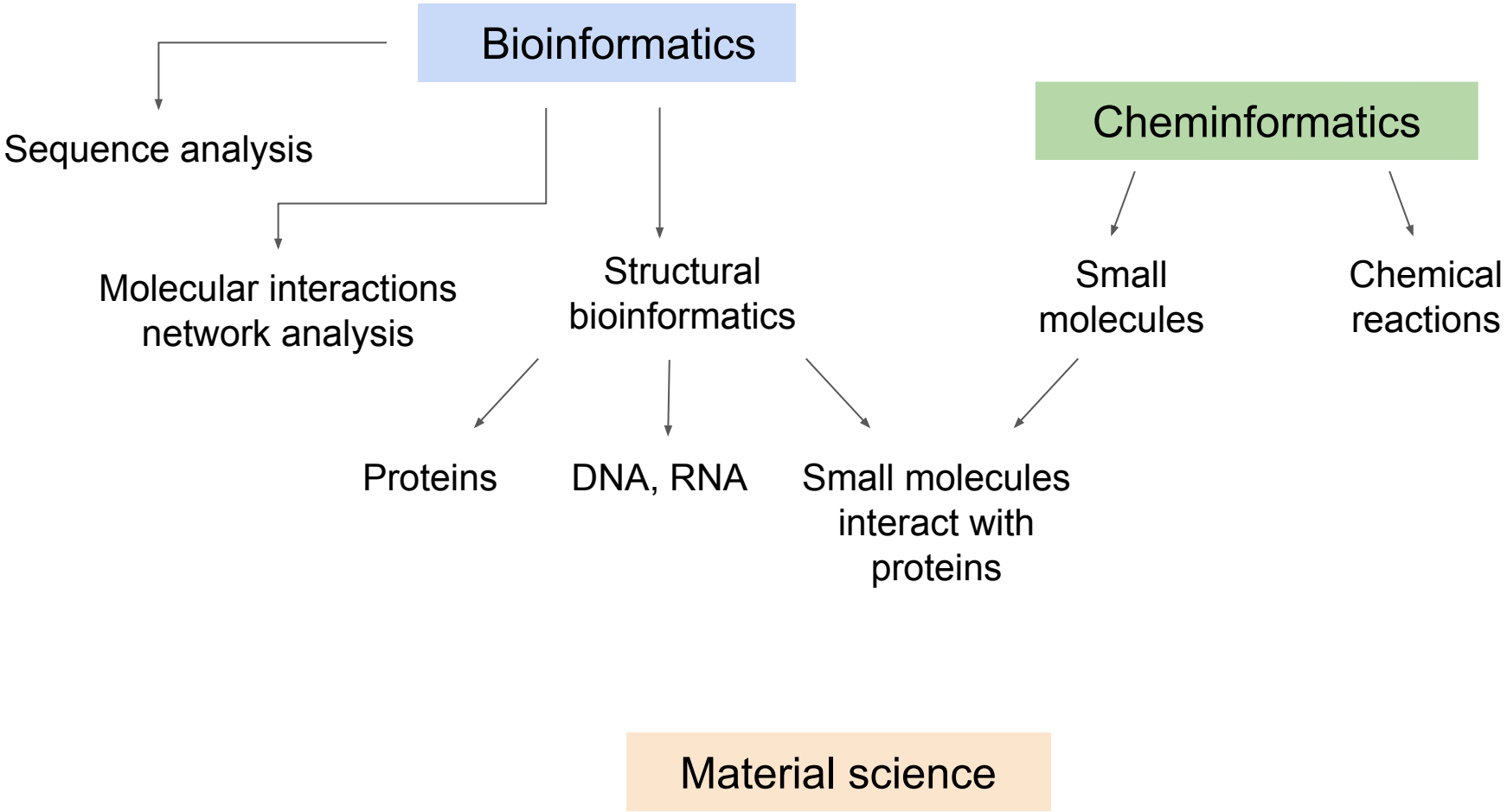
Machine Learning in Structural Bioinformatics

Maria Kadukova

HML Reading Group
16.05.2019

Inria INVENTORS FOR THE DIGITAL WORLD
MIPT MOSCOW INSTITUTE OF PHYSICS AND TECHNOLOGY
UNIVERSITÉ Grenoble Alpes

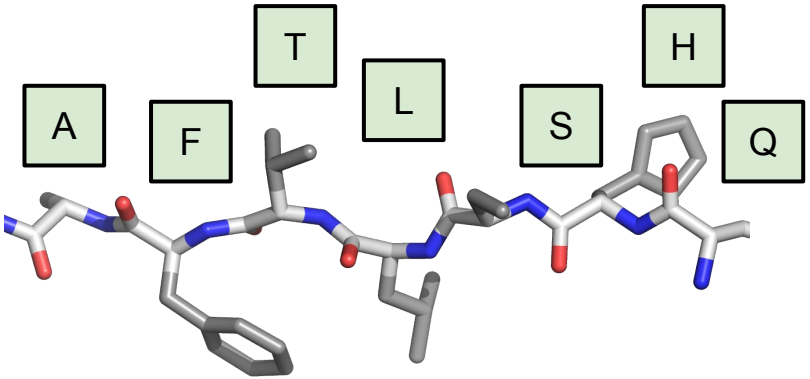

Bioinformatics
Sequence analysis
Molecular interactions network analysis
Structural bioinformatics
Proteins
DNA, RNA
Small molecules interact with proteins
Cheminformatics
Small molecules
Chemical reactions
Material science

Outline

- How do molecules look and function?
- Protein structure and interactions prediction
- Protein-ligand interactions
- Cheminformatics
- Other applications

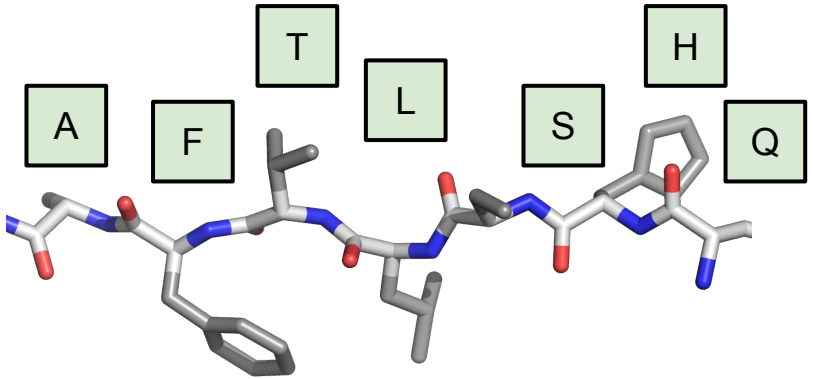
Macromolecules: proteins

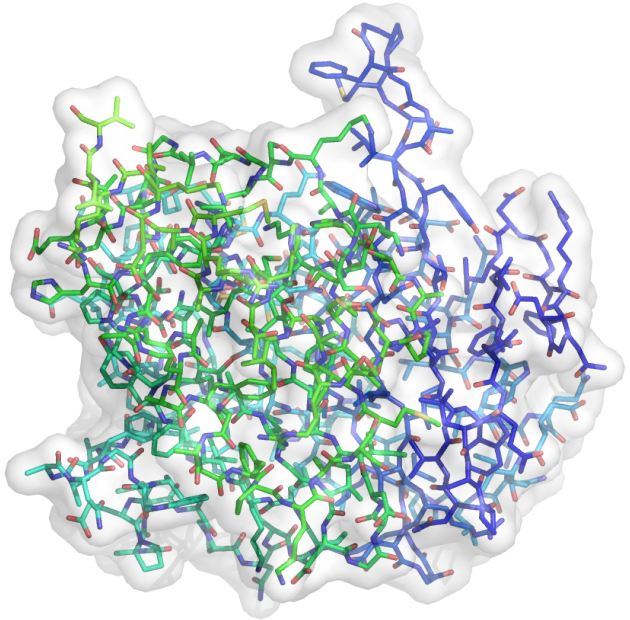
GCT	TTT	ACT	TTA	TCT	CAT	CAA



Macromolecules: proteins

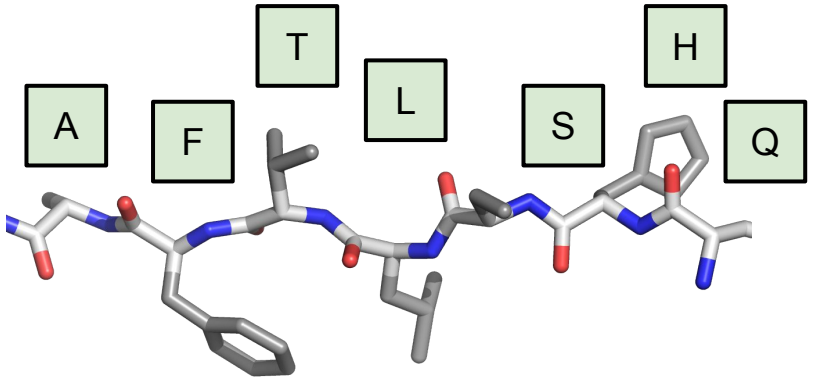
GCT	TTT	ACT	TTA	TCT	CAT	CAA

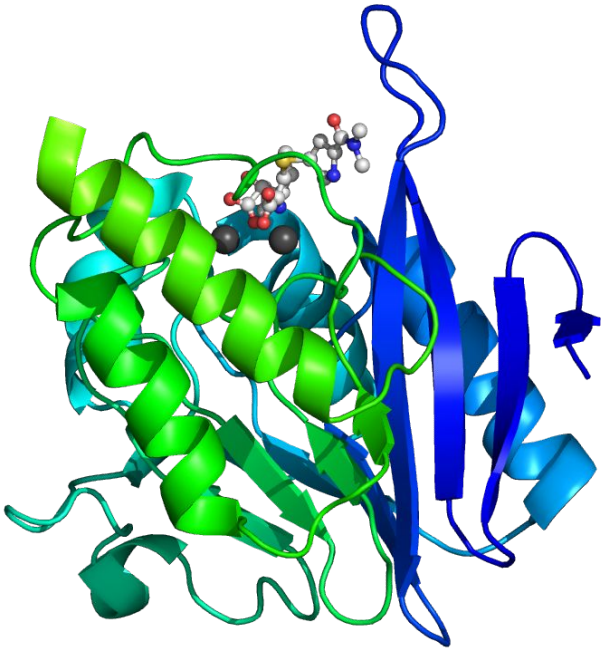




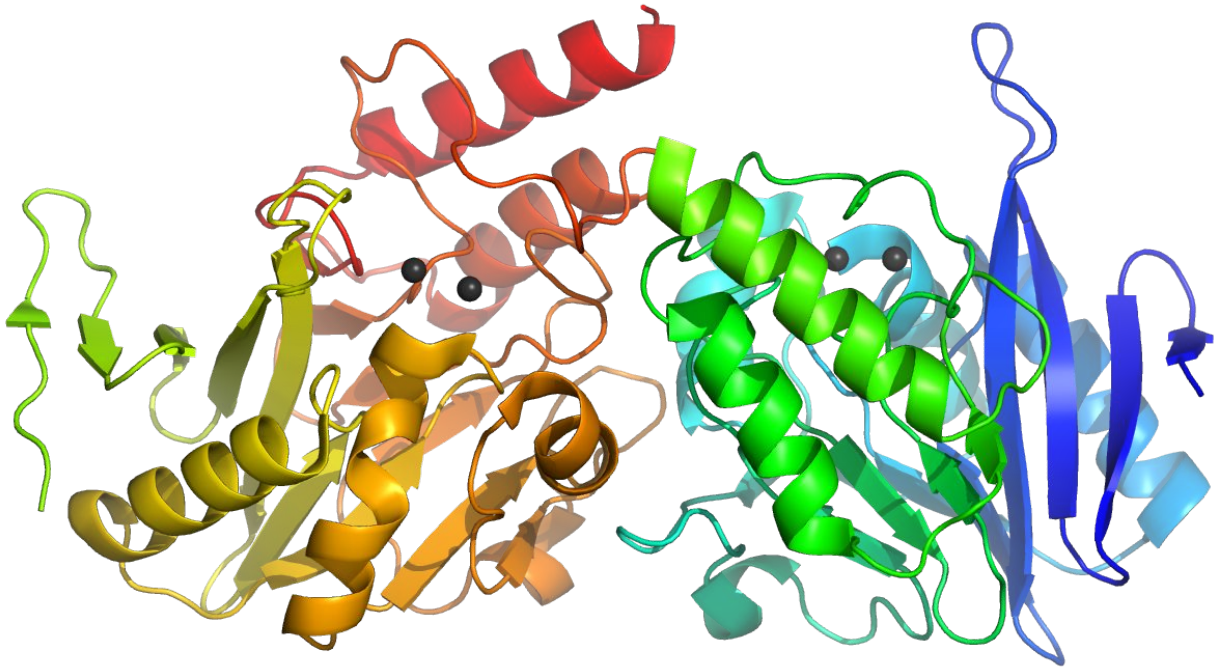
Macromolecules: proteins

GCT	TTT	ACT	TTA	TCT	CAT	CAA

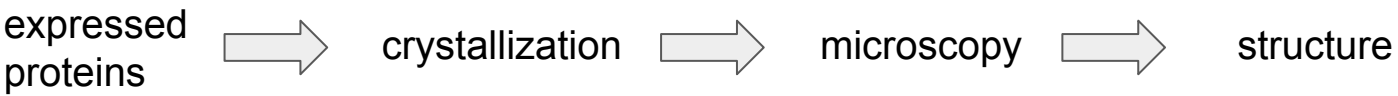




Macromolecules: proteins

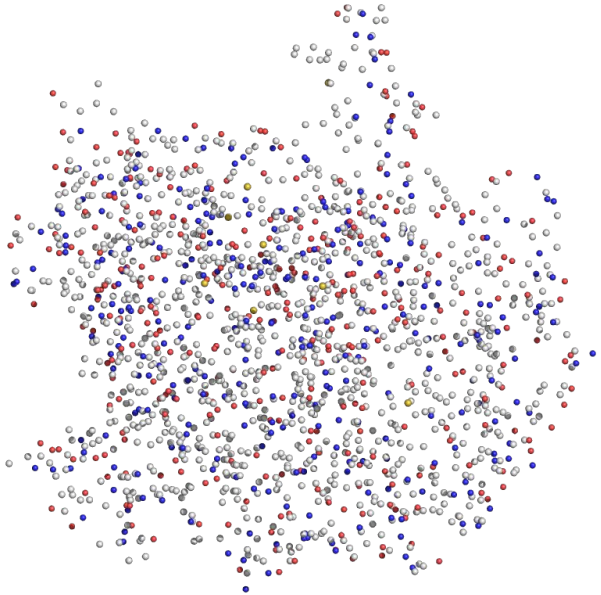


Proteins


expressed proteins
crystallization
microscopy
structure

Proteins: structure prediction

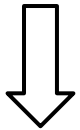
What we have



```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

What we want

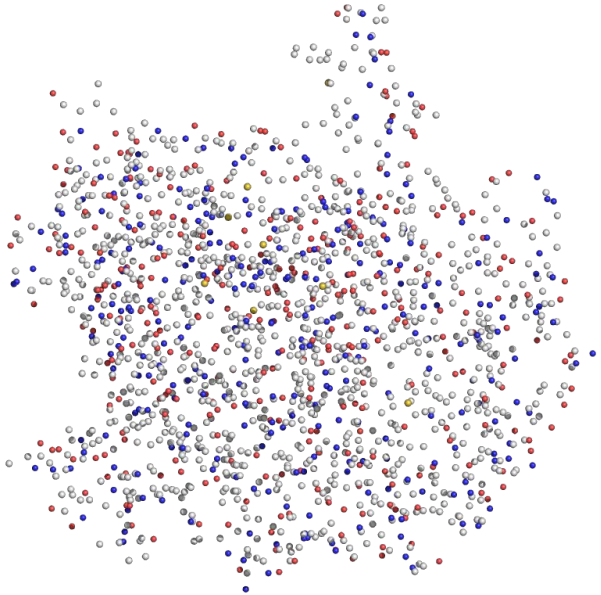
Sequence string



3D protein structure

Proteins: structure prediction

What we have



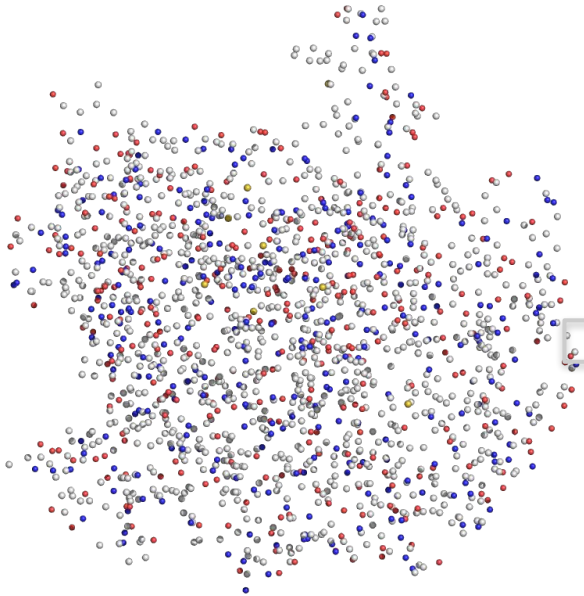
```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

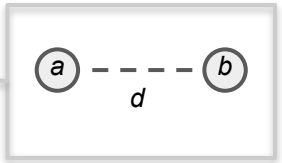
What we can do

- Use pieces of structures with similar sequence as building blocks

Proteins: structure prediction

What we have





MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV

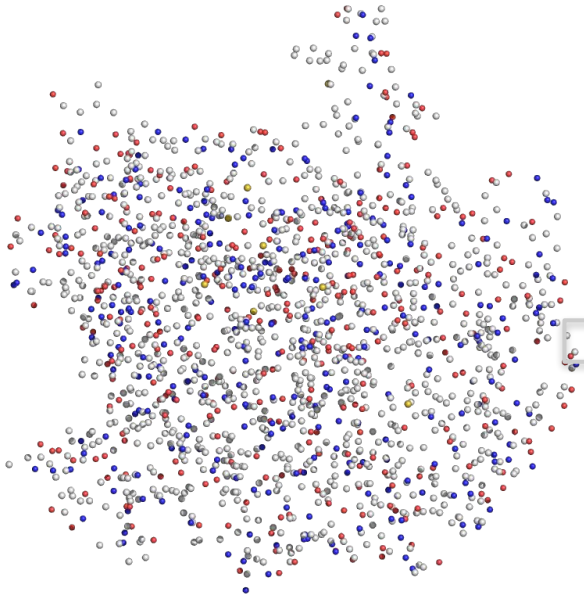
What we can do

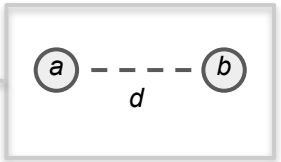
- Use pieces of structures with similar sequence as building blocks
- Define an energy function to minimize

$$U(d|a,b) = -kT \ln \frac{P(d|a,b)}{Q(d)}$$

Proteins: structure prediction

What we have



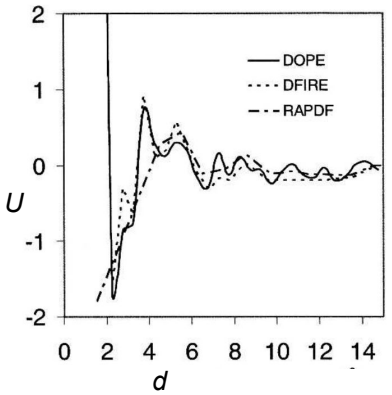



```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

What we can do

- Use pieces of structures with similar sequence as building blocks
- Define an energy function to minimize

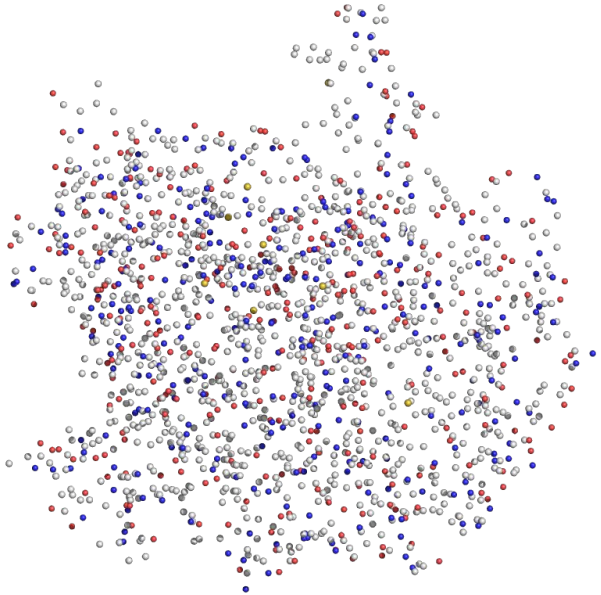
$$U(d|a,b) = -kT \ln \frac{P(d|a,b)}{Q(d)}$$



Shen, Sali. Statistical potential for assessment and prediction of protein structures, 2006

Proteins: structure prediction

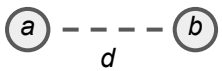
What we have



```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

What we can do

- Use pieces of structures with similar sequence as building blocks
- Define an energy function to minimize

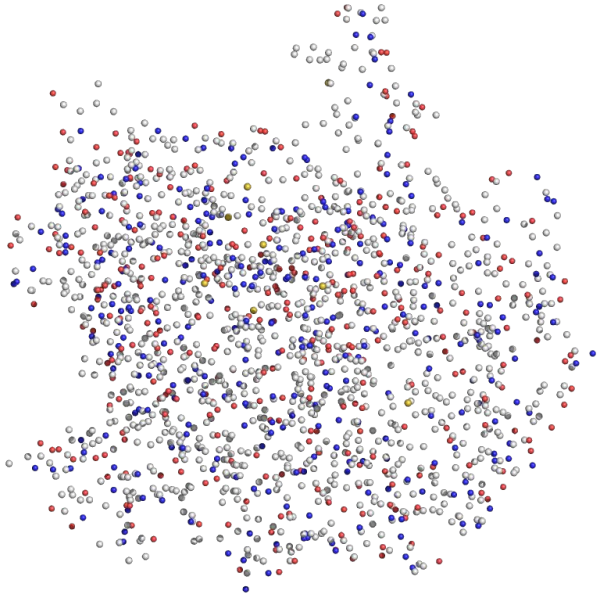


$$U(d|a,b) = -kT \ln \frac{P(d|a,b)}{Q(d)}$$

- Extract more features!

Proteins: structure prediction

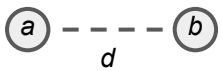
What we have



```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

What we can do

- Use pieces of structures with similar sequence as building blocks
- Define an energy function to minimize



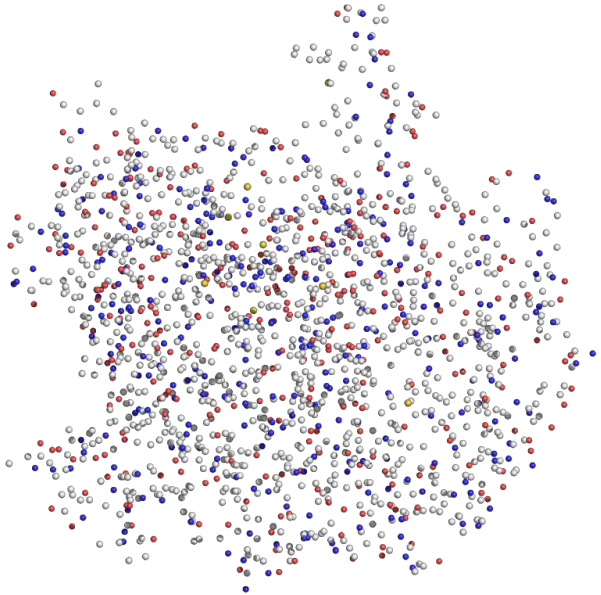
$$U(d|a,b) = -kT \ln \frac{P(d|a,b)}{Q(d)}$$

- Extract more features!

Other objectives?

Proteins: structure prediction

What we have



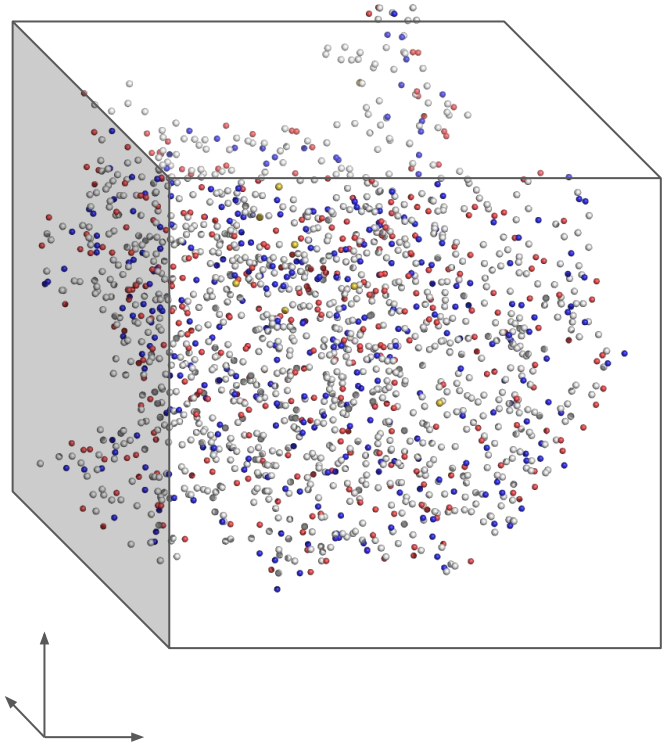
```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

What we can do

- Use false, but high-quality structures to learn

Proteins: structure prediction

What we have



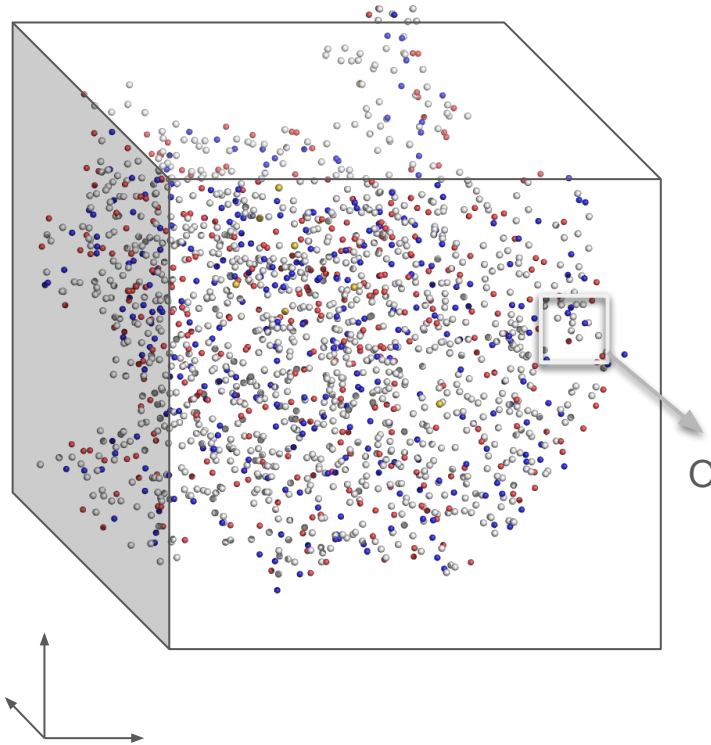
What we can do

- Use false, but high-quality structures to learn

R,G,B = [119, 172, 225]

Proteins: structure prediction

What we have



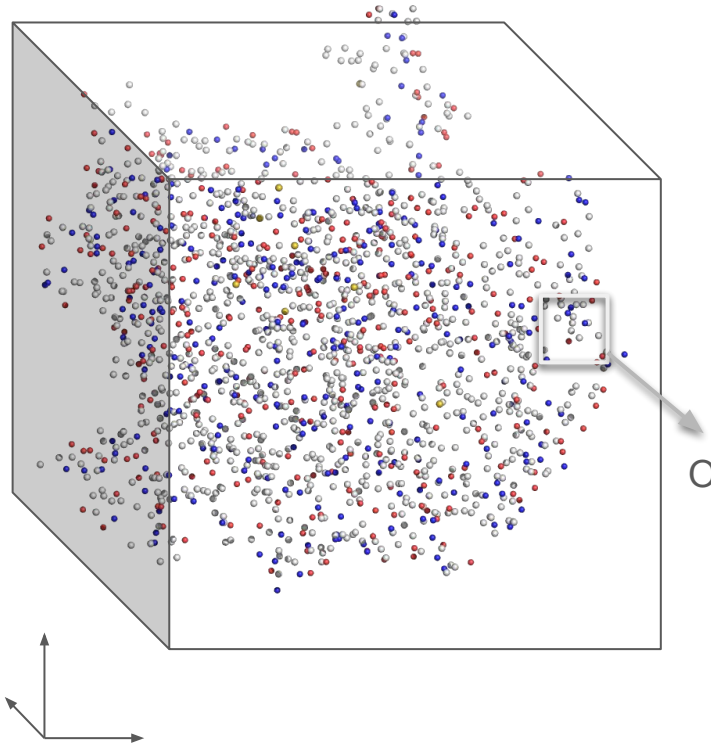
O_2, N_am, C_aro, ... =
[0.89, 0.55, 0.02, …]

What we can do

- use false, but high-quality structures to learn
- map densities of atoms around of each atom, atom types are the channels
- train a CNN!

Proteins: structure prediction

What we have

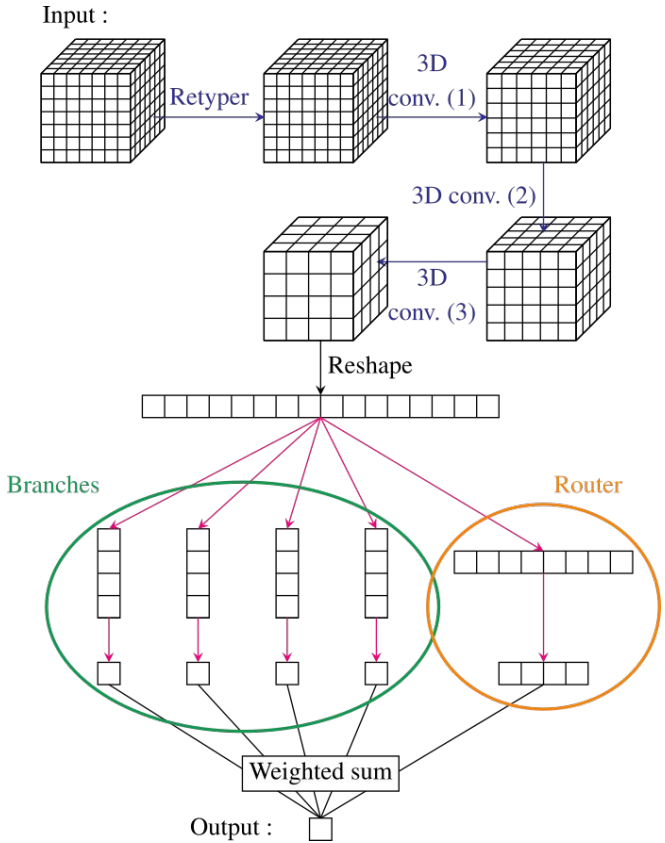


O_2, N_am, C_aro, ... =
[0.89, 0.55, 0.02, …]

What we can do

- use false, but high-quality structures to learn
- map densities of atoms around of each atom, atom types are the channels
- train a CNN!

Proteins: structure prediction



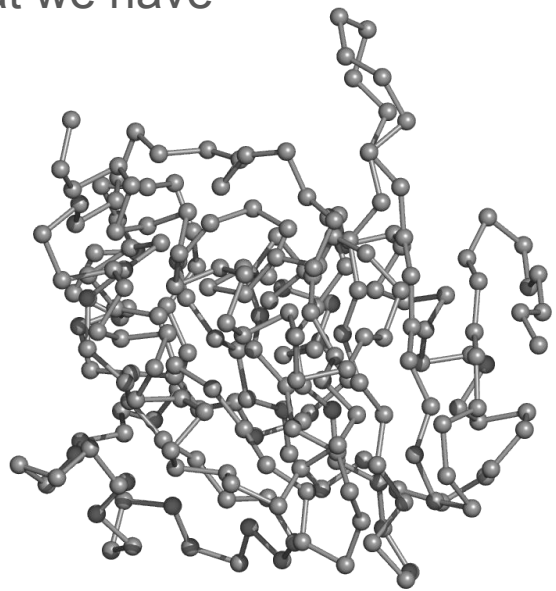
What we can do

- use false, but high-quality structures to learn
- map densities of atoms around of each atom, atom types are the channels
- train a CNN!

Pagès, Charmettant, Grudinin. Protein model quality assessment using 3D oriented convolutional neural networks. Submitted

Proteins: structure prediction

What we have



```
Q5E940_BOVIN -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_HUMAN -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_MOUSE -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
    RLA0_RAT -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_CHICK -----------MPREDRATWKSNYFMKIIQLLDDYPKCFVVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
```

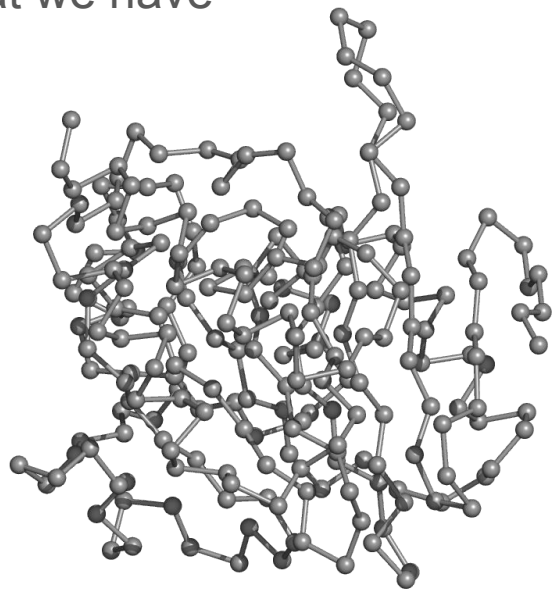
Multiple sequence alignment (MSA)

What we can do

- use MSA to find regions that **co-evolved** and thus can be spatially proximate
- make a matrix of contacts between these residues

Proteins: structure prediction

What we have



```
Q5E940_BOVIN -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_HUMAN -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_MOUSE -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
    RLA0_RAT -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_CHICK -----------MPREDRATWKSNYFMKIIQLLDDYPKCFVVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
```

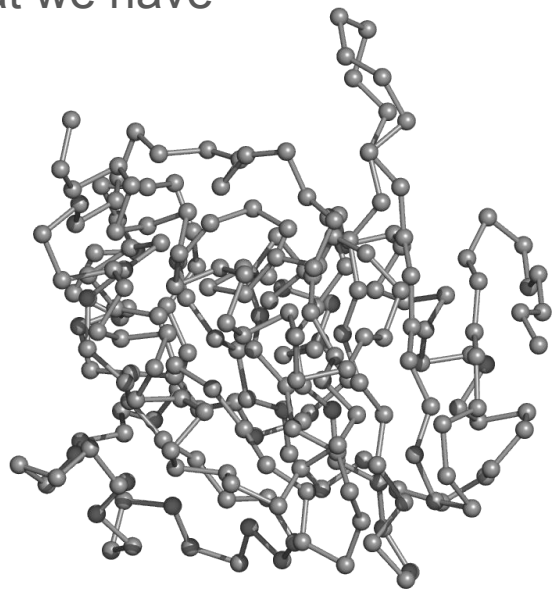
Multiple sequence alignment (MSA)

What we can do

- use MSA to find regions that **co-evolved** and thus can be spatially proximate
- make a matrix of contacts between these residues
- fulfil the constraints of contacts!

Proteins: structure prediction

What we have



```
Q5E940_BOVIN ----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE   76
  RLA0_HUMAN ----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE   76
  RLA0_MOUSE ----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE   76
    RLA0_RAT ----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE   76
  RLA0_CHICK ----------MPREDRATWKSNYFMKIIQLLDDYPKCFVVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE   76
```

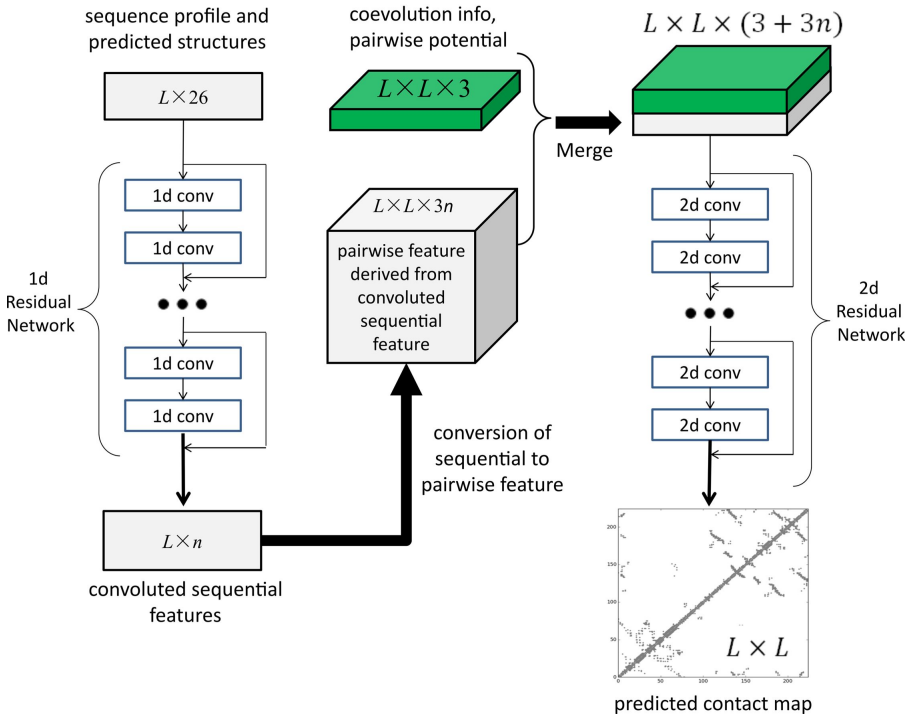
Multiple sequence alignment (MSA)

What we can do

- use MSA to find regions that **co-evolved** and thus can be spatially proximate
- make a matrix of contacts between these residues
- fulfil the constraints of contacts!
- or use them as features
- or train to **predict** them...

Proteins: structure prediction

Contacts prediction and residual neural networks

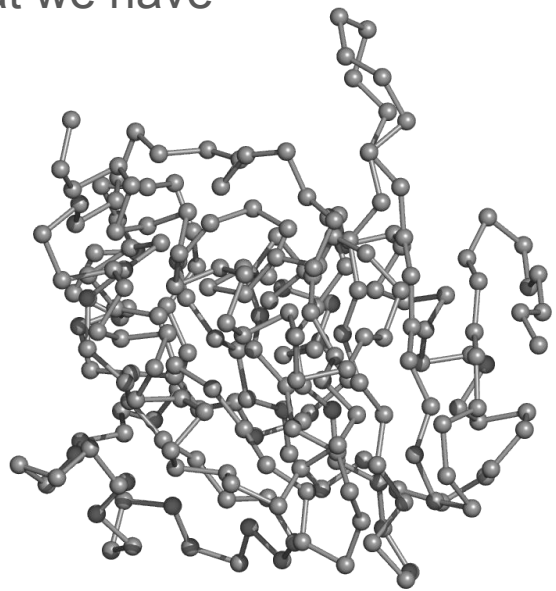


- diverse features (geometrical, sequence)
- very deep to find high-order correlations

Wang, Xu et al. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model, 2017

Proteins: structure prediction

What we have



```
Q5E940_BOVIN -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_HUMAN -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_MOUSE -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
    RLA0_RAT -----------MPREDRATWKSNYFLKIIQLLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
  RLA0_CHICK -----------MPREDRATWKSNYFMKIIQLLDDYPKCFVVGADNVGSKQMQQIRMSLRGK-AVVLMGKNTMMRKAIRGHLENN--PALE    76
```

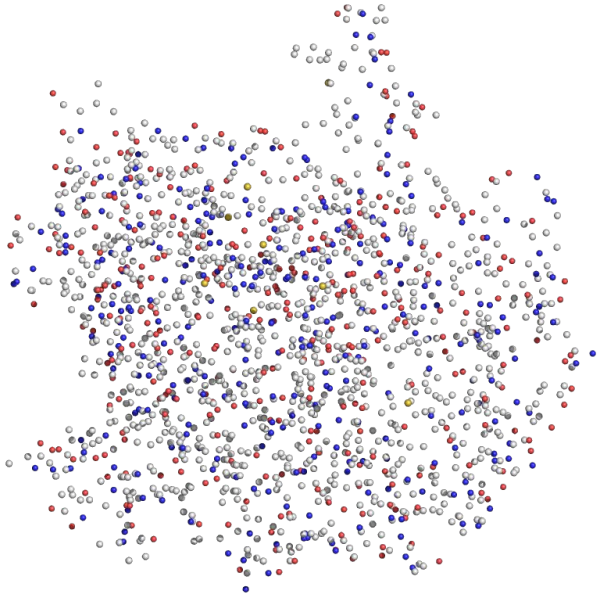
AlphaFold (and not only)

- binary contact matrix → contact distances
- additional scoring CNN
- AlphaFold uses the whole distribution of contact distances to compute likelihood
- this + NN-based scoring can be minimized

AlphaFold abstract and other abstracts: http://predictioncenter.org/casp13/doc/CASP13_Abstracts.pdf, 2018
earlier approach, similar to AlphaFold: Xu. Distance-based Protein Folding Powered by Deep Learning, 2018
nice review: https://moalquraishi.wordpress.com/2018/12/09/alphafold-casp13-what-just-happened

Proteins: structure prediction

What we have



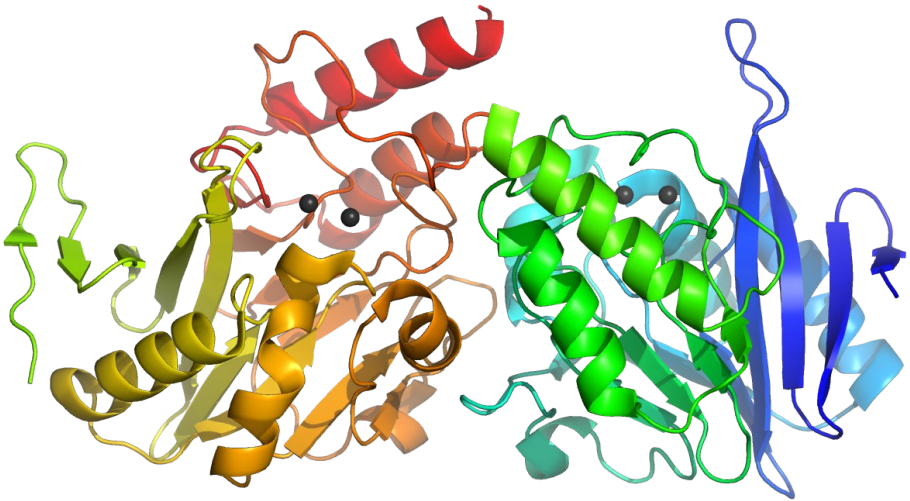
```
MELPNIMHPVAKLSTALAAALMLSGCMPGEIRPTIGQQMETGDQRFGDLVFRQLAPNVWQHTSYLDMP
GFGAVASNGLIVRDGGRVLVVDTAWTDDQTAQILNWIKQEINLPVALAVVTHAHQDKMGGMDALHAAG
IATYANALSNQLAPQEGMVAAQHSLTFAANGWVEPATAPNFGPLKVFYPGPGHTSDNITVGIDGTDIA
FGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAFPKASMIVMSHSAPDSRAAITHTARMADKLRLV
```

What we get

- template-based methods
- scoring with statistical potentials
- learning on decoys how to score
- co-evolution-based methods

Proteins: protein-protein interactions

What we have



What we want

- Interaction interface?
- Interaction energy?

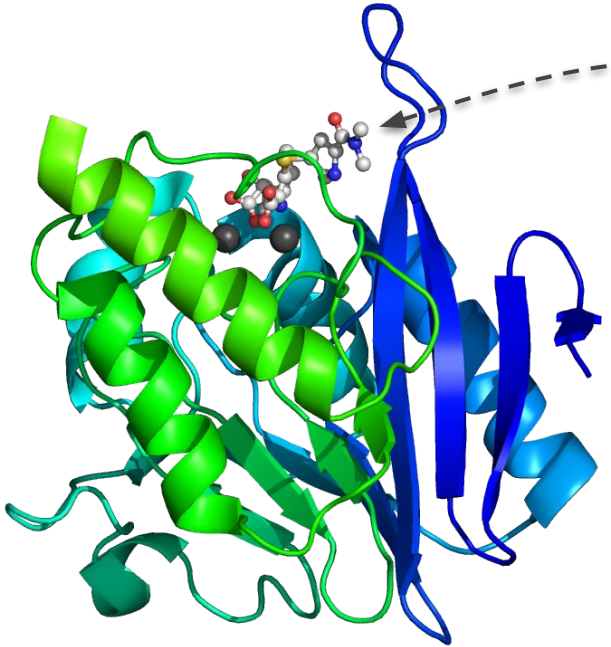
Outline

- ✓ How do molecules look and function?
- ✓ Protein structure and interactions prediction
- Protein-ligand interactions
- Cheminformatics
- Other applications

Proteins: protein-ligand interactions

What we have

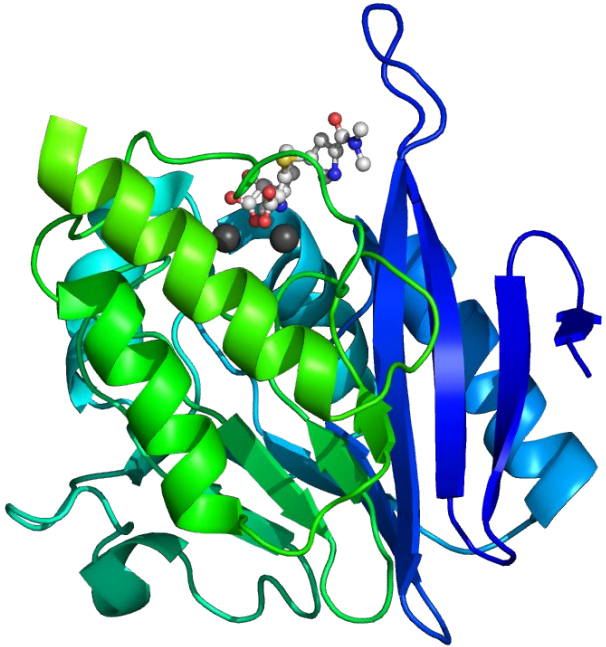
What we want

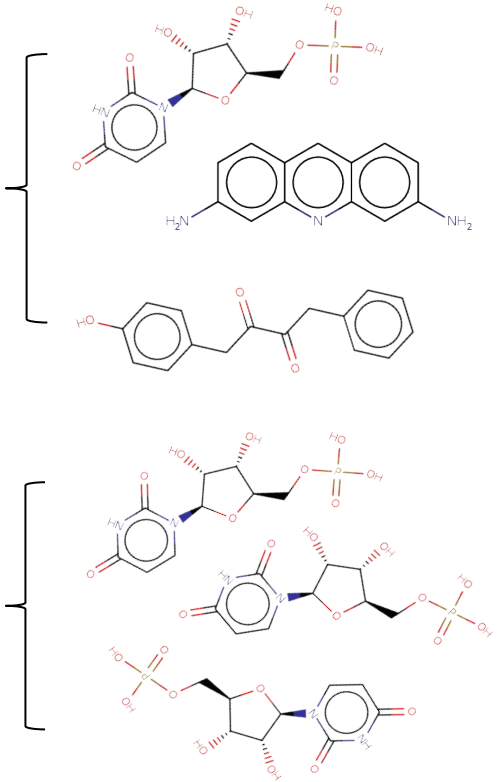


Put there another compound?

Proteins: protein-ligand interactions

What we have





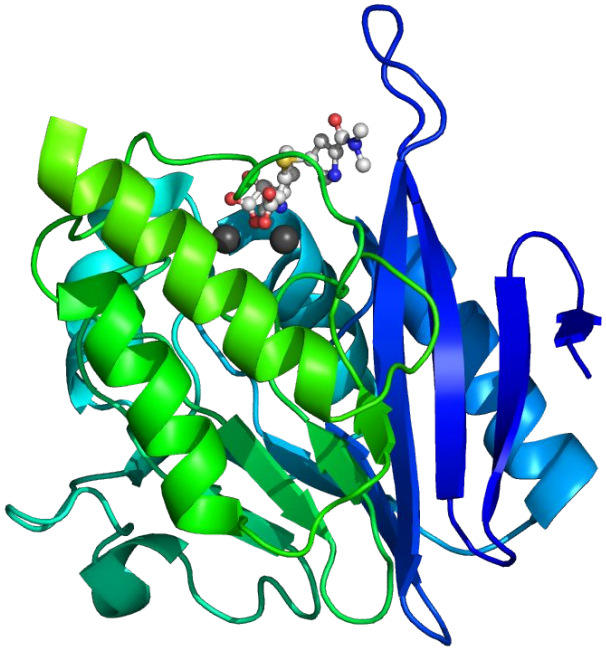
What we want

Who will bind?

With what strength?

Correct binding pose?

Proteins and ligands: data



$$\Delta G \sim -RT \ln K_b$$

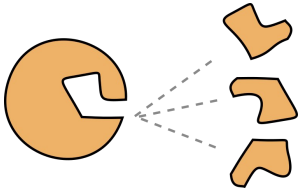
Features:

- extract them from 3D coordinates

Objectives:

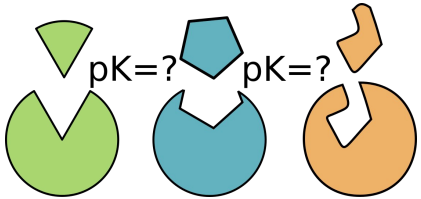
- K_b (affinities) are known for regression
- easier ‘’false’’ structures generation to do classification

Proteins and ligands: data and problems



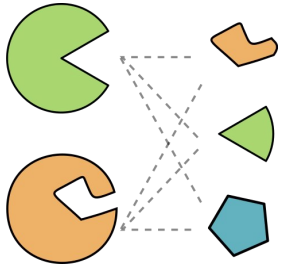
pose prediction

find the best 3D coordinates for the known ligand



scoring

affinities prediction, energy prediction for known ligands



virtual screening

which ligand binds a compound?

Proteins and ligands: molecular docking and scoring

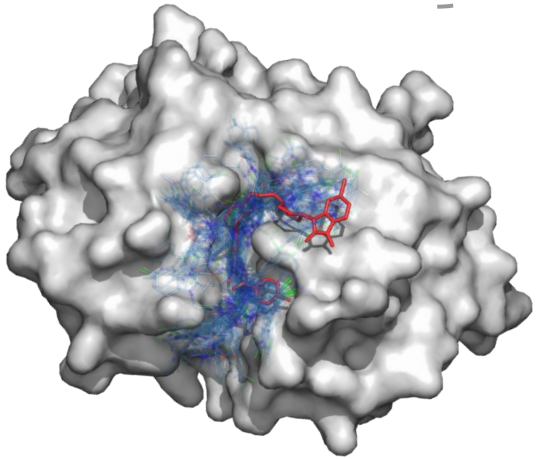
- Markov chain Monte-Carlo
- genetic algorithms
- molecular dynamics
- ...

sampling

scoring

$$\Delta G \sim -RT \ln K_b$$

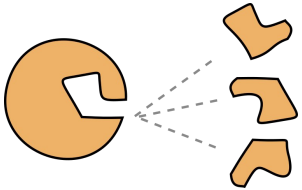
Lower free binding energy - more affine ligands.



A scoring function

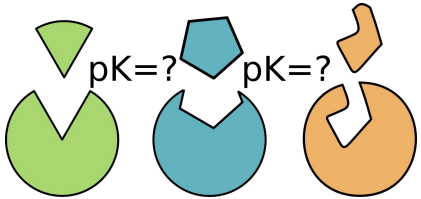
- predicts the binding free energy
- or scores the affinity
- can be a part of sampling

Proteins and ligands: data and problems



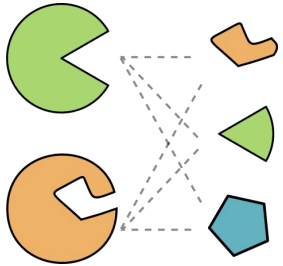
pose prediction

find the best 3D coordinates for the known ligand



scoring

affinities prediction, energy prediction for known ligands



virtual screening

which ligand binds a compound?

Proteins and ligands: scoring functions

- Physics-based

Energy terms, often trained with use of force fields, robust and rather slow

- Empirical

A combination of energy terms trained on affinities data (regression)

- Knowledge-based

Radial and angular distributions of atoms → statistical potentials

- Descriptor-based

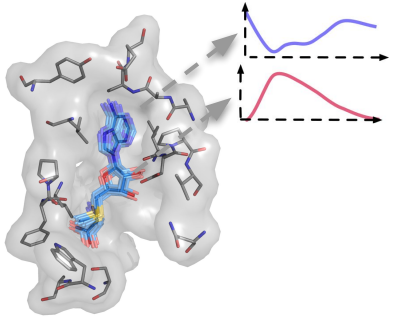
Various descriptors, sophisticated machine learning methods

Proteins and ligands: descriptors

- energy terms $\Delta g_i = f(r_{ab})$

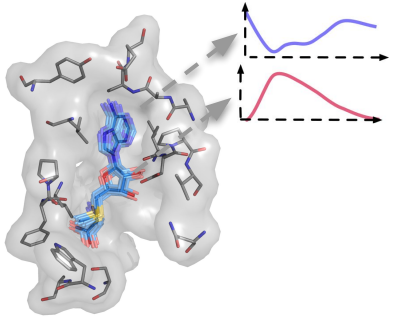
Proteins and ligands: descriptors

- energy terms $\Delta g_i = f(r_{ab})$
- radial, angular distributions of atoms



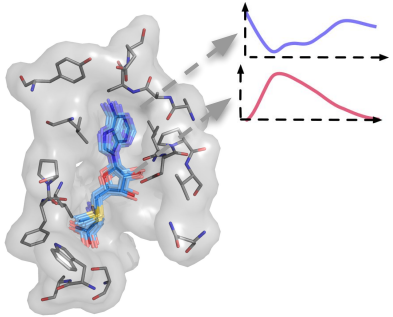
Proteins and ligands: descriptors

- energy terms $\Delta g_i = f(r_{ab})$
- radial, angular distributions of atoms
- 2D descriptors (molecule is a graph!)



Proteins and ligands: descriptors

- energy terms $\Delta g_i = f(r_{ab})$
- radial, angular distributions of atoms
- 2D descriptors (molecule is a graph!)
- surface descriptors
- score as descriptor (“meta” scoring function)
- ...



Convex-PL

$$\Delta G = \Delta H - T\Delta S$$

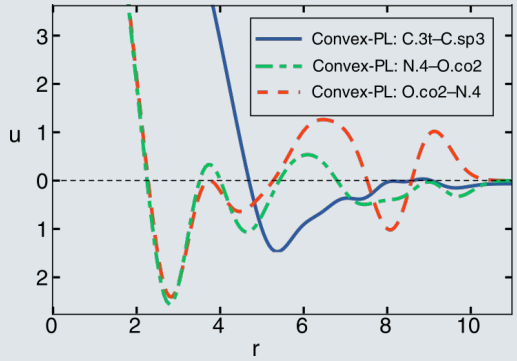
$$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$$

Convex-PL

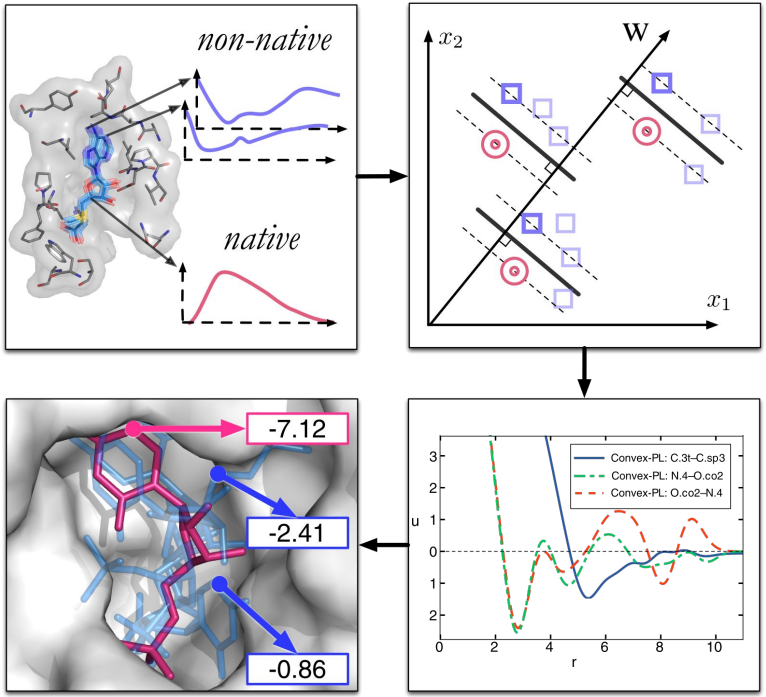
$$\Delta G = \Delta H - T\Delta S$$

$$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$$

knowledge-based
distance-dependent potential



Convex-PL: knowledge-based potential



- radial distribution functions as descriptors
- **w** is an unknown vector of interactions

no reference states → solve classification problem instead

train **w** to separate natives and decoys

$$\min: \frac{1}{2}\mathbf{w}\cdot\mathbf{w} + \sum_{ij} C_{ij}\xi_{ij}$$

$$\text{s.t.}: y_{ij}\left[\mathbf{w}\cdot\mathbf{x}_{ij} + b_i\right] - 1 + \xi_{ij} \geq 0, \xi_{ij} \geq 0$$

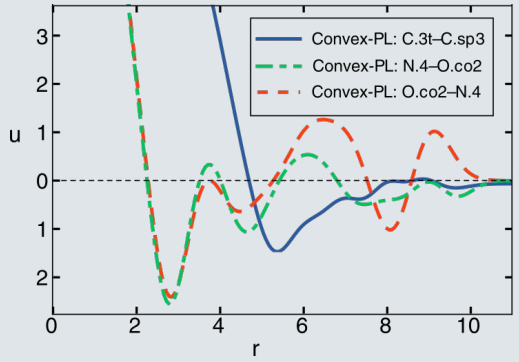
- perfect for pose prediction
- average at affinities prediction

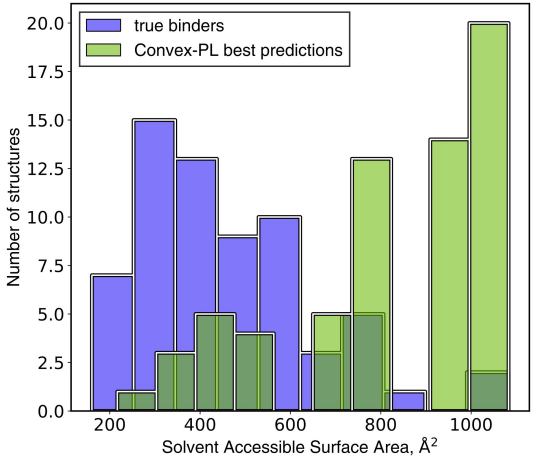
Convex-PL ≠ statistical potentials

Convex-PL: knowledge-based potential

$$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$$

knowledge-based
distance-dependent potential





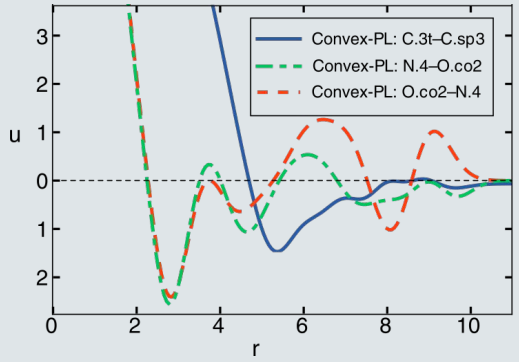
CASF Benchmark 2013

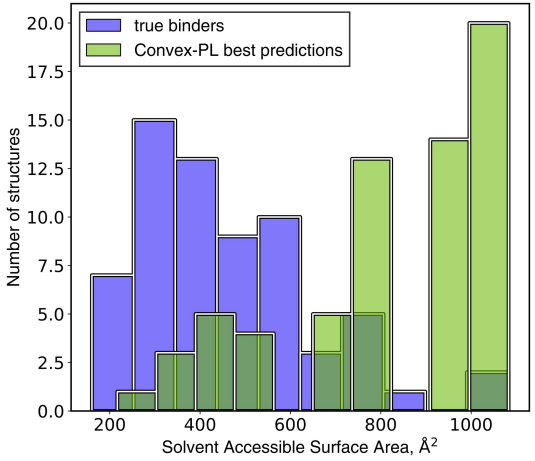
big interfaces bias

Convex-PL: knowledge-based potential

$$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$$

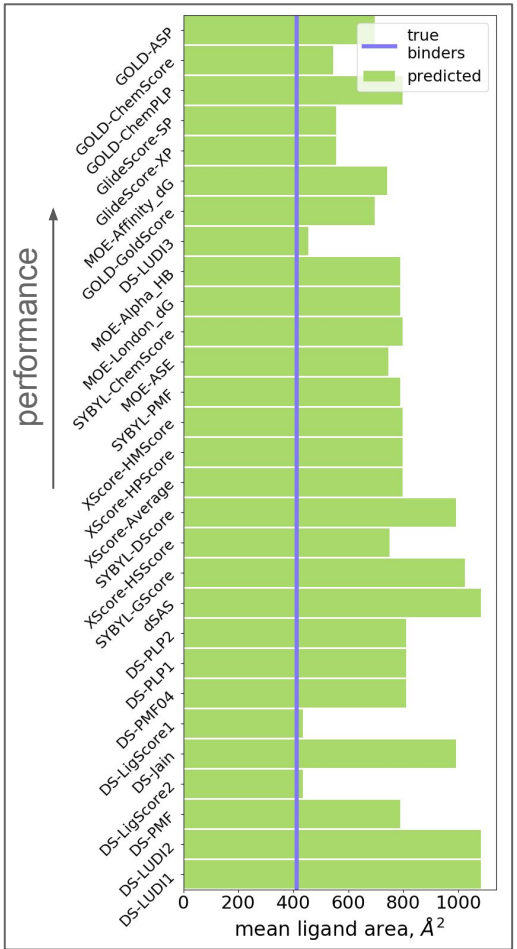
knowledge-based
distance-dependent potential





CASF Benchmark 2013

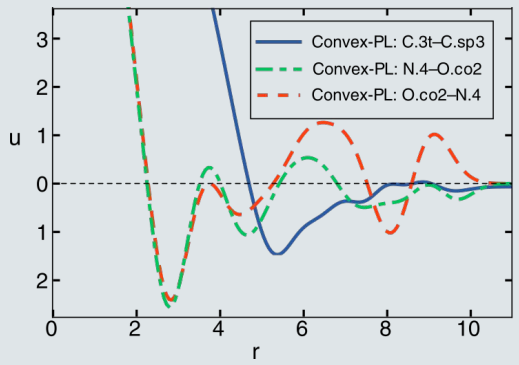
big interfaces bias



Convex-PL: more descriptors

$$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$$

knowledge-based
distance-dependent potential



approximated with a regression model

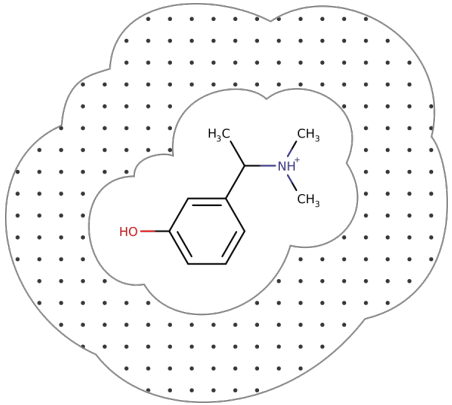
- solvent descriptors
- conformational entropy descriptors

☆ empirical scoring functions-style ☆

Convex-PL: more descriptors

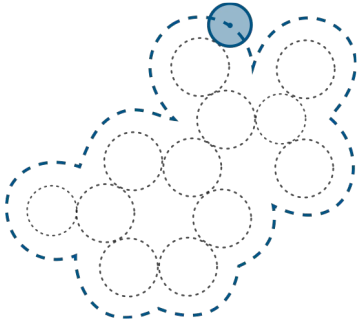
$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$

~ volume



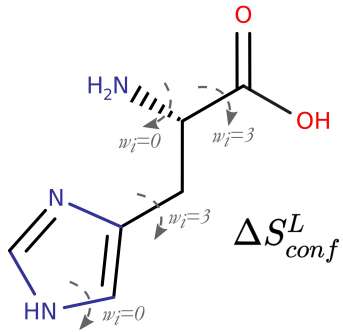
- 3D grid representation
- discrete

~ solvent accessible surface area



- continuous
- better hydrophobic effects representation

ligand flexibility



$$\Delta S^{L}_{conf} = \log\left(\prod_{i}^{\#\ bonds} w_i\right)$$

Convex-PL: regression-based model

$$\Delta G = \Delta H^{PL} + \Delta H^{solvent} - T(\Delta S^{solvent} + \Delta S^{conf} + \dots)$$

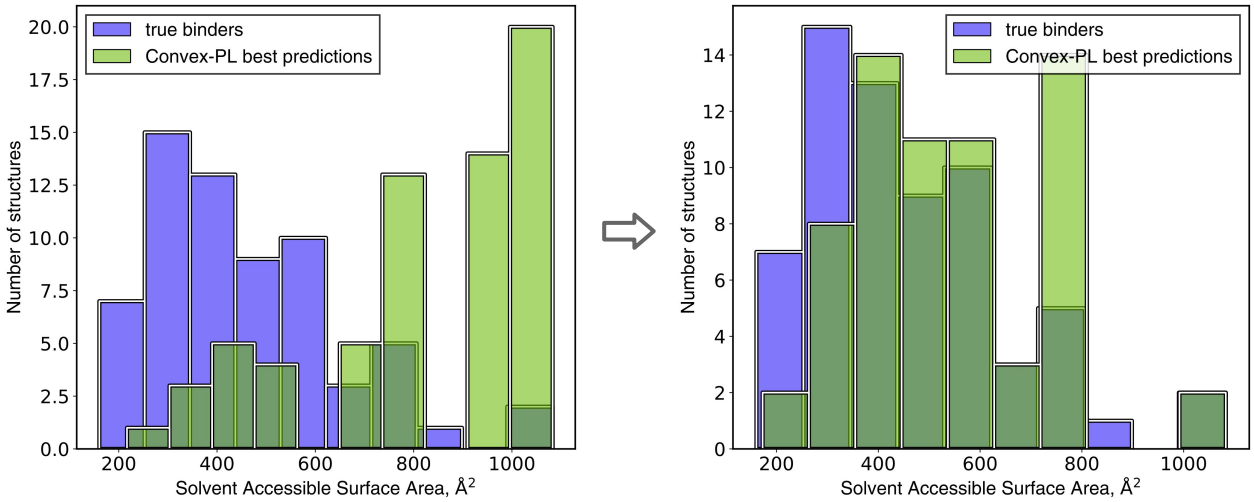
$$X = \begin{bmatrix} \text{knowledge-based score} & , & \text{solvent rdfs SASA descriptors} & , & \text{ligand flexibility} \end{bmatrix}$$

$$Y = [pK]$$

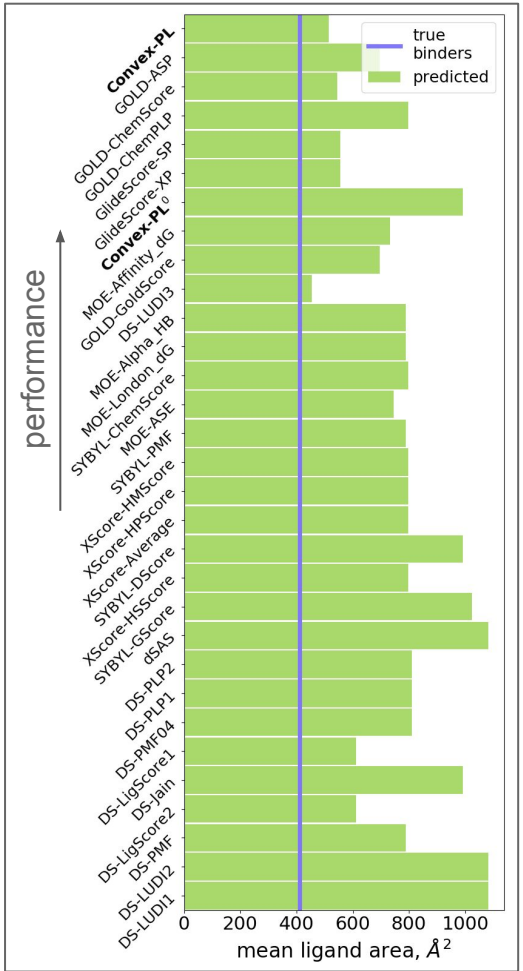
Ridge Regression model

- ✓ better score predictions
- ✓ smaller bias towards big interfaces

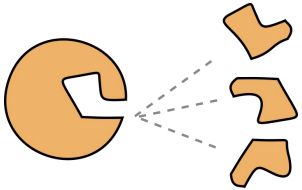
Convex-PL: regression-based model



CASF Benchmark 2013

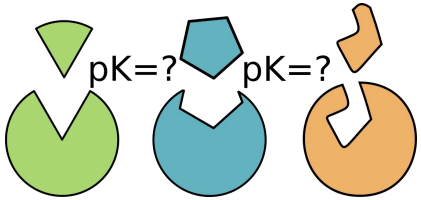


Proteins and ligands: problems



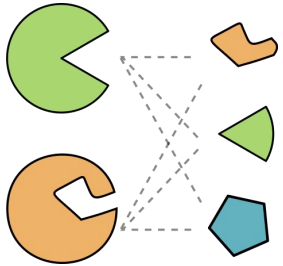
pose prediction

✓ we move the ligand
× receptor moves as well



scoring

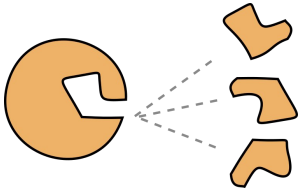
× low data quality



virtual screening

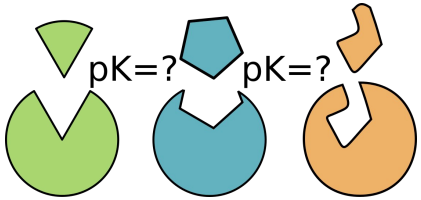
× is almost always 1-label classification

Proteins and ligands: problems



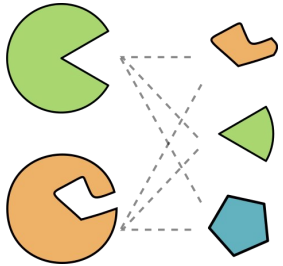
pose prediction

✓ we move the ligand
× receptor moves as well



scoring

× low data quality



virtual screening

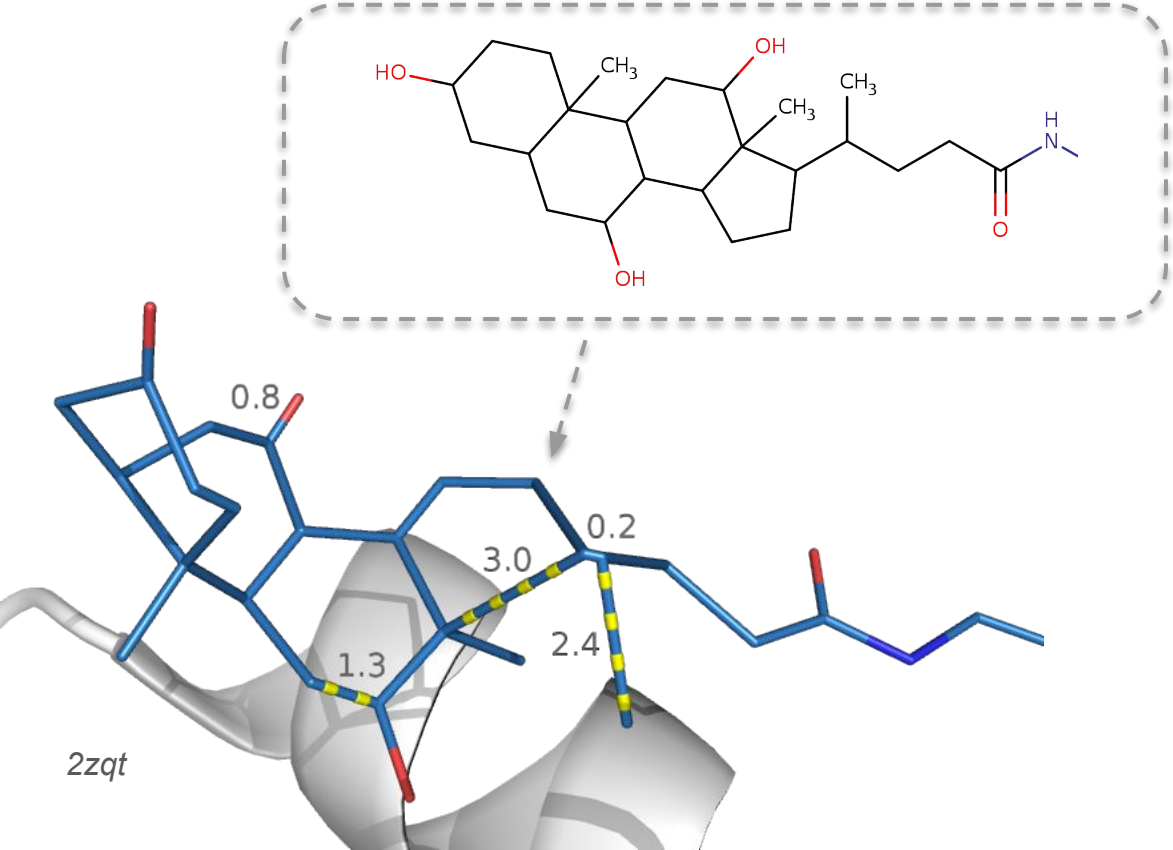
× is almost always 1-label classification

× receptor flexibility

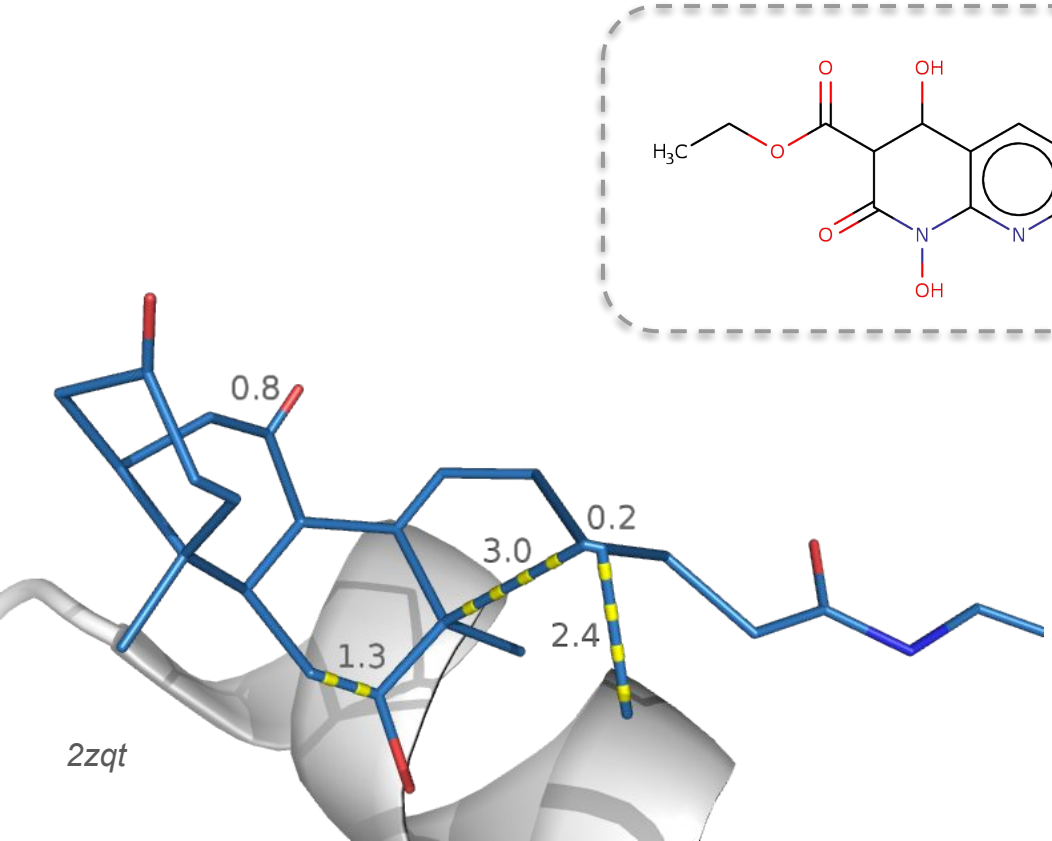
× temperature, solvent, entropy

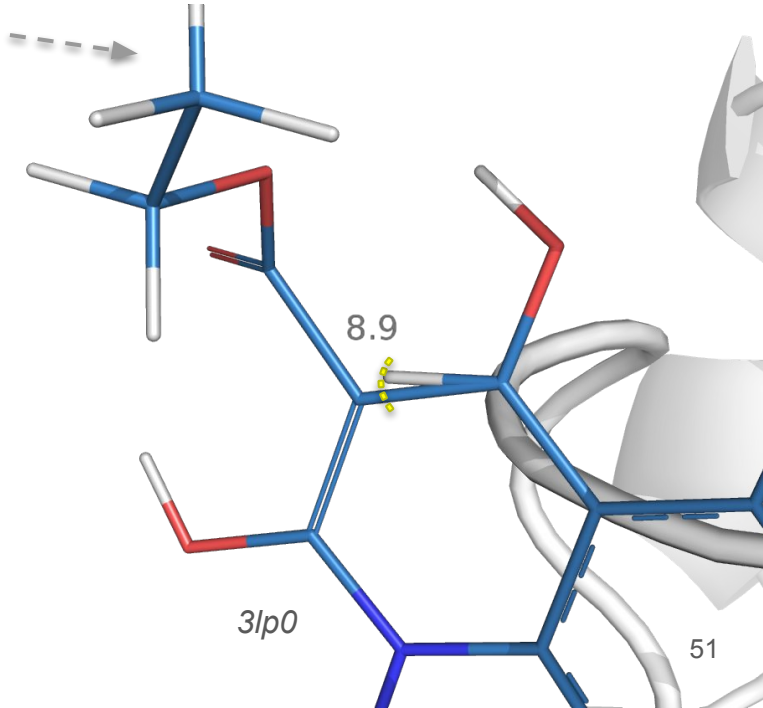
× ligand coordinates are **less accurate** than amino acid ones

Proteins and ligands: problems



Proteins and ligands: problems



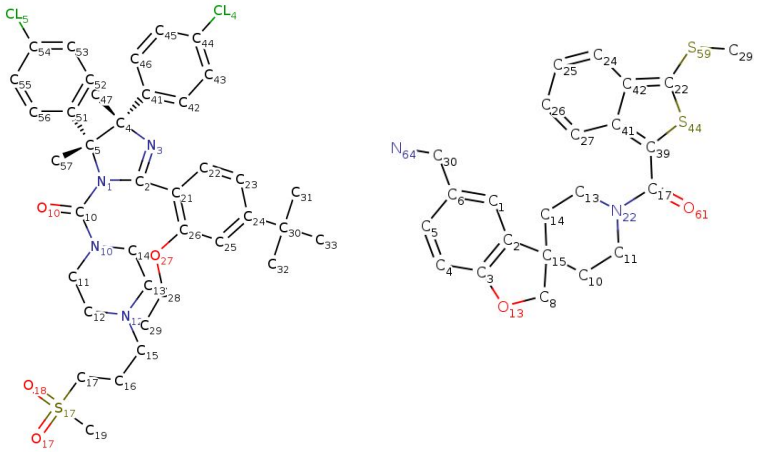


Outline

- ✓ How do molecules look and function?
- ✓ Protein structure and interactions prediction
- ✓ Protein-ligand interactions
- Cheminformatics
- Other applications

Cheminformatics: studying small molecules

What we have



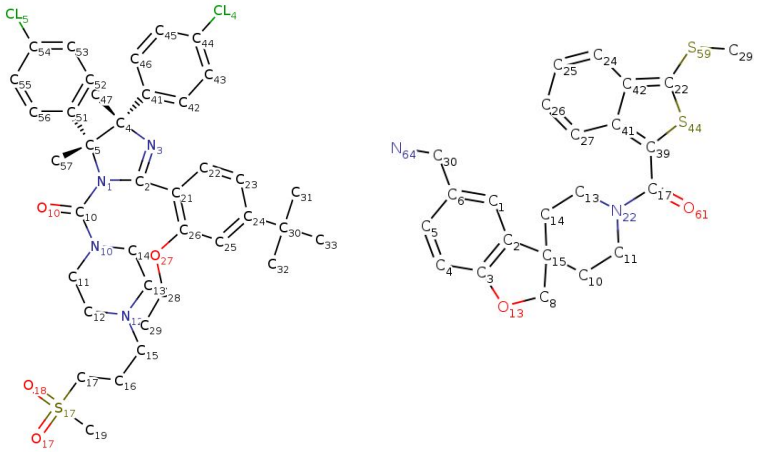
- millions of compounds
- partially labeled

What we want

- virtual screening
- chemical properties mapping
- regression towards binding energy, toxicity, etc
- generate new molecules
- generate synthesis pathways

Cheminformatics: descriptors in 2D

What we have

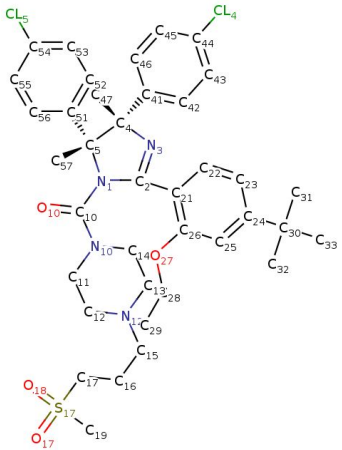


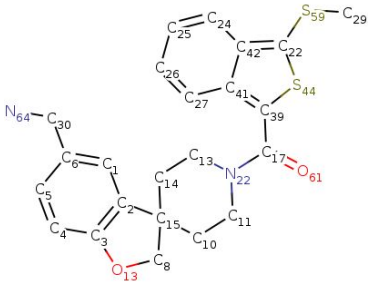
string representation (SMILES)

```
CCOc1cc(ccc1C1=N[C@@](C)(c2ccc(Cl)cc2)[C@](C)(N1C(
=O)N1CCN(CCCS(C)(=O)=O)CC1)c1ccc(Cl)cc1)C(C)(C)C

CSc1sc(C(=O)N2CCC3(COc4ccc(CN)cc34)CC2)c2ccccc12
```

Cheminformatics: descriptors in 2D

What we have





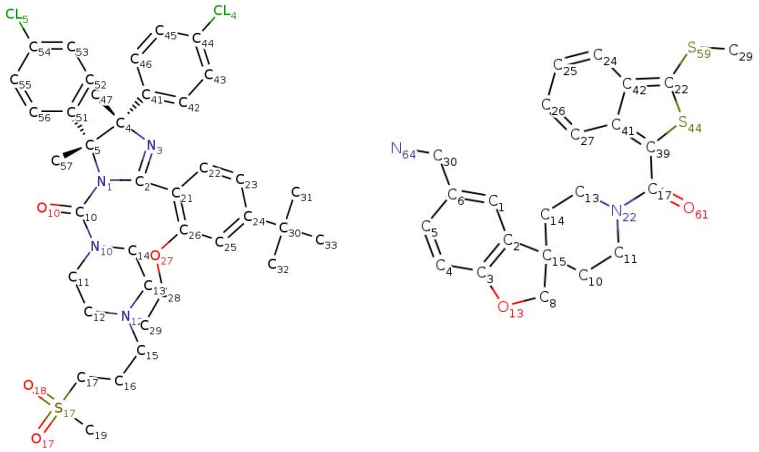
string representation (SMILES)

```
CCOc1cc(ccc1C1=N[C@@](C)(c2ccc(Cl)cc2)[C@](C)(N1C(
=O)N1CCN(CCCS(C)(=O)=O)CC1)c1ccc(Cl)cc1)C(C)(C)C

CSc1sc(C(=O)N2CCC3(COc4ccc(CN)cc34)CC2)c2ccccc12
```

graph representation

Cheminformatics: descriptors in 2D

What we have



string representation (SMILES)

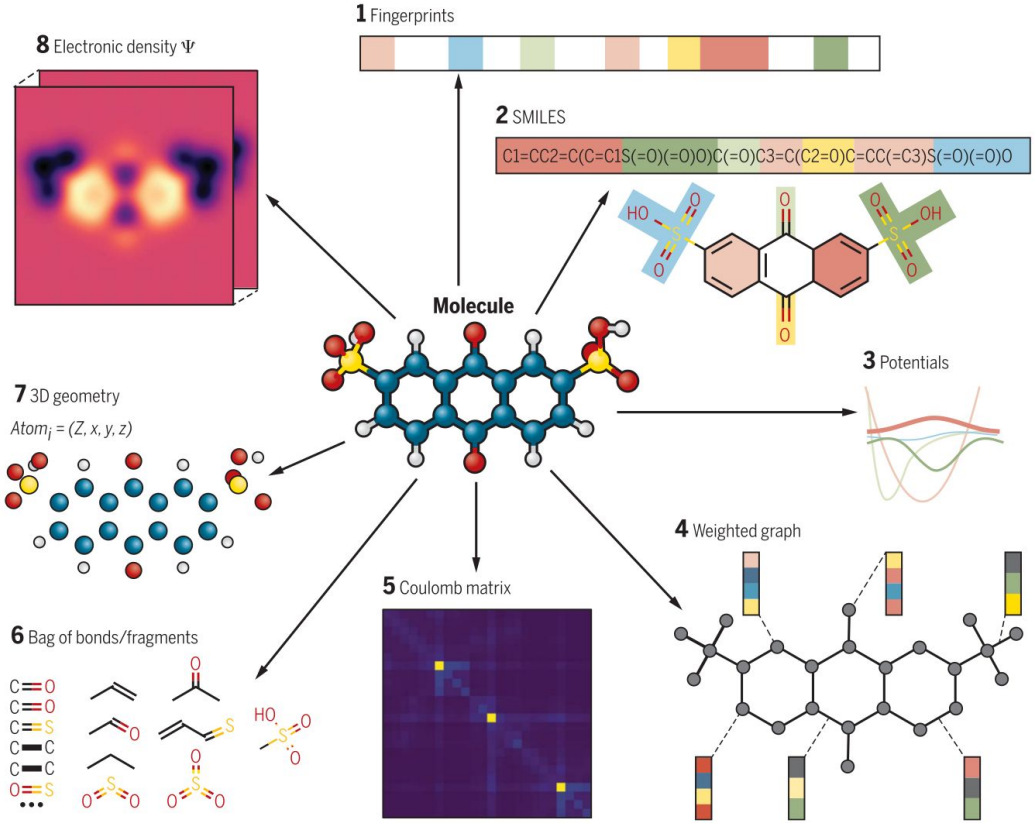
```
CCOc1cc(ccc1C1=N[C@@](C)(c2ccc(Cl)cc2)[C@](C)(N1C(
=O)N1CCN(CCCS(C)(=O)=O)CC1)c1ccc(Cl)cc1)C(C)(C)C

CSc1sc(C(=O)N2CCC3(COc4ccc(CN)cc34)CC2)c2ccccc12
```

graph representation

fingerprints

- presence of particular fragments
- local environment of each atom in the graph

Cheminformatics: descriptors



Sanchez-Lengeling, Aspuru-Guzik. Inverse molecular design using machine learning: Generative models for matter engineering. 2018

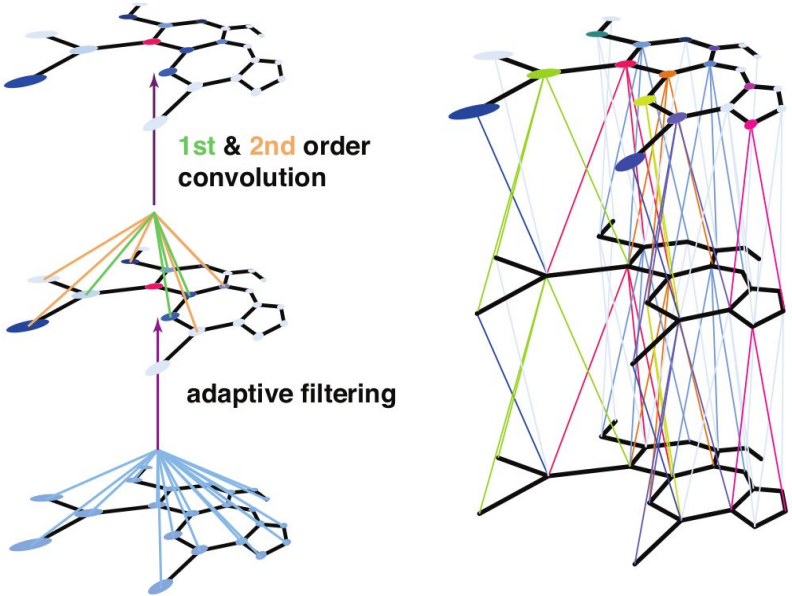
Cheminformatics: better generalization

RNNs on strings

- grammatical validity?

graph nets

- features are associated with weighted nodes
- several ways to define convolution operation
- recurrent architectures

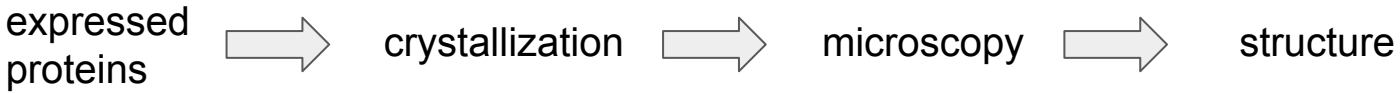


$$L^{(k)} = (W_k \odot \tilde{A}^k) + B_k$$

Zhou, Li. Convolution on Graph: A High-Order and Adaptive Approach, 2017
There are tons of recent papers!

More applications

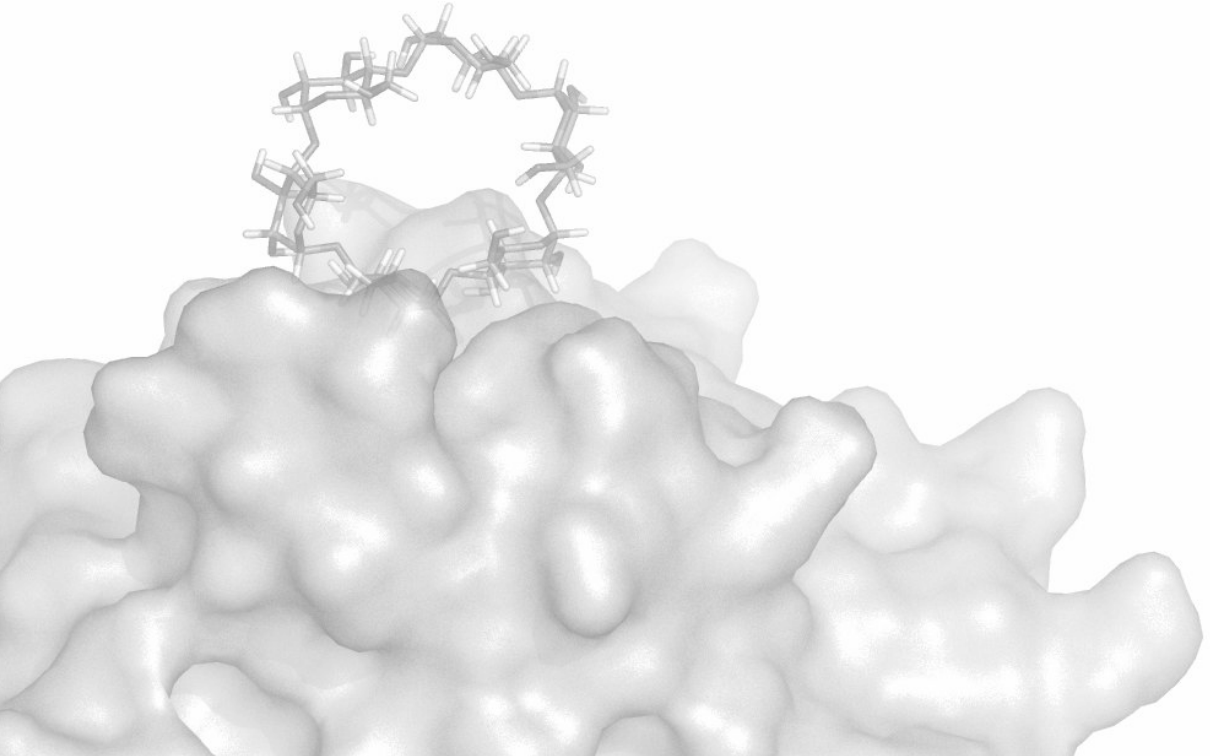
- these steps can be enhanced



- simulations (molecular dynamics, quantum chemistry) can be either “learned”, or analyzed

Thank you for the attention!

Nano-D team of Inria
MIPT



Sergei Grudinin
Vladimir Chupin
Stephan Redon
Leonard Jaillet
Ilya Igashov
Petr Popov
Andreas Eisenbarth