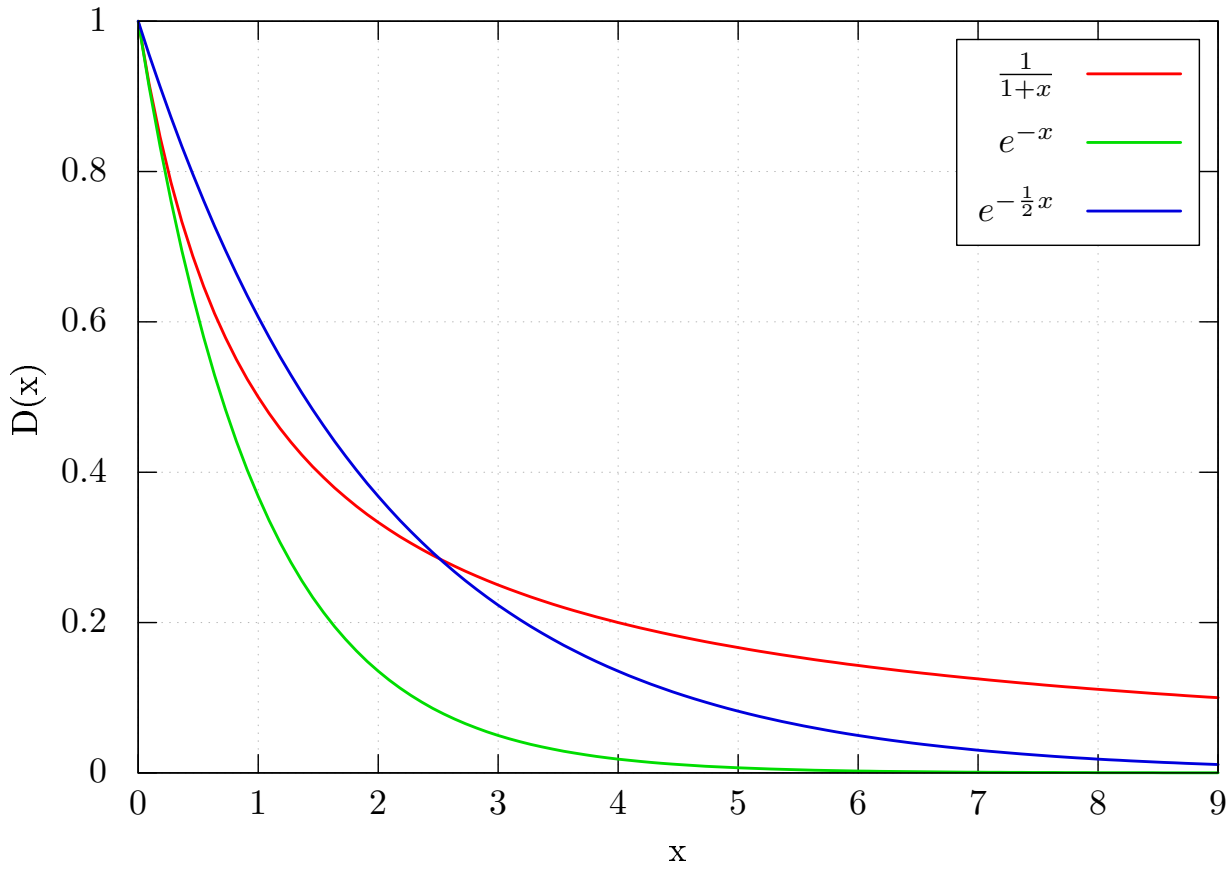
Distance function

Audry form

$$D(x) = \frac{1}{1+x}$$

Exponential decay form

$$D(x) = e^{-\alpha x}$$

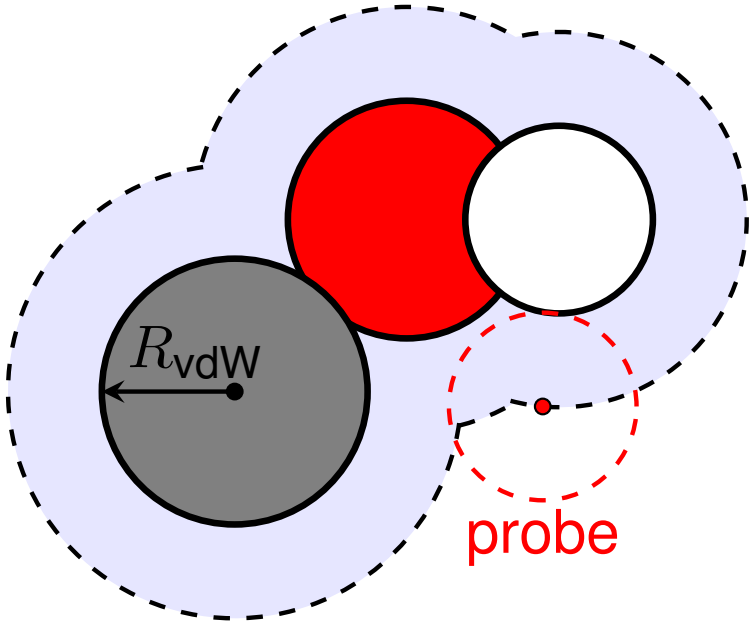


Solvent Accesible Surface

- The surface around a molecule accesible to solvent molecules

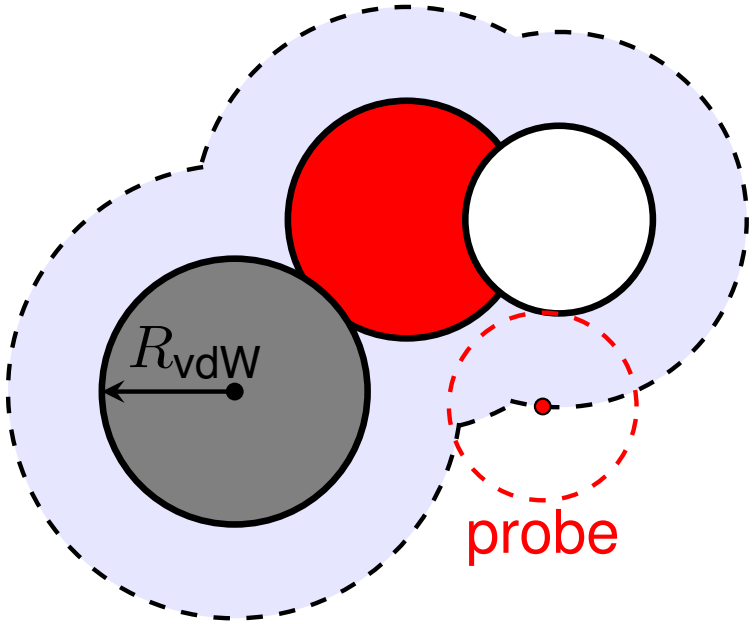
Solvent Accesible Surface

- The surface around a molecule accesible to solvent molecules



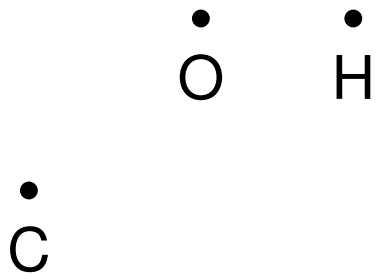
Solvent Accesible Surface

- The surface around a molecule accesible to solvent molecules

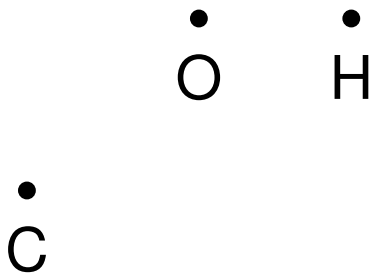


(For water molecules usually $r = 1.4\,[\text{Å}]$)

How to Create the Solvent Accesible Surface?

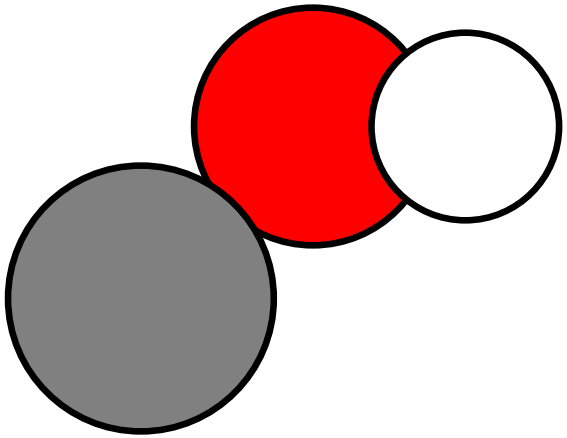


How to Create the Solvent Accesible Surface?



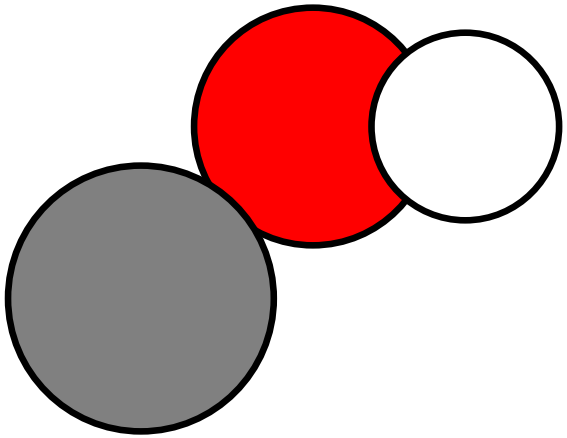
1. Take all atoms with their vdW-radii

How to Create the Solvent Accesible Surface?



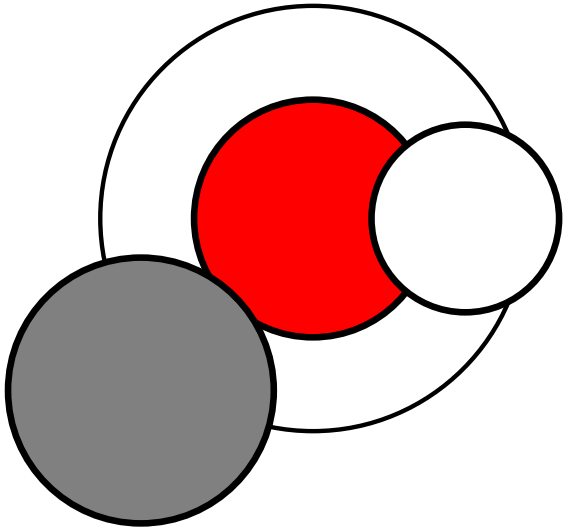
1. Take all atoms with their vdW-radii

How to Create the Solvent Accesible Surface?



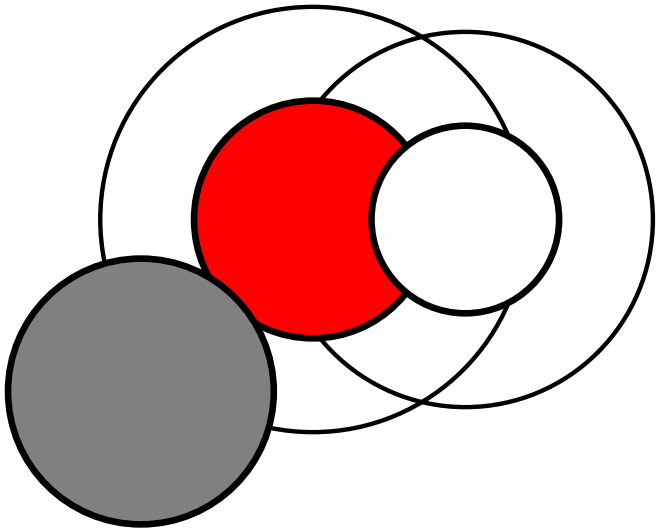
1. Take all atoms with their vdW-radii
2. Create spheres around all atoms with $R^i = R^i_{\text{vdw}} + R_{\text{probe}}$

How to Create the Solvent Accesible Surface?



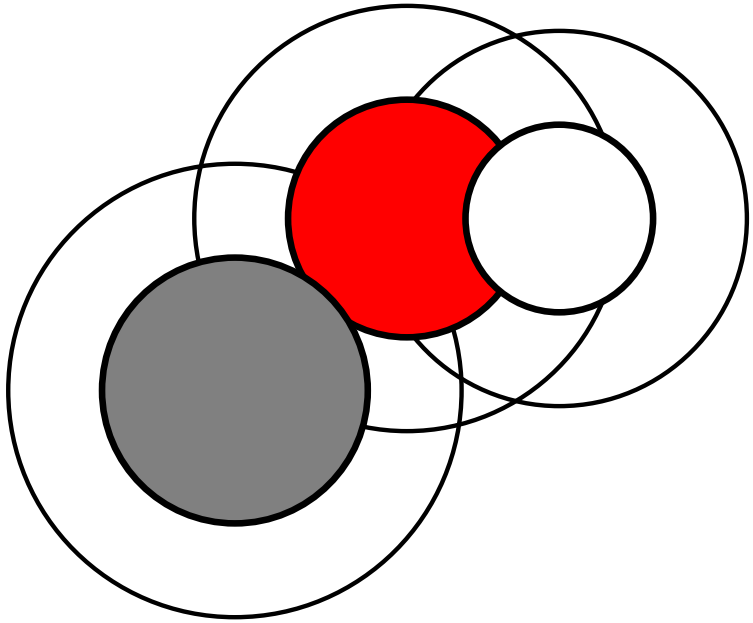
1. Take all atoms with their vdW-radii
2. Create spheres around all atoms with $R^i = R^i_{\text{vdw}} + R_{\text{probe}}$

How to Create the Solvent Accesible Surface?



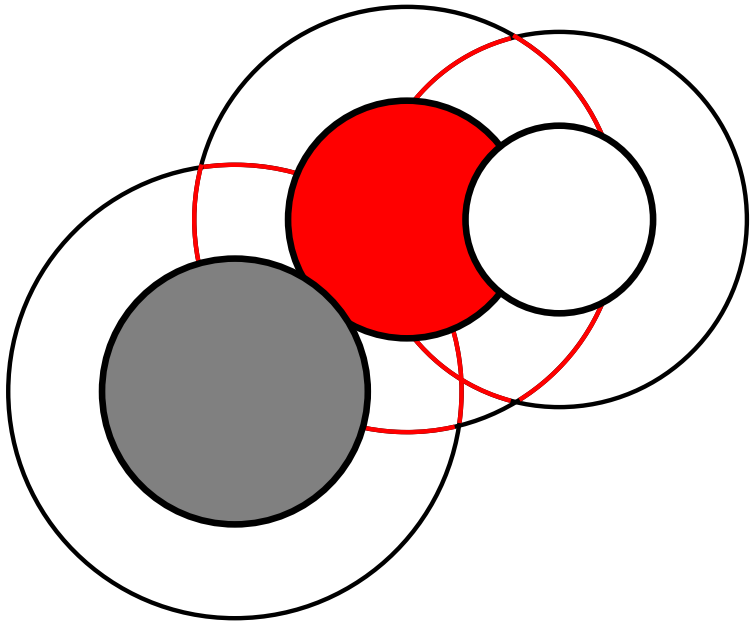
1. Take all atoms with their vdW-radii
2. Create spheres around all atoms with $R^i = R^i_{\text{vdw}} + R_{\text{probe}}$

How to Create the Solvent Accesible Surface?



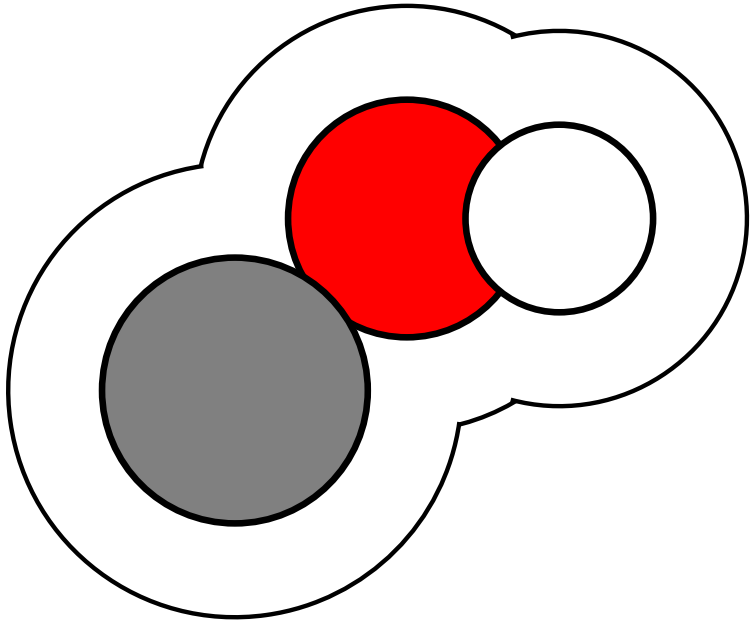
1. Take all atoms with their vdW-radii
2. Create spheres around all atoms with $R^i = R^i_{\text{vdw}} + R_{\text{probe}}$

How to Create the Solvent Accesible Surface?



1. Take all atoms with their vdW-radii
2. Create spheres around all atoms with $R^i = R^i_{\text{vdw}} + R_{\text{probe}}$
3. Delete all points that are "burried" in other extended spheres (i.e. $\Delta\left(p^i, c^j\right) \leq R^j + R_{\text{probe}}$)

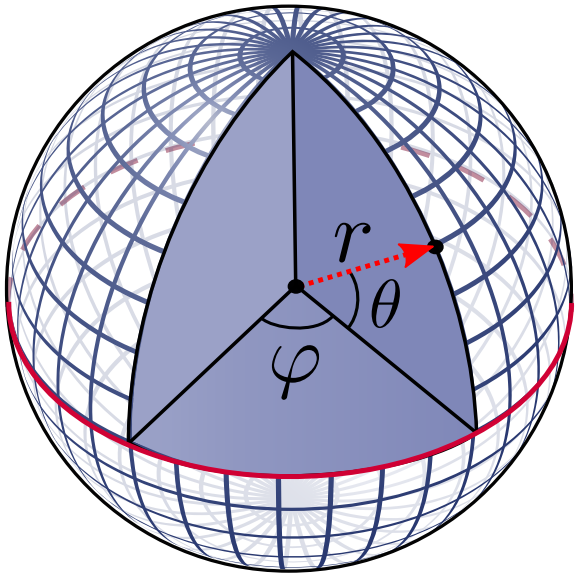
How to Create the Solvent Accesible Surface?



1. Take all atoms with their vdW-radii
2. Create spheres around all atoms with $R^i = R^i_{\text{vdw}} + R_{\text{probe}}$
3. Delete all points that are "burried" in other extended spheres (i.e. $\Delta\left(p^i, c^j\right) \leq R^j + R_{\text{probe}}$)
4. The remaining surface is the solvent-accesible surface of the molecule

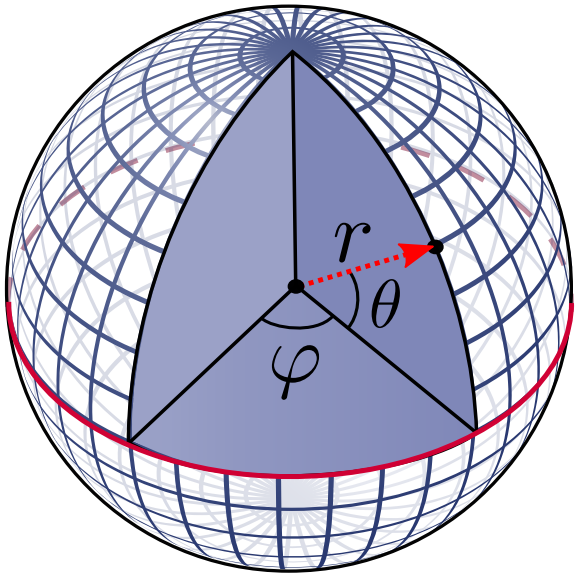
Evenly Distributed Points

How to distribute N points on a surface of a sphere?



Evenly Distributed Points

How to distribute N points on a surface of a sphere?

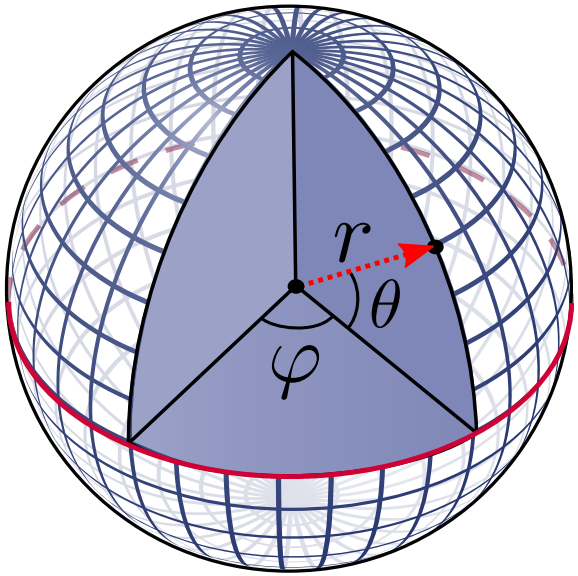


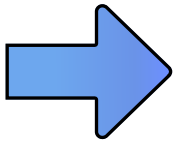
$$\varphi_i = i \cdot \frac{2\pi}{N}$$

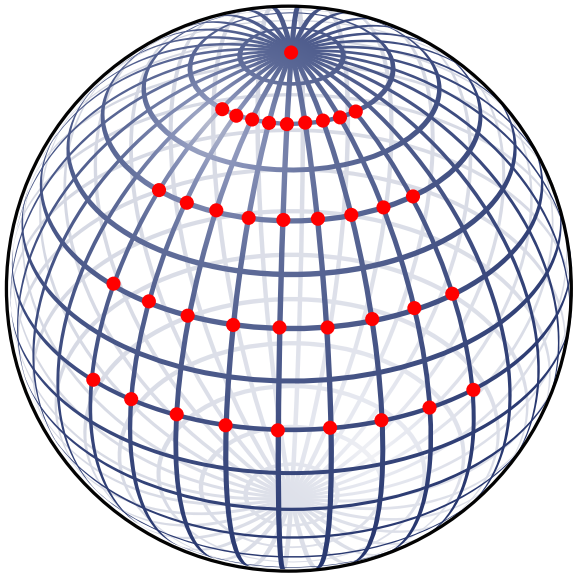
$$\theta_j = j \cdot \frac{\pi}{N}$$

Evenly Distributed Points

How to distribute N points on a surface of a sphere?





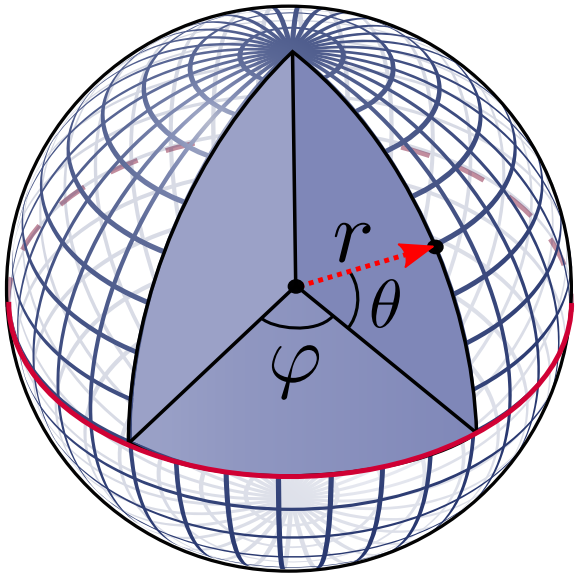


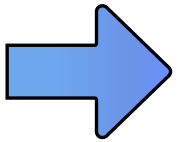
$$\varphi_i = i \cdot \frac{2\pi}{N}$$

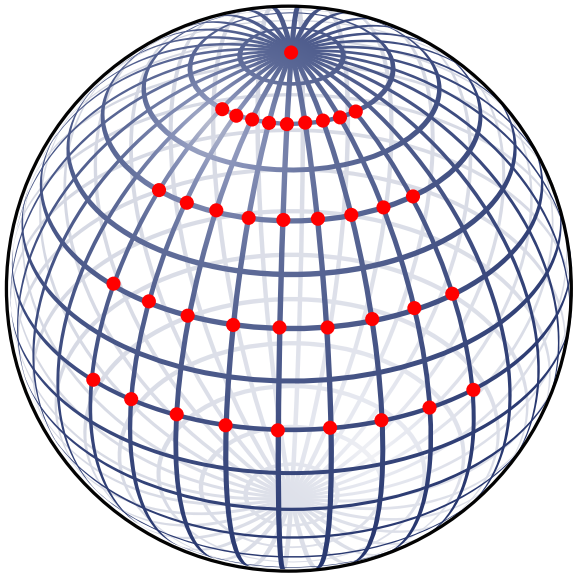
$$\theta_j = j \cdot \frac{\pi}{N}$$

Evenly Distributed Points

How to distribute N points on a surface of a sphere?







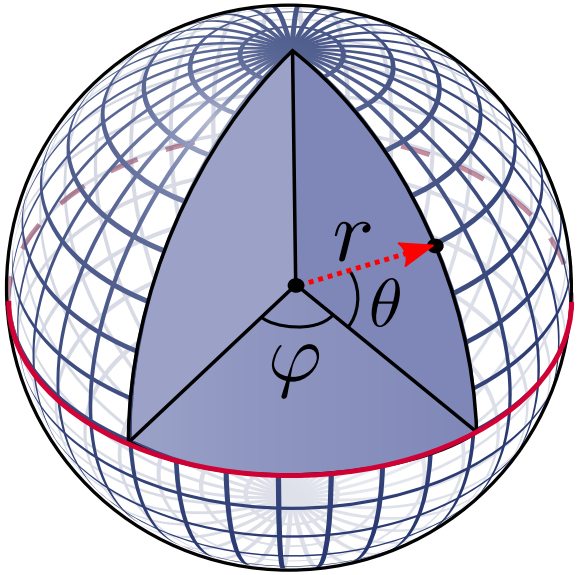
$$\varphi_i = i \cdot \frac{2\pi}{N}$$

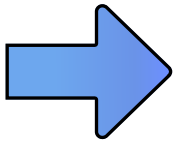
$$\theta_j = j \cdot \frac{\pi}{N}$$

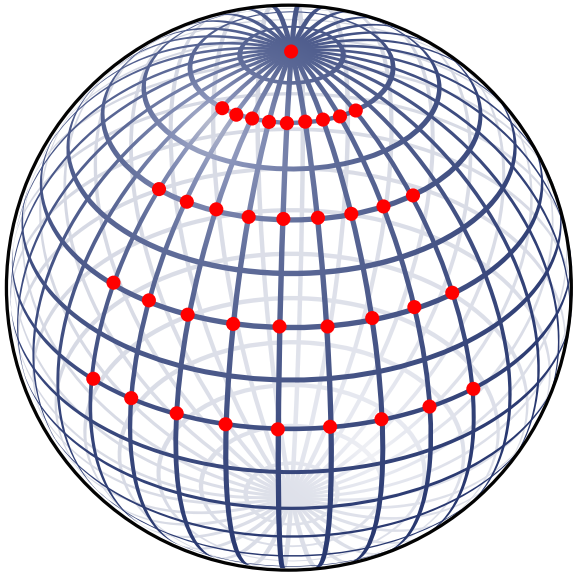
- Points are not evenly distributed

Evenly Distributed Points

How to distribute N points on a surface of a sphere?







$$\varphi_i = i \cdot \frac{2\pi}{N}$$

$$\theta_j = j \cdot \frac{\pi}{N}$$

- Points are not evenly distributed
- Several points overlap at poles

Evenly Distributed Points

Solution: Vogel's method

In 2 dimensions:

Evenly Distributed Points

Solution: Vogel's method

In 2 dimensions:

- Distances: $r_i = \sqrt{\frac{i}{N}}$
- Angle: $\theta_i = \varphi i$

  (φ is the golden ratio!)

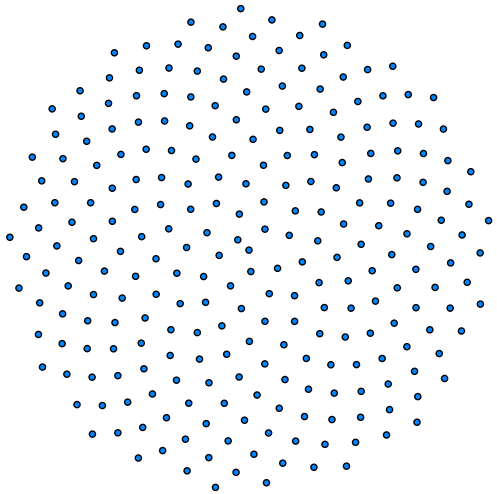
Evenly Distributed Points

Solution: Vogel's method

In 2 dimensions:

- Distances: $r_i = \sqrt{\frac{i}{N}}$
- Angle: $\theta_i = \varphi i$

  (φ is the golden ratio!)



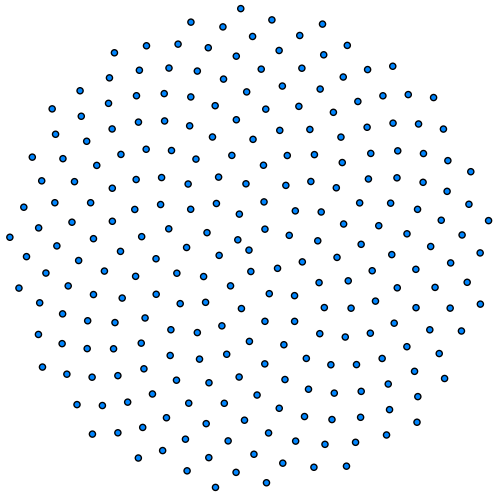
Evenly Distributed Points

Solution: Vogel's method

In 2 dimensions:

- Distances: $r_i = \sqrt{\frac{i}{N}}$
- Angle: $\theta_i = \varphi i$

  (φ is the golden ratio!)



In 3 dimensions (cylindrical coordinates):

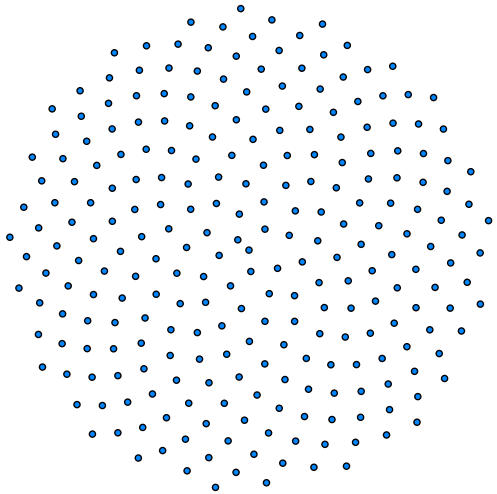
Evenly Distributed Points

Solution: Vogel's method

In 2 dimensions:

- Distances: $r_i = \sqrt{\frac{i}{N}}$
- Angle: $\theta_i = \varphi i$

  (φ is the golden ratio!)



In 3 dimensions (cylindrical coordinates):

- Distances: $z_i = \left(1 - \frac{1}{N}\right)\left(1 - \frac{2i}{N-1}\right)$
- Angles:

  $\theta_i = \varphi i,\ \rho_i = \sqrt{1 - z_i^2}$

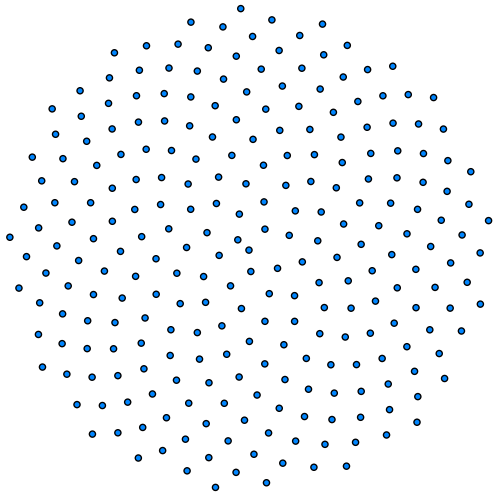
Evenly Distributed Points

Solution: Vogel's method

In 2 dimensions:

- Distances: $r_i = \sqrt{\frac{i}{N}}$
- Angle: $\theta_i = \varphi i$

  (φ is the golden ratio!)



In 3 dimensions (cylindrical coordinates):

- Distances: $z_i = \left(1 - \frac{1}{N}\right)\left(1 - \frac{2i}{N-1}\right)$
- Angles:

  $\theta_i = \varphi i,\ \rho_i = \sqrt{1 - z_i^2}$

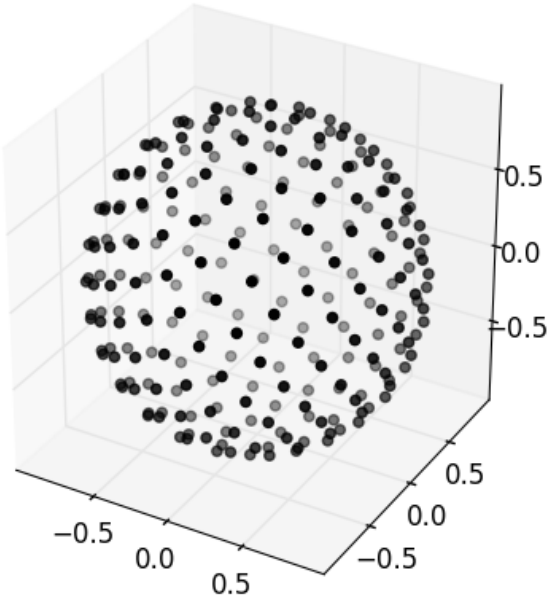


Image source: Marmakoide's Blog

Integration

- Each atom's total surface area:
  $V^a = 4\pi \left(R^a_{\text{vdW}} + R_{\text{probe}}\right)^2$

Integration

- Each atom's total surface area:
  $V^a = 4\pi \left(R^a_{\text{vdW}} + R_{\text{probe}}\right)^2$
- The surface is represented by N points

Integration

- Each atom's total surface area:
  $V^a = 4\pi \left(R^a_{\text{vdW}} + R_{\text{probe}}\right)^2$
- The surface is represented by N points
- Meaning: each point has $V^a_j = \frac{4\pi}{N} \left(R^i_{\text{vdW}} + R_{\text{probe}}\right)^2$

Integration

- Each atom's total surface area:
  $V^a = 4\pi \left(R^a_{\text{vdW}} + R_{\text{probe}}\right)^2$
- The surface is represented by N points
- Meaning: each point has $V^a_j = \frac{4\pi}{N} \left(R^i_{\text{vdW}} + R_{\text{probe}}\right)^2$
- In addition, each point has: MHP^a_j

Integration

- Each atom's total surface area:
  $V^a = 4\pi \left(R^a_{\text{vdW}} + R_{\text{probe}}\right)^2$
- The surface is represented by N points
- Meaning: each point has $V^a_j = \frac{4\pi}{N} \left(R^i_{\text{vdW}} + R_{\text{probe}}\right)^2$
- In addition, each point has: MHP^a_j

Therefore, each atom has a total MHP of:

$$\text{MHP}^a = \frac{4\pi}{N} \sum_{j=0}^{M} \text{MHP}^a_j$$

Program Specifications

- Written in **Python3**, utylizing **ProDy**

Program Specifications

- Written in **Python3**, utylizing **ProDy**
- Heavy calculation written in **Cython**

Program Specifications

- Written in **Python3**, utylizing **ProDy**
- Heavy calculation written in **Cython**
- Uses neighbor cells implementation for faster calculation

Program Specifications

- Written in **Python3**, utylizing **ProDy**
- Heavy calculation written in **Cython**
- Uses neighbor cells implementation for faster calculation
- Uses PSF, PDB and DCD files

Program Specifications

- Written in **Python3**, utylizing **ProDy**
- Heavy calculation written in **Cython**
- Uses neighbor cells implementation for faster calculation
- Uses PSF, PDB and DCD files
- Generates a PDB output, MHP values in **beta** column

Program Options

- Input: PSF + PDB or DCD

Program Options

- Input: PSF + PDB or DCD
- Subselection (optional): Atomic selection (like vmd)

Program Options

- Input: PSF + PDB or DCD
- Subselection (optional): Atomic selection (like vmd)
- Number of points per atom (default: 64)

Program Options

- Input: PSF + PDB or DCD
- Subselection (optional): Atomic selection (like vmd)
- Number of points per atom (default: 64)
- Solvent probe radius (defalt: $1.4\mathring{A}$)

Program Options

- Input: PSF + PDB or DCD
- Subselection (optional): Atomic selection (like vmd)
- Number of points per atom (default: 64)
- Solvent probe radius (defalt: $1.4\mathring{A}$)
- Cutoff distance for distance function (default: $4\mathring{A}$)

Program Options

- Input: PSF + PDB or DCD
- Subselection (optional): Atomic selection (like vmd)
- Number of points per atom (default: 64)
- Solvent probe radius (defalt: $1.4\mathring{A}$)
- Cutoff distance for distance function (default: $4\mathring{A}$)
- Frame range (if DCD)

Validation via Known $\log P$ Values

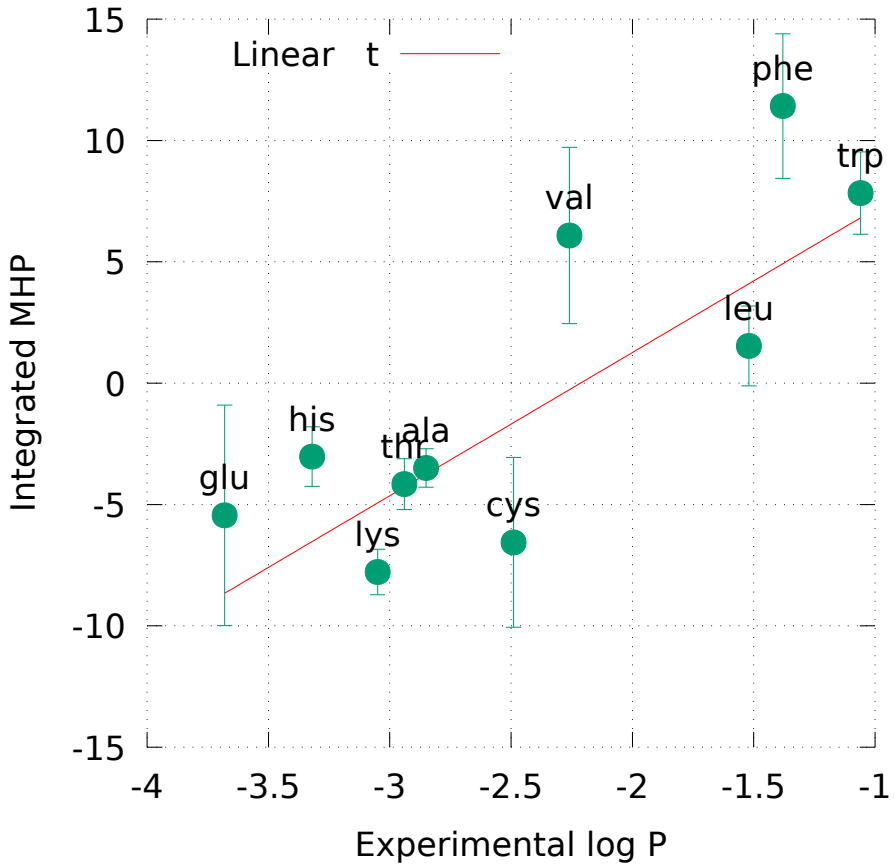
- By integrating and comparing to known $\log P$ values, a correlation can be measured.

Validation via Known $\log P$ Values

- By integrating and comparing to known $\log P$ values, a correlation can be measured.
- A groups of amino acids of varying hydrophobicity where simulated and their MHP calculated.

Validation via Known $\log P$ Values

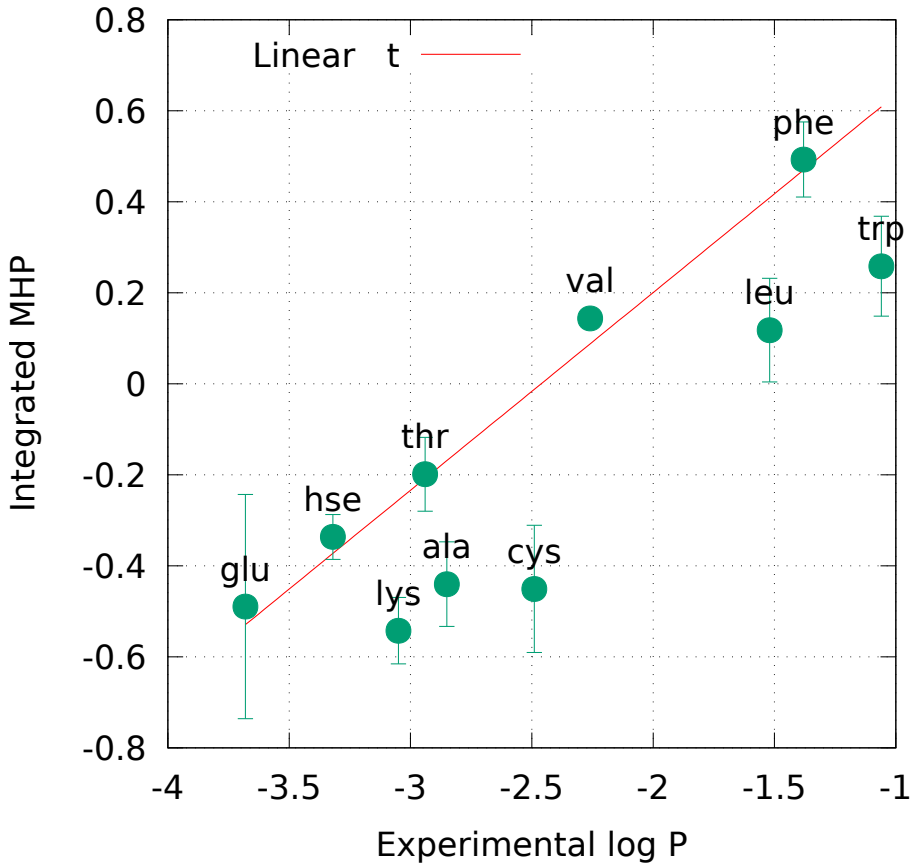
Validation in vacuum (5 frames per molecule)[6], $R^2 = 0.668$



[6]MD simulation using NAMD, performed by Dr. Tillmann Utesch

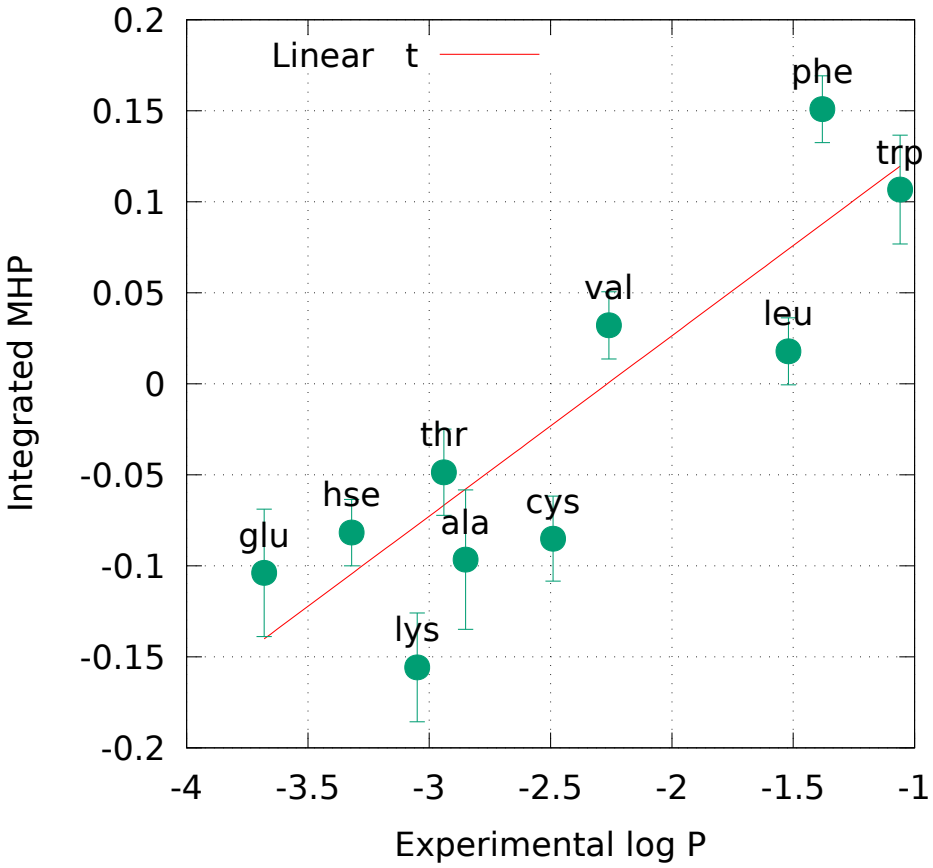
Validation via Known $\log P$ Values

Validation in water + structural optimization (10 frames per molecule), $R^2 = 0.748$



Validation via Known $\log P$ Values

Validation in water + structural optimization + SAS normalization (10 frames per molecule), $R^2 = 0.760$



Validation via Known $\log P$ Values

- The validation shows a reasonable qualitative correlation to real data.

Validation via Known $\log P$ Values

- The validation shows a reasonable qualitative correlation to real data.
- Performed in water (with structural optimization), the results became more accurate.

Validation via Known $\log P$ Values

- The validation shows a reasonable qualitative correlation to real data.
- Performed in water (with structural optimization), the results became more accurate.
- The environment did not match experiments, which could affect the accuracy.

Validation via Known $\log P$ Values

- The validation shows a reasonable qualitative correlation to real data.
- Performed in water (with structural optimization), the results became more accurate.
- The environment did not match experiments, which could affect the accuracy.
- Amino acids are small molecules, each error becomes more significant.

Validation via Known $\log P$ Values

- The validation shows a reasonable qualitative correlation to real data.
- Performed in water (with structural optimization), the results became more accurate.
- The environment did not match experiments, which could affect the accuracy.
- Amino acids are small molecules, each error becomes more significant.
- Larger trajectories will sample conformational space better.

An Example System

An existing trajectory (100 frames) of a protein-membrane system[7] was analazyed.

- The peptide: OP-145, a Cathelicidin derivative with improved properties.

[7] Trajectory provided by Dr. Alejandra de Miguel Catalina

An Example System

An existing trajectory (100 frames) of a protein-membrane system[7] was analazyed.

- The peptide: OP-145, a Cathelicidin derivative with improved properties.
- The interaction mechanism pathway was studied by means of all-atom simulation.

[7] Trajectory provided by Dr. Alejandra de Miguel Catalina

An Example System

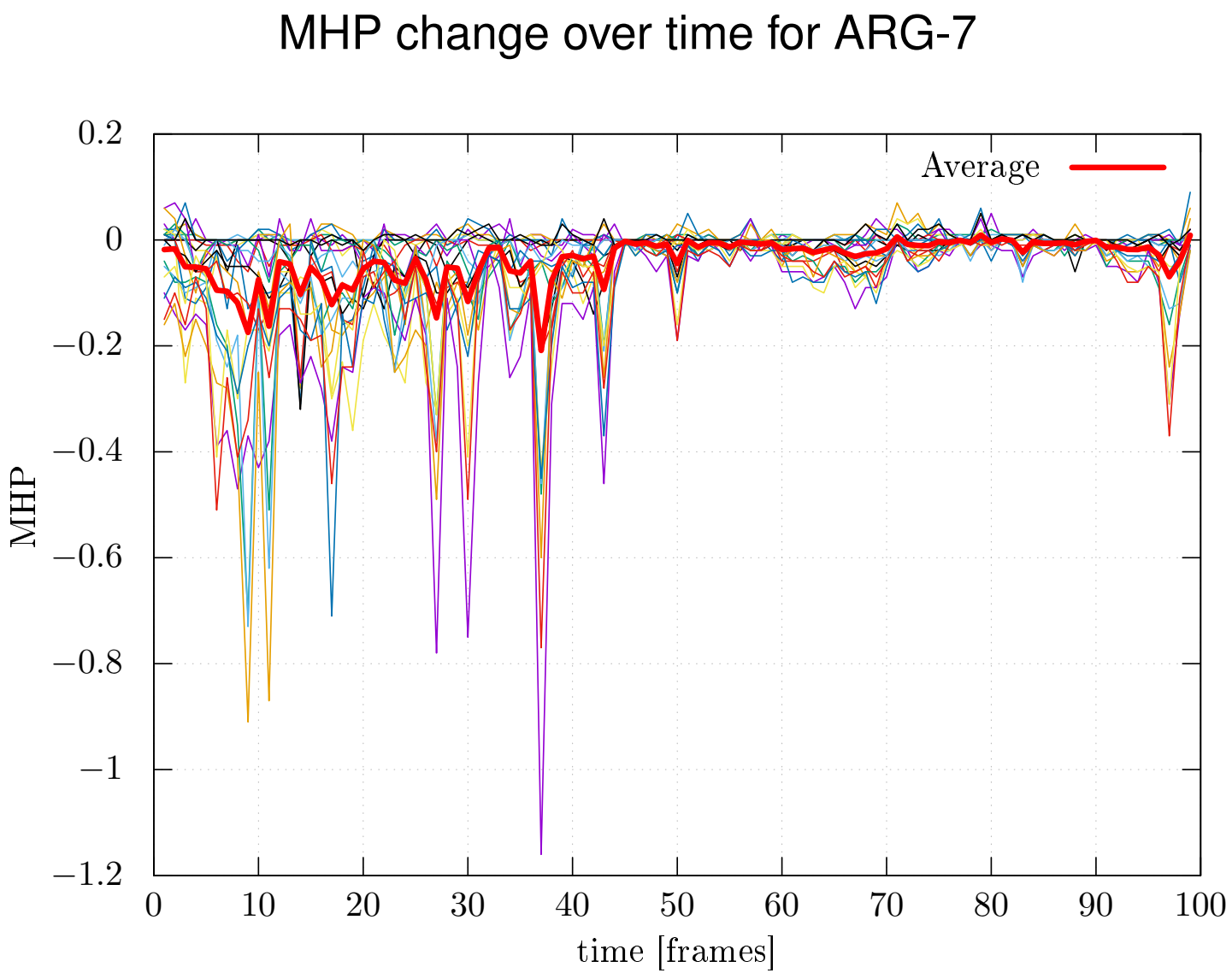
An existing trajectory (100 frames) of a protein-membrane system[7] was analazyed.

- The peptide: OP-145, a Cathelicidin derivative with improved properties.
- The interaction mechanism pathway was studied by means of all-atom simulation.
- The membrane used for the study consists of a mixture of two lipids, PG and PE, in agreement with experimental measurements.

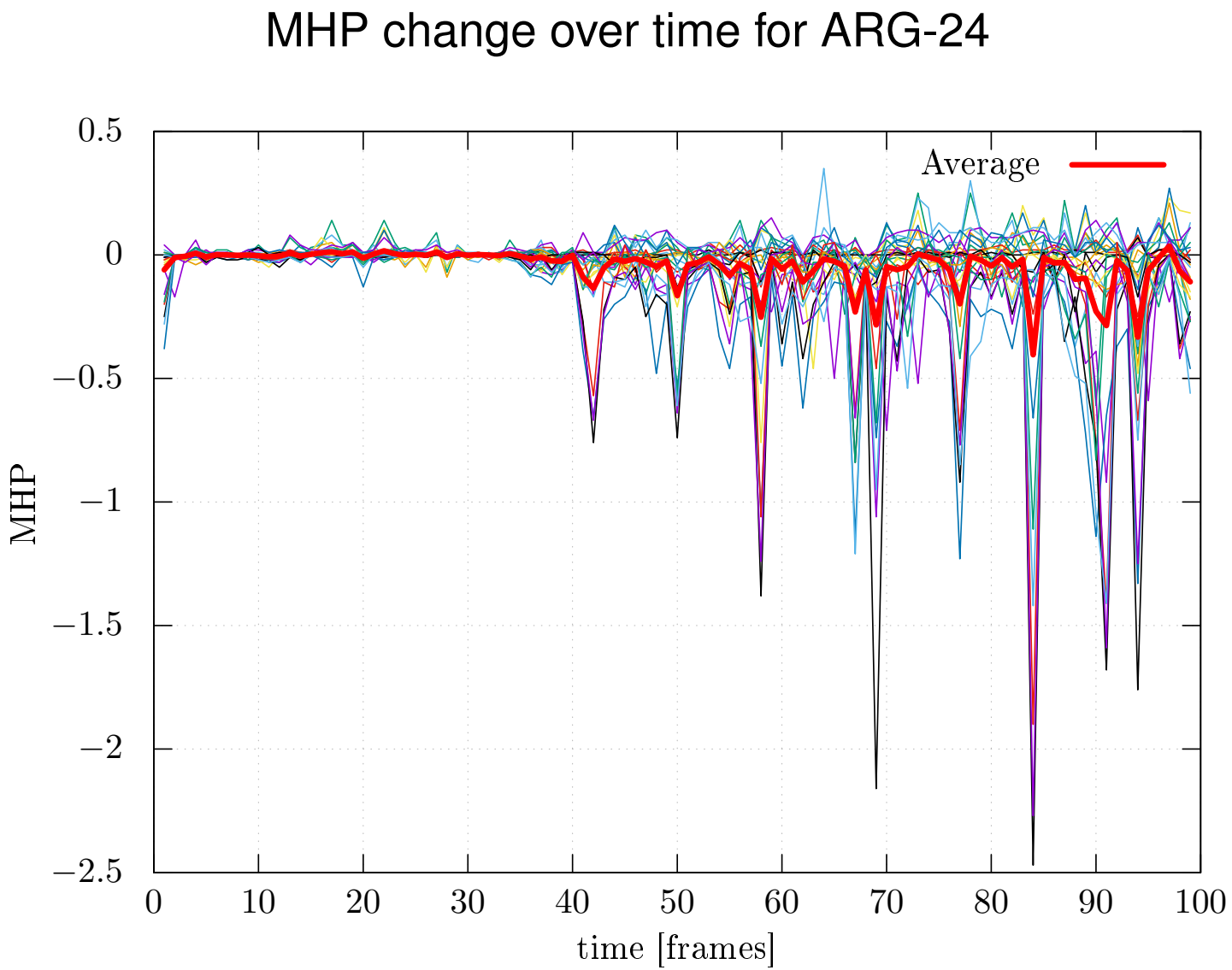
[7] Trajectory provided by Dr. Alejandra de Miguel Catalina

A video of the system

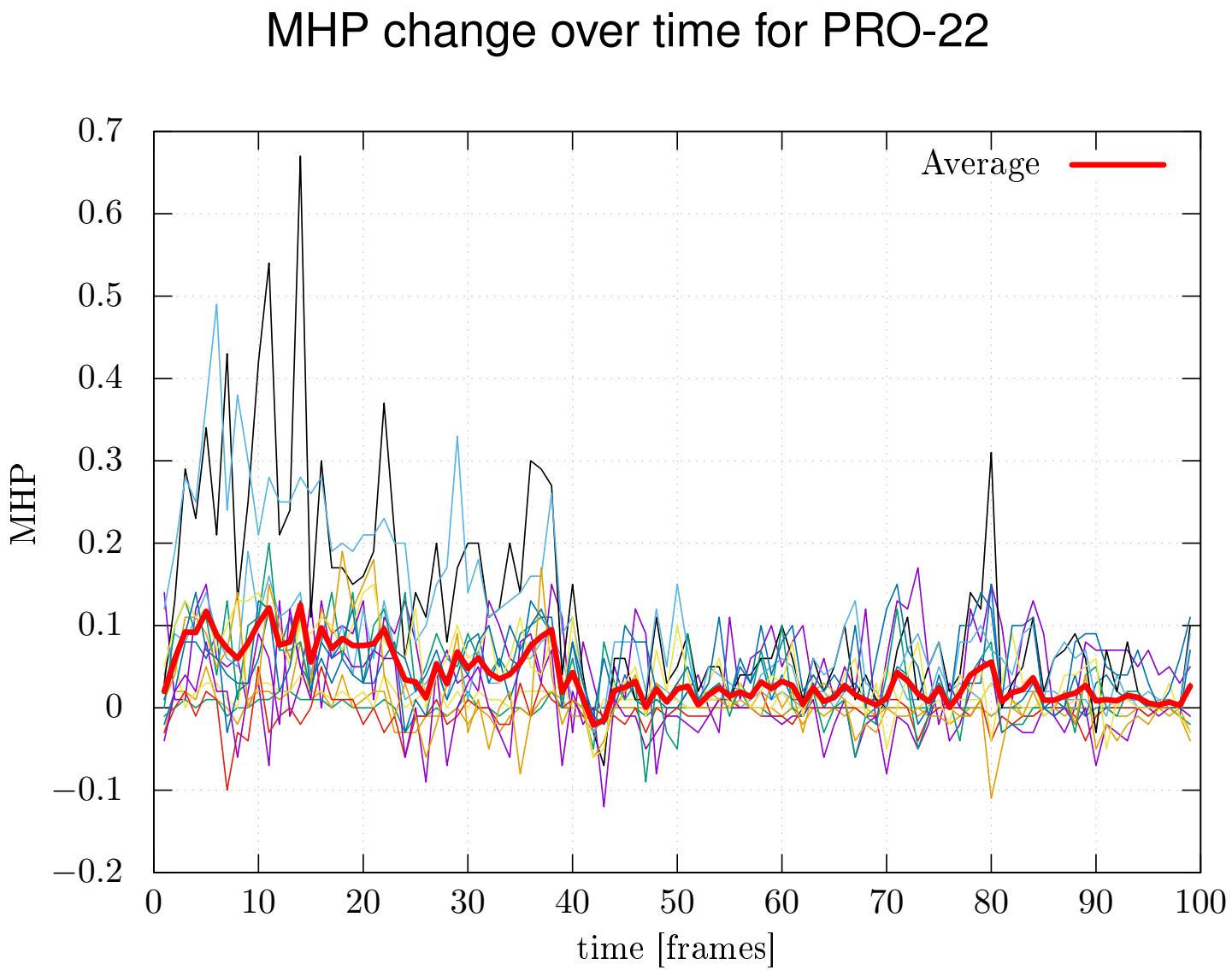
MHP Change Over Time



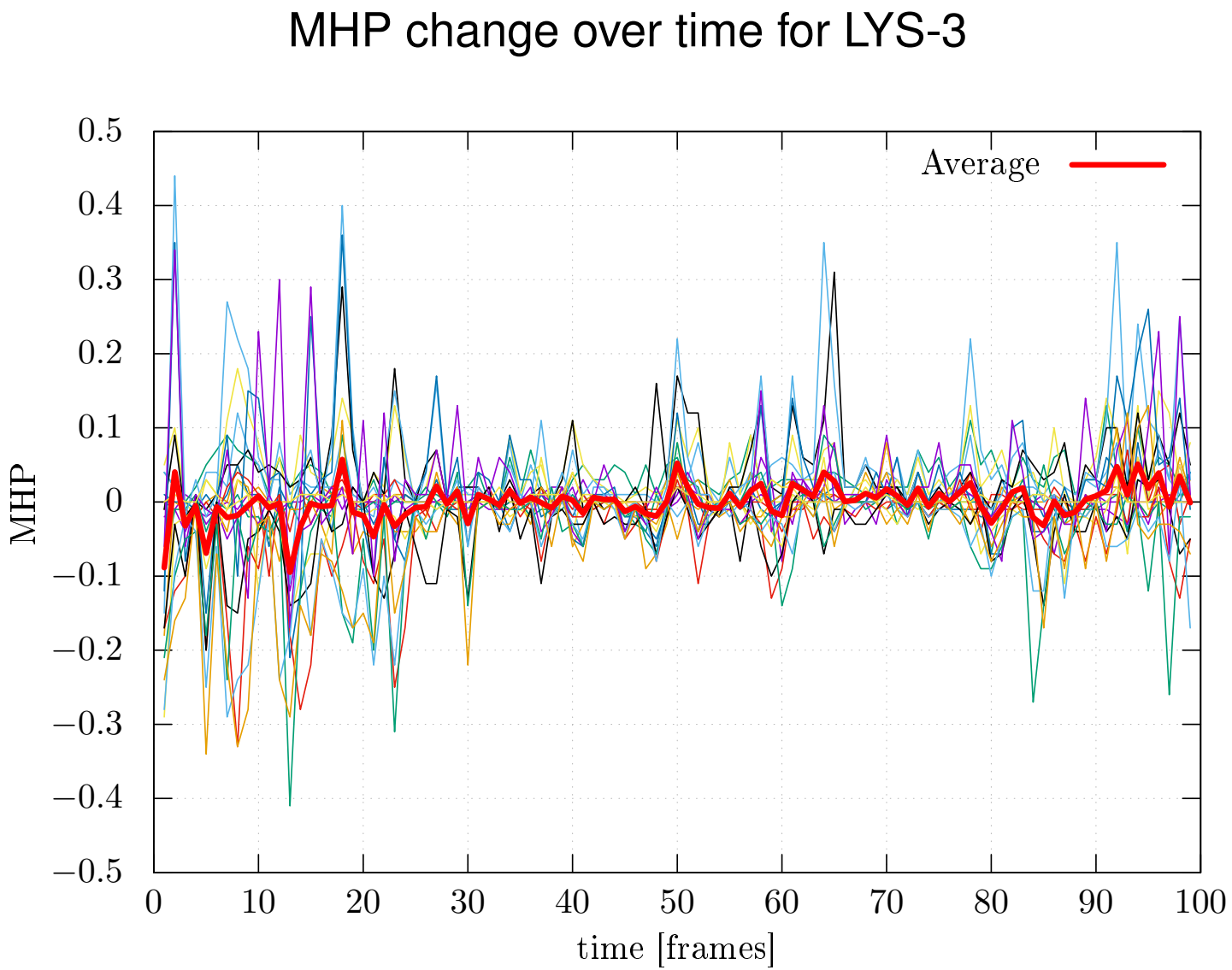
MHP Change Over Time



MHP Change Over Time



MHP Change Over Time



MHP Change Over Time

- We again get qualitative correlation to expected results (i.e. interior Hydrophobic).

MHP Change Over Time

- We again get qualitative correlation to expected results (i.e. interior Hydrophobic).
- The amino acid residues have the correct hydrophobicity.

MHP Change Over Time

- We again get qualitative correlation to expected results (i.e. interior Hydrophobic).
- The amino acid residues have the correct hydrophobicity.
- Observing changes in MHP of fragments of interest is possible.

MHP Change Over Time

- We again get qualitative correlation to expected results (i.e. interior Hydrophobic).
- The amino acid residues have the correct hydrophobicity.
- Observing changes in MHP of fragments of interest is possible.
- More systems could (and should!) be analyzed using this method.

Bibliography

[1] Timothy V. Pyrkovand Anton O. Chugunovand Nikolay A. Krylovand Dmitry E. Noldeand Roman G. Efremov. "PLATINUM: a web tool for analysis of hydrophobic/hydrophilic organization of biomolecular complexes". In: *Bioinformatics* 25.9 (2009), pp. 1201–1202.

[2] Patrick Gaillard, Pierre-Alain Carrupff, and Alain Boudon Bernard Testa. "Molecular Lipophilicity Potential, a tool in 3D QSAR: Method and applications". In: *Journal of Computer-Aided Molecular Design* 8.235 (1994), pp. 83–96.

[3] Arup K. Ghose, Vellarkad N. Viswanadhan, and John J. Wendoloski. "Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An Analysis of ALOGP and CLOGP Methods". In: *The Journal of Physical Chemistry A* 102 (1998), pp. 3762–3772.

[4] Vladimir PliSka, Bernard Testa, and Han van de Waterbeemd. *Lipophilicity in Drug Action and Toxicology*. VCH Verlagsgesellschaft mbH, 1996.

Thank You for Your Attention!