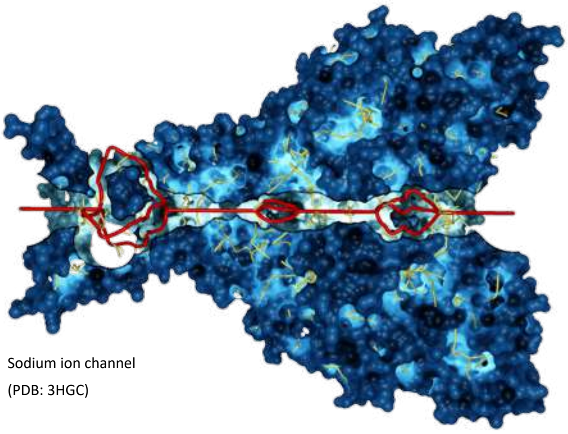
Visual Analytics of Biomolecular Systems



Sodium ion channel
(PDB: 3HGC)

Hans-Christian Hege

10th Annual SeRC Meeting, Söderköping, May 16, 2019

Motivation

- *Visualization* — Long tradition in molecular visualization
  - Advanced visualization & interaction techniques developed
  - However: lack of insight into
    - the molecular sciences
    - the practical needs in molecular sciences
- *Molecular Science* — Researchers are building visualization tools *on their own*
  - Tools grow in practice; satisfy many practical needs
  - However:
    - software is limited to a certain complexity level
    - knowledge transfer from vis research is delayed

➔ intensify communication & cooperation between the fields

SeRC is the perfect place for this!

Overview

- Data Visualization / Visual Analytics
  - Visualization, Reasoning and Knowledge Advancement
  - Our Mindset, Objectives and Approach
- Advances in Interactive Biomolecular Visualization
  - Visualization of metastable conformations
  - Molecular Surfaces: Ligand accessible surface
  - Analysis of dynamic channels in biomolecular systems
  - Multi-scale visualization: bridging the molecular & biological world
- Future Challenges

3

Data Visualization / Visual Analytics

0.2098471	1.0287235	-7.0923847
0.2340987	1.4209827	-8.5098723
0.2374923	1.6809283	-6.2570975
0.3981029	1.7092871	-7.1409287
0.2482091	1.5081965	-9.2408712
0.3891032	1.6098725	-6.2864081
0.3021934	1.9208027	-5.7269401
0.4809381	1.4270982	-5.9081726
0.5860298	1.4075263	-8.4209817
0.6892071	1.6082671	-4.2408971
0.6109827	1.5789430	-5.2240989
0.7301928	1.6398569	-2.2408574
0.6829301	1.5432788	-8.9724131
0.7109813	1.5987612	-2.4089746
...	...	...

Visualization, Reasoning and Knowledge Advancement

How does visualization work?

Why is visualization so powerful?

Human Visual System

- Humans perceive the world via various senses; major sense: visual sense
- Humans evolved as visual creatures
  - 70% of our sensory receptors are in the eyes
  - Nearly 50% of our brain is involved in visual processing
  - 90% of information transmitted to our brain is visual
  - We can recall 80% of what we see (and do)
  - A highly capable pattern recognition apparatus included in the visual system
  - Fast: 150 ms to process a visual symbol + 100 ms to give it a meaning

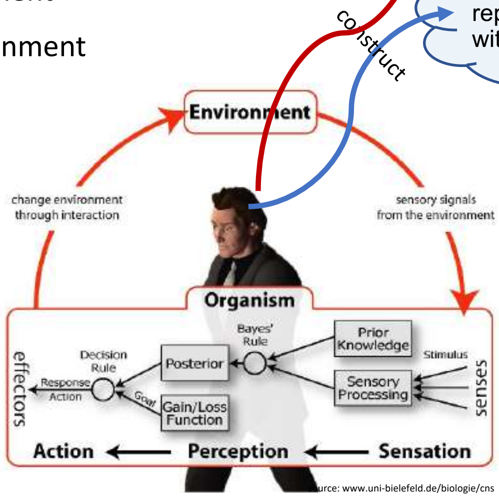
Sensation-Perception-Action Loop

mental model of the world

All human needs & activities rely on two basic ingredients

reliable percept, representing the world with which we interact

- Perception of the environment
- Interaction with the environment



Herrmann von Helmholtz:
human perception is
a problem of inference

Understanding & Reasoning

From perceived information we form

mental images & sets of rules

- Bringing facts of the world into meaningful relationships
- Organizing facts mentally

source: http://angelabrook.com

Understanding

- Kind of compression of mental images & sets of rules
- "Understanding X" = being able to figure out
  a simple mental image or a simple set of rules that "explains" X

Reasoning

- Visualizing mentally and applying rule sets mentally

Why is Visualization so Powerful?

External images encode information – in the first place *spatially* – that can be looked up very quickly if needed for reasoning.

External images facilitate

- Understanding by supporting the formation of mental images
- Reasoning by extending the limited capacity of working memory

The use of external images is evolutionarily supported in 2 ways:

- In Human Perception: pictorial transmission & visual processing is the most effective way of evaluating external information
- In Human Cognition: spatial / visual reasoning is one of the strongest cognitive abilities

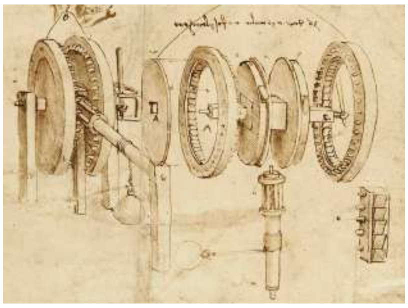
Types of Computer-Based Visualization

1. Pictorial Representation of Imagined Scenes

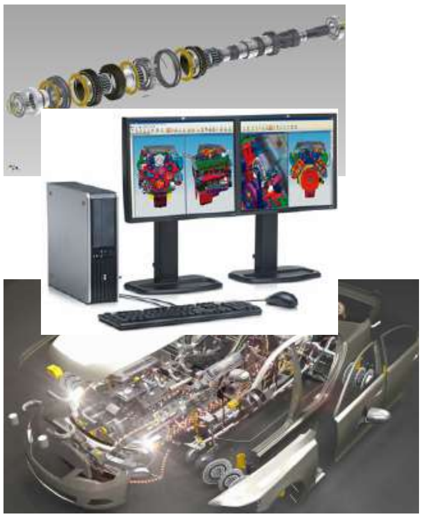
Leonardo da Vinci,
Study of a Tuscan Landscape, Valley of the Arno,
1473, pen and ink

Epic Games,
Study on physically-based rendering,
2015, Unreal Engine 4

2. Pictorial Representation of Imagined Constructions



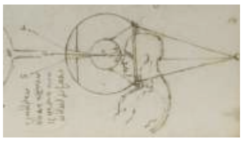
Leonardo da Vinci,
Transformation of alternating to continuous motion
Codex Atlanticus f.8v-b, 1504/5
pen and ink



Moh.Visuals, Inc
Vehicle exploded view animation
Computer animation, 2011

3. Pictorial Representation of Physical Information

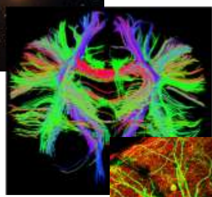
Roger Bacon, 1250-60
Spherical water lens



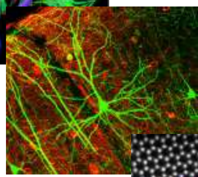
Leonardo da Vinci, 1490-95
The eye as camera obscura

Galilei Galilei, 1609
Refracting telescope

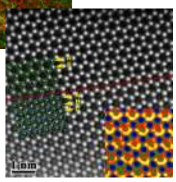
Stephan's quintett;
group of five galaxies,
Hubble telescope,
ERA Team, 2000



Neuronal pathways
in human brain,
DTI,
Alexander AL et al, 2007

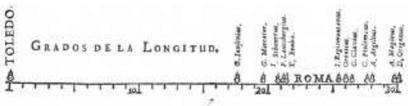


Pyramidal neuron
from the hippocampus,
CFM
Lee WCA, et al, 2005

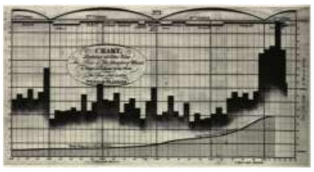


Atomic structure of
zinc-antimony nanowires,
TEM,
Shahbazian-Yassar R et al, 2014

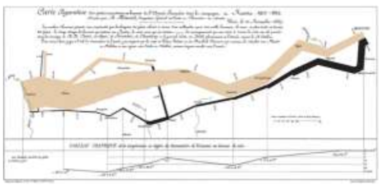
4. Pictorial Representation of Digital Information



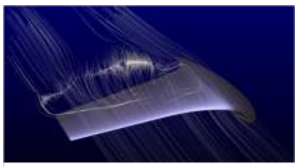
M.F. Van Langren, 1664, Estimates of longitudes



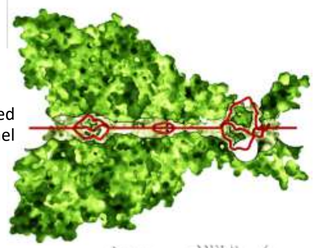
W. Playfair, 1821, price of wheat, wages,
reigning monarch between 1565-1820



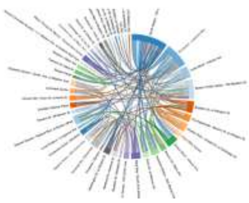
C. Minard, 1869, no. of men in Napoleon's
1812 army; army location; temperatures



Simulated flow
over wing



Simulated
ion channel



Relations
between sets

Reconstructed
antique site

App for museum
information

Our Mindset, Objectives and Approach

Data Visualization

Starting point: Data, which represent certain aspects of an object or a phenomenon

Task: From the data,

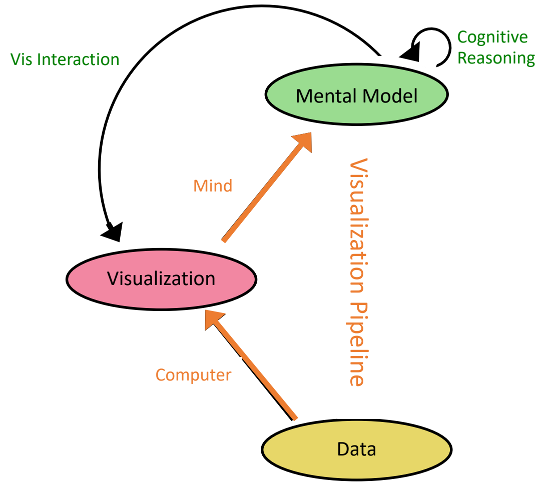
create external visualizations

such that the

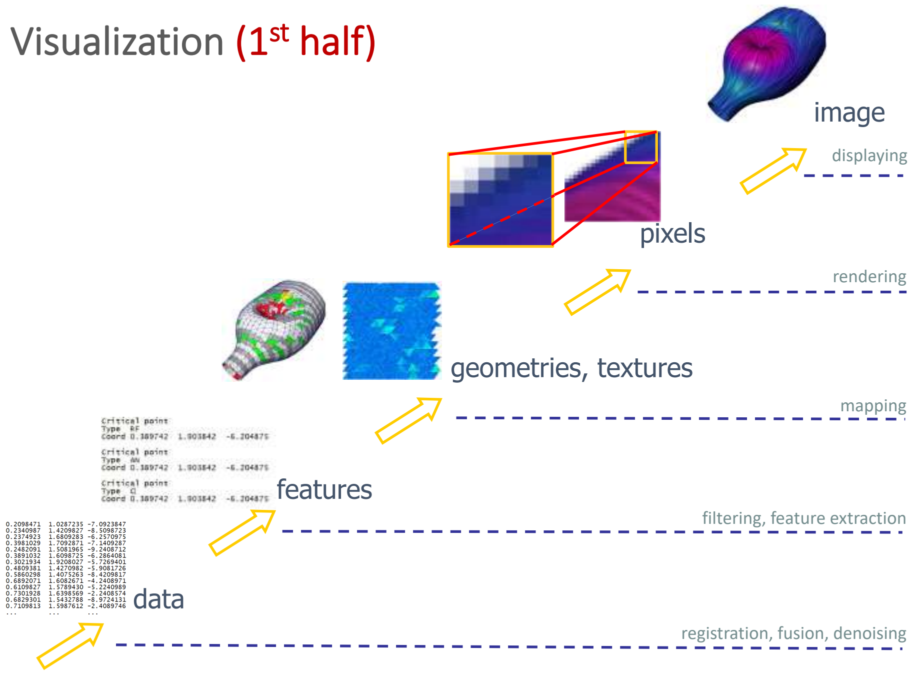
resulting internal visualizations

convey insight.

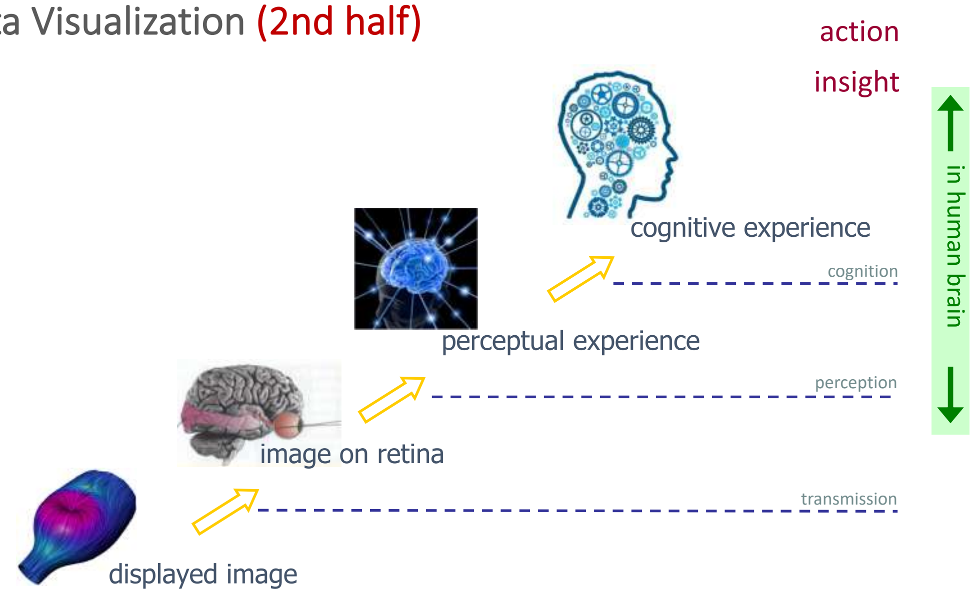
Conceptual Core of Data Visualization



Data Visualization (1st half)

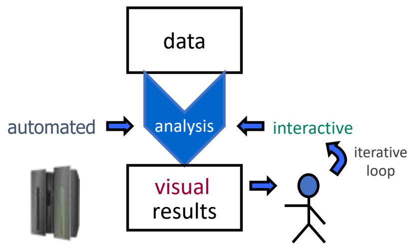


Data Visualization (2nd half)



Visual Analytics

Aim: Combine human strengths with automated data analysis



- Quickly detect complex patterns in data and relate them to existing knowledge
- Analyze data quantitatively
- Iteratively improve analysis & understanding

Analysis is boosted by interactive techniques

Data Visualization: Place Information Visibly in Space

Space is nothing but a […] set of relations among bodies […].

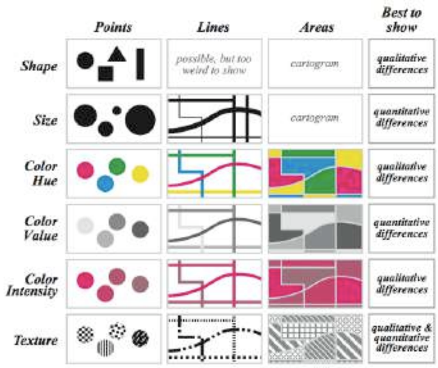
Gottfried Wilhelm Leibniz

1716, in a letter to Samuel Clarke

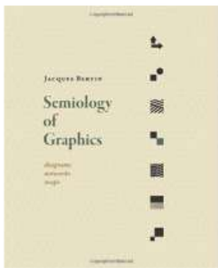
[Data] Visualization is made by a series of basic components that are determined by 6 basic variables …

Jaques Bertin

1967, Sémiologie Graphique. Les diagrammes, les réseaux, les cartes



Visual Design



- J. Bertin: Proposal for effective visualization by scale-dependent use of variables (1967)
- Rediscovered by computer scientists (1991 for GUI Design, 1994 for data visualization)
- Today
  - Visual design – more and more on a rational basis
  - Empirically evaluated by user studies
  - Considering cognitive aspects – sometimes
- Future
  - Theory of design?
  - Simulate the user?

Data Visualization – Major Ingredients – Past, Present, Future

• Time sharing computers	~ 50's
• I/O devices	~ 60's
• Interactive computer graphics	~ 70's
• Graphics hardware	~ 80's
• Virtual & augmented reality	~ 90's
• Advanced visualization techniques	~ 00's
• Visual Analytics	~ 10's
• VA + ML, Cognitive Computing	~ 20's

How We Typically Proceed

• Requirement analysis	types of data ?, types of "features" ?, analysis tasks ?
• Features	formal definition + extraction algorithm
• Visual design	using a large body of knowledge
• Implementation	analysis + graphics
• Practical testing	
• Evaluation	technical; perceptual & cognitive
• Publication	

Visualization: When is it Helpful?

When

- Aiming at communication of facts (to laypeople and professionals)
- Seeking overview, no definite questions a-priori
- Aiming at consistency checking: internal, mental model ⟷ real model
- Data too complex for automatic processing / analysis
- Programming effort too high for automatic processing / analysis
- Data features are qualitative, not yet quantified/quantifiable
- Aiming at discovery of unknown facts, relations, or correlations (including errors!)
- Aiming at creation of meaningful hypotheses

Advances in Interactive Biomolecular Visualization

Thanks to:

Norbert Lindow

Daniel Baum

Nicoleta Bondar

27

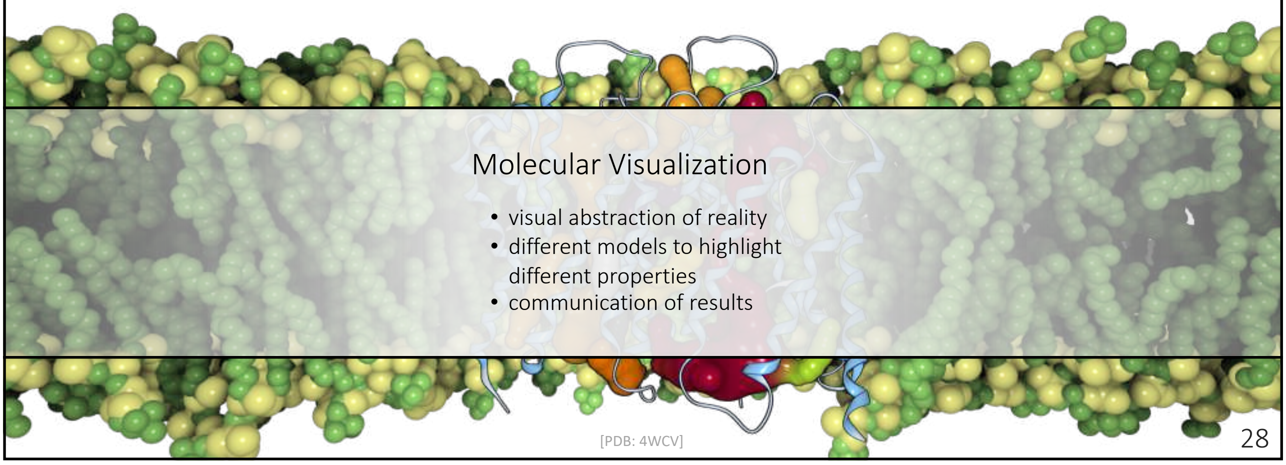
Understanding the Biomolecular World

Experimental Structure Determination

- X-ray crystallography
- nuclear magnetic resonance spectroscopy (NMR)
- cryo-electron microscopy

Simulation of Dynamics

- NAMD
- Gromacs
- AMBER



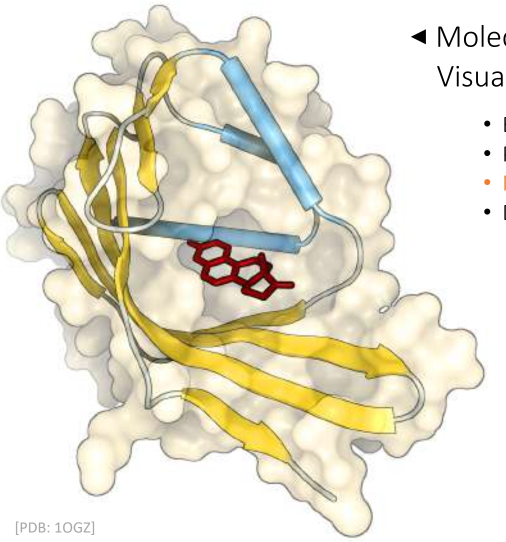
Molecular Visualization

- visual abstraction of reality
- different models to highlight different properties
- communication of results

[PDB: 4WCV]

28

Biomolecular Visualization



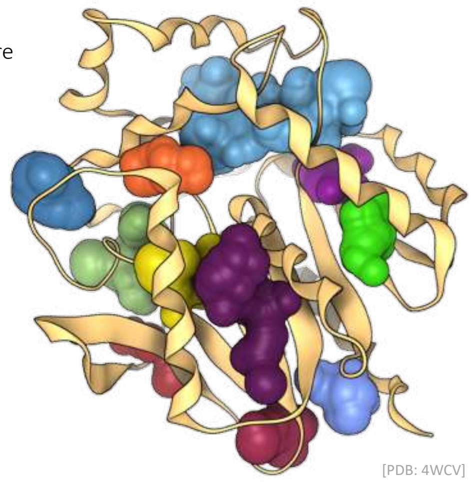
[PDB: 1OGZ]

◄ Molecular Structure Visualization

- Ball-and-stick
- Protein Secondary structure
- Molecular surfaces
- DNA & RNA

Molecular Cavity Visualization ►

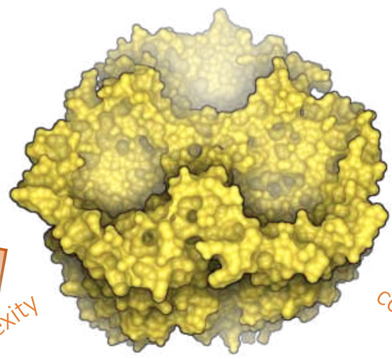
- Definition & Classification
- Computation
- Visualization
- Dynamics analysis



[PDB: 4WCV]

29

Molecular Structure Visualization



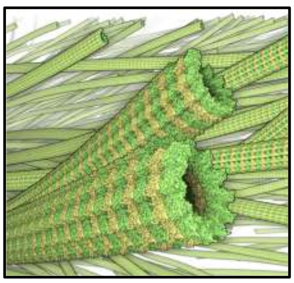
[Lindow et al. 2010]

performance

complexity

performance

complexity



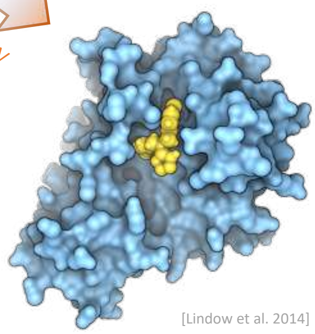
[Lindow et al. 2012]

◄ Van der Waals Surface / Solvent Excluded Surface / Ligand Excluded Surface ►

Models the van der Waals surface

Shows the accessibility according to a probe sphere's radius

Shows the accessibility according to the size of a ligand



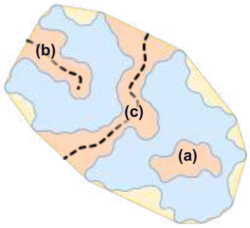
[Lindow et al. 2014]

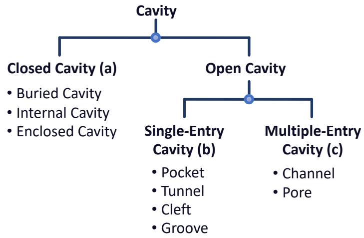
30

Molecular Cavity Visualization

Definition & Classification

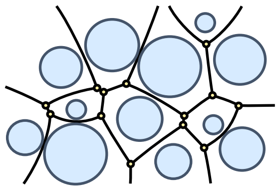
[Krone et al. 2016]

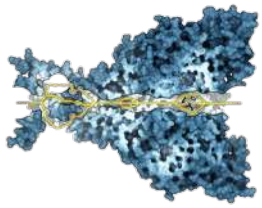


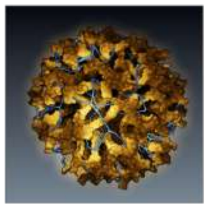


Computation & Visualization

[Lindow et al. 2011]

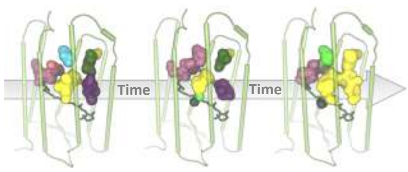


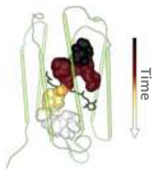


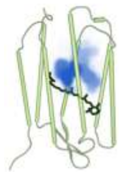


Tracing & Analysis

[Lindow et al. 2012, 2013]







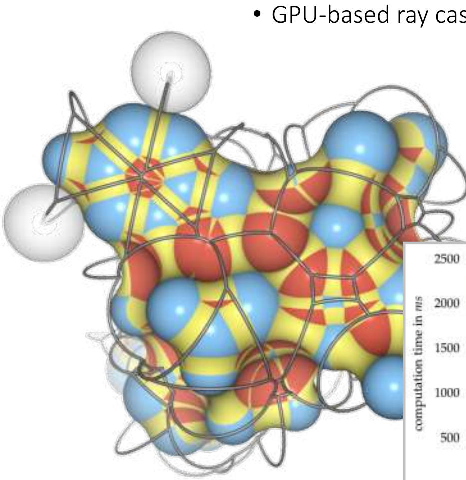
31

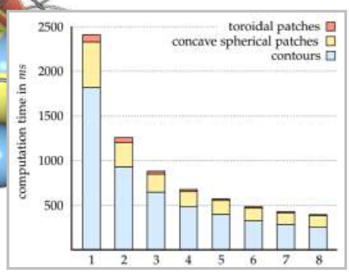
Solvent Excluded Surface

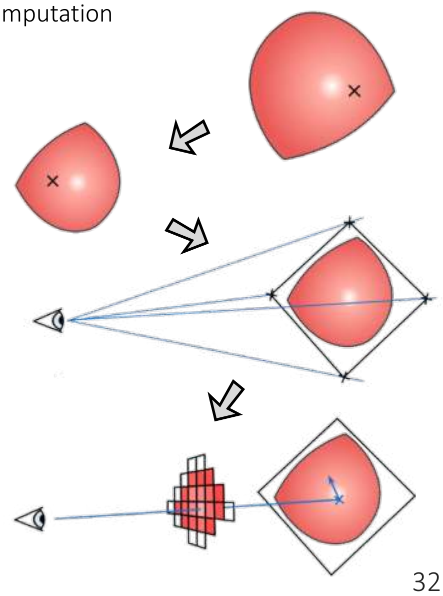
[Lindow et al. 2010]

Problem: Real-time visualization of molecular surfaces for time-dependent data

Contributions:
- Fast detection of local neighborhood for patch computation
- Parallel computation of the surface description
- GPU-based ray casting per patch





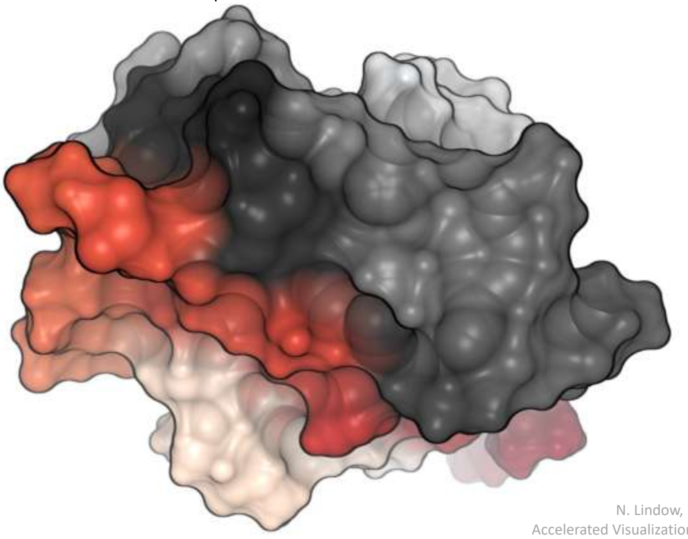


[PDB: 1UGT]

32

Solvent Excluded Surface

Real-time visualization of molecular surfaces for proteins



[PDB: 1UGT]
[Data: B. Kallies]

N. Lindow, D. Baum, S. Prohaska, H.-C. Hege
Accelerated Visualization of Dynamic Molecular Surfaces
Comput. Graph. Forum, (2010)

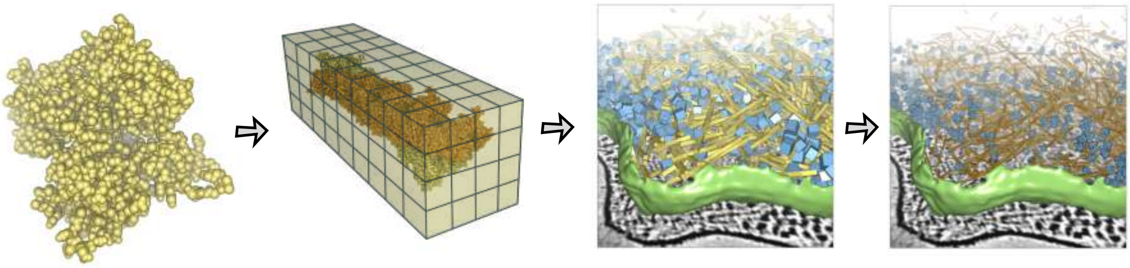
33

Van der Waals Surface

[Lindow et al. 2012]

Problem: How to render biological structures with billions of atoms

Method:



create recurrent components, each consisting of multiple proteins

render boxes of many instances

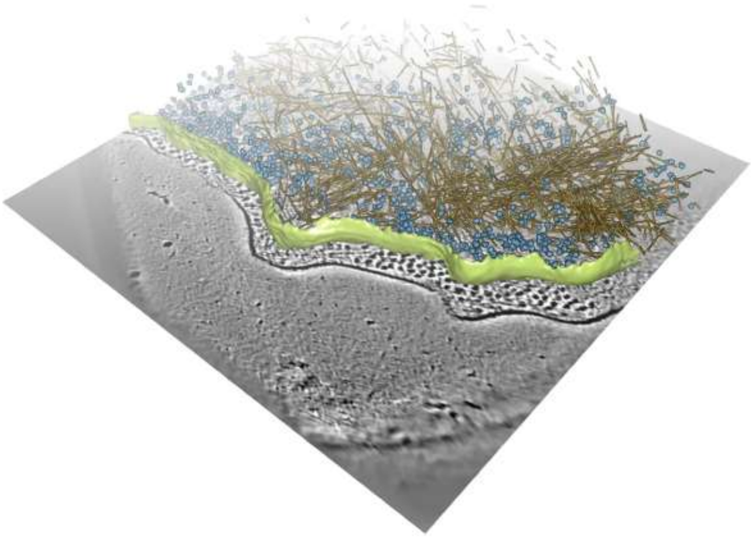
ray casting within each box

Contributions:
- A rendering technique that allows bridging 5 orders of magnitude in length scale without explicit LoD
- Deferred shading with a three-step normal computation to avoid artifacts and to improve the structure perception

34

Van der Waals Surface

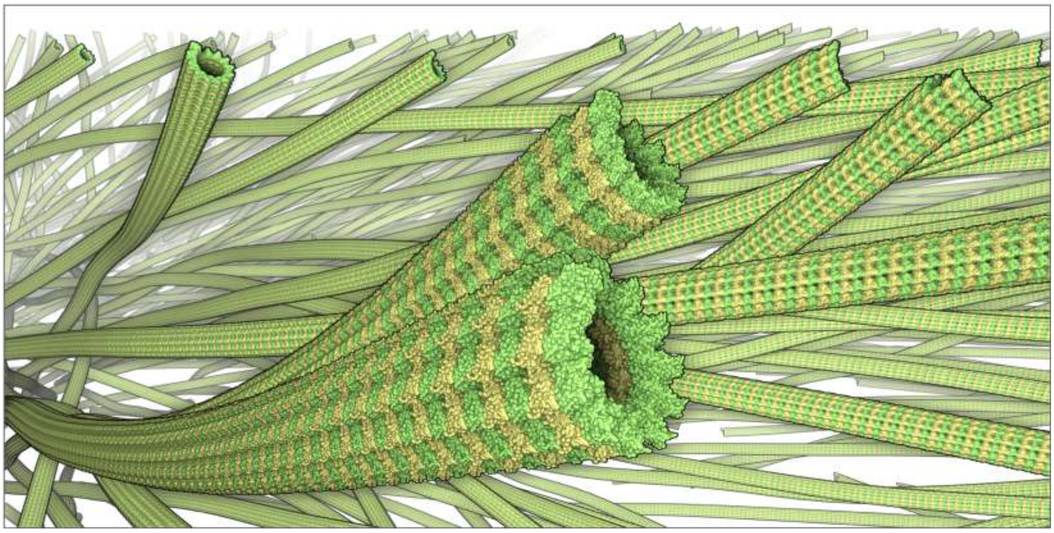
[Lindow et al. 2012]



[Data: A. Rigort, W. Baumeister, Reconstruction: D. Günther, A. Rigort] 35

Van der Waals Surface

[Lindow et al. 2012]



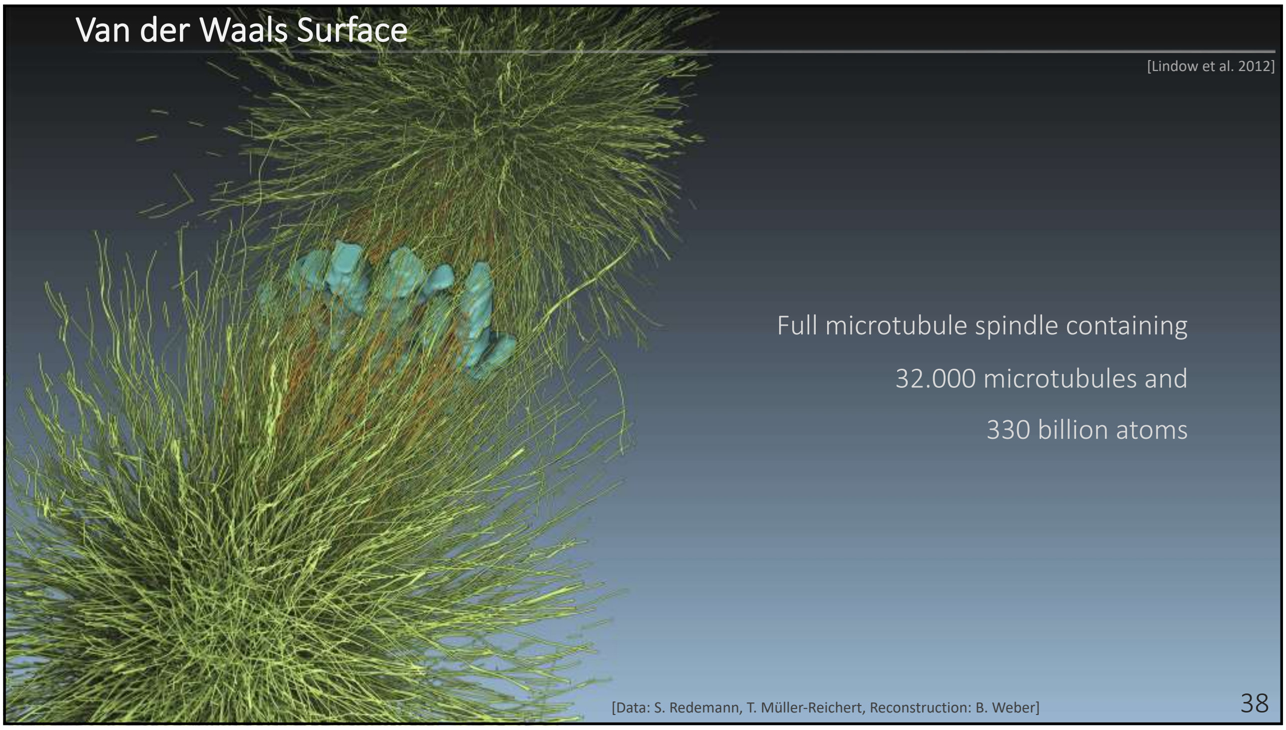
[Data: S. Redemann, T. Müller-Reichert, Reconstruction: B. Weber]

Microtubules with 10 billion atoms at 3 fps

37

Van der Waals Surface

[Lindow et al. 2012]



Full microtubule spindle containing
32.000 microtubules and
330 billion atoms

[Data: S. Redemann, T. Müller-Reichert, Reconstruction: B. Weber]

38

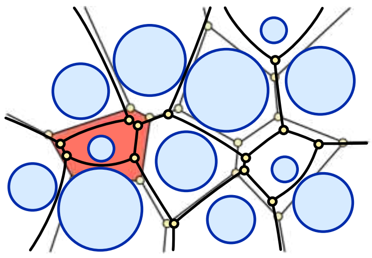
Molecular Cavity Computation

[Lindow et al. 2011]

Problem: Molecular cavity computation with high geometrical accuracy

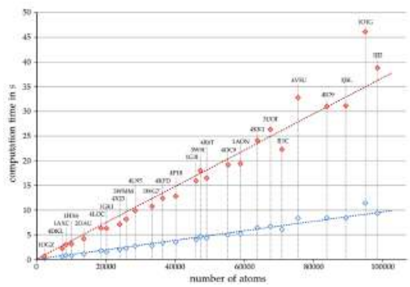
Method: Computation of Voronoi diagrams of atom spheres for geometrically optimal paths

$$V_i = \left\{ p \in \mathbb{R}^3 \;\middle|\; \|p - p_i\| \leq \|p - p_j\|,\ \forall j \in \{1, \ldots, m\},\ j \neq i \right\}$$



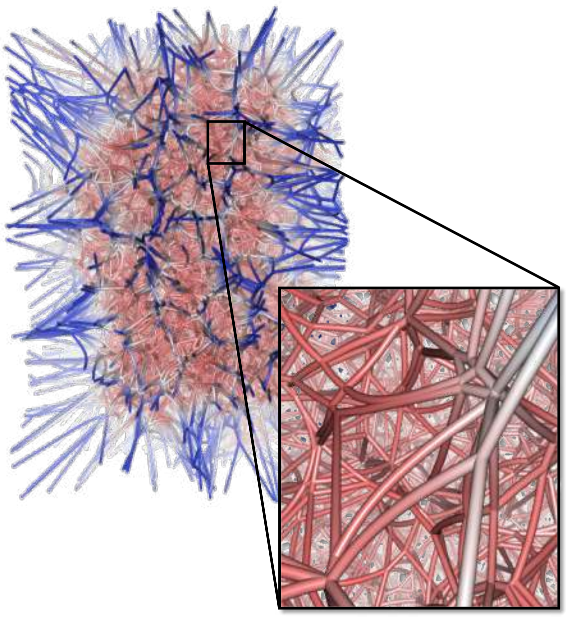
Contributions:

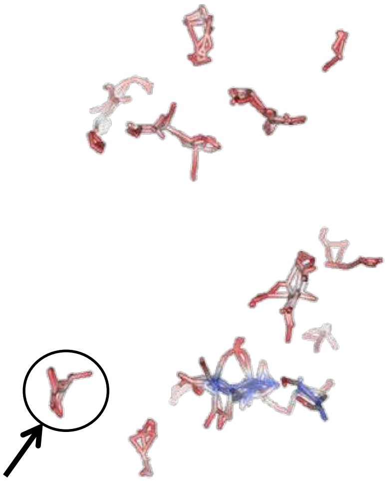
- New algorithms that scales linearly for molecular data
- It is more than twice as fast as the method by Manak et al. (2010)
- Efficient filtering to get the skeleton of all cavities



39

Molecular Cavity Computation





40

Molecular Path & Cavity Visualization

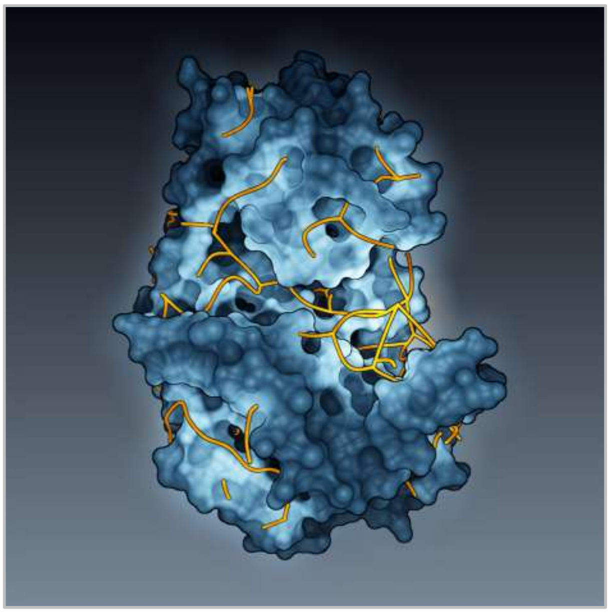
Problem:

Visualization of molecular structure in combination with paths or cavities

Contribution:

A novel path illumination technique using many small point lights

- direct illumination
- ambient occlusion
- light placement
- glow

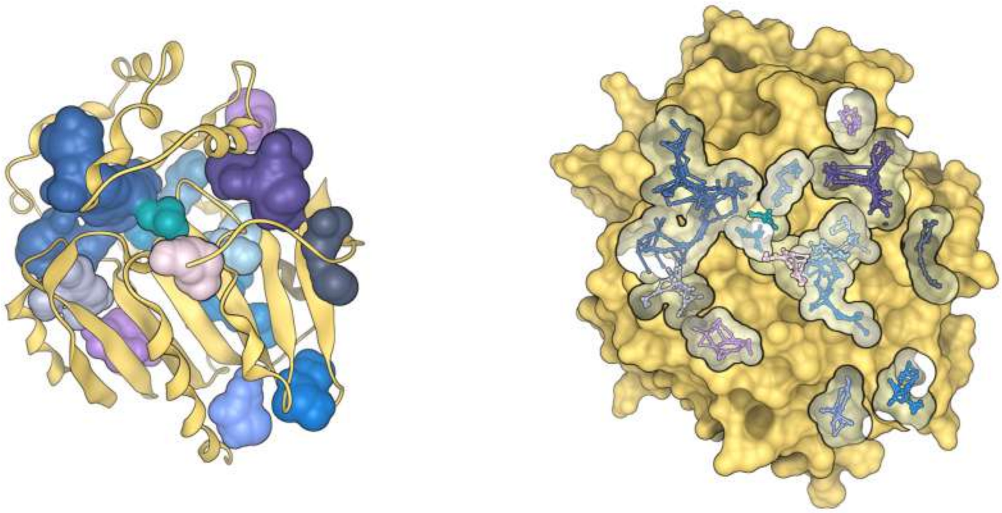


[PDB: 5PEP]

41

Molecular Path & Cavity Visualization

Contributions:
- Cavity visualization using the skin surface
- A view-dependent molecular surface clipping



[PDB: 4WCV]

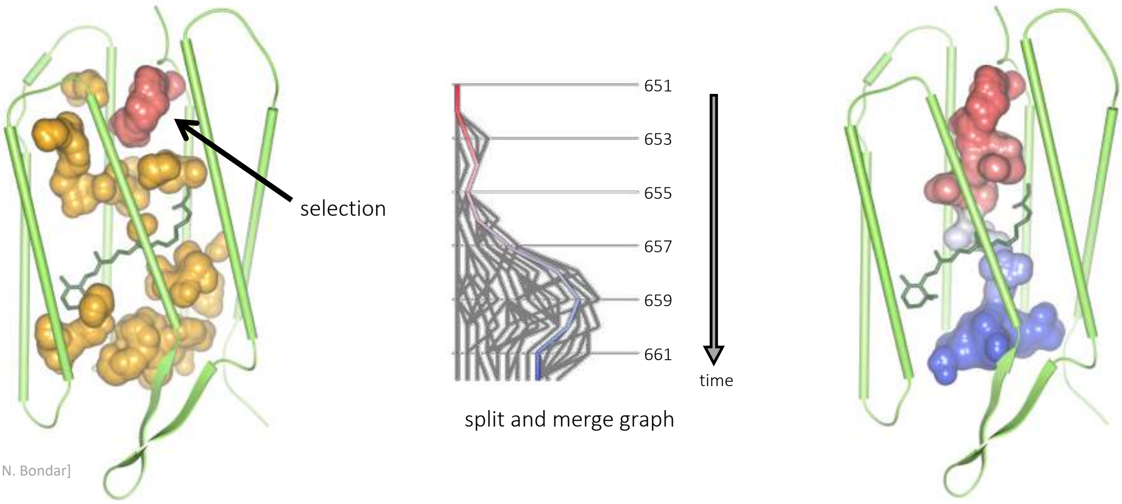
42

Cavity Dynamics

[Lindow et al. 2012, 2013]

Problem: How to trace cavities over time and how to investigate their dynamics

Method:
- Computation of cavities for each time step
- Interactive tracing of cavities over time
- Selection and visualization of dynamic cavities



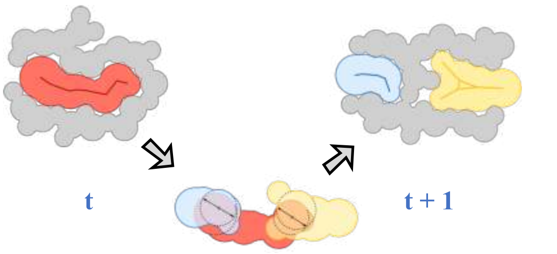
[Data: N. Bondar]

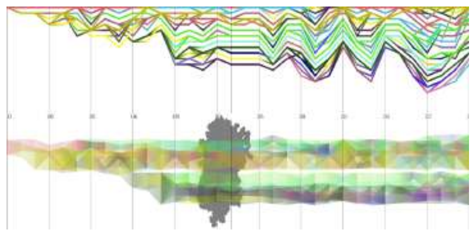
44

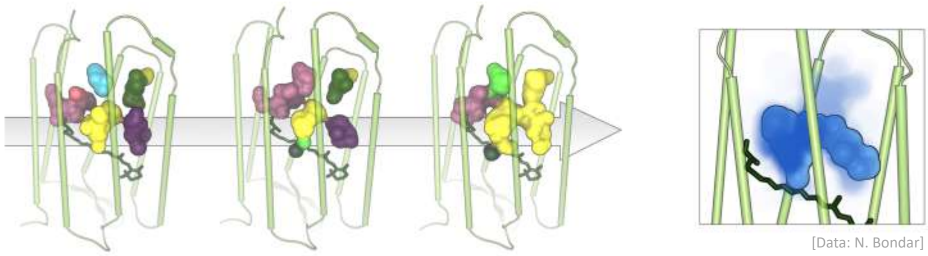
Cavity Dynamics

[Lindow et al. 2012, 2013]

Contributions:
- Fast overlap approximation for interactive tracing
- Graph visualization for cavity dynamics
- 3D dynamics visualizations







[Data: N. Bondar]

45

Ligand Excluded Surface

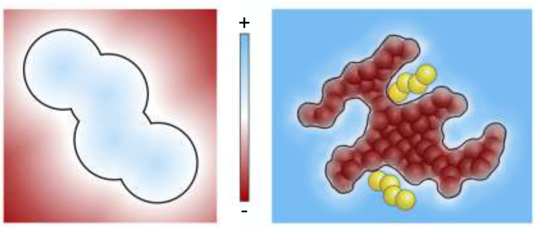
[Lindow et al. 2014]

Problem: The shape of most ligands significantly differs from a sphere

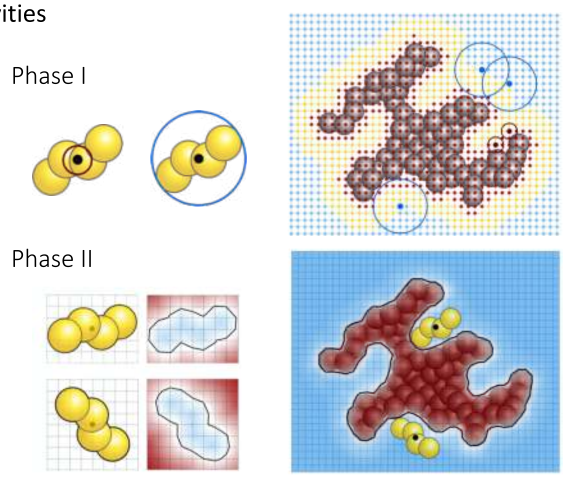
Contributions:
- Definition of a new molecular surface that considers the ligand's geometry and dynamics
- Development of an efficient discrete algorithm to compute the surface and the complementary cavities

$$I_{r,l}(k,T,R) = \begin{cases} 1, & \left\| p_i^r - (Rp_{jk}^l + T) \right\| \geq r_i^r + r_j^l, \forall i = 1,\dots,n, \forall j = 1,\dots,m \\ 0, & otherwise. \end{cases}$$

$$d_{r,l}(p) = \max_{\substack{k=1\dots c, \\ T\in\mathbb{R}^3, R\in SO^3}} \begin{cases} \max_{j=1\dots m} r_j^l - \left\| p - (Rp_{jk}^l + T) \right\|, & I_{r,l}(k,T,R) = 1 \\ -\infty, & otherwise. \end{cases}$$



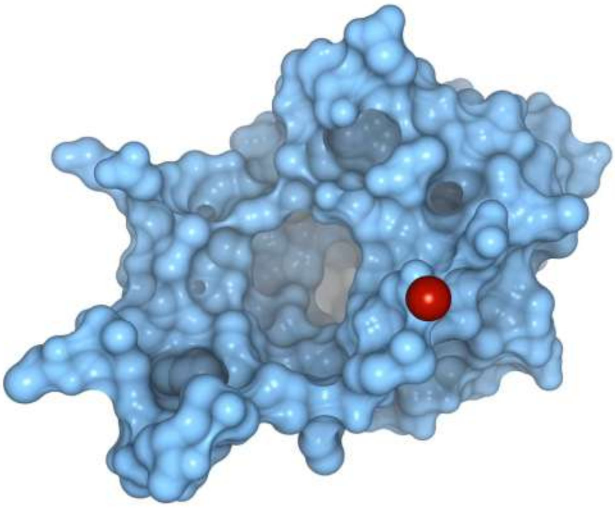
Implicit surface definition of distance function



46

Ligand Excluded Surface

[Lindow et al. 2014]



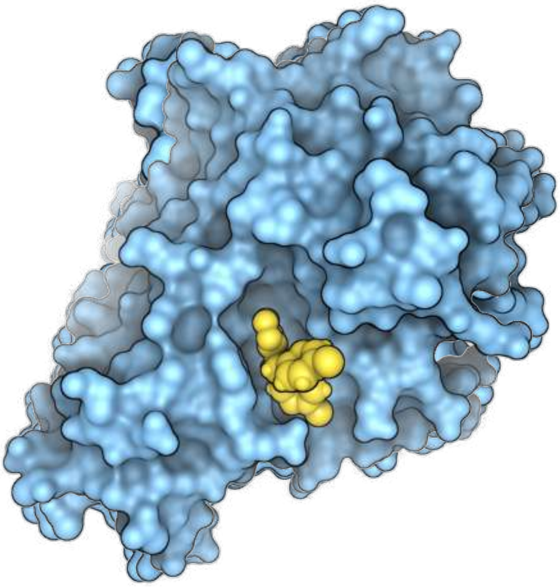
47

Ligand Excluded Surface

[Lindow et al. 2014]

Outlook:

Connection of valid ligand states over time



[Data: G. Bowman]

48

Future Challenges

Outlook

Within the classical model:

- Depict more physics (depending on context)
- Tools for interactive specification of start conditions (simulation set up)
- Use atomic accessibility radii (reflecting the accessibility of atoms)
- Multi-scale approaches for visualization and cavity analysis
- Further accelerations and visual improvements

Considering quantum effects:

- Orbital visualization (⇒ Orbkit)
- Depict electron / nuclear densities / fluxes (equilibrium geometries, reaction energetics)
- Multielectron wave functions (?)

State of the Art Reports

B Kozlíková, M Krone, M Falk, N Lindow, M Baaden, D Baum, I Viola, J Parulek and H-C Hege.
Visualization of biomolecular structures: State of the art revisited.
Comput. Graph. Forum, 36:8, pp. 178-204, 2017. DOI: 10.1111/cgf.13072

M Krone, B Kozlíková, N Lindow, M Baaden, D Baum, J Parulek, H-C Hege and I Viola.
Visual analysis of biomolecular cavities: State of the art.
Comput. Graph. Forum, 35:3, pp. 527-551, 2016. DOI: 10.1111/cgf.12928

Thank YOU !