



Visual Analysis of protein-ligand interactions

P. Vázquez¹, P. Hermosilla², J. Estrada³, V. Guallar³, A. Vinacua¹

¹ ViRVIG Group, UPC

² U. Ulm

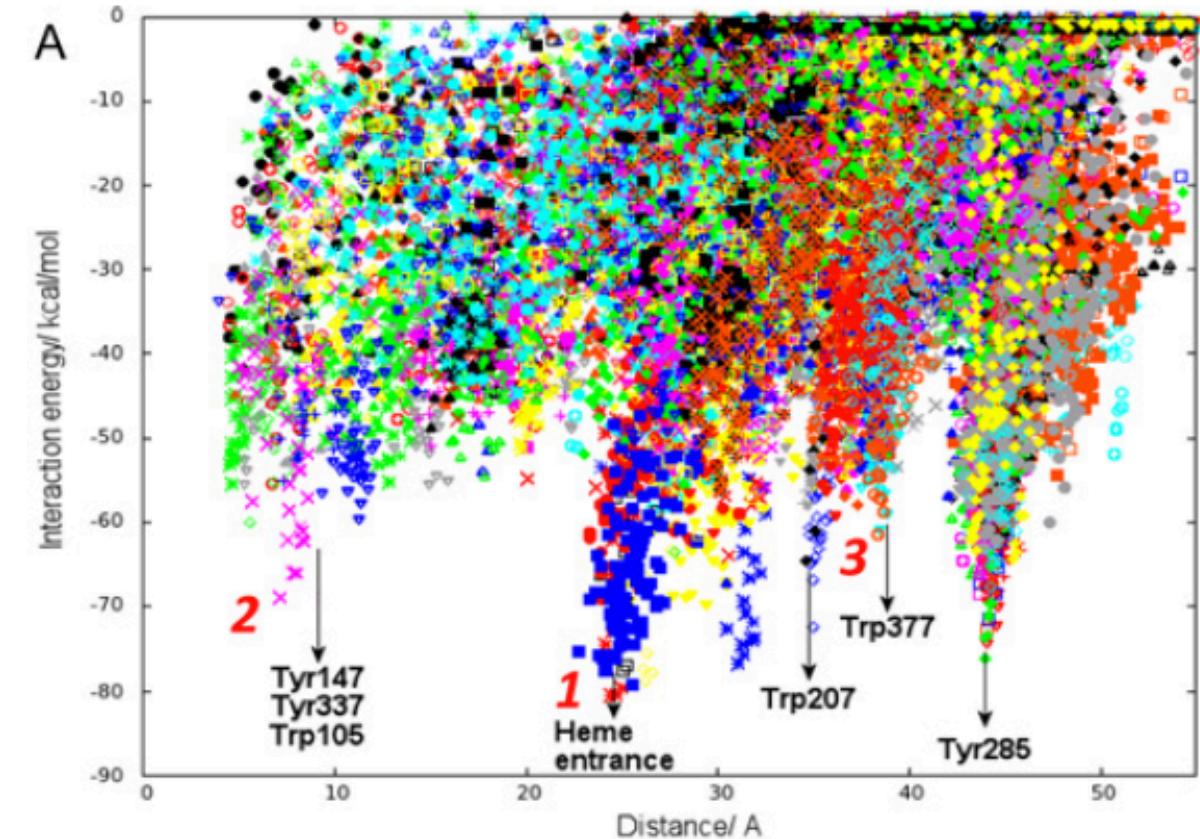
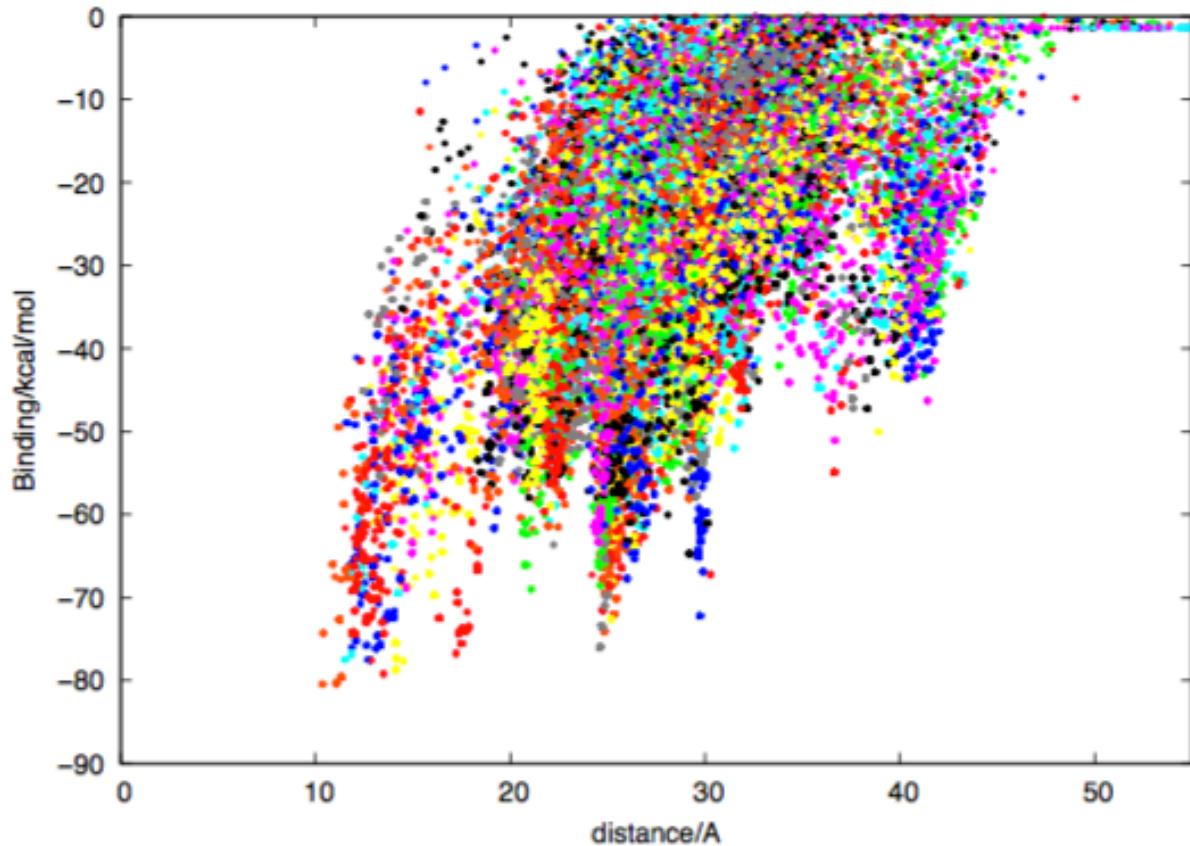
³ Barcelona SuperComputing Center

Motivation

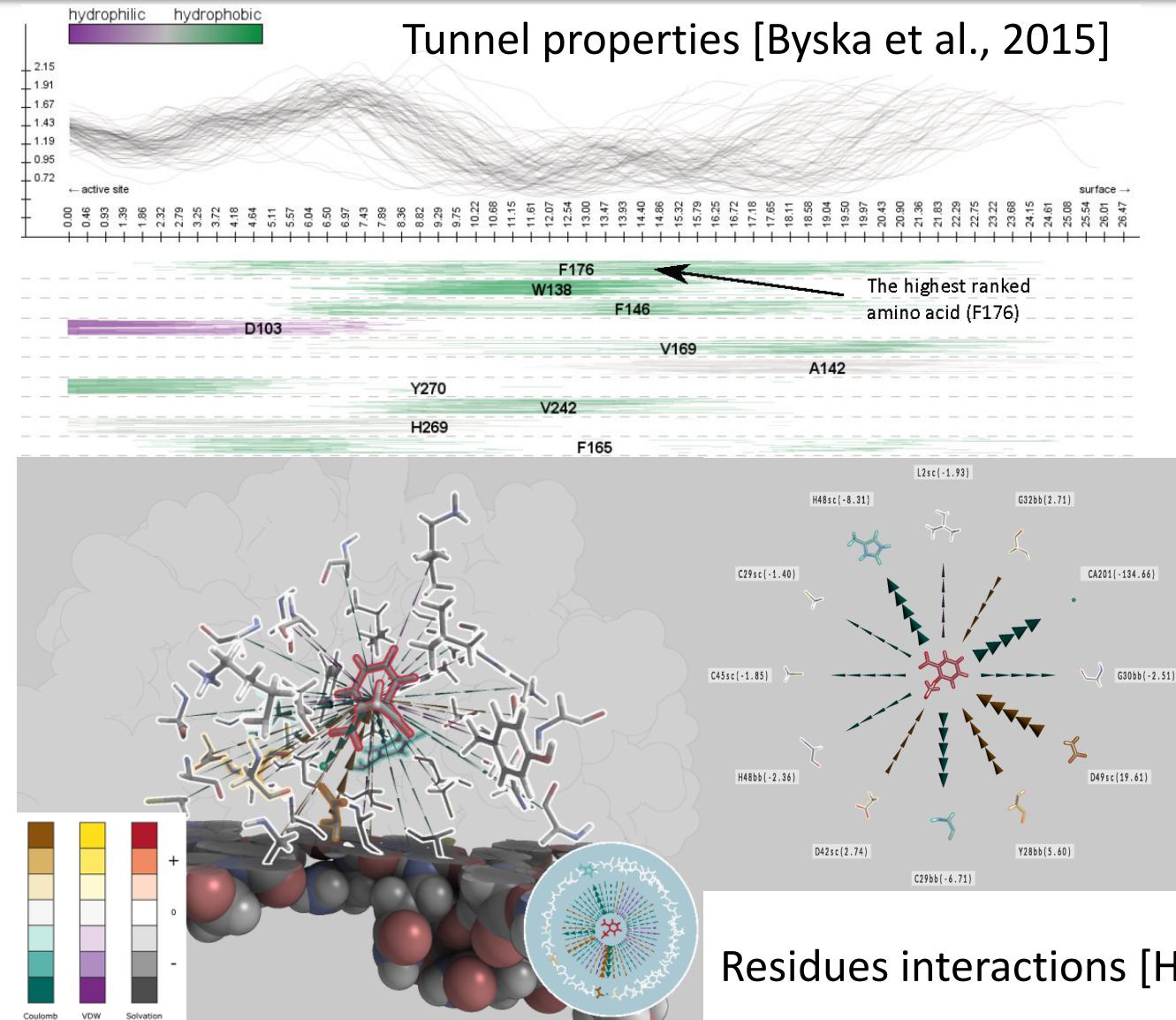
- Molecular Simulations
 - Evaluate binding affinity of ligands to proteins
 - Great utility for drug design
 - Result is a set of snapshots (trajectory)

Motivation

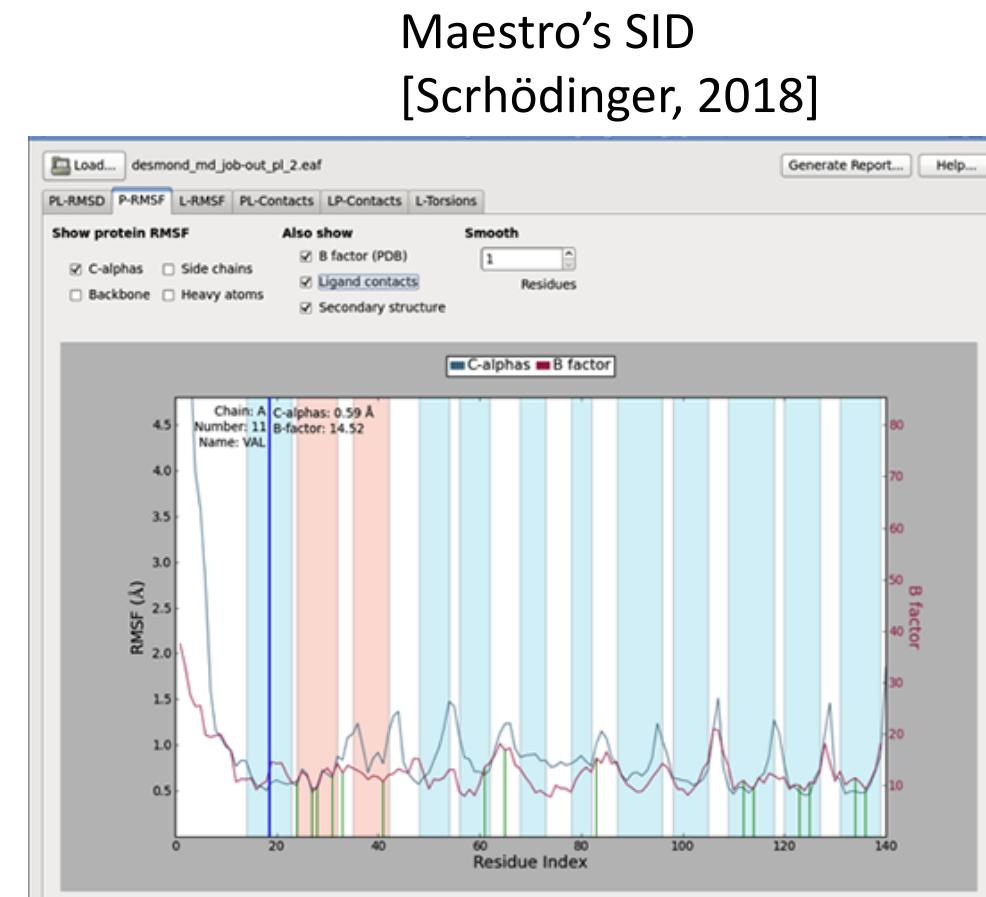
- Molecular Simulations generate a lot of data



Motivation



Tunnel properties [Byska et al., 2015]



Maestro's SID
[Schrödinger, 2018]

Motivation

- Data analysis is time consuming
 - No specific tools for overview simulation
- Requirements
 - Whole trajectory analysis
 - Including physico-chemical properties, e.g. per-residue energy interactions, h-bonds, fluctuations

Objectives

- Facilitate trajectory analysis
 - Most flexible residues?
 - Residues interacting with the ligand at a certain distance?
- Facilitate trajectories comparison
 - Which residues intervene in the docking? What if we change the ligand?
 - How differ these simulations?

Objectives

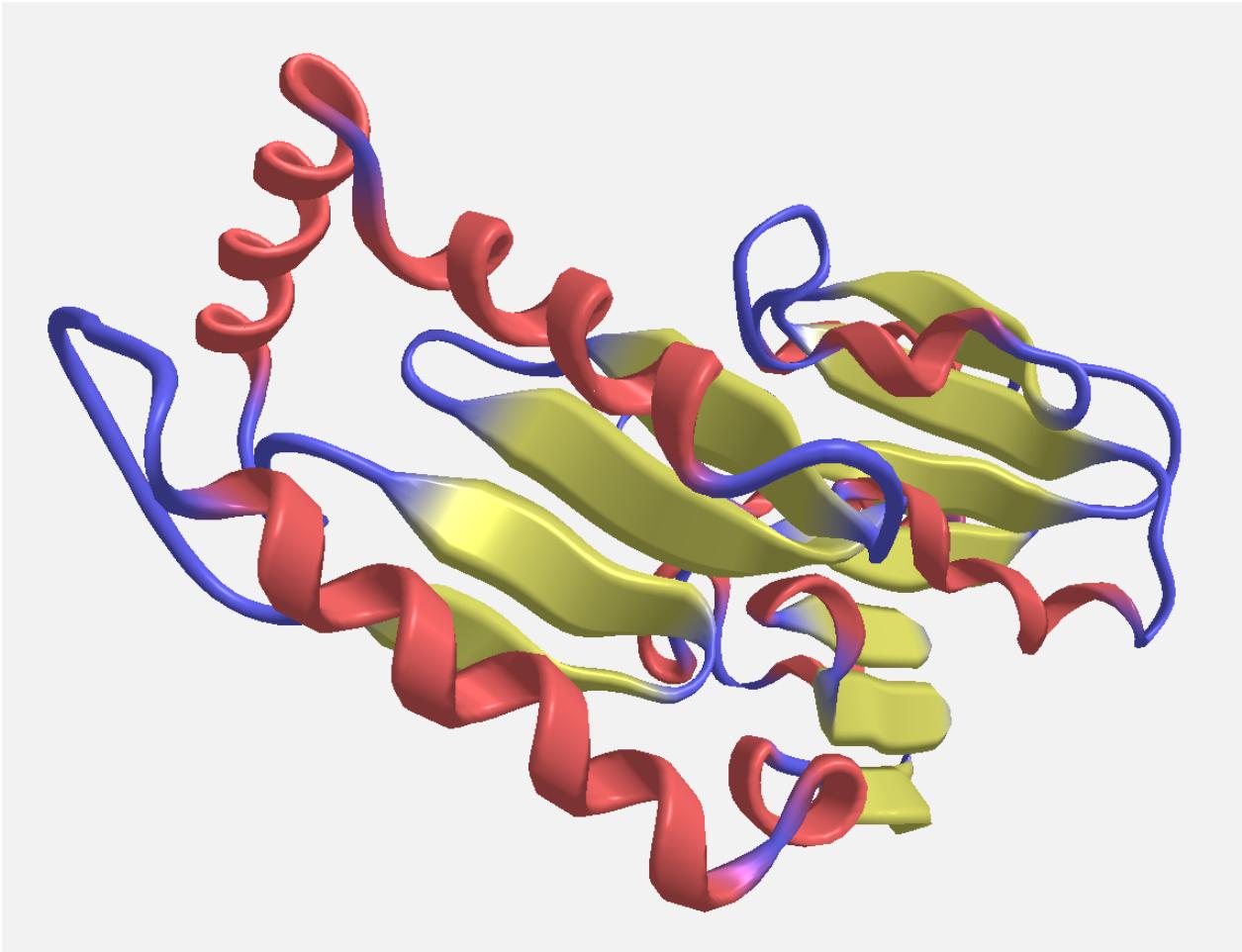
- Our solution
 - Discards 3D structure information
 - Preserve secondary structures and residue distribution
 - Reduce screen footprint
 - Individual trajectory snapshots
 - Leave room to show simulation-related data (RMSF, h-bonds...)
 - Interactive, compact visual analysis of multiple data

Outline

- Background
- Design: Visually supported trajectory analysis
- Use cases
- Conclusions & Future Work

Proteins

- Chains of amino-acids



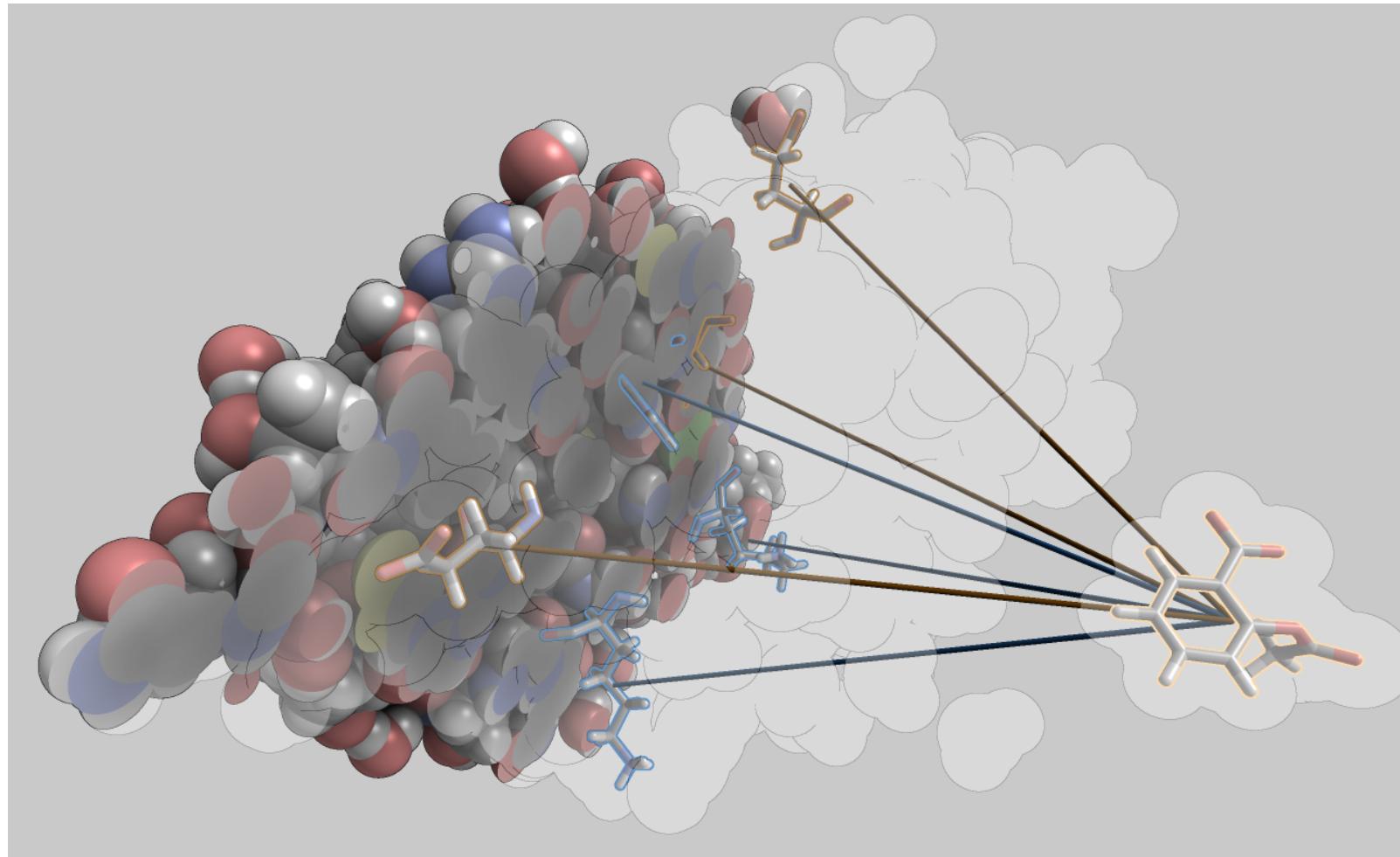
Simulation energy

- Free energy of binding
 - Low energy levels are more stable
 - Summation of terms
- Total energy

$$\Delta G_{bind} \approx E_{vdW} + E_{vac} + E_{solv}$$

- E_{vdW} : van der Waals energy (short-range)
- E_{vac} : electrostatic energy in the vacuum (long-range)
- E_{solv} : electrostatic energy due to the solvent

Energy visualization

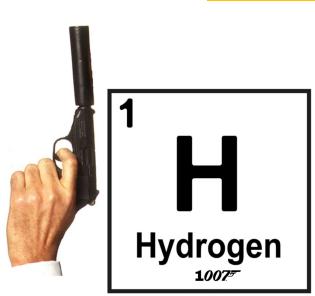


[Hermosilla et al., 2016]

Root Mean Square Fluctuation

- Root Mean Square Fluctuation: average fluctuation of a residue's atom around a reference position

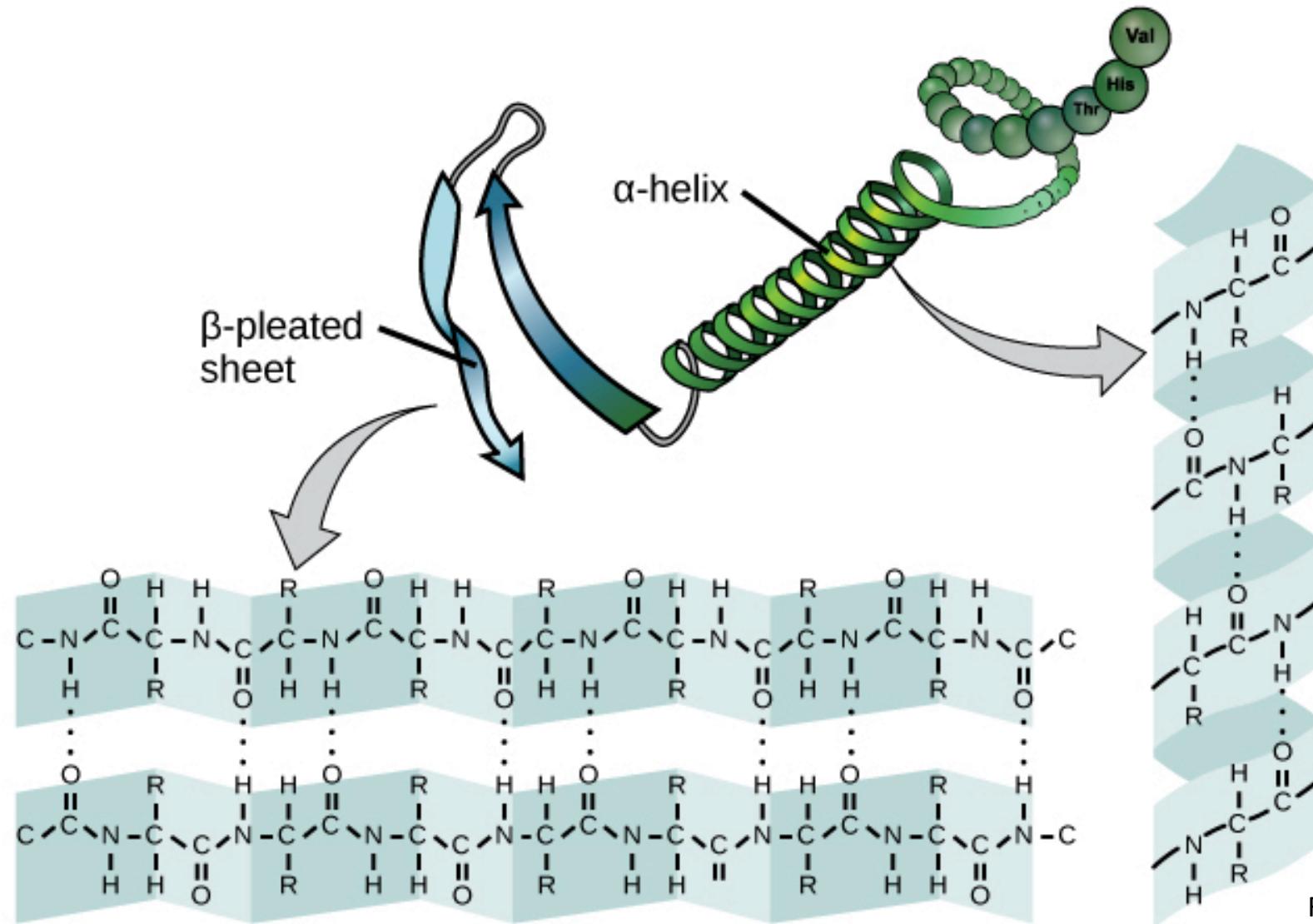
$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(r'_i(t) - r'_i(t_{ref}) \right)^2}$$



Bonds, hydrogen bonds

- Cause the molecule to fold into a specific shape
- Secondary structures are determined by H-bonds
 - Regularly occurring bonds between aminoacids → alpha helix
 - Aminoacids of different strands → beta sheets

Hydrogen bonds



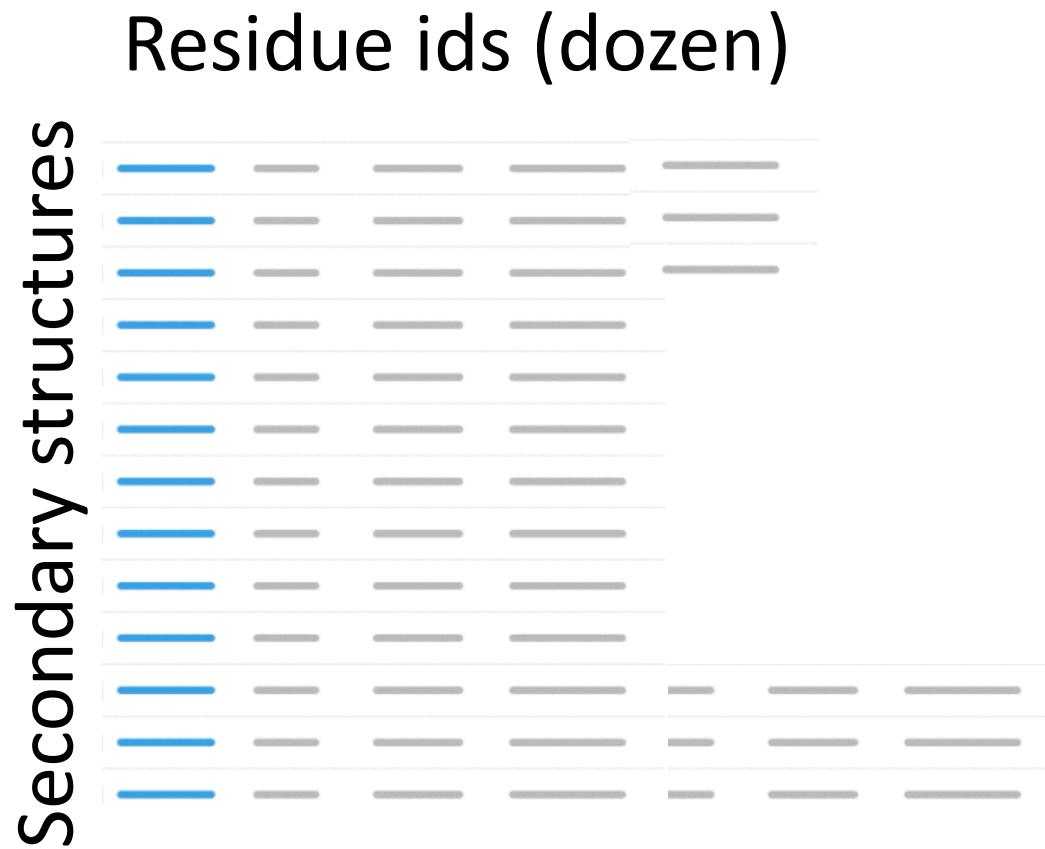
From: OpenStax Biology

Data

- Thousands of atoms x hundreds of steps
- Hundreds of residues
- A dozen or so of secondary structures
- Each step:
 - Atoms positions
 - System energy
 - Energy between ligand & residues (3 types)
 - H-Bonds (2 types)

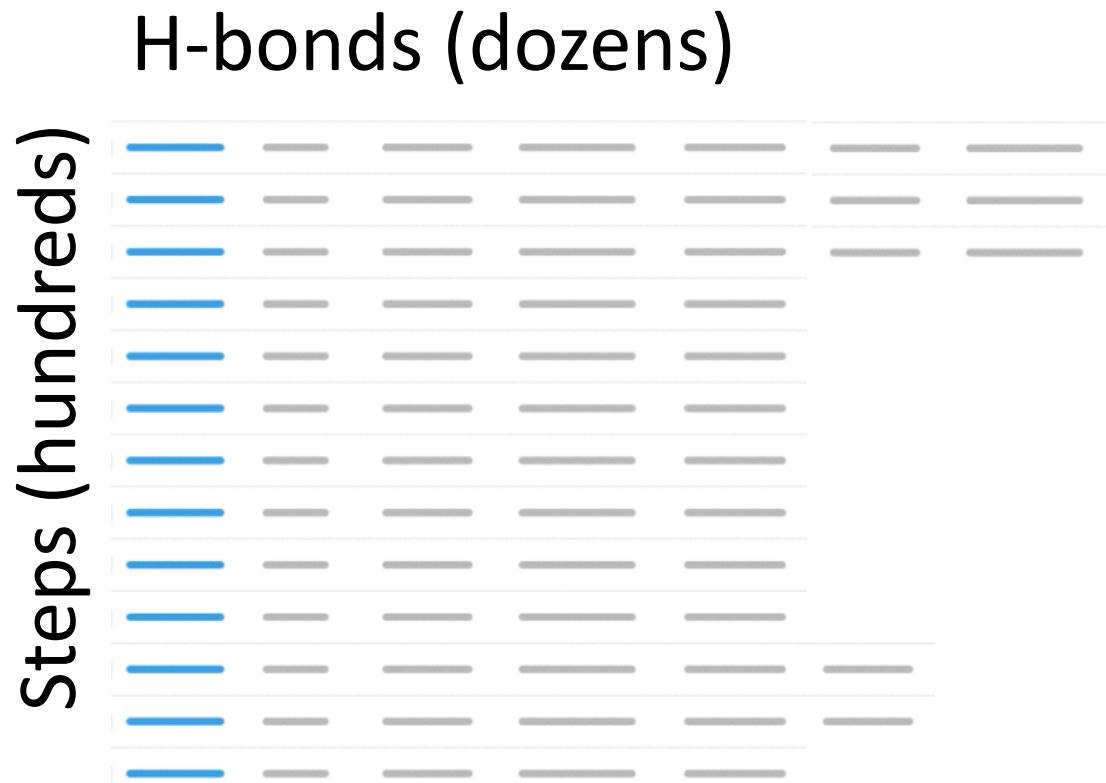
Data

- Molecular structure & fluctuation



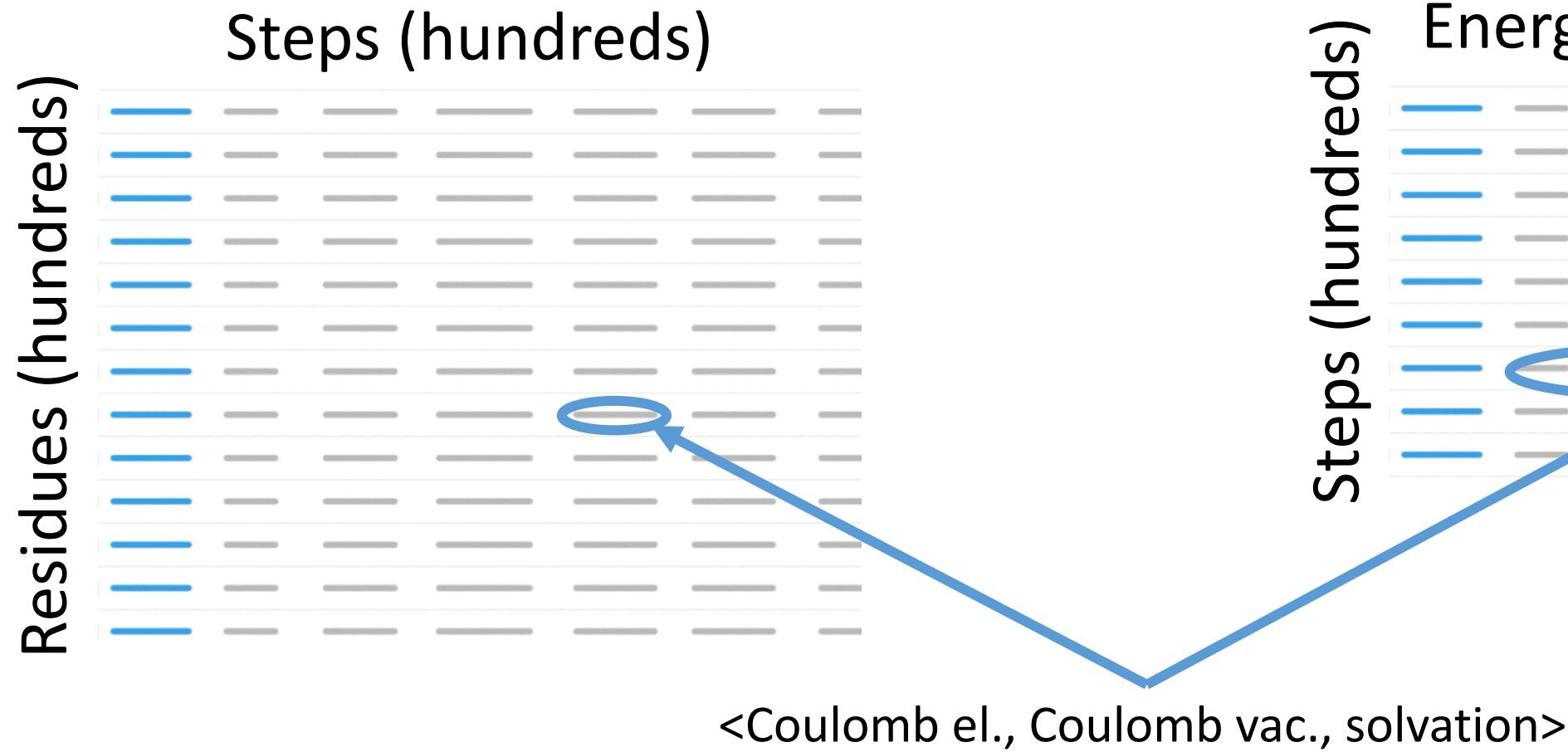
Data

- H-bonds



Data

- Energy



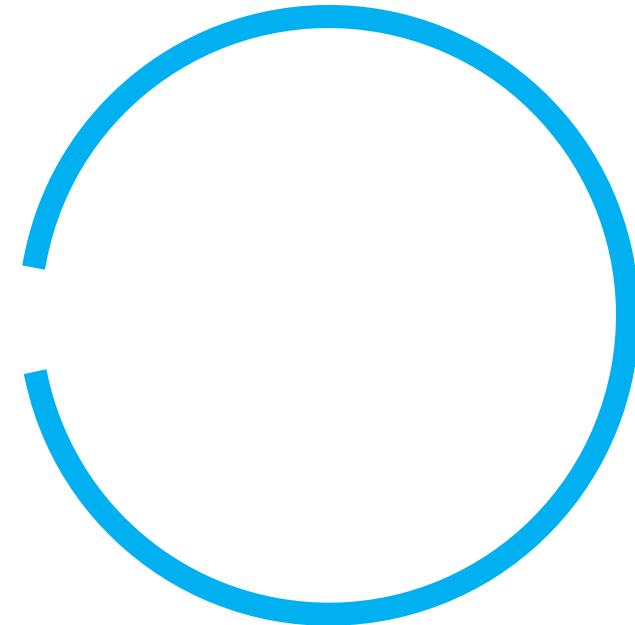
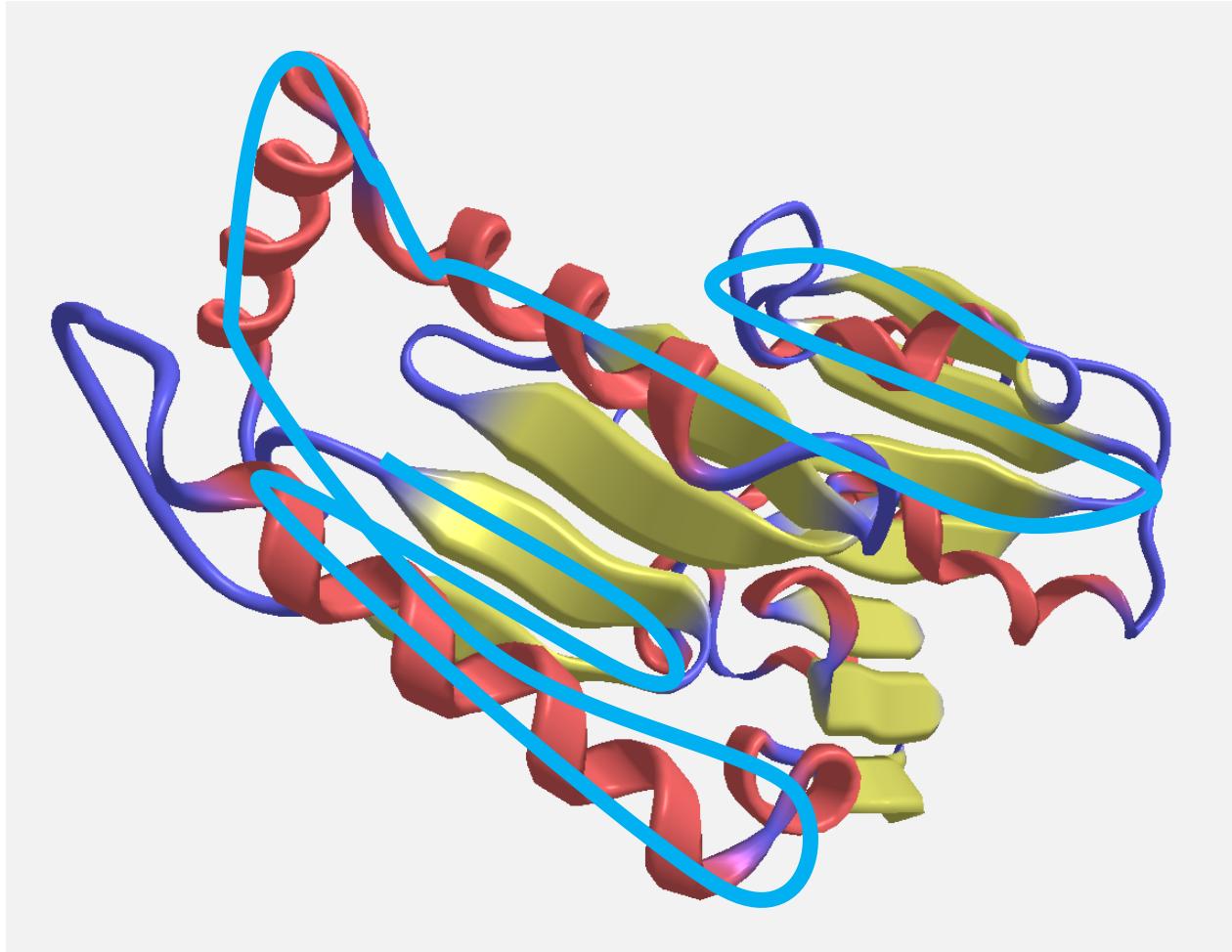
Data

- Atoms' positions per step, dynamic → REMOVE
- Residues (molecular structure), static → KEEP
- Energy (1 value) per step, dynamic → KEEP
- Inter-molecular h-bonds per step (connections to residues, backbone & side chain configuration), dynamic → KEEP
- Intra-molecular h-bonds per step (pairs of connections, donor & acceptor), dynamic → KEEP
- Residues vs ligand energies (3 values) per step, dynamic → KEEP
- RMSF per residue, static → KEEP

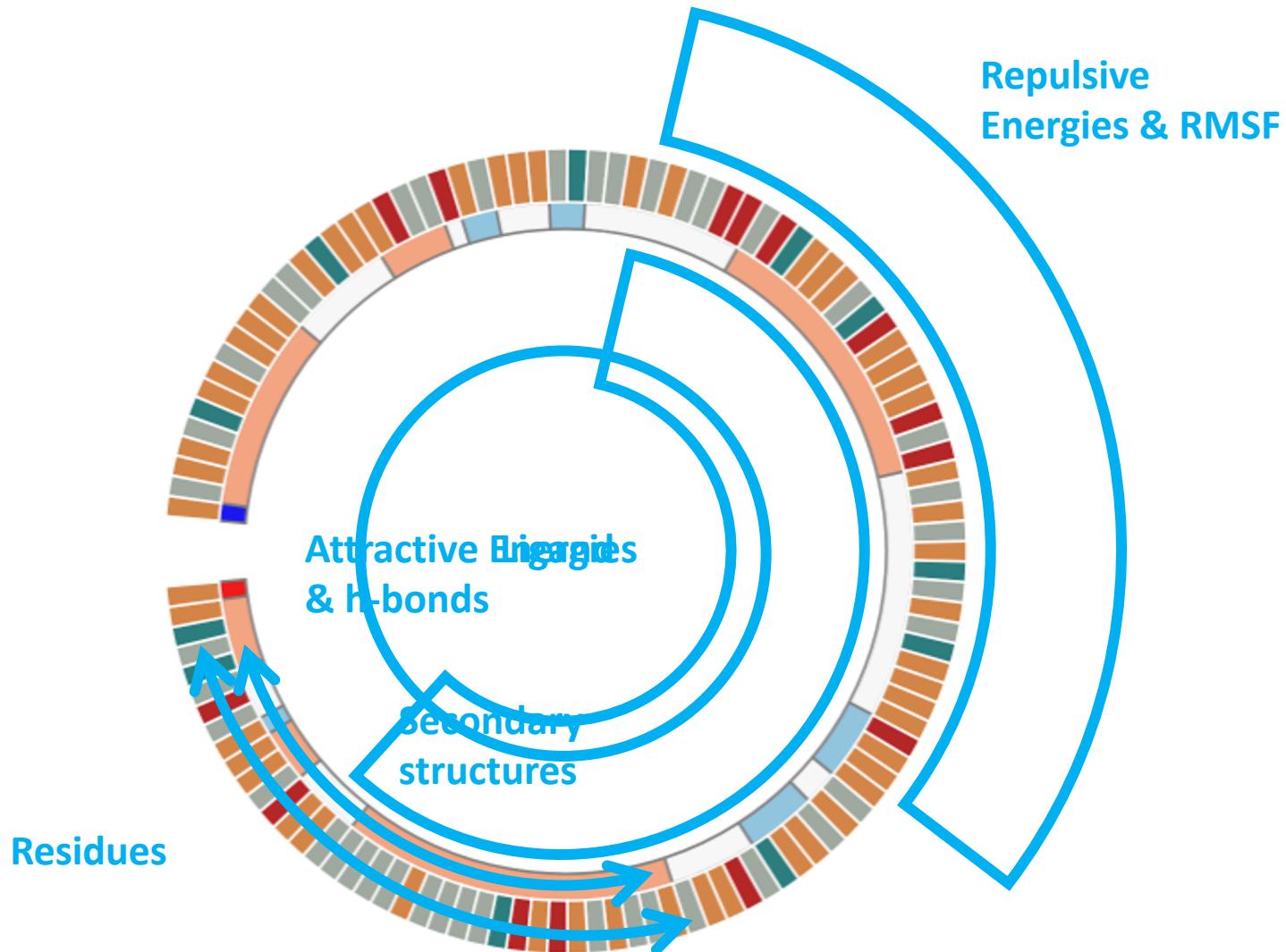
Data

- Derived values
 - Minimum energy per residue along simulation
 - Maximum energy per residue along simulation

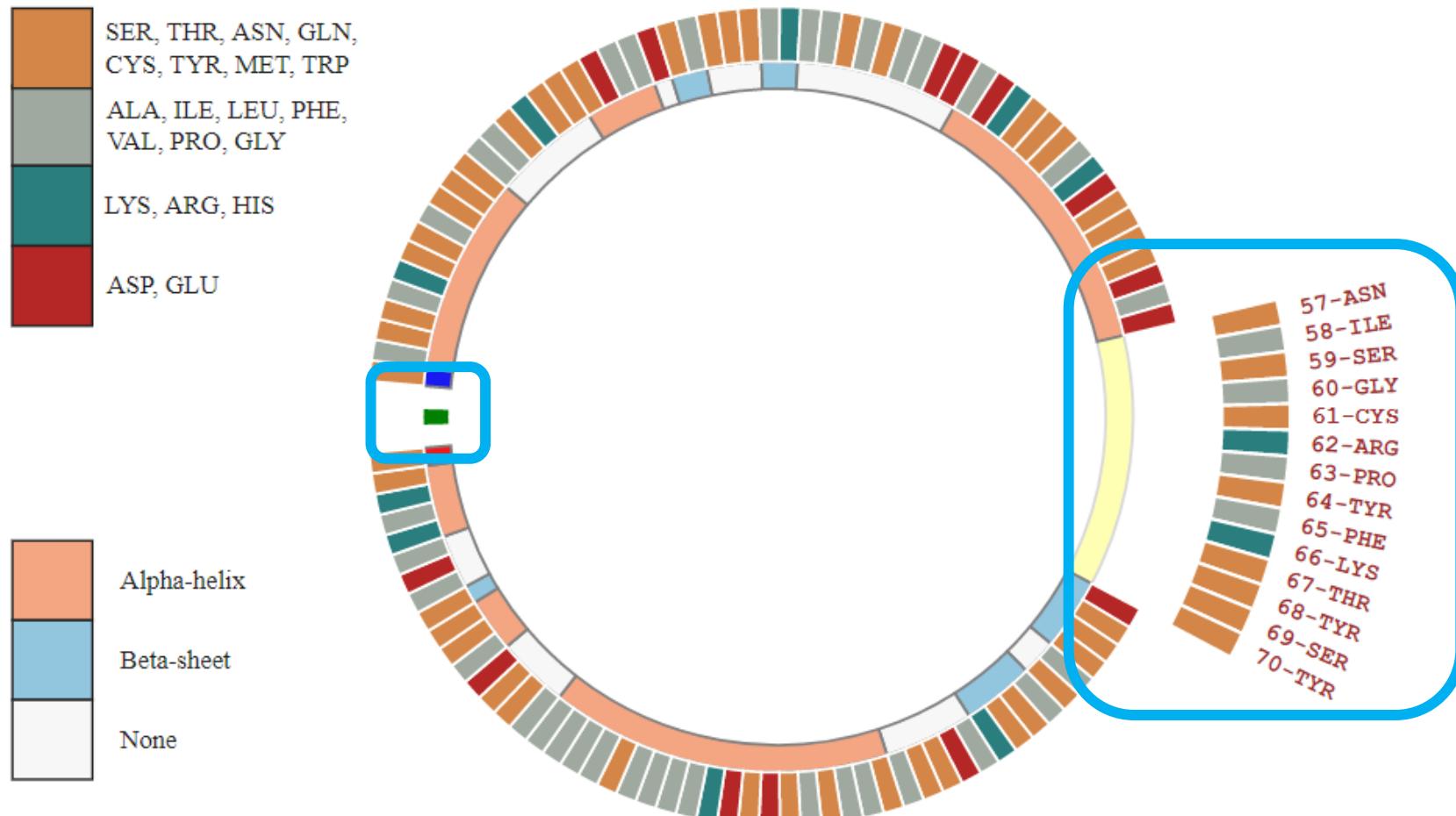
Main idea



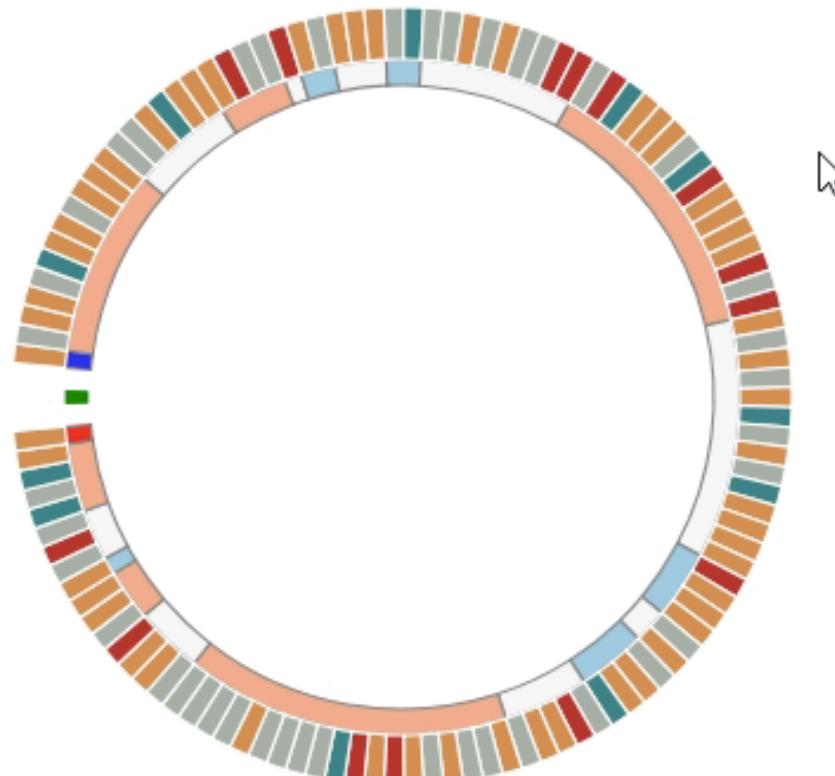
Overview



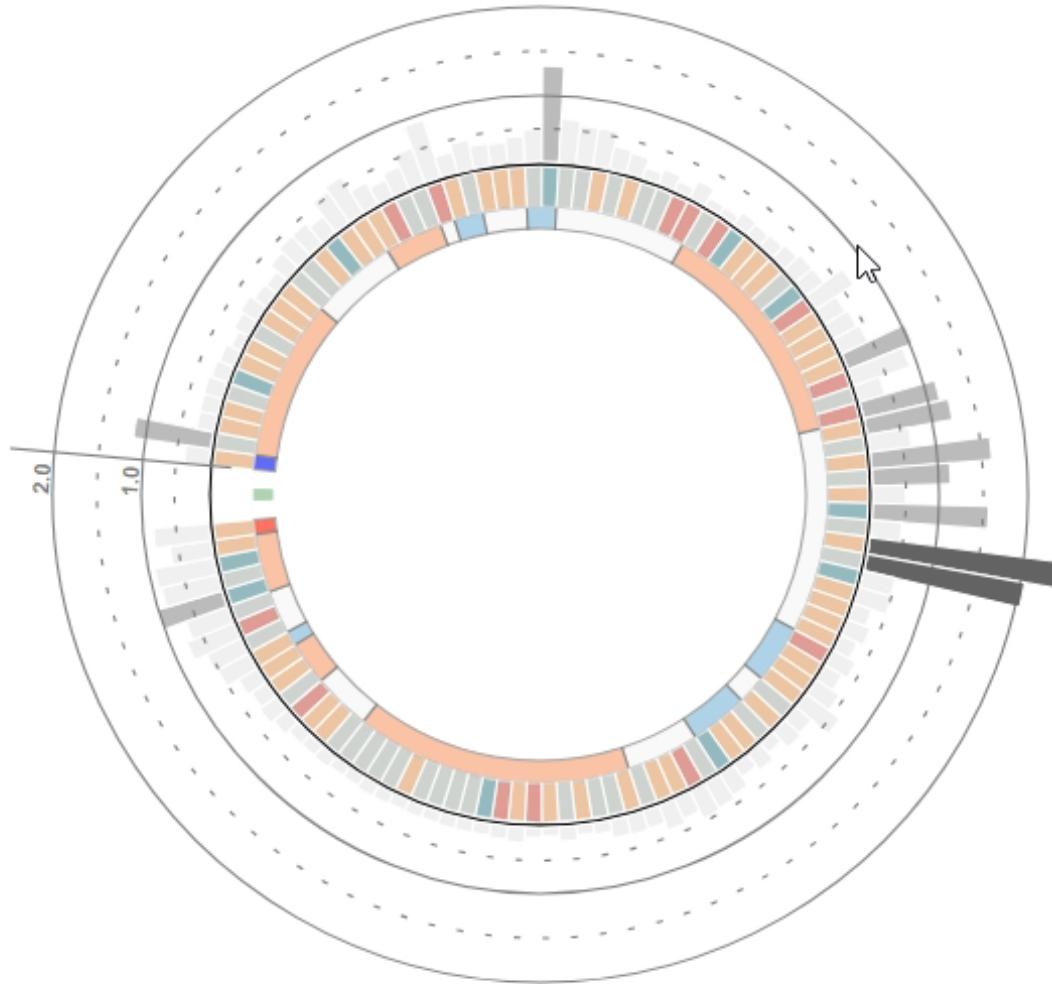
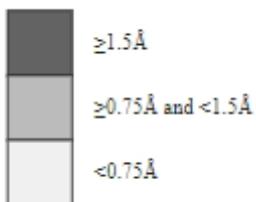
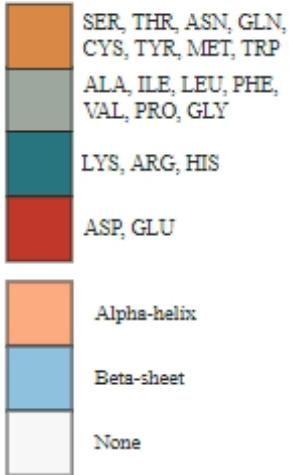
Molecular structure



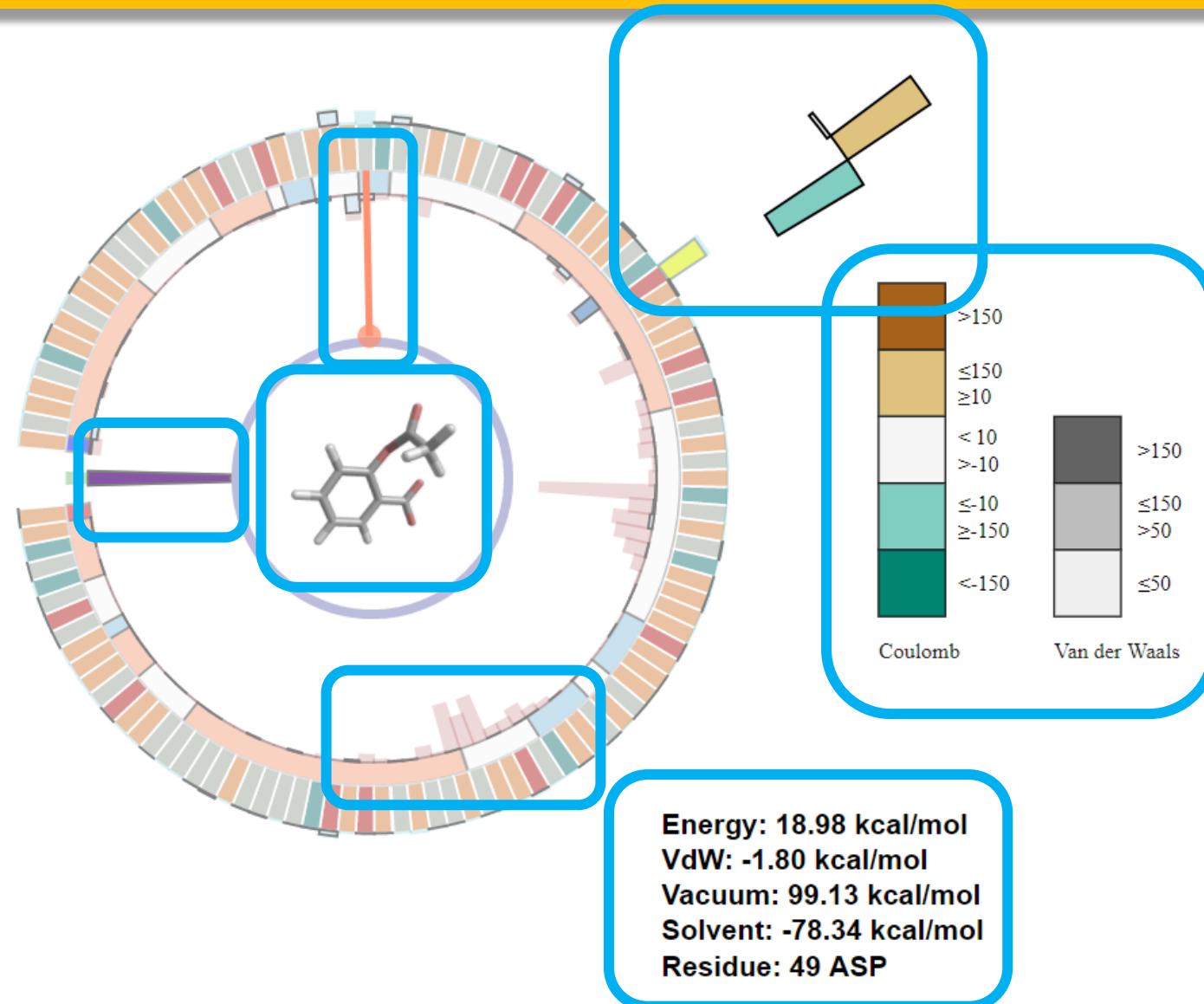
Molecular structure



RMSF



Energy



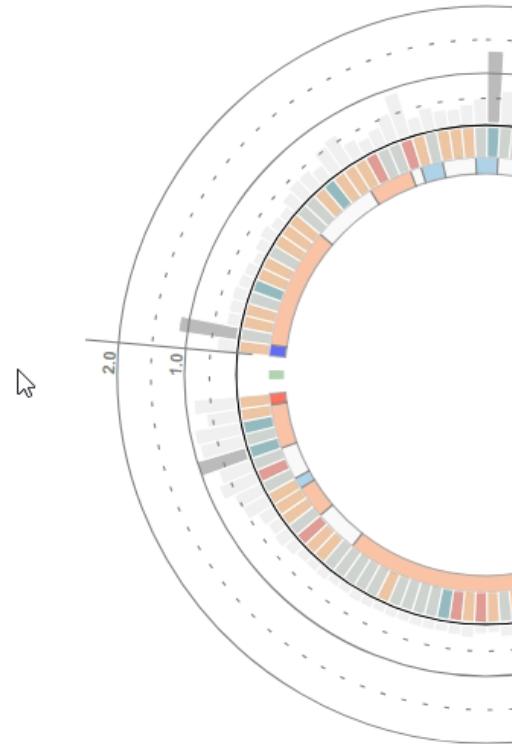
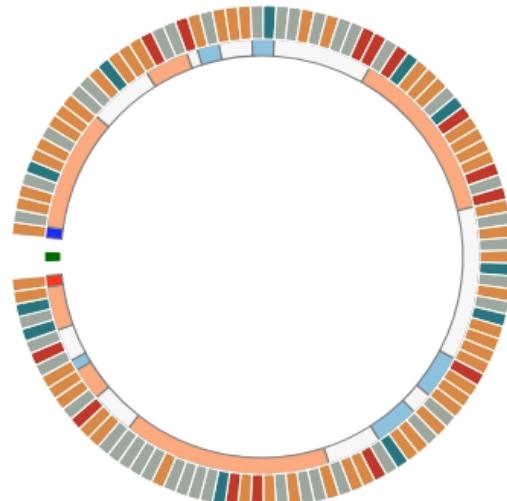
Energy

Visual Analysis of Protein-Ligand Interactions

Residues ▾

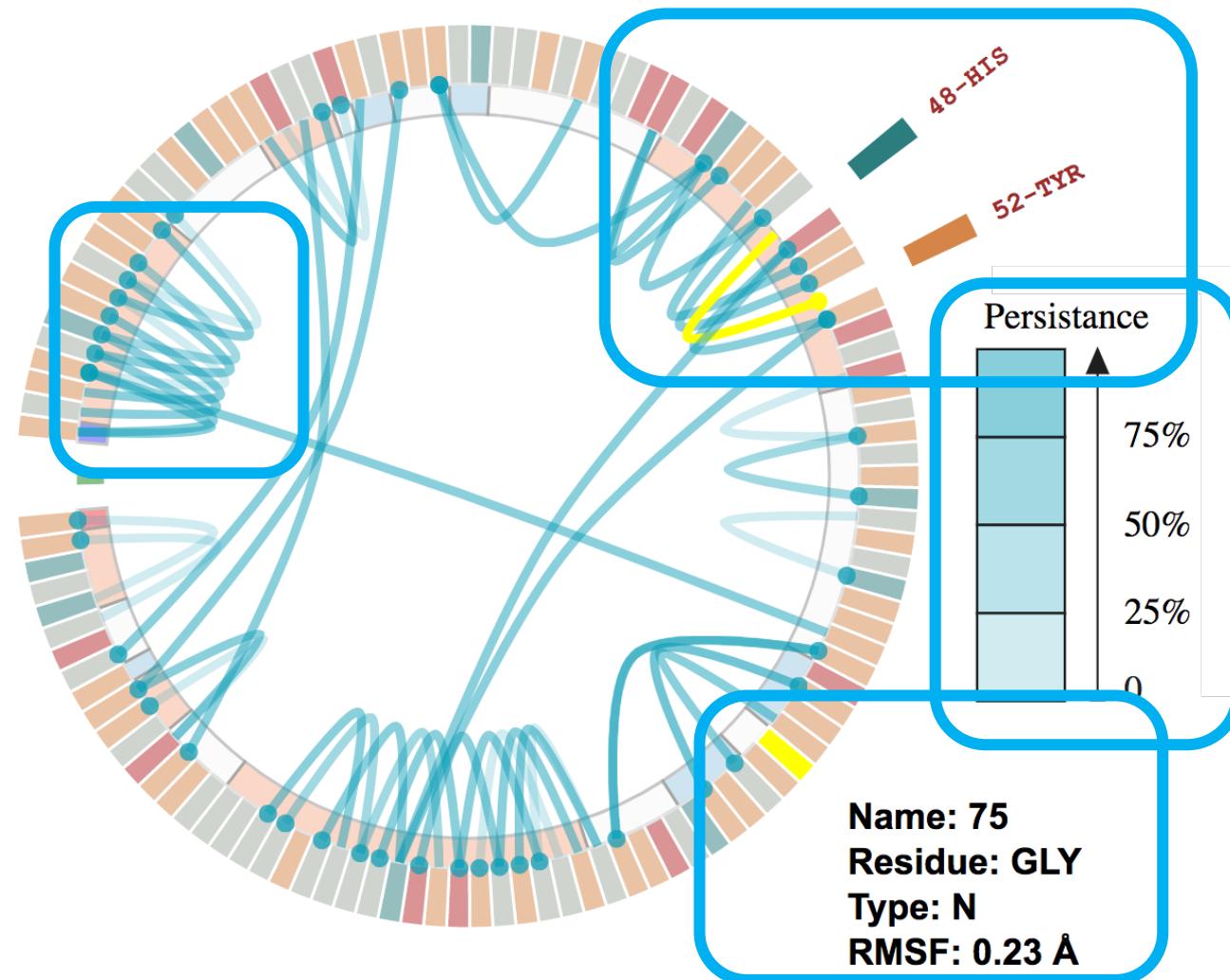
RMSF ▾

SER, THR, ASN, GLN, CYS, TYR, MET, TRP
ALA, ILE, LEU, PHE, VAL, PRO, GLY
LYS, ARG, HIS
ASP, GLU
Alpha-helix
Beta-sheet
None

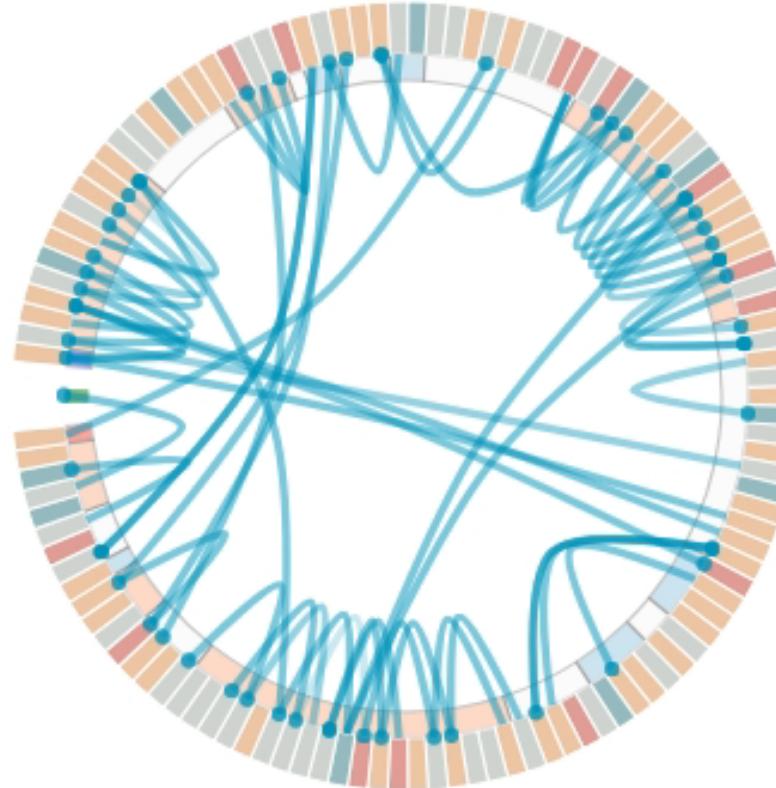


Frame 0

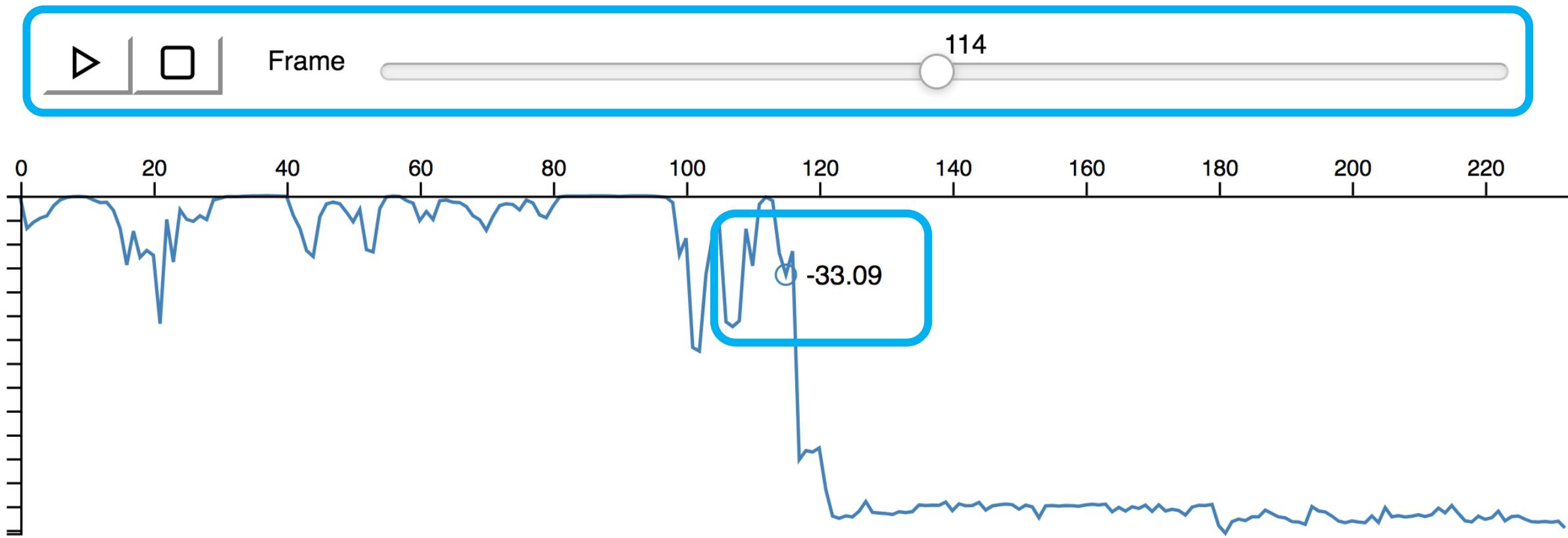
Hydrogen bonds



Hydrogen bonds



Energy chart



Visual Analysis of Protein-Ligand Interactions

Energies ▾

- VdW
- Max Energies
- Coul. Vacuum
- Overlay bonds
- Coul. Solvent

Pos [0..500] 0

Neg [-500..0] -500

Distance 50.0

SER, THR, ASN, GLN,
CYS, TYR, MET, TRP
ALA, ILE, LEU, PHE,
VAL, PRO, GLY

LYS, ARG, HIS

ASP, GLU

Alpha-helix

Beta-sheet

None

>150

≤150

≥10

<10

≥-10

≤-10

≥-150

≤-150

Coulomb

Van der Waals

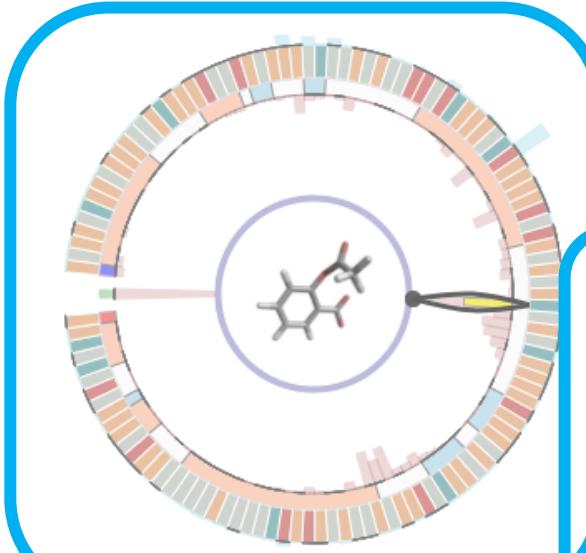
Dynamic legends and
details on demand

Energy chart

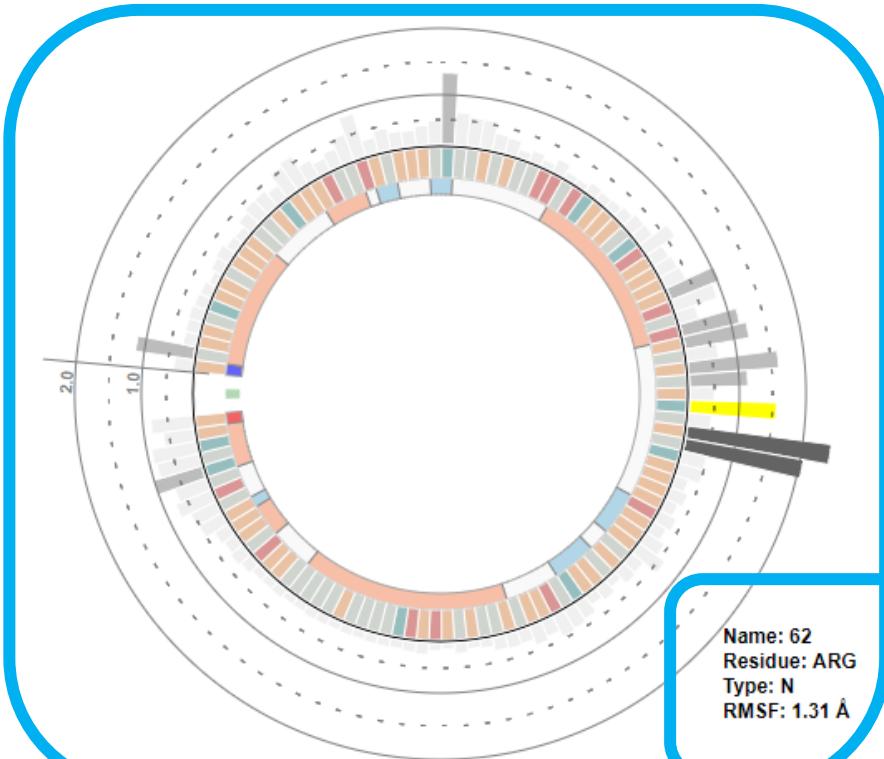
RMSF ▾

Controls

Property Views

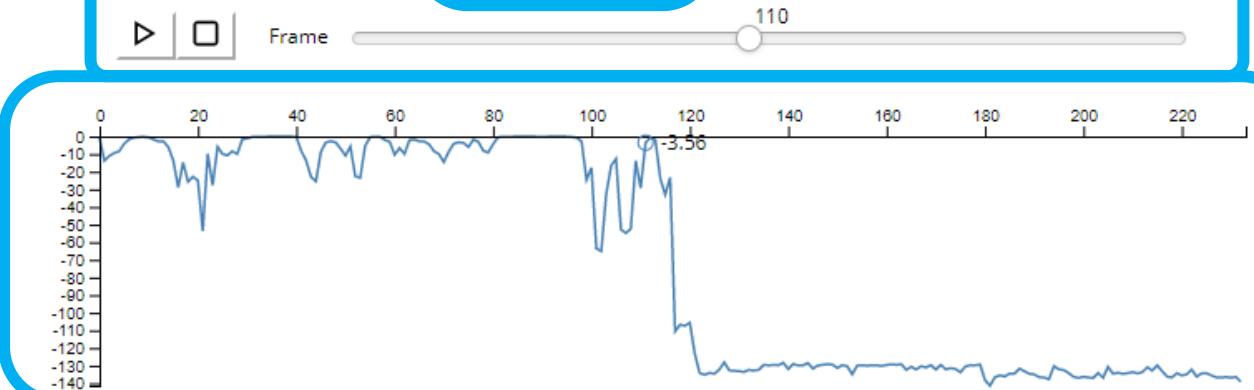


Energy: -29.52 kcal/mol
VdW: 3.97 kcal/mol
Vacuum: -105.59 kcal/mol
Solvent: 72.09 kcal/mol
Residue: 62 ARG



Name: 62
Residue: ARG
Type: N
RMSF: 1.31 Å

≥1.5 Å
≥0.75 Å and <1.5 Å
<0.75 Å



Visual Analysis of Protein-Ligand Interactions

Bonds ▾

- Inter-Molecular Bonds
- Intra-Molecular Bonds

SER, THR, ASN, GLN,
CYS, TYR, MET, TRP
ALA, ILE, LEU, PHE,
VAL, PRO, GLY

LYS, ARG, HIS

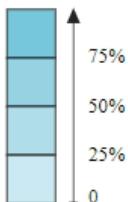
ASP, GLU

Alpha-helix

Beta-sheet

None

Persistance



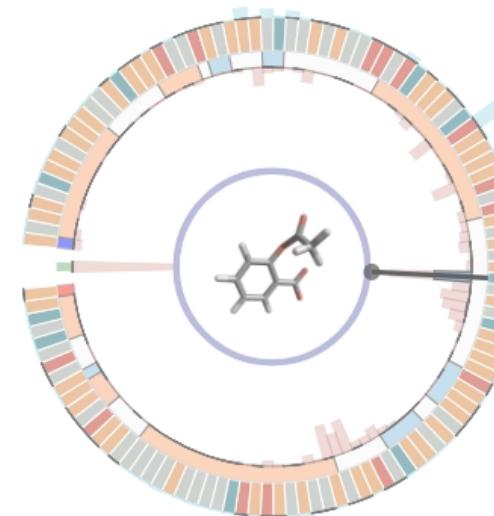
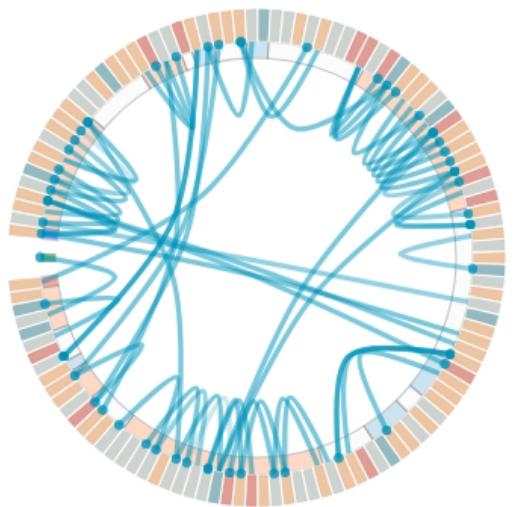
Energies ▾

- VdW
- Coul. Vacuum
- Coul. Solvent
- Max Energies
- Overlay bonds

Pos [0..500] 0

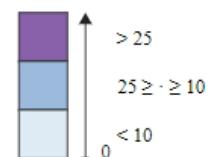
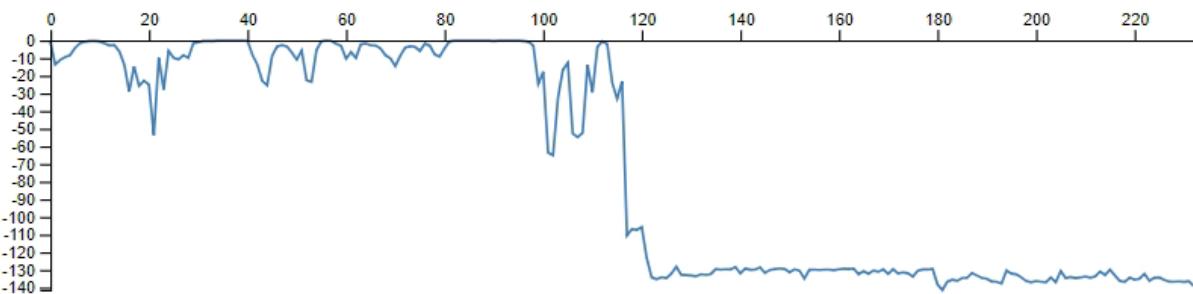
Neg [-500..0] -500

Distance 50.0

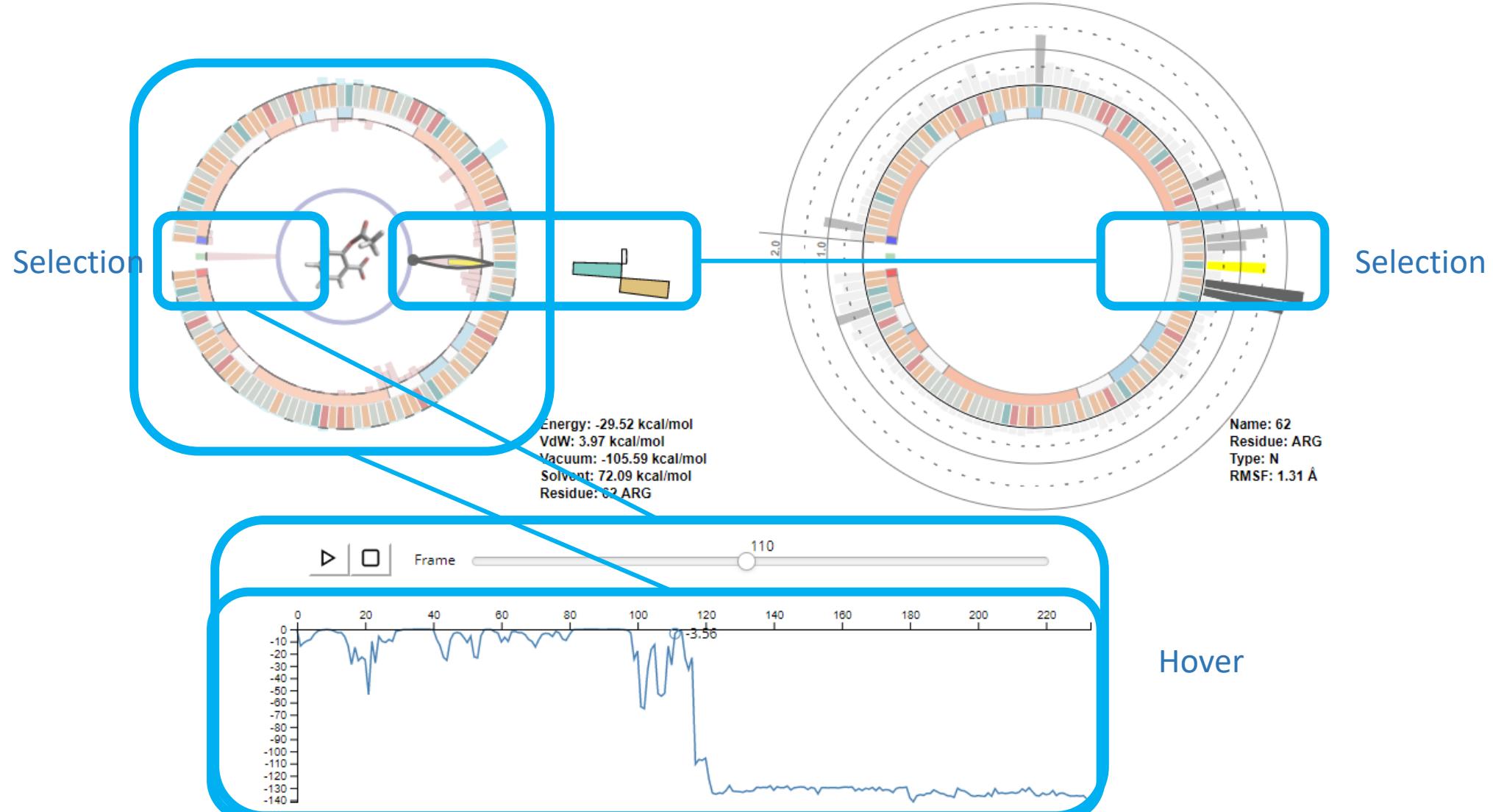


Frame

114



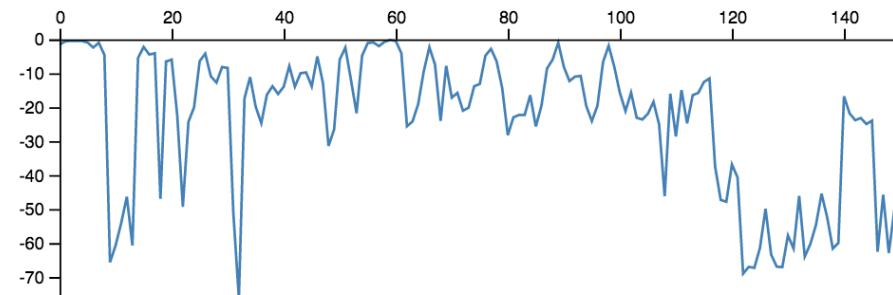
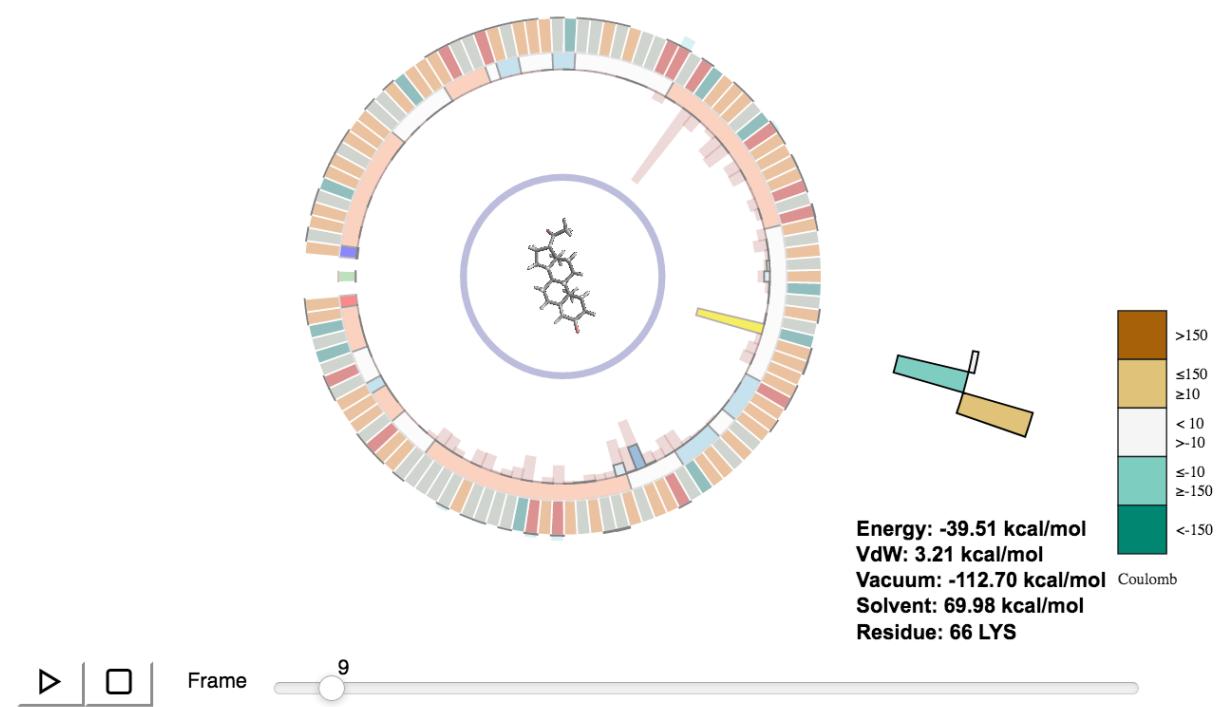
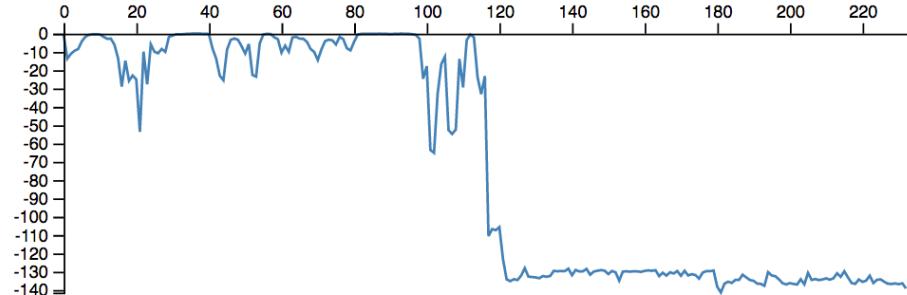
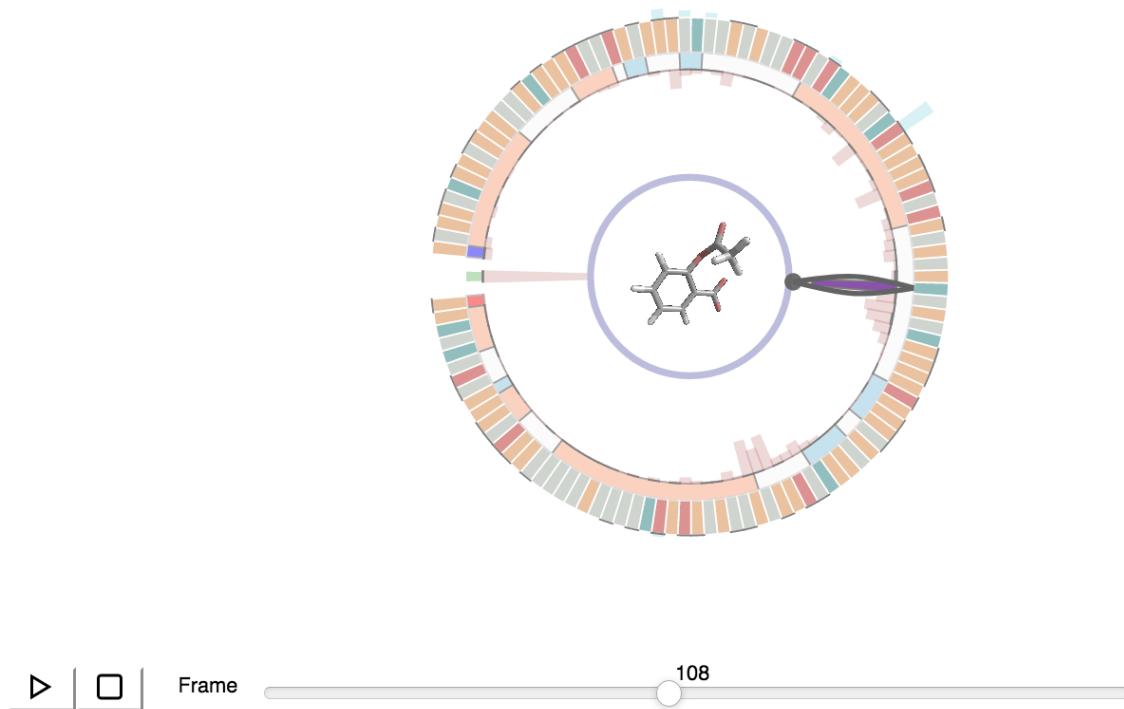
Interaction



Implementation

- Input:
 - Real data from Monte Carlo-based Molecular Simulations
 - Stored as JSON files
- JavaScript + D3
 - Any capable browser
 - Example in: www.cs.upc.edu/~ppau/Visualization/BondsVis.html

Implementation



Shape changes

- Analyzing shape changes induced by protein-ligand interaction
 1. Infer shape changes from RMSF view
 2. Analyze trajectory range with high degree of interaction of the implied residues
 3. Inspect h-bonds

Visual Analysis of Protein-Ligand Interactions

Residues ▾

RMSF ▾

SER, THR, ASN, GLN,
CYS, TYR, MET, TRP
ALA, ILE, LEU, PHE,
VAL, PRO, GLY

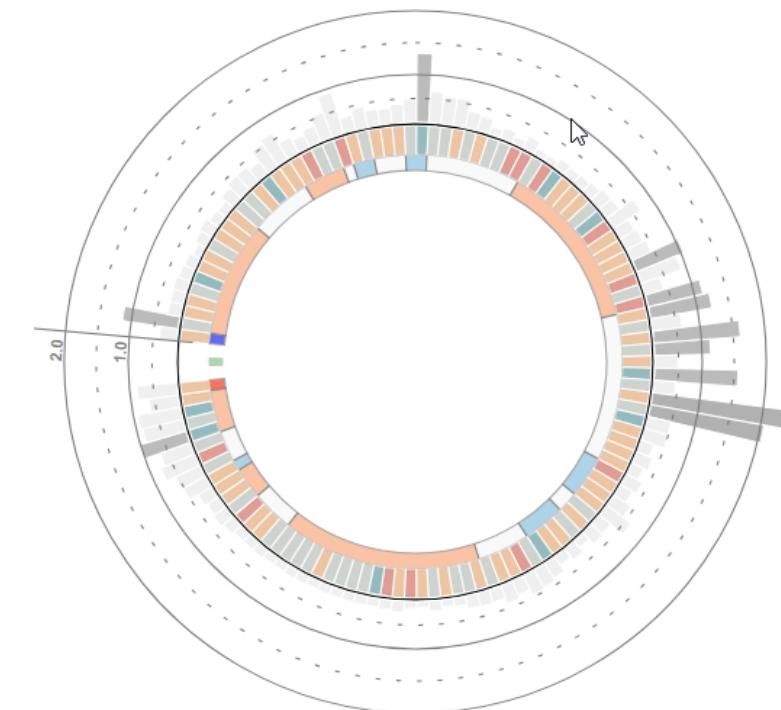
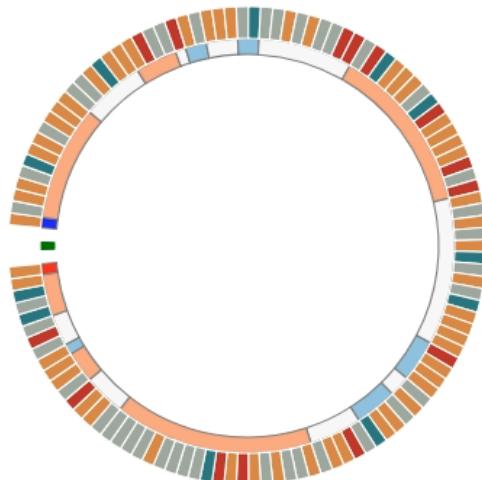
LYS, ARG, HIS

ASP, GLU

Alpha-helix

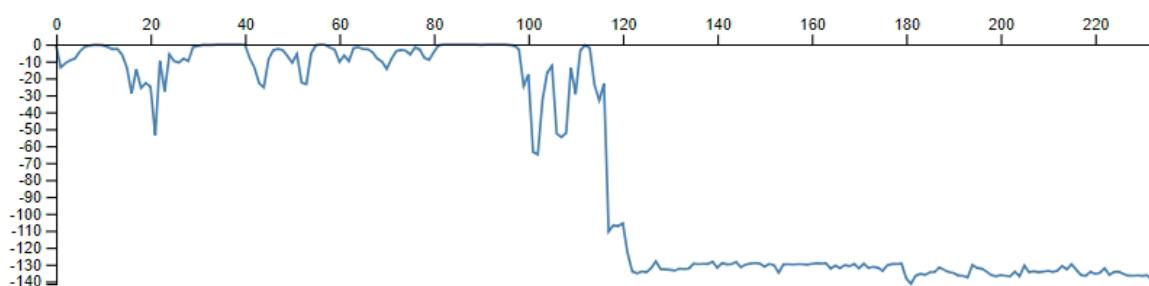
Beta-sheet

None

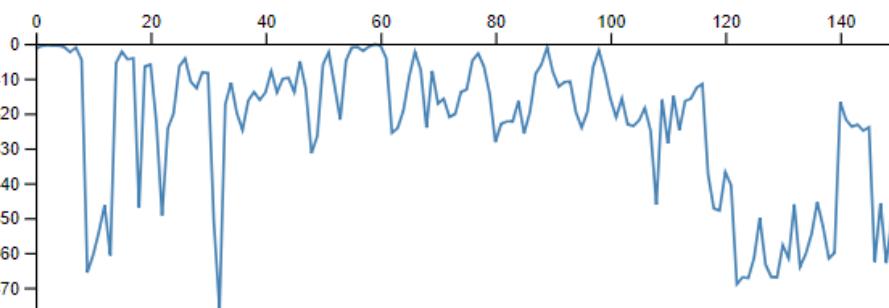
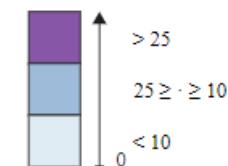
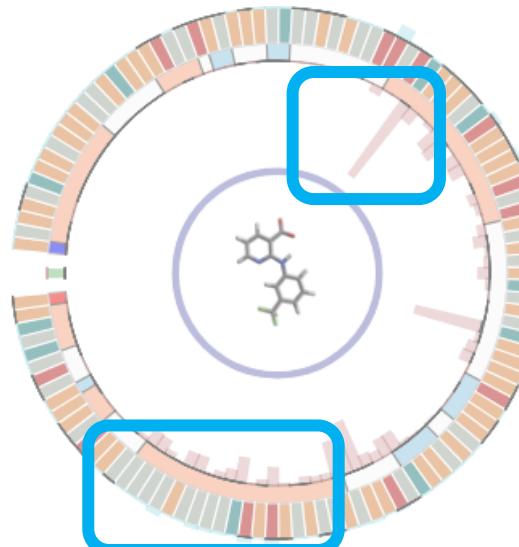


$\geq 1.5 \text{ \AA}$
 $\geq 0.75 \text{ \AA} \text{ and } < 1.5 \text{ \AA}$
 $< 0.75 \text{ \AA}$

▶ | □ | Frame 0



Binding analysis



Domain experts' input

- Informal user study
 - Practical session (20-30 minutes) + questionnaire
 - Experts were interested to use it in their work
 - Helps them to better understand the overall behavior of a simulation, and to infer h-bonds formation.
 - Particularly useful to summarize trajectories
 - Features not available in known software
- Extra features
 - Would add some more simulation properties and a coupled 3D view

Conclusions

- Compact way to display MS information
 - Removing 3D structure
 - But backbone configuration
 - Analysis of whole MS trajectory
 - On-site comparison of different simulations

Future Work

- Add multiple simulations
 - Maybe large screen
 - Add 3D view on demand



Visual Analysis of protein-ligand interactions

P. Vázquez¹, P. Hermosilla², J. Estrada³, V. Guallar³, A. Vinacua¹

Acknowledgments

Spanish Ministry of Economy and Competitividad under Grants TIN2014- 52211-C2-1-R and TIN2017-88515-C2-1-R (GEN3DLIVE), by CTQ2016-79138-R, and by DFG under grants RO 3408/2-1 and RO 3408/3-1.

Also thanks to my son Marçal.