

# Voronoi diagrams for molecules

Dr. Klement Olechnovič

Vilnius University Life Sciences Center

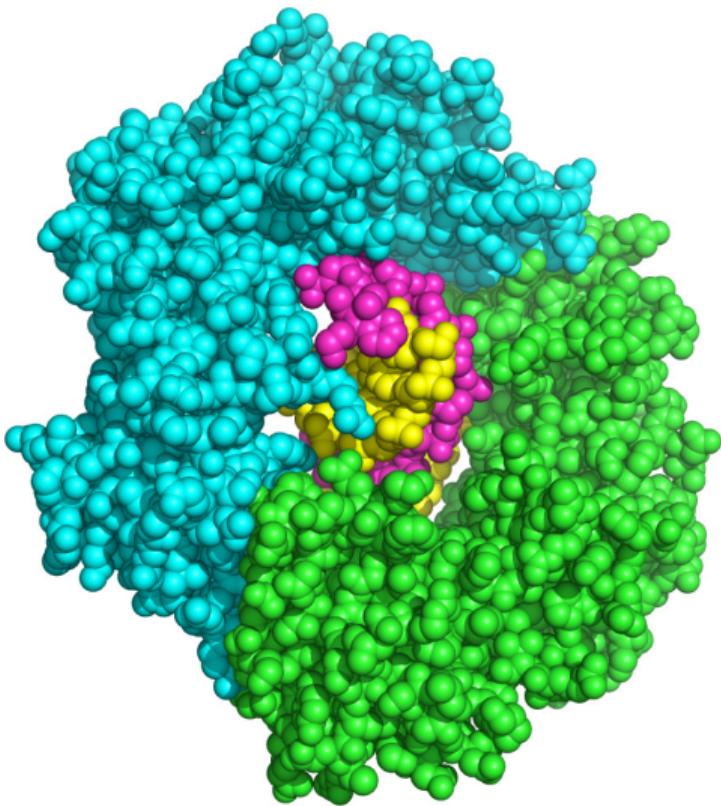
2025-09-16



Life Sciences  
Center



Institute of  
Biotechnology



When studying biological macromolecules, some common problems are:

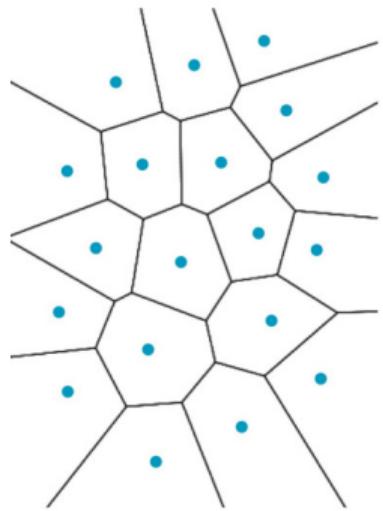
- ▶ analyzing how different parts in a molecule interact
- ▶ selecting the best prediction of a multimeric complex

Some possible solutions involve:

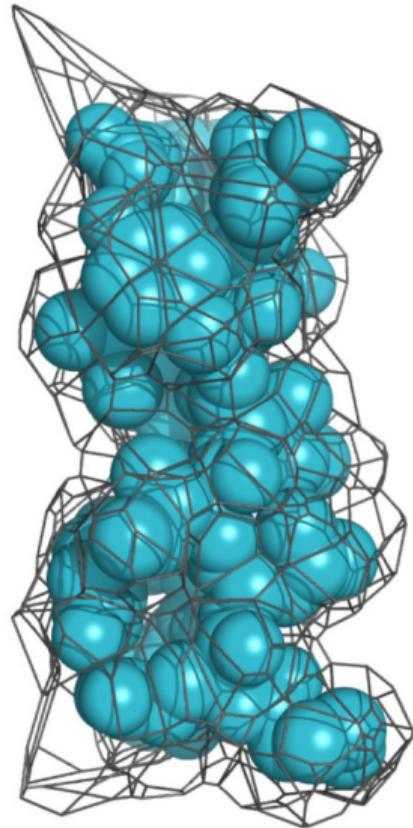
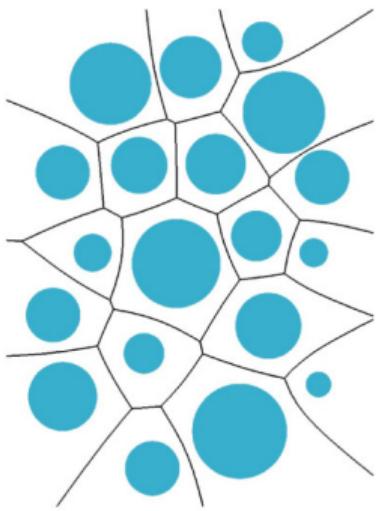
- ▶ computational geometry
- ▶ machine learning using **open data**
- ▶ developing free software

# Voronoi diagram of points and balls

"Classic" Voronoi diagram  
of points

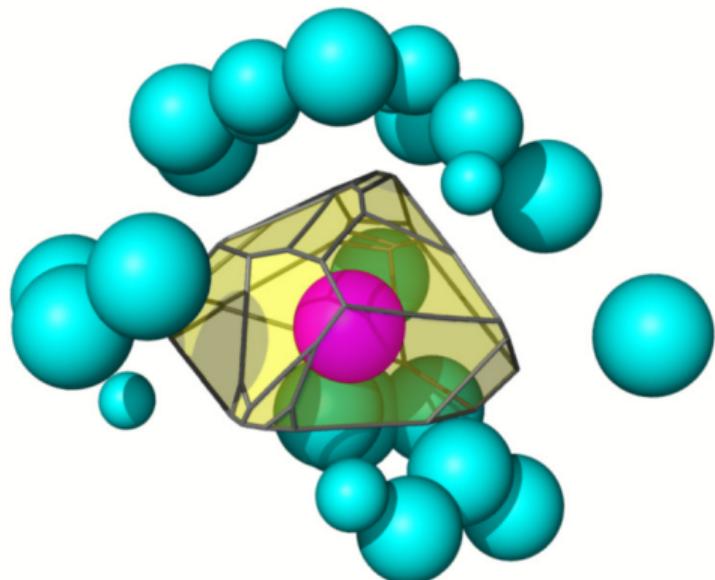


Voronoi diagram  
of balls

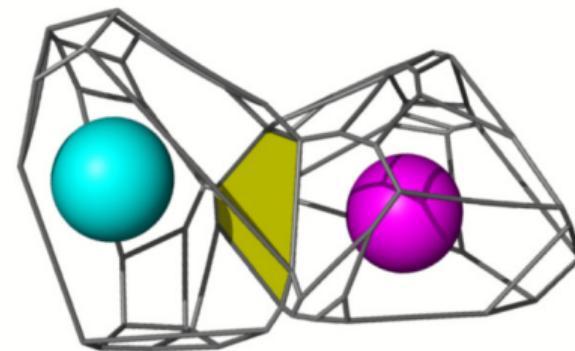


# Voronoi tessellation-based analysis of structures

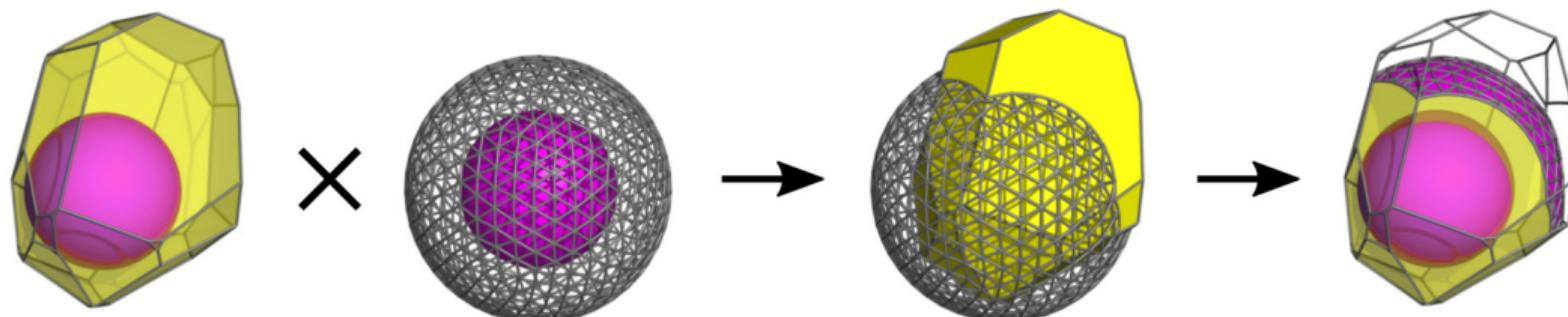
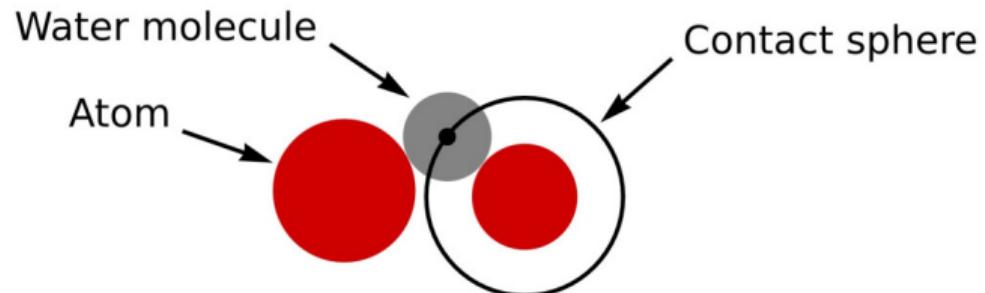
Voronoi cell of an atom surrounded by its neighbors



Atom-atom contact surface defined as the face shared by two adjacent Voronoi cells.

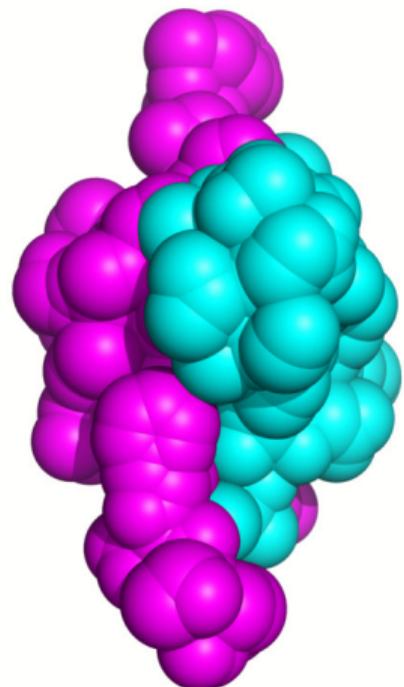


## Constrained contacts

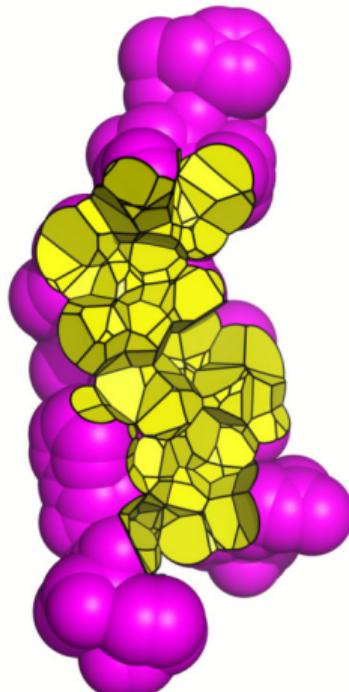


# Inter-chain contacts

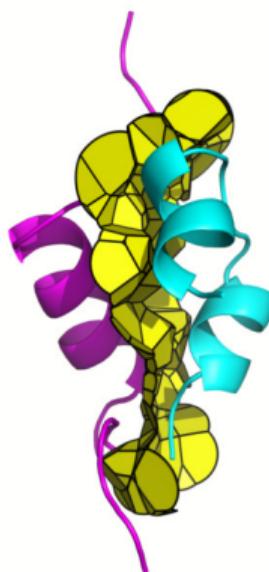
Solvent-accessible surface  
of an insulin heterodimer  
PDB:4UNG colored by subunit



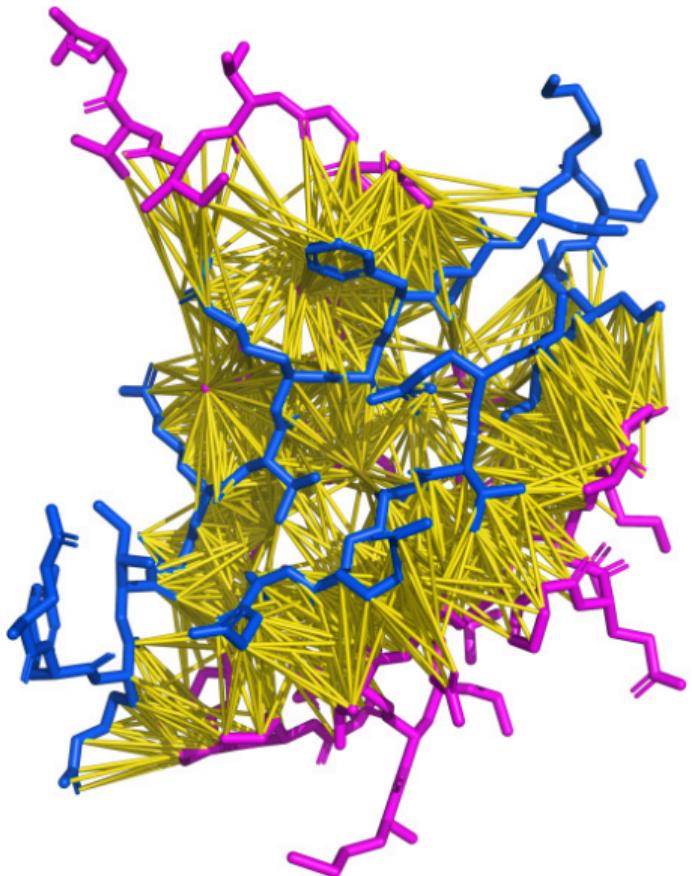
The intersubunit interface  
shown together with the  
SAS of one subunit



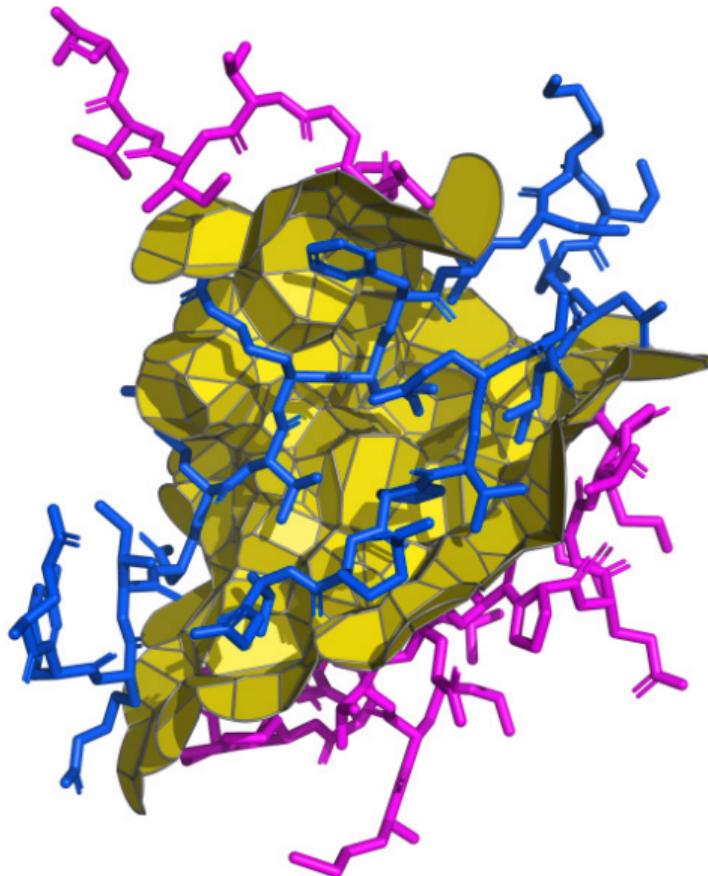
The intersubunit interface  
shown together with  
both subunits represented  
as cartoons



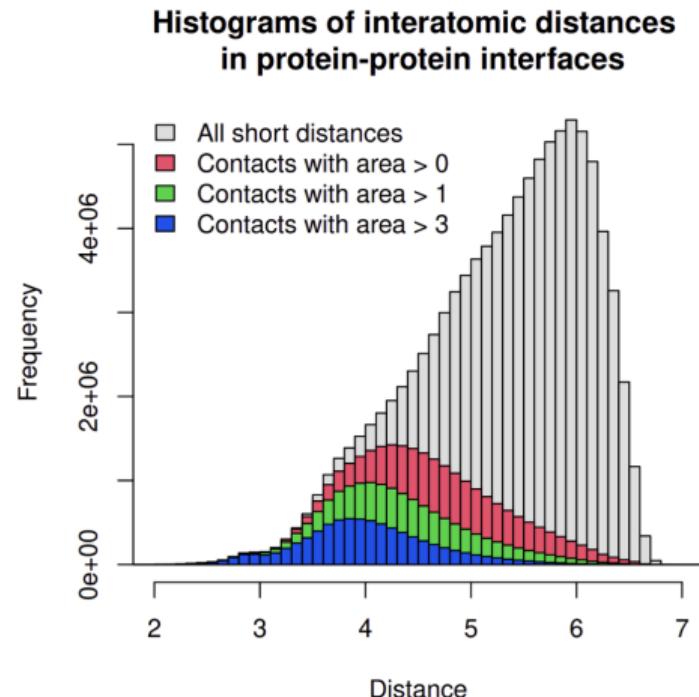
## Inter-chain contact areas vs distances



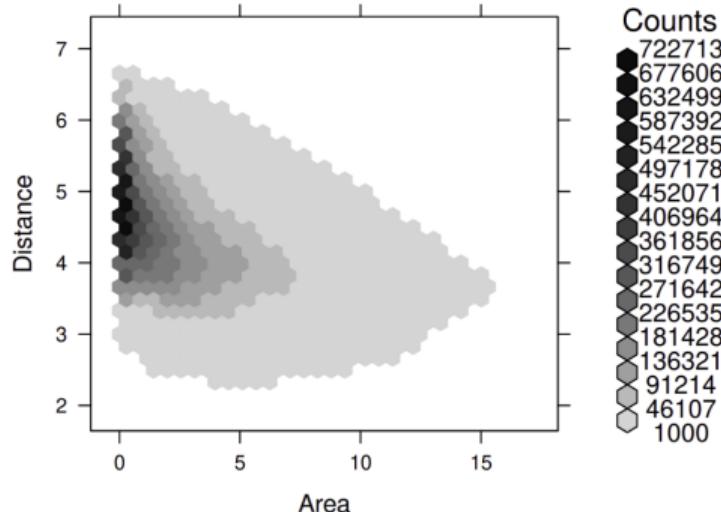
VS



# Inter-chain contact areas vs distances, PDB-based statistics



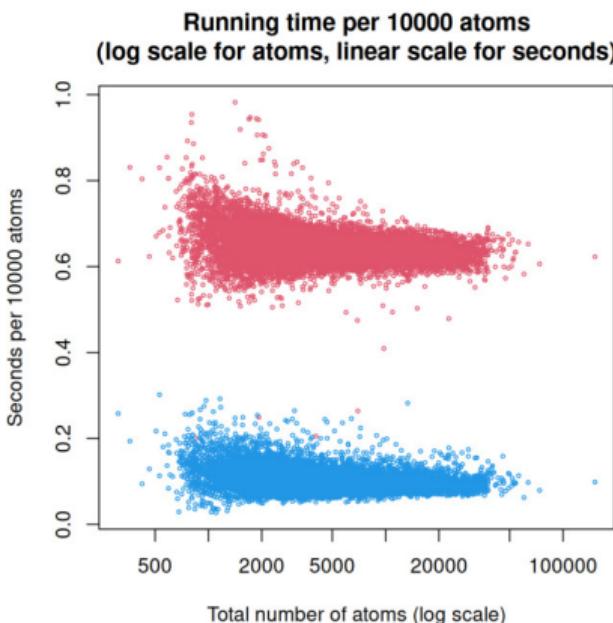
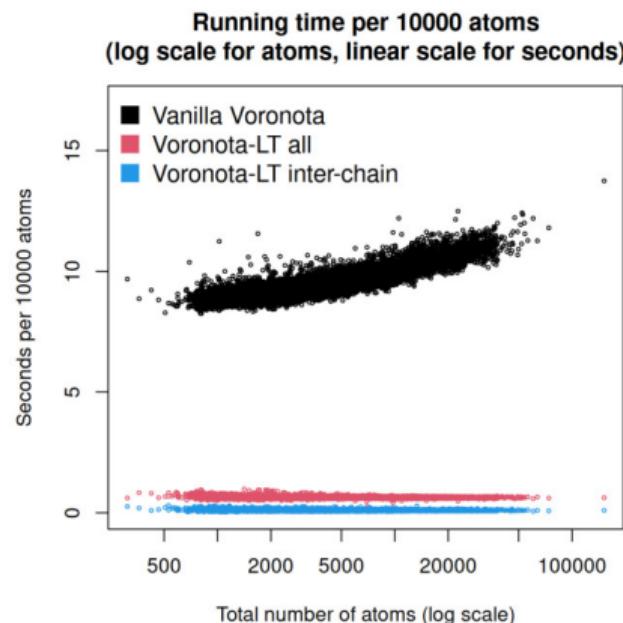
**Tessellation-derived contact areas vs interatomic distances in protein-protein interfaces**



$$\text{corr}(\text{area}, \text{distance}) \approx -0.43$$

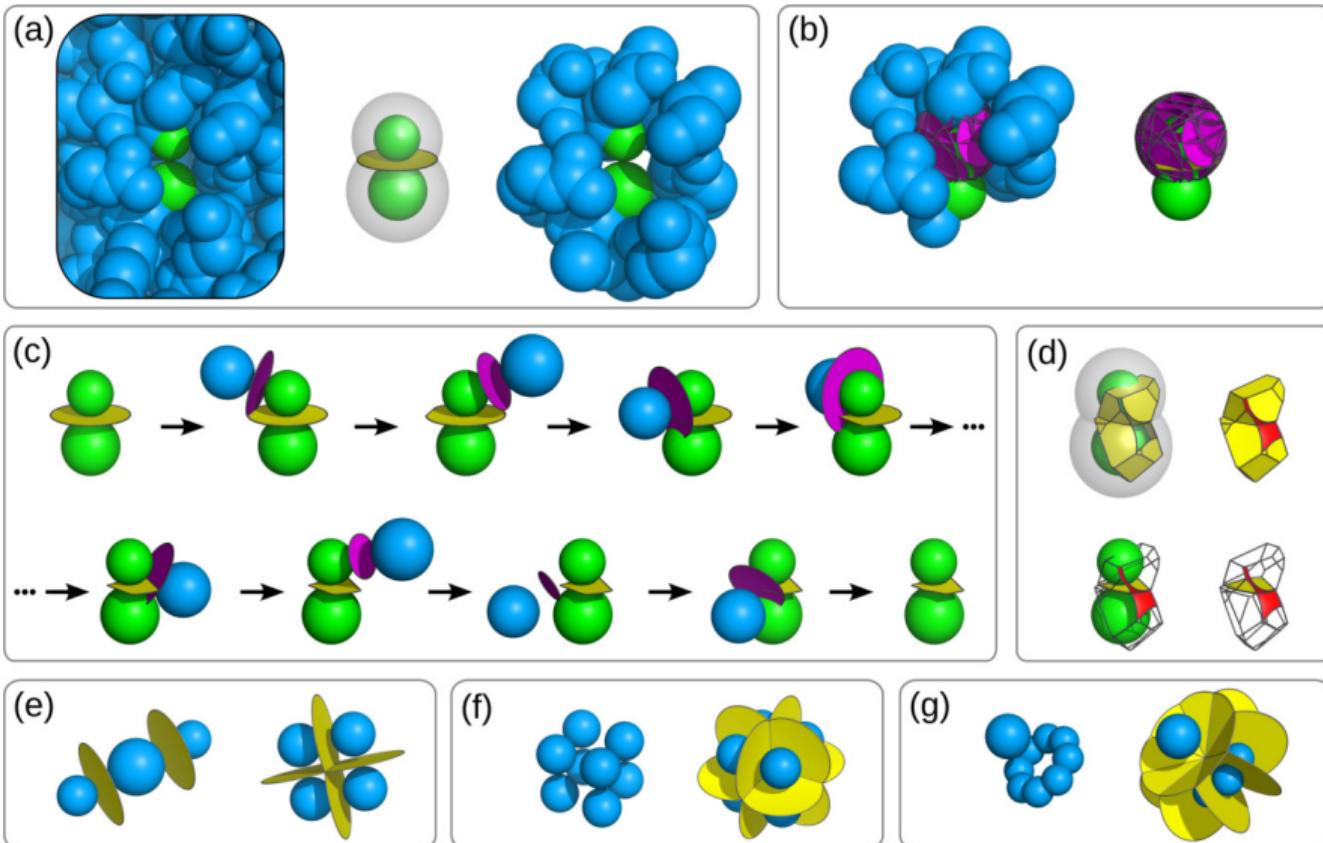
# Voronota-LT

Voronota-LT is a new fast software for constructing tessellation-derived atomic contact areas and volumes. It is significantly faster than its predecessor, Voronota:



Olechnovic and Grudinin. *Voronota-LT: efficient, flexible and solvent-aware tessellation-based analysis of atomic interactions.* JCC (2025)

# Vorononata-LT



# Getting COD crystal structure data using OPTIMADE API

```
#!/bin/bash

CODID="7250851"

JSONFILE="../optimade_cod_${CODID}.json"

[ -s "$JSONFILE" ] || \
curl -s "https://www.crystallography.net/cod/optimade/structures/${CODID}?
response_fields=cartesian_site_positions,species_at_sites,lattice_vectors,species" \
> "$JSONFILE"
```

# Getting COD crystal structure data using OPTIMADE API

```
kliment@local:/tmp$ jq '.data.attributes.cartesian_site_positions[]' "${JSONFILE}"
```

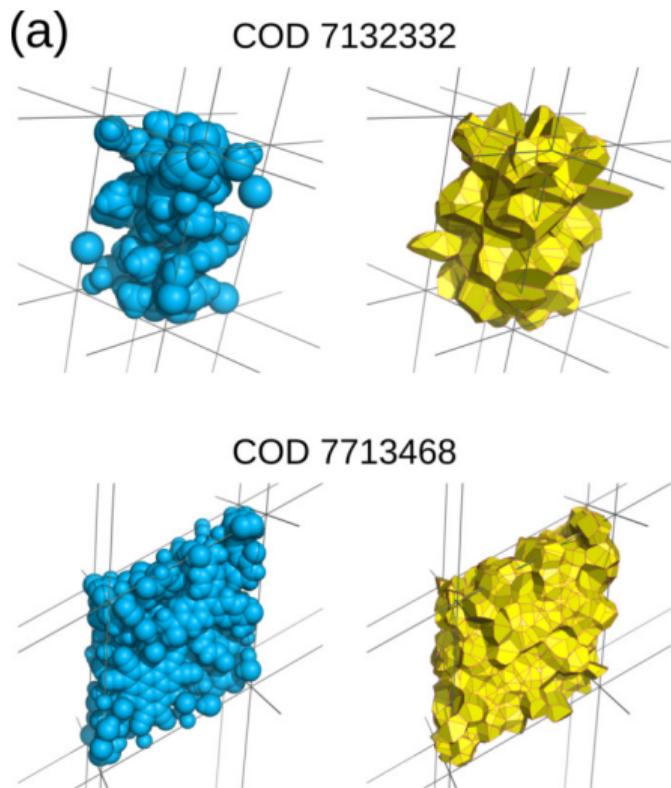
```
[  
  2.05265319791184,  
  5.890757274,  
  4.58437338884256  
]  
[  
  4.39151320737323,  
  13.277657274,  
  3.14201180150793  
]  
[  
  2.7797796126583,  
  8.883042726,  
  10.8683969918584  
]  
[  
  0.440919603196916,  
  1.496142726,  
  12.310758579193  
]  
...
```

# Getting COD crystal structure data using OPTIMADE API

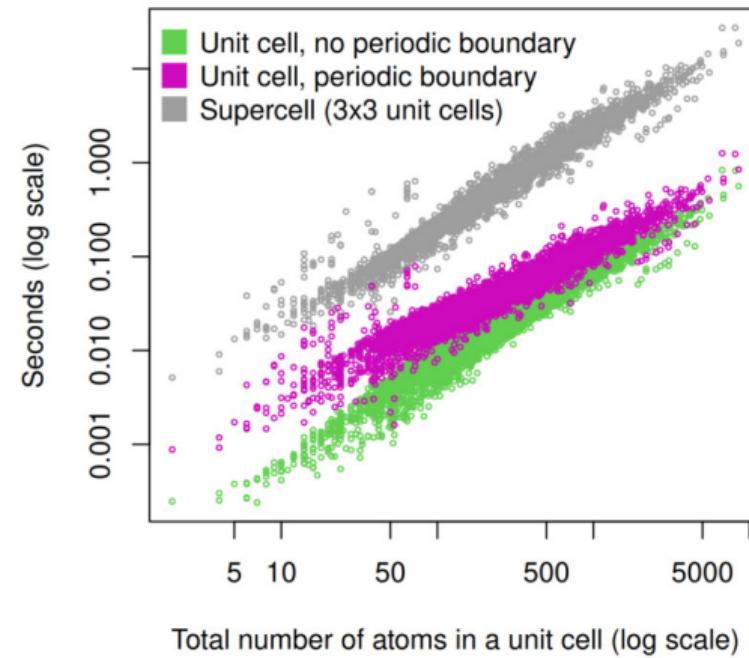
```
kliment@local:/tmp$ jq '.data.attributes.lattice_vectors' "${JSONFILE}"
```

```
[  
  [  
    8.0559,  
     0,  
     0  
  ],  
  [  
    0,  
   14.7738,  
     0  
  ],  
  [  
    -3.22346718942986,  
     0,  
  15.452770380701  
  ]  
]
```

# Running Voronota-LT on COD data

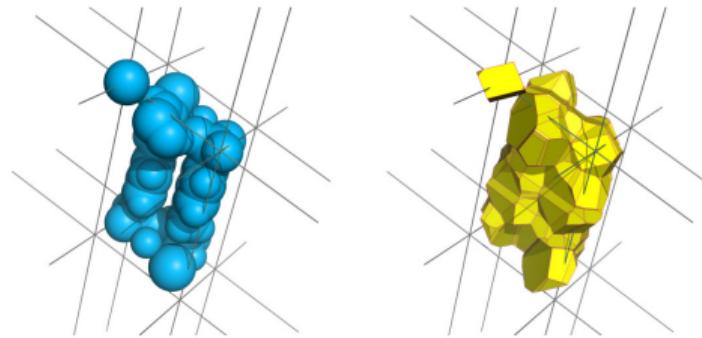


(b) **Running time for molecular crystals**  
**(log scale for atoms, log scale for seconds)**

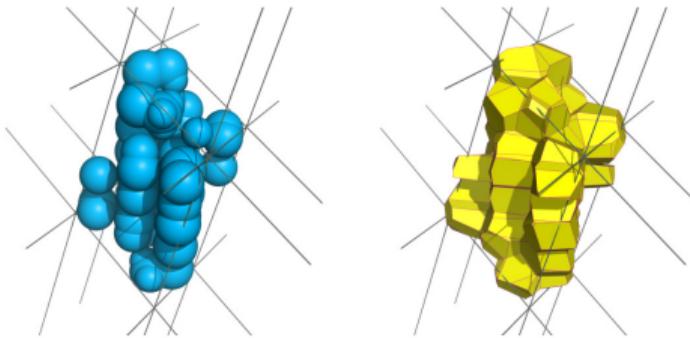


# Running Voronota-LT on COD data — some examples

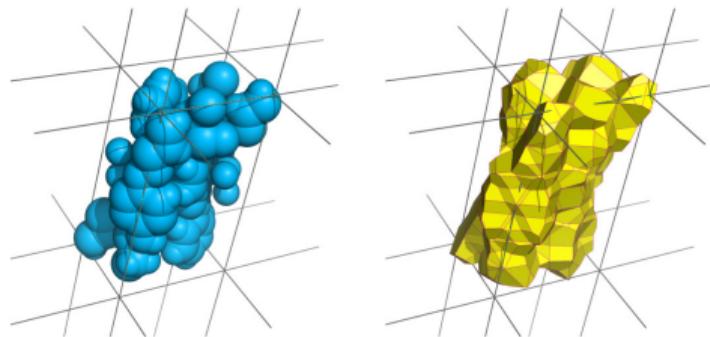
COD 7247424



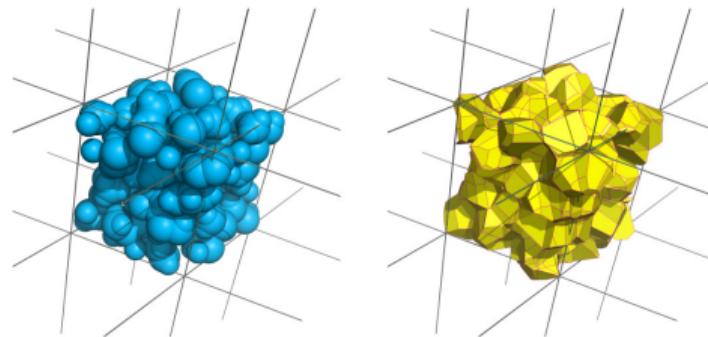
COD 7246672



COD 7132033

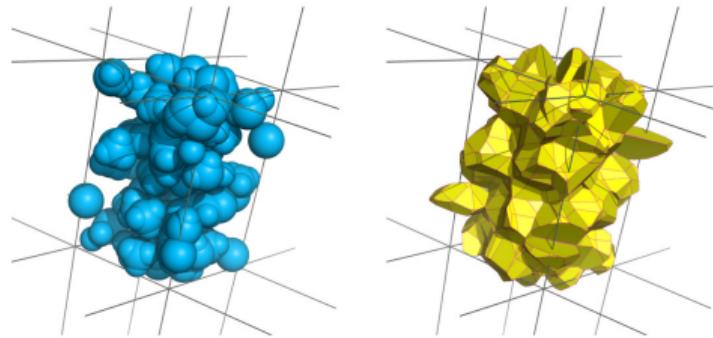


COD 7712772

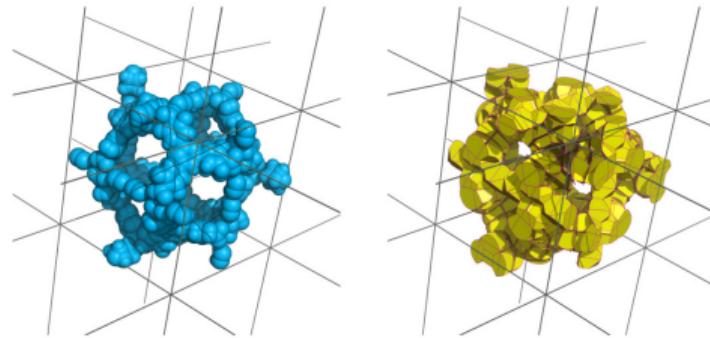


# Running Voronota-LT on COD data — some examples

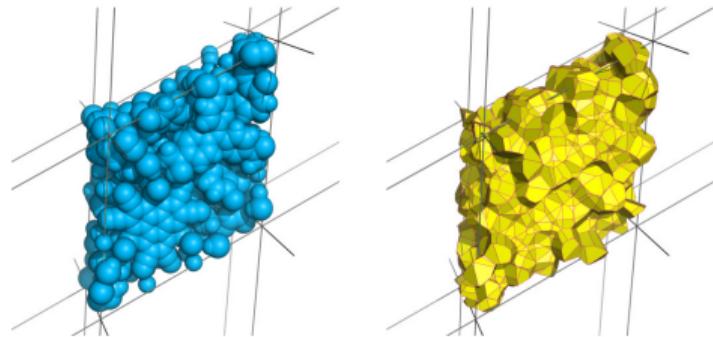
COD 7132332



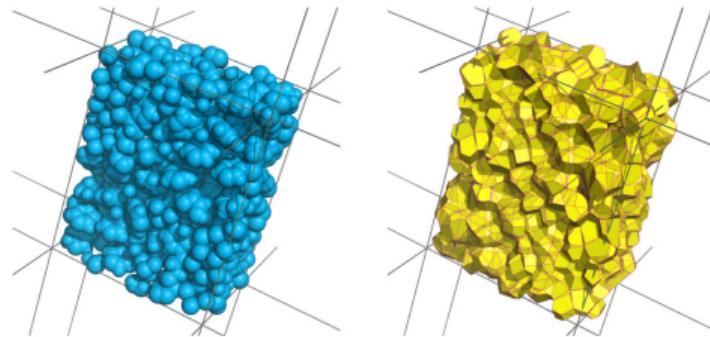
COD 7131232



COD 7713468



COD 1569733



## Some questions

Is it a good idea to provide an output of a tool (or a computational service) in an OPTIMADE-compatible way?

For example, the contacts output can be viewed as a table that references atoms. Is OPTIMADE API good for querying and summarizing pairwise interactions? Multibody interactions?

ID1_chain	ID1_residue	ID1_atom	ID2_chain	ID2_residue	ID2_atom	ID1_index	ID2_index	area	arc_legh
A	3 LYS	CD	B	41 TYR	CE1	16	2133	0.229169	0
A	3 LYS	CE	B	41 TYR	OH	17	2136	3.45511	0.460282
A	3 LYS	NZ	B	41 TYR	OH	18	2136	1.9173	3.36683
A	12 PRO	C	B	23 ILE	CD1	71	1980	2.62514	0
A	12 PRO	O	B	19 TYR	CE1	72	1939	2.47794	0
A	12 PRO	O	B	23 ILE	CD1	72	1980	0.736106	0
A	12 PRO	CB	B	23 ILE	CD1	73	1980	4.69368	0
A	12 PRO	CB	B	23 ILE	CG1	73	1978	6.86848	0
A	12 PRO	CG	B	35 TYR	CB	74	2077	4.55369	0
A	12 PRO	CG	B	22 TYR	CE2	74	1970	4.64473	0
A	13 GLU	N	B	23 ILE	CD1	76	1980	1.42152	0
A	13 GLU	N	B	23 ILE	CG2	76	1979	0.0194025	0
*	13 GLU	CA	B	23 ILE	CD1	77	1980	1.18001	0.158003

# Thanks

Thank you!

Useful links:

- ▶ <https://www.kliment.lt>
- ▶ <https://www.bioinformatics.lt>
- ▶ <https://grulab.imag.fr>



Funded by  
the European Union