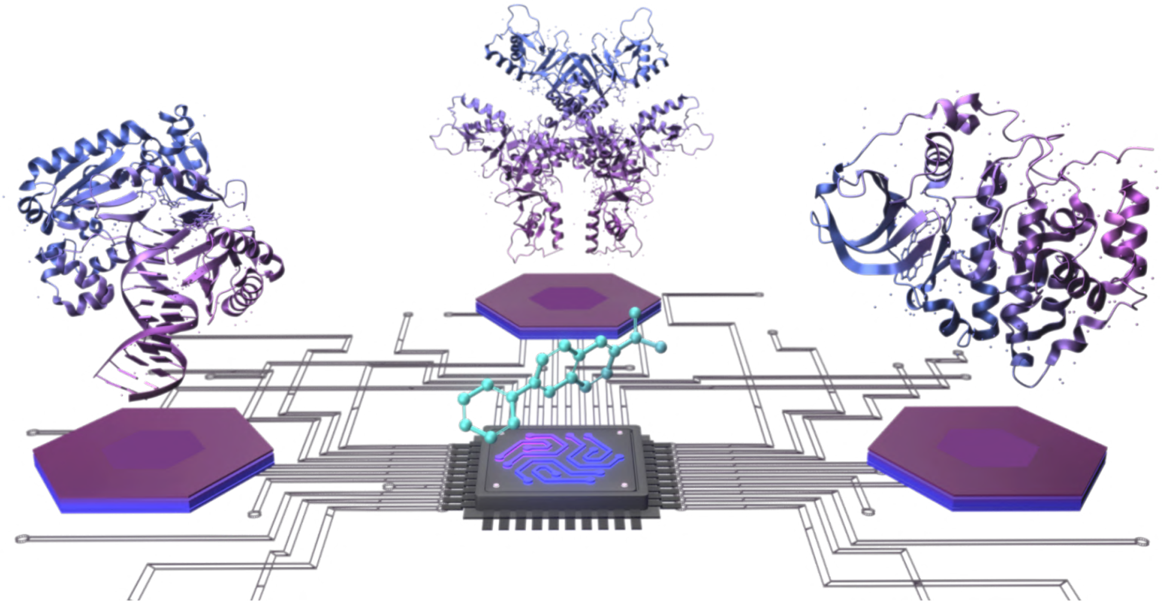
RECEPTOR.AI



AI-based prediction of small molecules' selectivity to highly similar protein variants

Table of contents

Table of contents

Introduction

General scheme of pipeline

General characteristics of the pipeline

Benchmarks

- Target characterisation
- Binding sites prediction
- Chemical space

Main experiment

- Molecular dynamics simulation
- Generation and comparison of the binding pockets pharmacophores
  - 3D pharmacophore fingerprint
  - Shape fingerprints
  - Tanimoto similarities between fingerprints
  - Clustering of similarity matrix and choosing “selective” frames
- Virtual screening using diffusion-based AI model
- Comparison with conventional docking
- AI model benchmarking
- Results

Conclusions

- Appendix: Performance metrics cheat sheet
  - Condition types
  - Result types
  - Metrics’ description

Introduction

The ability to discriminate between several highly similar protein targets is crucial for modern precision and individualised medicine. Modern drugs are expected to be laser-focused on disease-related protein variants, which are specific to a particular tumour, tissue, cell type or the specific patient genotype while having no adverse off-target effects.

In order to achieve the goal of ultra-selectivity to similar protein variants, we created a unique technology of selectivity prediction for the candidate compounds. It is based on molecular dynamics simulations of the target of interest (on-target) and its off-targets; the algorithm for selecting "selective" frames of the on-target trajectory (it finds most dissimilar ones based on comparing pharmacophores and shape of on/off-binding pockets; diffusion model for AI-docking of compounds and AI-rescoring model for estimation of binding poses (it was trained using compounds with known biological activities).

This stack of technologies is incorporated into the Receptor.AI drug discovery platform and is available for any protein for which target and off-target variants could be determined.

Our platform is able to design ultra-selective small molecules for active and allosteric sites alike and operates even for the most challenging targets which lack known ligands or have poorly resolved structures.

In order to test our platform we selected a subset of highly similar proteins from a popular family of drug targets: Janus tyrosine kinases (JAKs). These proteins are involved in a multitude of diseases, from cancer to cardiovascular diseases, inflammation and metabolic disorders. There is a large amount of known ligands with reliable activity data for these proteins, which allows us to perform comprehensive and unbiased benchmarking of our technologies.

General scheme of pipeline

1. The workflow starts with defining the target protein and any number of explicit similar off-target proteins.
2. The target and all off-targets are subject to all-atom MD simulations in their native environment.
   - The MD trajectories are processed by the proprietary clustering algorithm, which extracts the sets of representative protein

conformations which account for overall protein flexibility and the local dynamics of the binding pocket of interest.

   - For each protein, the ensemble of the most relevant conformations is formed.

3. For each conformation of the target and off-target proteins, the pharmacophore model of the binding pocket is generated. These models are then aligned and analysed by our proprietary feature detection algorithm, which emphasises the pharmacophore bits that are unique and specific to the target protein.

   - The resulting model is called a differential pharmacophore and represents both structural and dynamical distinctive features of the binding pocket in the target protein.

   - The differential pharmacophore is then used in a diffusion-based AI model.

4. The diffusion-based AI model is trained to predict 3D ligand conformers that match the binding pocket.

   - The model starts from the random distribution of the ligand poses and estimates the best matching pose by modelling the process of inverted diffusion.

   - The model works in the same manner as image-generating AI that "emerges" the picture from the random pixels.

   - Diffusion-based model is trained on ~900M molecules from the Zinc database. For each molecule, ~100 3D conformers are generated, and for each conformer, an artificial binding pocket is simulated.

5. This model is used for the high-throughput screening of the virtual chemical space. For each compound, the model "emerges" the conformer with the best fit to the binding pocket.

6. Finally, the results of screening with a diffusion-based model rescore using the set of custom AI scoring functions, and a set of active and selective compounds is returned along with the best binding pose for each of them.

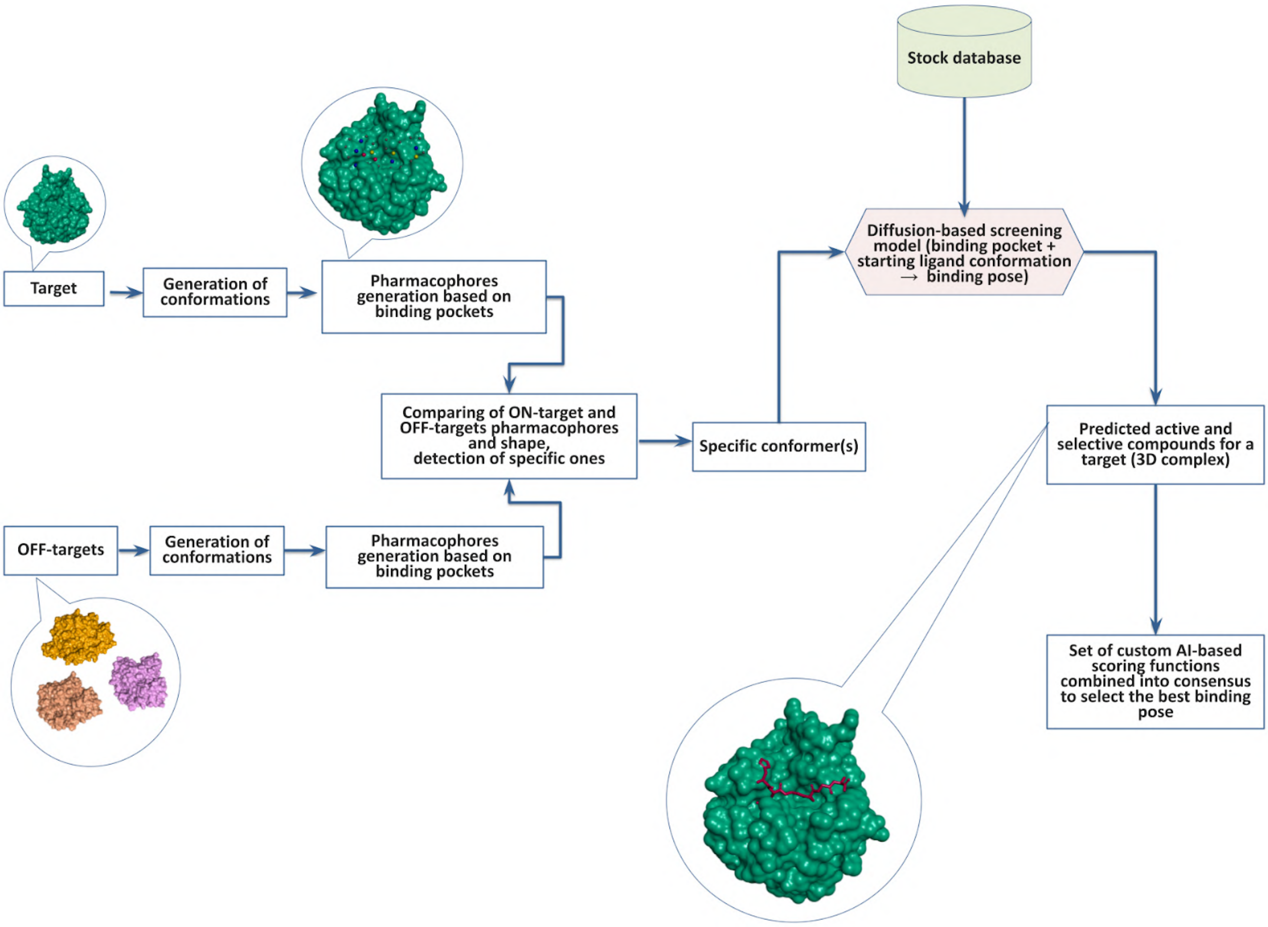

Stock database
Target
Generation of conformations
Pharmacophores generation based on binding pockets
Diffusion-based screening model (binding pocket + starting ligand conformation → binding pose)
Comparing of ON-target and OFF-targets pharmacophores and shape, detection of specific ones
Specific conformer(s)
Predicted active and selective compounds for a target (3D complex)
OFF-targets
Generation of conformations
Pharmacophores generation based on binding pockets
Set of custom AI-based scoring functions combined into consensus to select the best binding pose

Figure 1. The scheme of the pipeline used in this study.

General characteristics of the pipeline

- Length of molecular dynamics simulation of on/off-target: ~500 ns.
- Number of frames in the equilibrated parts of trajectories: ~1000 per trajectory.
- Number of frames after initial clustering: ~200 per trajectory.
- Number of frames after pharmacophore and shape comparison and secondary clustering: ~7 "selective" frames subject to ensembled docking using the diffusion-based model.
- Chemical space: stock compounds ~5M (~2.5M after basic ADMET filtering).
- Docking runs with AI-rescoring: ~5M
- Visual inspection ~2000 compounds per frame (~14000 altogether).

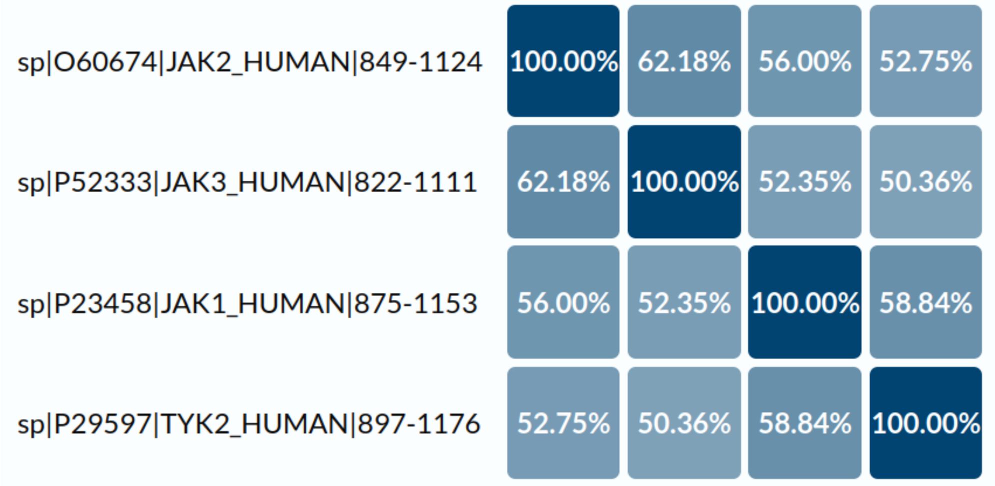
- Final compounds for biological testing: ~500.

Benchmarks

Target characterisation

The selectivity of compounds was determined against JAKs - JAK1, JAK2, JAK3, TYK2 (kinase domains, JH1).

The overall sequence identity between the family members is rather small in the case of JAKs (50-60%)(Fig. 2).



sp|O60674|JAK2_HUMAN|849-1124	100.00%	62.18%	56.00%	52.75%
sp|P52333|JAK3_HUMAN|822-1111	62.18%	100.00%	52.35%	50.36%
sp|P23458|JAK1_HUMAN|875-1153	56.00%	52.35%	100.00%	58.84%
sp|P29597|TYK2_HUMAN|897-1176	52.75%	50.36%	58.84%	100.00%

Figure. 2. Sequence identity matrices of JAK proteins used in the study.

Despite these differences, the family could be characterised as highly similar proteins when comparing their functional binding pockets.

The active site of JAKs contains 18 functionally important residues, but only 2 of them are variable, while the rest is either identical or highly conservative. These residues are shown in Fig. 3 and 4.

```
JAK2        LKFLQQLGKGNFGSVEMCRYDPLQDNTGEVVAVKKLQ-HSTEEHLRDFEREIEILKSLQH
JAK3        LKYISQLGKGNFGSVELCRYDPLGDNTGALVAVKQLQ-HSGPDQQRDFQREIQILKALHS
JAK1        LKRIRDLGEGHFGKVELCRYDPEGDNTGEQVAVKSLKPESGGNHIADLKKEIEILRNLYH
TYK2        LKKIRDLGEGHFGKVSLYCYDPTNDGTGEMVAVKALKADCGPQHRSGWKQEIDILRTLYH
            ** : :**:*:**.*.:  ***  *.**  **** *: ..  ::  . ::**:**: *

JAK2        DNIVKYKGVCYSAGRRNLKLIMEYLPYGSLRDYLQKHKERIDHIKLLQYTSQICKGMEYL
JAK3        DFIVKYRGVSYGPGRQSLRLVMEYLPSGCLRDFLQRHRARLDASRLLLYSSQICKGMEYL
JAK1        ENIVKYKGICTEDGGNGIKLIMEFLPSGSLKEYLPKNKNKINLKQQLKYAVQICKGMDYL
TYK2        EHIIKYKGCCEDQGEKSLQLVMEYVPLGSLRDYLPRH--SIGLAQLLLFAQQICEGMAYL
            : *:**:* .   * ..::*:**::* *.*:::* ::   :.  : * :: ***:** **

JAK2        GTKRYIHRDLATRNILVENENRVKIGDFGLTKVLPQDKEYYKVKEPGESPIFWYAPESLT
JAK3        GSRRCVHRDLAARNILVESEAHVKIADFGLAKLLPLDKDYYVVREPGQSPIFWYAPESLS
JAK1        GSRQYVHRDLAARNVLVESEHQVKIGDFGLTKAIETDKEYYTVKDDRDSPVFWYAPECLM
TYK2        HAQHYIHRDLAARNVLLDNDRLVKIGDFGLAKAVPEGHEYYRVREDGDSPVFWYAPECLK
             ::: :*****:**:*::.:  ***.****:* :  .::** *::  :**:******.*

JAK2        ESKFSVASDVWSFGVVLYELFTYIEKSKSPPAEFMRMIGNDKQGQMIVFHLIELLKNNGR
JAK3        DNIFSRQSDVWSFGVVLYELFTYCDKSCSPSAEFLRMMGCERDVP-ALCRLLELLEEGQR
JAK1        QSKFYIASDVWSFGVTLHELLTYCDSDSSPMALFLKMIGPT-HGQMTVTRLVNTLKEGKR
TYK2        EYKFYYASDVWSFGVTLYELLTHCDSSQSPPTKFLELIGIA-QGQMTVLRLTELLERGER
            :  *   ********.*:**:*: :.. ** : *:.::*   .    : :* : *:.. *

JAK2        LPRPDGCPDEIYMIMTECWNNNVNQRPSFRDLALRVD---------------
JAK3        LPAPPACPAEVHELMKLCWAPSPQDRPSFSALGPQLDMLWSGSRGCETHAFT
JAK1        LPCPPNCPDEVYQLMRKCWEFQPSNRTSFQNLIEGFEALL------------
TYK2        LPRPDKCPCEVYHLMKNCWETEASFRPTFENLIPILKTVHEKY---------
            ** *  ** *:: :*  **  . . * :*  *   ..
```

🟪 - identical AA residues

🟩 - highly conservative AA residues

🟦 - variable AA residues

Figure 3. Sequence alignment of JAKs used in this study.

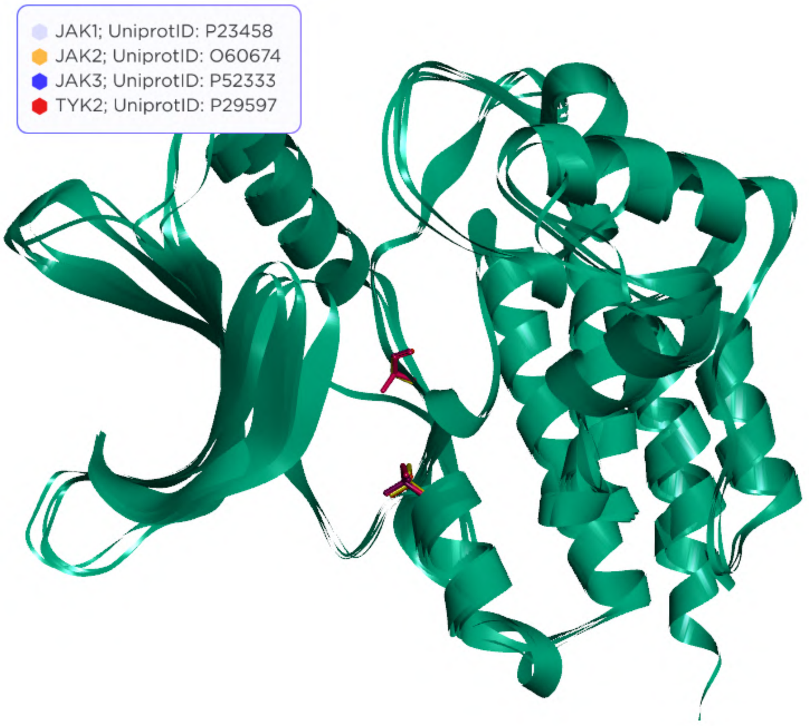


Figure 4. Structural alignment of JAKs used in this study. Sidechains of the variable residues are shown.

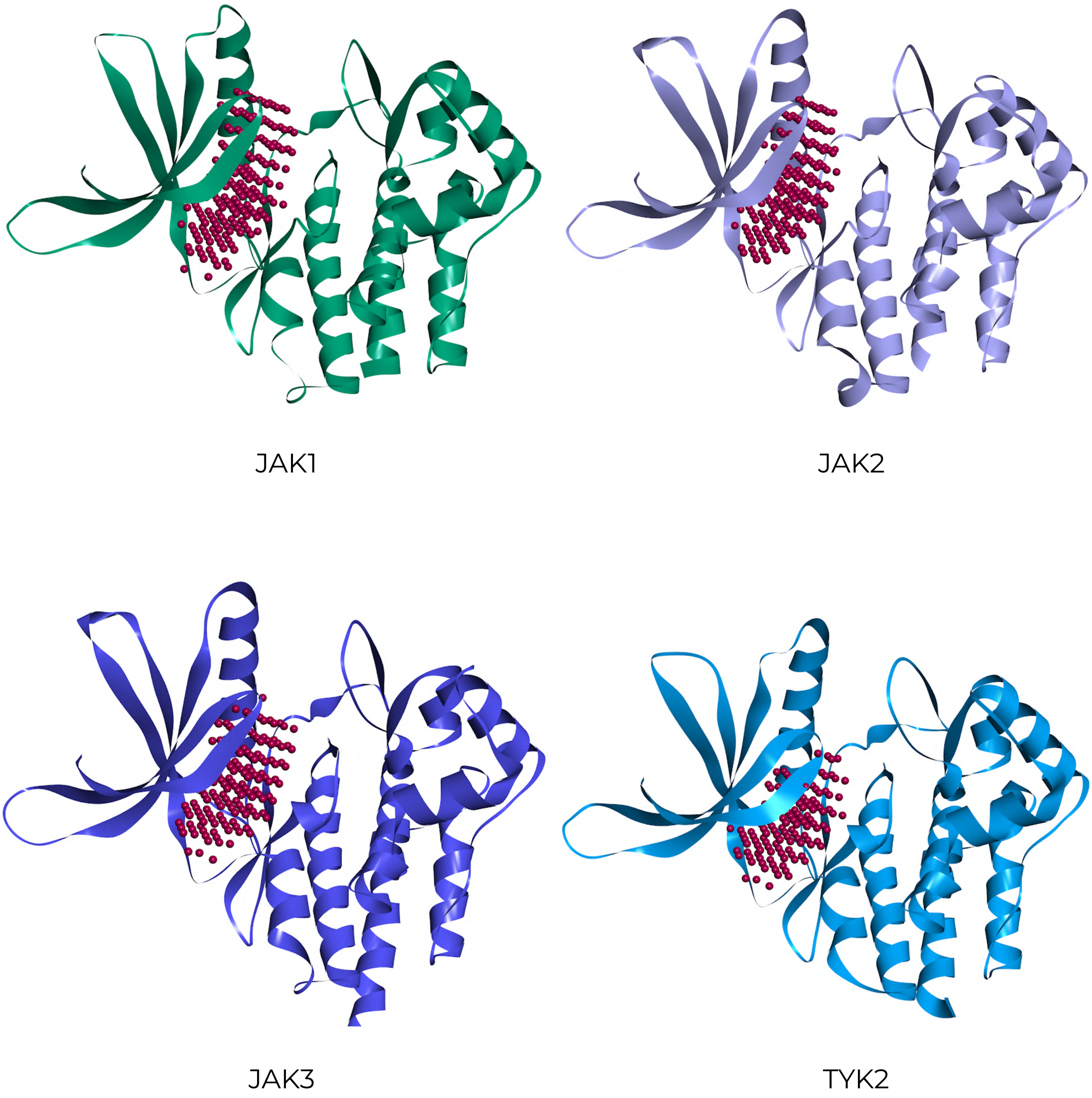
Our selectivity prediction technique emphasises the differences in a few variable residues automatically based on sequence and structural similarity between target and off-target proteins.

Binding sites prediction

For binding site prediction, we used several algorithms:

- PUResNet - predicting protein-ligand binding sites using deep convolutional neural networks.
- pyKVFinder - Python package for biomolecular cavity detection and characterisation.
- fpocket - protein pocket detection algorithm based on Voronoi tessellation.
- mdpocket - the extension of fpocket to analyse conformational ensembles of proteins in MD trajectories.

Compounds with known activities to JAKs, which are available in the public databases interact with ATP-binding sites of kinases only. That's why in this study we used ATP-binding sites predicted using PUResNet that coincide with the ATP-binding pocket (Fig. 5).



JAK1

JAK2

JAK3

TYK2

Figure 5. ATP-binding sites of JAKs used in this study. The volume of the predicted binding pocket is shown by spheres.

In addition to this, we also demonstrated that used techniques of pocket prediction are able to detect the allosteric binding sites for subsequent rational design of compounds which bind to them. However, due to the lack of publicly available data about allosteric kinase binders we didn't use allosteric sites for benchmarking.

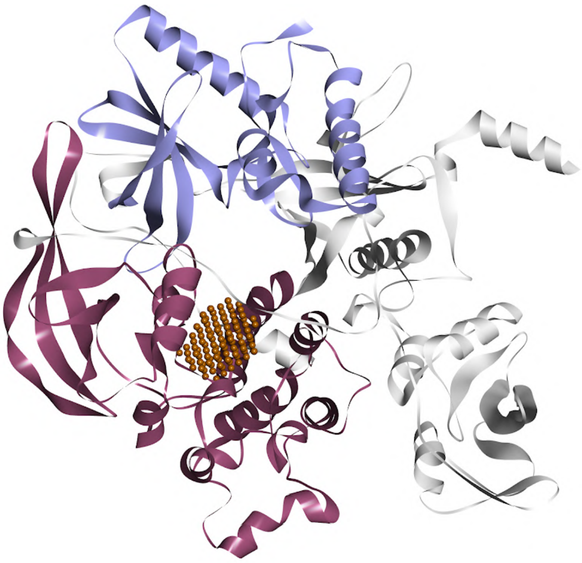


Figure 6. Predicted allosteric binding sites of JAK2. The Pseudokinase domain is shown in dark red, the kinase domain is blue, SH2 and FERM domains are grey. The allosteric site in the pseudokinase domain is shown.

Amino acid residues of all binding pockets are detailed in Table 1.

Table 1. Amino acid residues of all selected JAKs binding pockets.

Pocket	Residues
JAK1	881 882 884 885 887 889 906 908 925 938 956 957 958 959 960 962 963 966 1003 1007 1008 1010 1020 1021 1023 1024
JAK2	855 856 858 859 861 863 880 882 898 911 929 930 931 932 933 935 936 939 976 980 981 983 993 994 996 997
JAK2 (allosteric 1)	671 672 673 674 675 677 678 703 704 707 711 712 714 715 716 718 719
JAK2 (allosteric 2)	529 530 531 532 535 592 593 596 597 600 669 670 671 672 700 701 702 703 704 705 706 730

JAK3	828 829 831 832 834 836 853 855 871 884 902 903 904 905 906 908 909 912 949 953 954 956 966 967 969 970
TYK2	903 904 906 907 909 911 928 930 947 960 978 979 980 981 982 984 985 988 1023 1027 1028 1030 1040 1041 1043 1044

Chemical space

6830 compounds with known activity against the JAKs family were taken from the ChEMBL database. All these compounds have activity against at least two kinases from the family, which allows us to assess their selectivity.

The number of selective/non-selective compounds for each kinase according to this criterion in available experimental data is shown in Table 2.

Table 2. The number of selective/non-selective compounds for each target kinase.

	On-/Off-target
JAK1	1866 / 2321
JAK2	780 / 4147
JAK3	432 / 2329
TYK2	102 / 1242

It is clearly seen that the dataset is significantly skewed. The largest number of compounds is found for JAK1, which is the most commonly used as a primary drug target in its family.

In order to establish a reliable measure of compound selectivity, we plotted the ratio of the number of selective to non-selective compounds as a function of their experimental activity ratio (Figure 7).

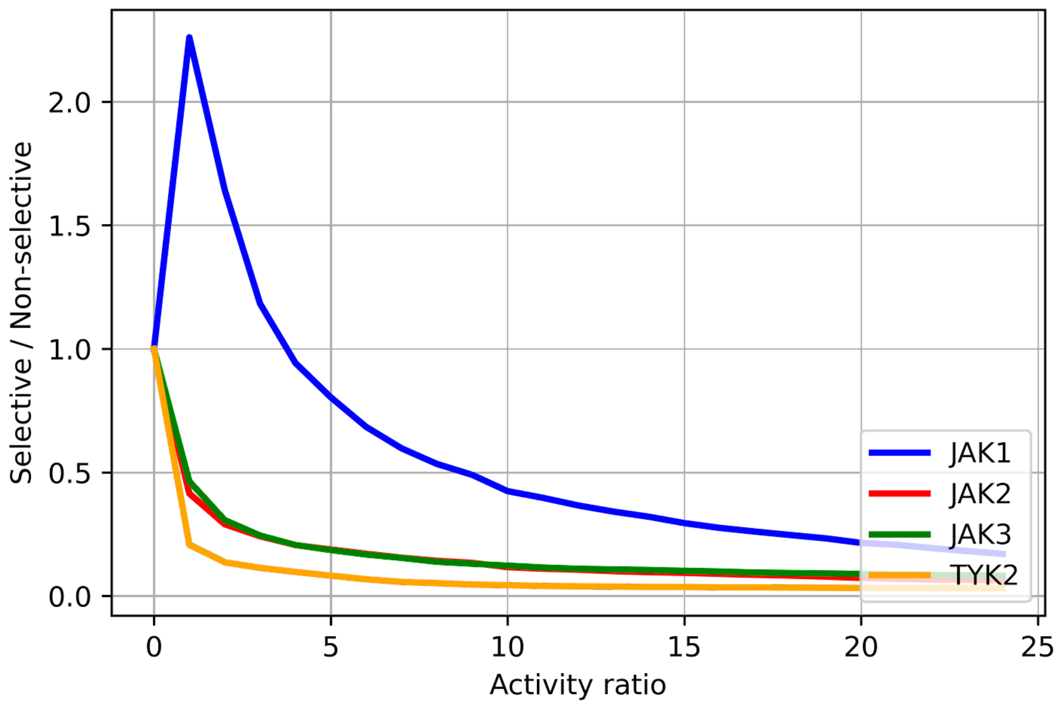


Figure 7. The plot Selective / Non-selective versus Activity ratio for JAKs.

An activity ratio of ~5 is a good threshold for establishing the compound's selectivity for JAKs.

Main experiment

Molecular dynamics simulation

All-atom molecular dynamics simulations were performed in GROMACS 2023 with CHARMM36 force field. A rigorous equilibration procedure with four stages of gradually reducing position restraints was utilised to keep the structure from undesirable conformational changes in the first stages of dynamics. Data was collected from at least ~400 ns of the equilibrated parts of trajectories.

Analysis of obtained trajectories

The equilibrated part of each trajectory was initially clustered by RMSD of selected amino acids (Table 1) forming the ATP-binding site using the agglomerative clustering while keeping only the cluster centres which are at least 0.1 nm apart.

In-house tools based on Pteros molecular modelling library were used. Table 3 summarises information about the clustering.

Table 3. Statistics of MD trajectories and their clusterisation.

	Number of clusters	Number of atoms	Number of residues
JAK1	155	4630	287
JAK2	138	4889	297
JAK3	89	4617	290
TYK2	168	4685	291

The Elbow method was used for selecting the optimal number of clusters for an agglomerative clustering algorithm. During this process, KMeans clustering is performed first with different parameter K. The point at which the sum of all distances between the cluster centres stops decreasing rapidly is the optimal number of clusters (Figure 8, the approach is described in detail here). The optimal number of clusters are: JAK1 - 4, JAK2, JAK3 and TYK2 - 5.

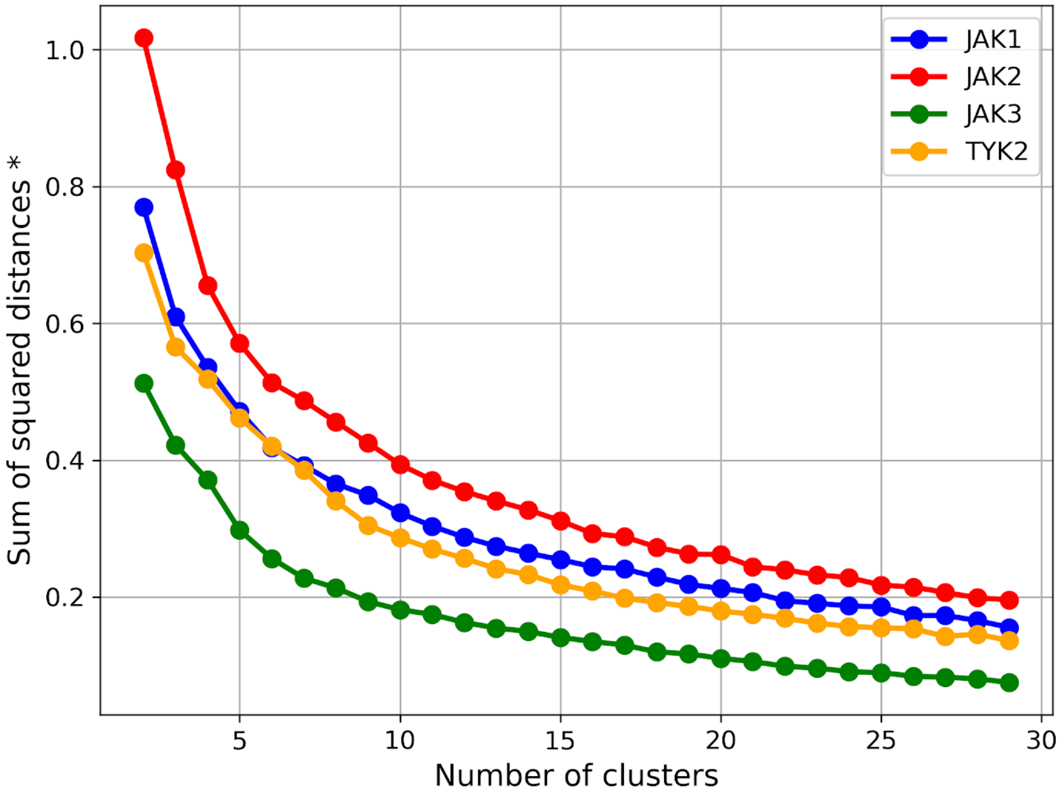
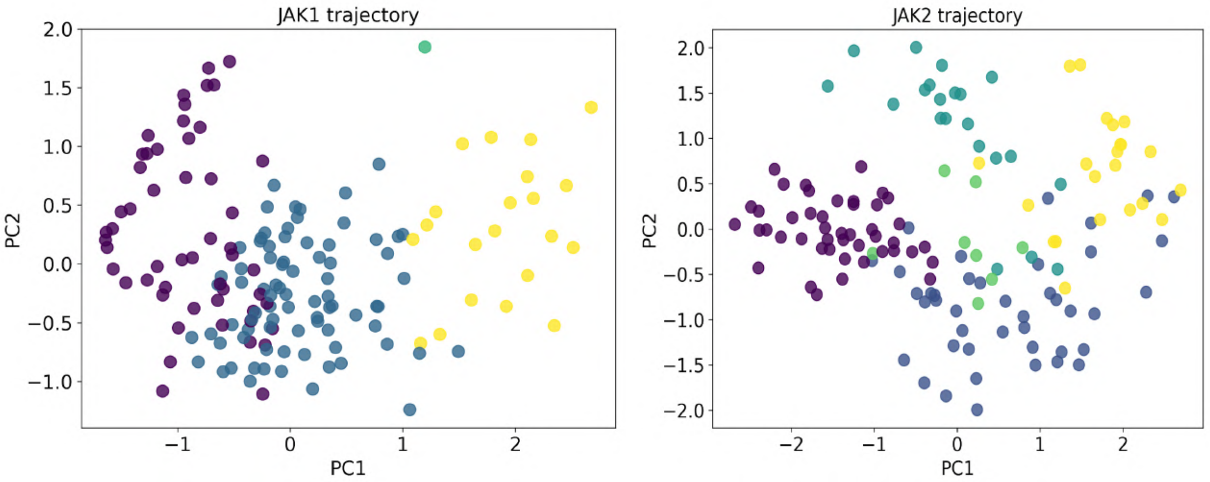


Figure 8. The plots for choosing the optimal number of clusters using the Elbow method. * - sum of squared distances of samples to their closest cluster centre.

Results of agglomerative clustering are shown in Figure 9 as projections on the first two principal components computed from the covariance matrices of the residues used for clustering.



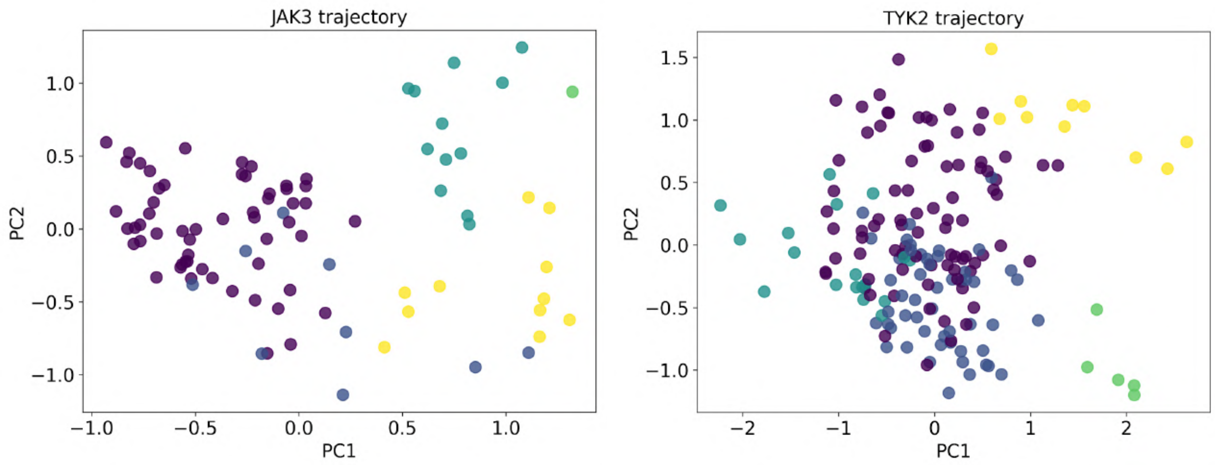


Figure 9. Results of clustering of JAKs trajectories. Each point on the plots corresponds to one initial cluster while the colours denote the final clusters obtained by the Elbow method.

Generation and comparison of the binding pockets pharmacophores

We performed 4 separate experiments for each of the JAKs where the other three kinases were considered as off-targets. For each experiment, the amino acid residues of the on-target and off-target binding sites in each frame were converted to custom 3D pharmacophores. In addition, two types of shape fingerprints were computed.

3D pharmacophore fingerprint

3D pharmacophore fingerprints encode seven types of features (Donor, Acceptor, Aromatic, Hydrophobic, Halogen, Basic, and Acidic) and the distances between them. All possible triangles and quartets of detected features are enumerated (Figure 10) and encoded into a bit vector (Figure 11).

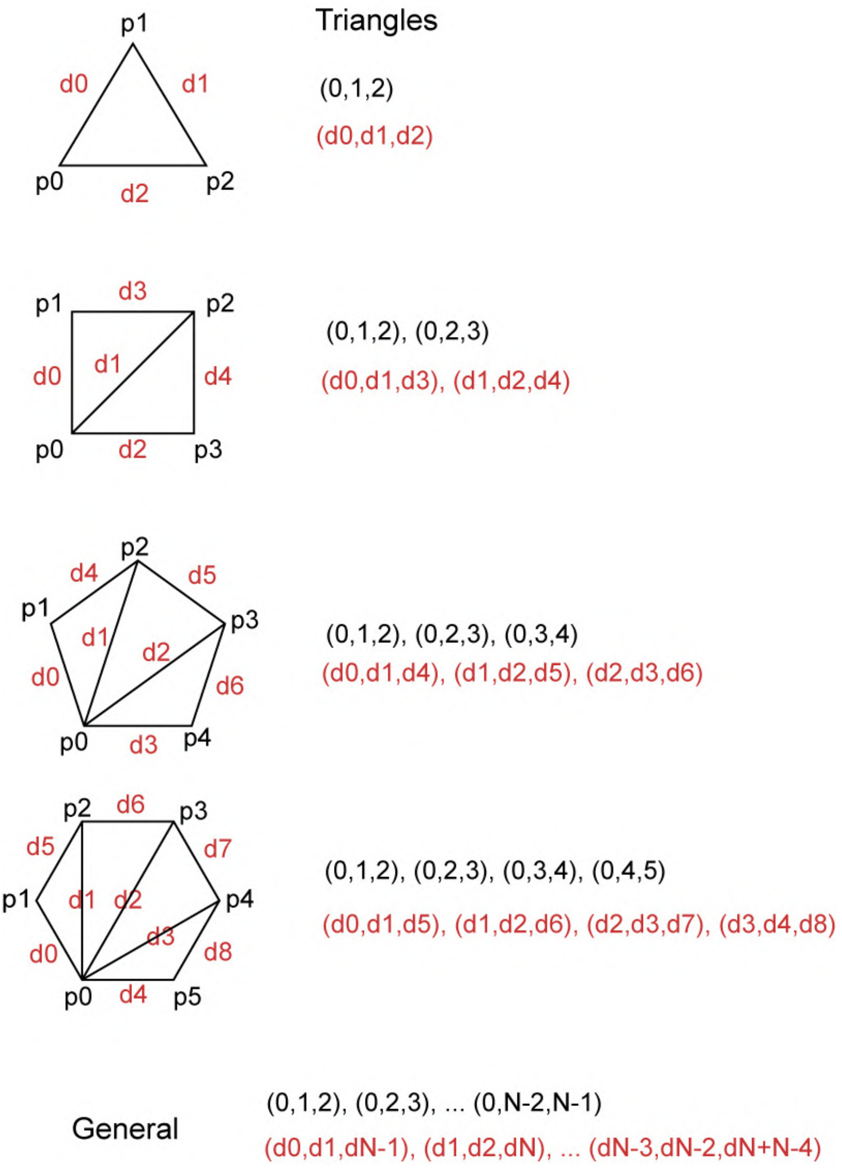
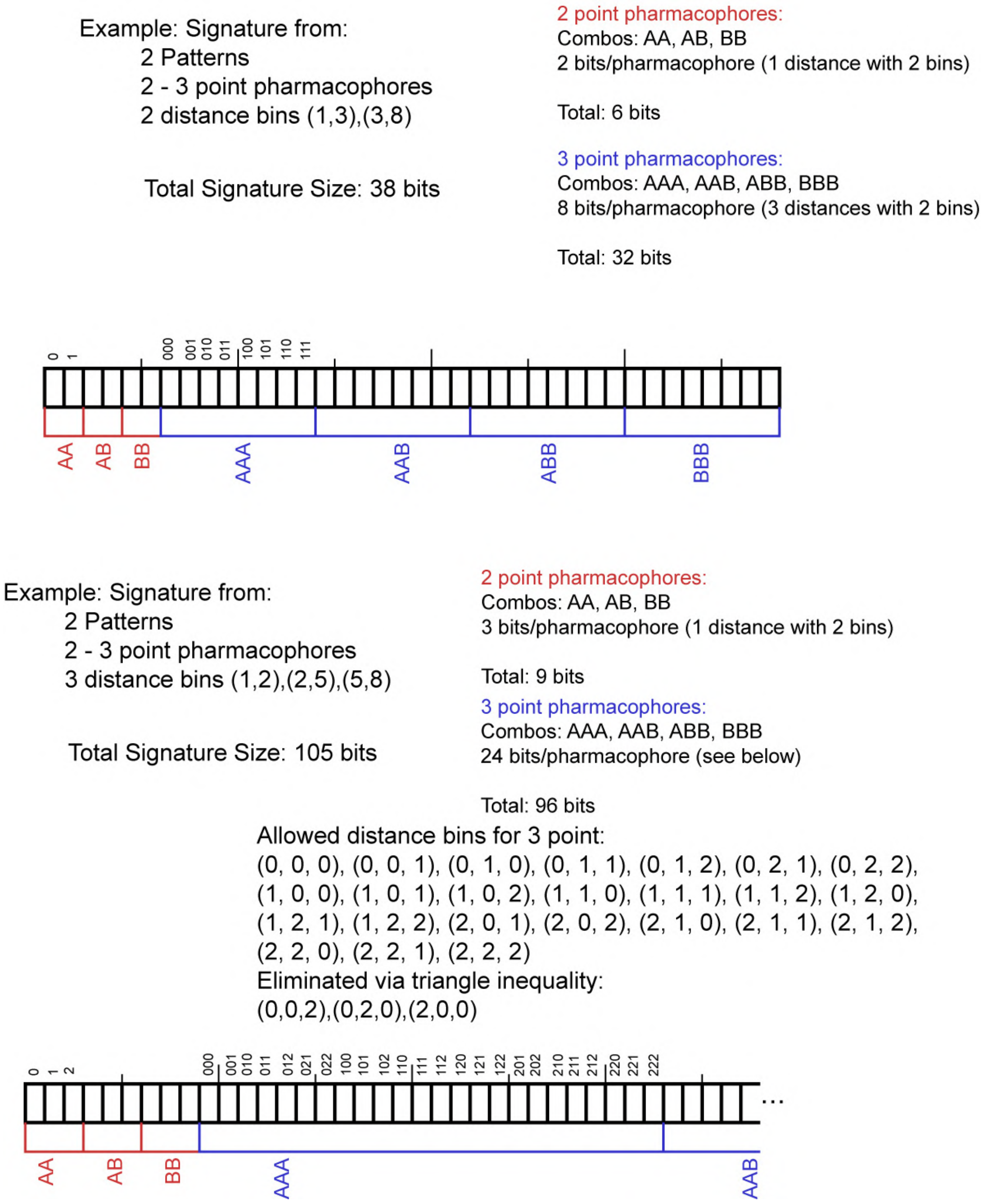


Figure 10. Example of forming triangles of features (only one of the possible combinations for molecules with 3-6 features is demonstrated).



Example: Signature from:
2 Patterns
2 - 3 point pharmacophores
2 distance bins (1,3),(3,8)

Total Signature Size: 38 bits

2 point pharmacophores:
Combos: AA, AB, BB
2 bits/pharmacophore (1 distance with 2 bins)

Total: 6 bits

3 point pharmacophores:
Combos: AAA, AAB, ABB, BBB
8 bits/pharmacophore (3 distances with 2 bins)

Total: 32 bits

Example: Signature from:
2 Patterns
2 - 3 point pharmacophores
3 distance bins (1,2),(2,5),(5,8)

Total Signature Size: 105 bits

2 point pharmacophores:
Combos: AA, AB, BB
3 bits/pharmacophore (1 distance with 2 bins)

Total: 9 bits

3 point pharmacophores:
Combos: AAA, AAB, ABB, BBB
24 bits/pharmacophore (see below)

Total: 96 bits

Allowed distance bins for 3 point:
(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (0, 1, 2), (0, 2, 1), (0, 2, 2), (1, 0, 0), (1, 0, 1), (1, 0, 2), (1, 1, 0), (1, 1, 1), (1, 1, 2), (1, 2, 0), (1, 2, 1), (1, 2, 2), (2, 0, 1), (2, 0, 2), (2, 1, 0), (2, 1, 1), (2, 1, 2), (2, 2, 0), (2, 2, 1), (2, 2, 2)
Eliminated via triangle inequality:
(0,0,2),(0,2,0),(2,0,0)

Figure 11. General principles of encoding of 3D pharmacophore fingerprints.

Shape fingerprints

Shape fingerprints of the binding pockets are encoded by a 3D vesicle model. The first fingerprint is used to describe the volume mismatch between the molecules, while the second - to describe their volume overlap.

Tanimoto similarities between fingerprints

After the calculation of fingerprints, pairwise Tanimoto similarity scores (between all selected conformations of on-target and off-targets) were calculated for each of the kinases. Three similarity matrices were obtained (one per fingerprint), which were averaged into the general similarity matrix.

Clustering of similarity matrix and choosing “selective” frames

Obtained similarity matrices for JAKs frames were clustered with a proprietary algorithm to choose the clusters of “selective” frames that differentiate target and off-targets the best. Those frames are the most distinct from all the off-target frames. An example of such clustering (for JAK1 as an on-target) is shown in Figure 12.

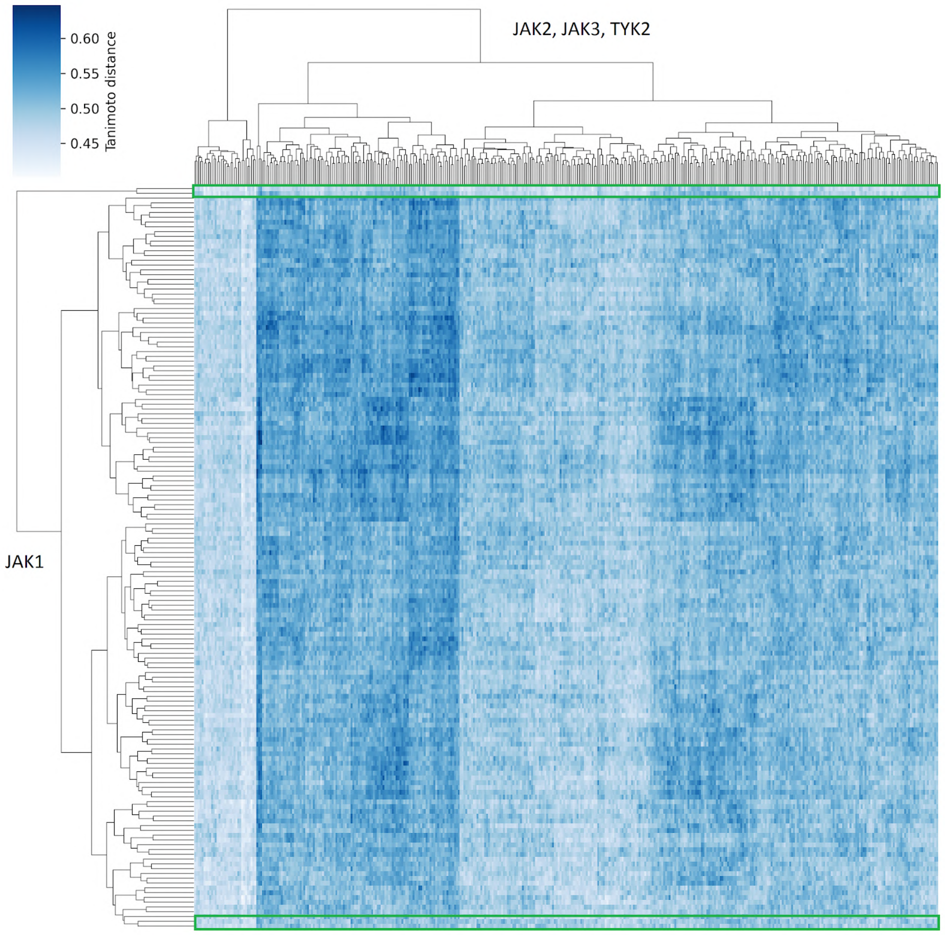
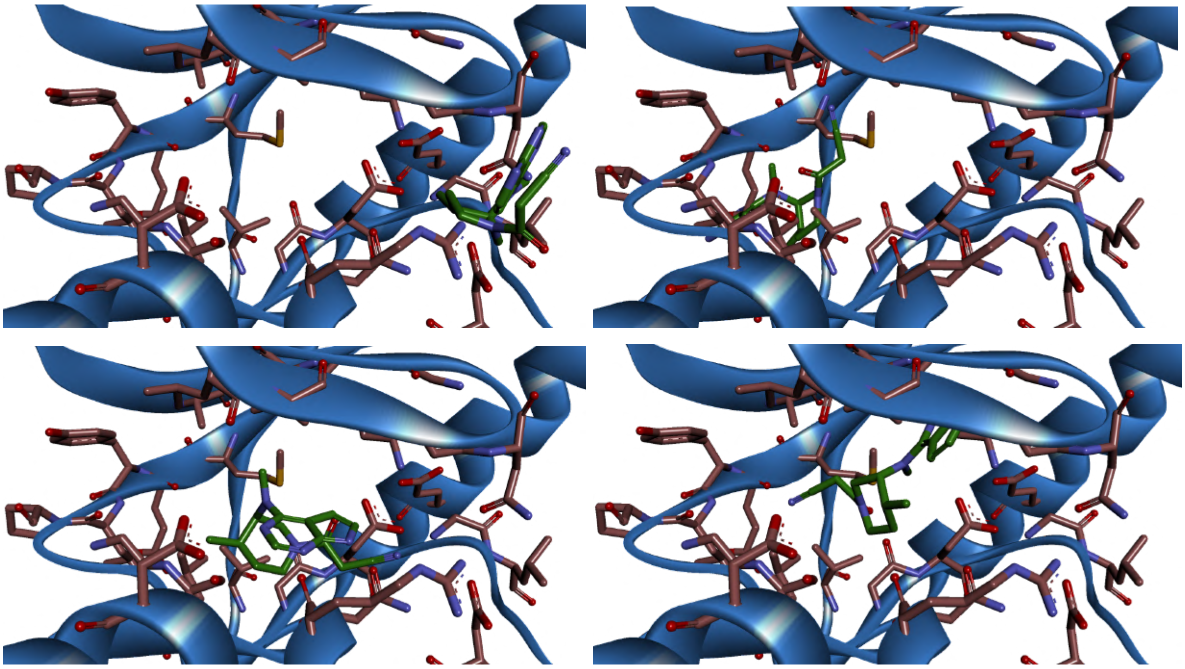


Figure 12. Results of similarity matrix clustering between JAK1 as an on-target and JAK2, LAK3 and TYK2 as off-targets. Green rectangles show the clusters of the most "selective" conformations.

From 5 to 6 "selective" conformations were obtained for each of the studied kinases, which were subject to the virtual screening.

Virtual screening using a diffusion-based AI model

Obtained "selective" conformations were used for docking of test compounds using the diffusion-based generative AI model, which is called ArtiDock. It was trained to fit a molecule into a defined binding pocket as a drop-in replacement of conventional docking. During the training phase, the position of each input ligand in a complex undergoes modifications caused by translational, rotational, and torsional noise, which can be referred to as "forward diffusion" (Figure 13). The model then learns how to reverse the diffusion process and reconstruct the ideal ligand pose from its random orientation and conformation. This method enables the generation of numerous alternative ligand poses throughout the inference process. The best pose is chosen based on the custom scoring function, which takes into account favourable non-covalent interactions and unfavourable contacts within the formed complex.



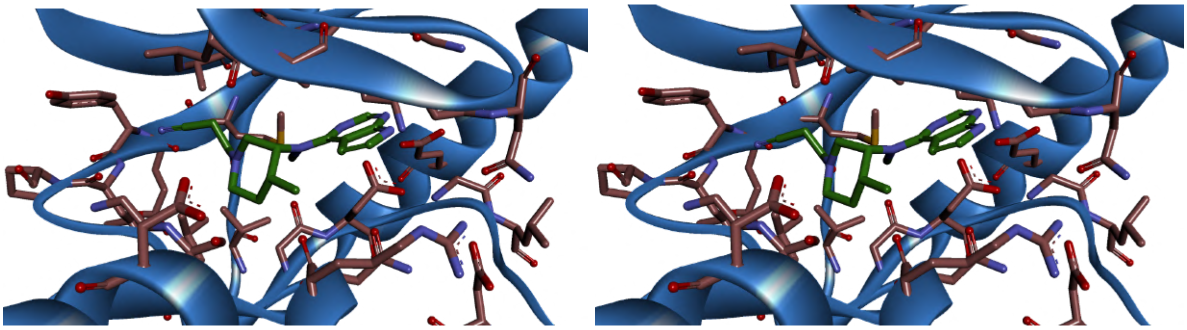


Figure 13. The reverse diffusion from a random ligand pose (top-left) to the final predicted ligand pose (bottom-right). The pocket carbon atoms are in pale red, and the ligand carbon atoms are in green.

Protein-ligand complexes that had been experimentally determined as well as the simulated conformer-specific "synthetic" binding pockets were included in the training dataset. The generation of synthetic data is performed with our proprietary SynProt algorithm, which mimics the statistical distribution of non-covalent interactions in real complexes protein-ligand complexes. The inclusion of the simulated complexes improved the model performance. For the PoseBuster dataset (*10.48550/arXiv.2308.05777)*, the ArtiDock outperforms all the best modern AI docking techniques, including DiffDock and AlphaFold-latest and conventional docking such as Glide, as shown in Figure 14.

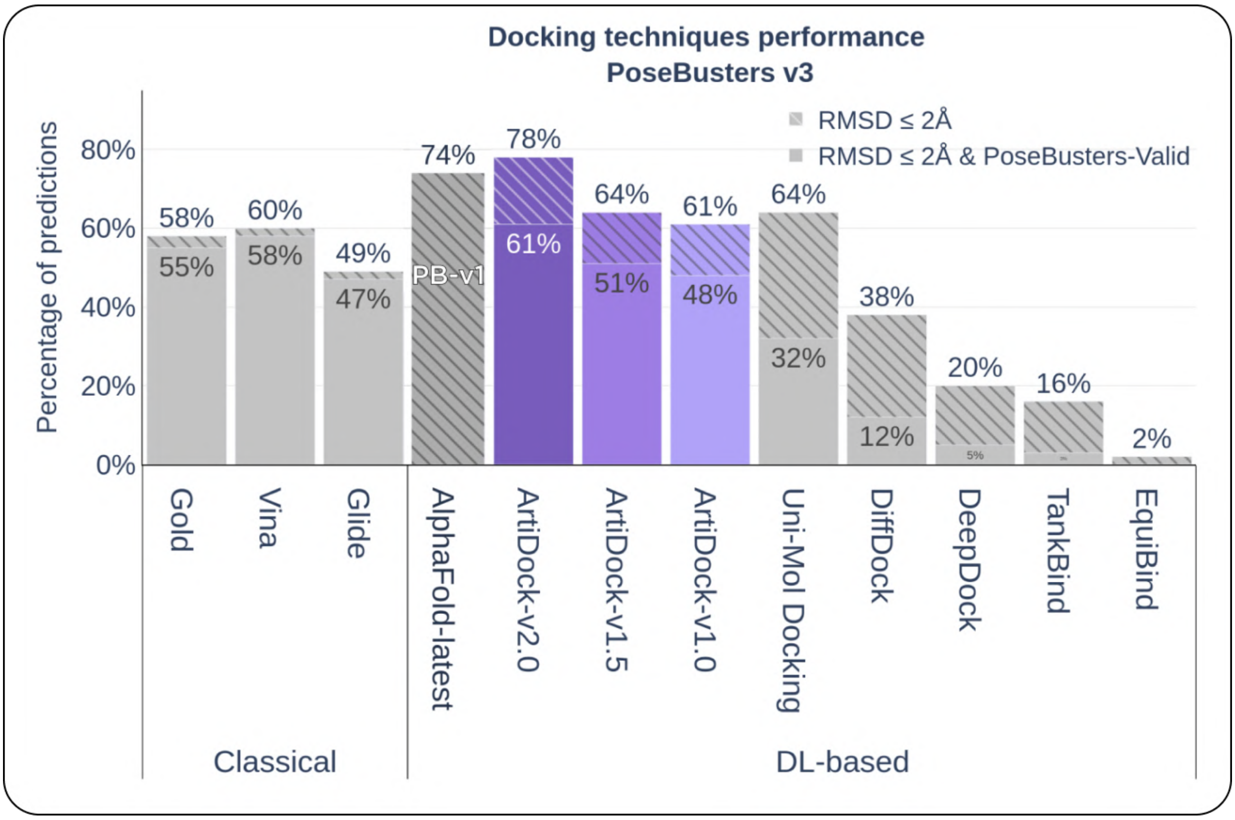


Figure 14. Performance comparison of classical and deep-learning-based docking techniques for the PoseBuster dataset.

Figure 15 shows the comparison of individual PoseBuster structural metrics for ArtiDock, DiffDock and AlphaFold-latest - the newly announced AlphaFold version, which is capable of predicting the protein-ligand complexes. It is clearly seen that ArtiDock significantly outperforms DiffDock for all the metrics, which are not predicted with 100% accuracy. ArtiDock even outperforms AlphaFold-latest for tetrahedral chirality.

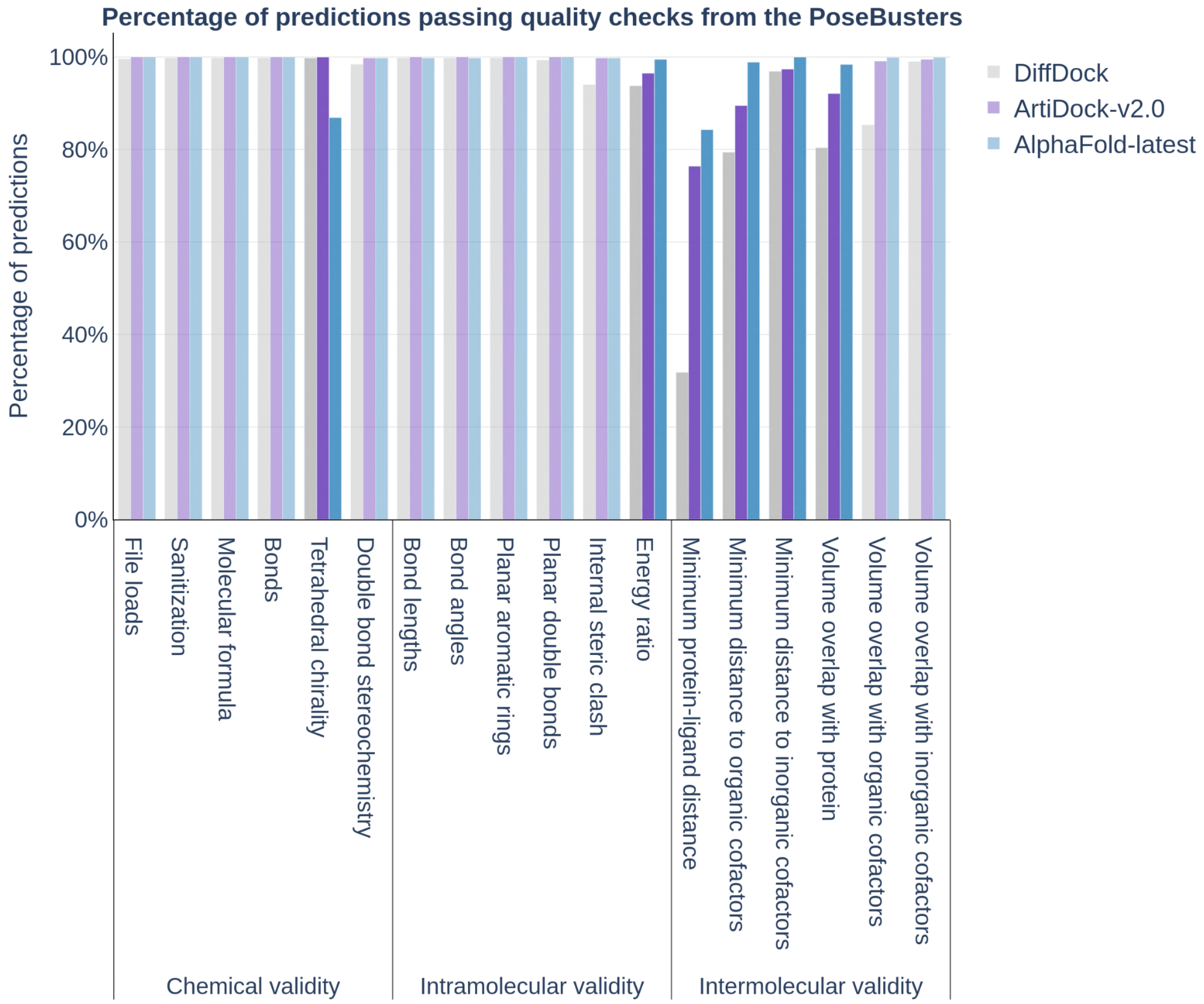


Figure 15. Comparison of individual PoseBuster metrics for ArtiDock, DiffDock and the AlphaFold-latest.

It is worth emphasizing that although AlphaFold-latest provides the best overall prediction quality, this precision comes at the cost of extremely slow inference speed.

Figure 16 shows the inference speeds of all studied techniques. ArtiDock is somewhat slower than some of the "quick and dirty" AI techniques but is still 2 orders of magnitude faster than DiffDock, Vina and Gold while being superior to them in quality. The AlphaFold-latest is expectedly the worst performed here. It is three orders of magnitude slower than ArtiDock, which makes it unusable for the virtual screening - the area where the technology of Receptor.AI shines.

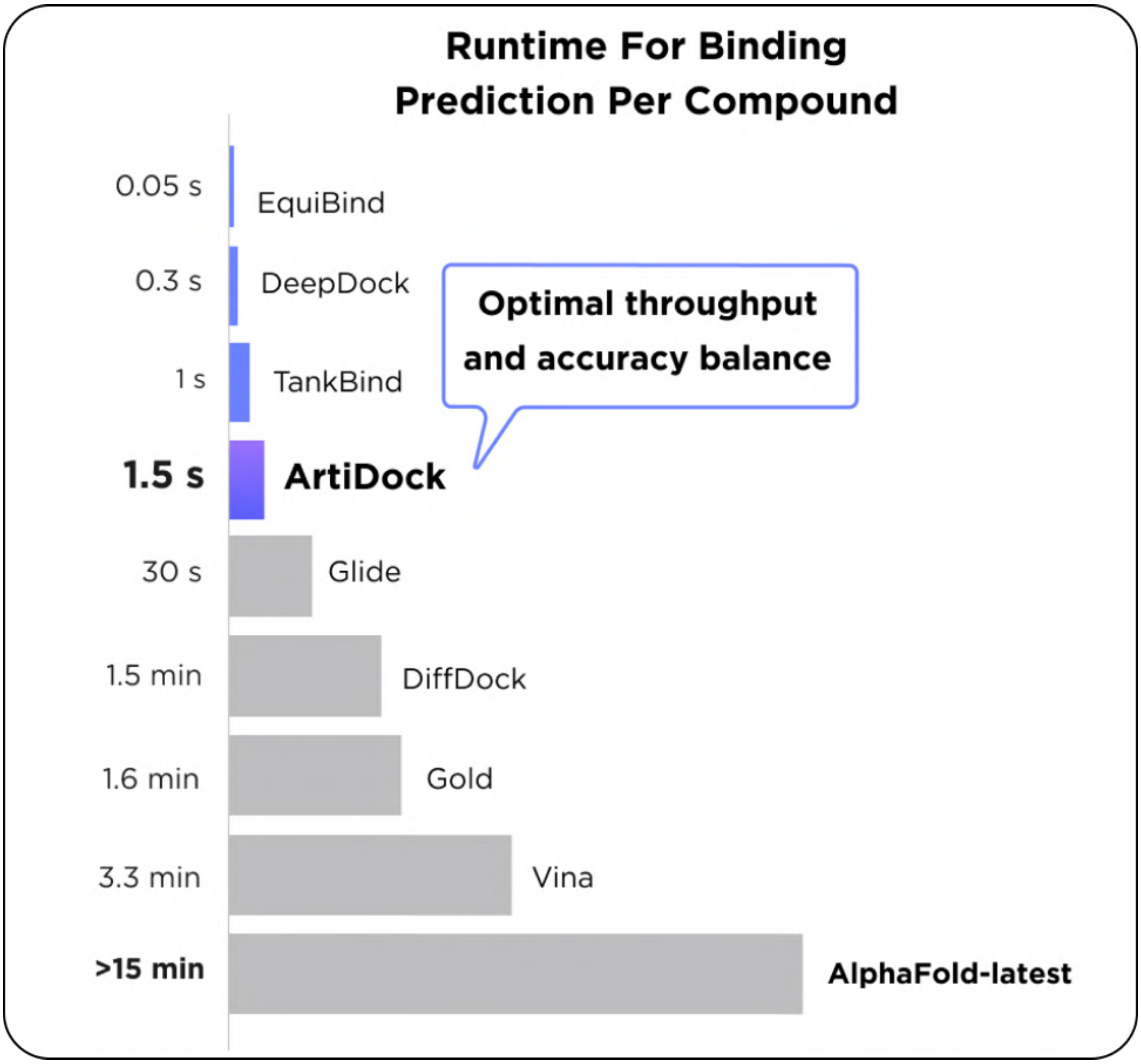


Figure 16. Comparison of the inference speed of classical and deep-learning-based docking techniques.

AI model benchmarking

A series of pairwise comparisons were performed when each kinase in a family was set as a target and all the rest as off-targets. A compound that is >5 times more active on the target kinase than on the off-target one was considered selective to the target.

The scores of all compounds were computed using the algorithm, setting each kinase consecutively as a target and the other one's pair as off-targets. The pairwise differences between scores were evaluated. If the difference is greater than the established cutoff, it indicates that the compound is selective and vice versa. This is a classical binary classification task, and accordingly, it is evaluated by the standard metrics for such problems. The main metrics are the Matthews correlation and the F1 score, but a number of secondary statistical metrics were also computed. The metrics were computed for each pair separately. Then the metrics for all pairs were averaged, and the cumulative plots were built for each family.

Results

The main performance metrics of selectivity prediction are shown in Table 4. The Receiver Operator Characteristic curves for the selectivity prediction (averaged for all kinases) are shown in Figure 17.

Table 4. Main performance metrics of selectivity prediction for JAKs.

Targets	MCC	F1-score	ROC-AUC	Accuracy	Recall	Precision	Specificity	NPV	PR-AUC
JAK1	0.74	0.83	0.86	0.88	0.78	0.89	0.94	0.87	0.88
JAK2	0.63	0.8	0.81	0.81	0.75	0.85	0.87	0.78	0.86
JAK3	0.69	0.75	0.83	0.9	0.71	0.79	0.95	0.93	0.78
TYK2	0.76	0.8	0.89	0.93	0.83	0.78	0.95	0.97	0.82
Average	0.71	0.8	0.85	0.88	0.77	0.83	0.93	0.89	0.83

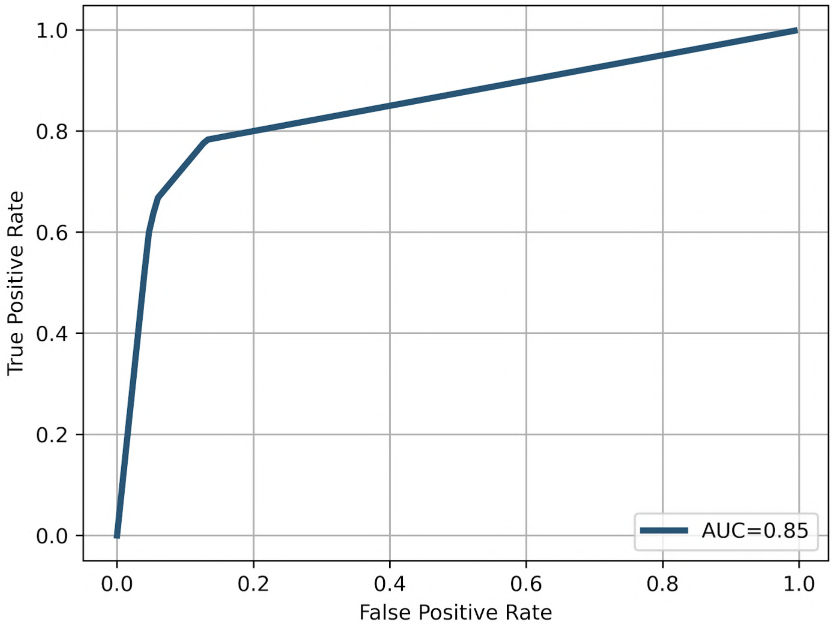


Figure 17. The **Receiver Operator Characteristic (ROC) curve** for the selectivity prediction (averaged for all kinases).

Conclusions

- Our selectivity prediction workflow shows excellent overall performance and a good balance between false positives and false negatives rate.
- The technique is able to successfully recognise and prioritise the most selective JAK ligands, which are present in the public chemical databases.
- Our technology allows leveraging very minor differences between the proteins (2-3 amino acids in the functionally important pocket).
- Explicit accounting for the protein conformational mobility eliminates the bias of the non-native crystal structures and embraces the transient dynamics of the binding pockets, which contributes significantly to selectivity.

Appendix: Performance metrics cheat sheet

Condition types

Condition positive (P): the number of real positive cases in the data.

Condition negative (N): the number of real negative cases in the data.

Result types

True positive (TP): A test result that correctly indicates the presence of a condition or characteristic.

True negative (TN): A test result that correctly indicates the absence of a condition or characteristic.

False positive (FP): A test result which wrongly indicates that a particular condition or attribute is present.

False negative (FN): A test result which wrongly indicates that a particular condition or attribute is absent.

Metrics' description

Accuracy (ACC):

$$ACC = \frac{TP+TN}{P+N} = \frac{TP+TN}{TP+TN+FP+FN}$$

Accuracy is how close a given set of measurements (observations or readings) are to their true value.

Precision or positive predictive value (PPV):

$$PPV = \frac{TP}{TP+FP}$$

Precision is how close the measurements are to each other.

Recall or true positive rate (TPR):

$$TPR = \frac{TP}{P} = \frac{TP}{TP+FN}$$

True positive rate is the probability of a positive test result, conditioned on the individual truly being positive.

F1-score:

$$F_1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$$

F-score is a measure of a test's accuracy. It is calculated from the precision and recall of the test, where the precision is the number of true positive results divided by the number of all positive results, including those not identified correctly, and the recall is the number of true positive results divided by the number of all samples that should have been identified as positive. The F1 score is the harmonic mean of precision and recall. It thus symmetrically represents both precision and recall in one metric.

Matthews correlation coefficient (MCC):

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TN+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}}$$

MCC is used as a measure of the quality of binary (two-class) classifications. The MCC takes values between -1 and 1. A score of 1 indicates perfect agreement.

Specificity or true negative rate (TNR):

$$TNR = \frac{TN}{N} = \frac{TN}{TN+FP}$$

Specificity (true negative rate) is the probability of a negative test result, conditioned on the individual truly being negative.

NPV:

$$NPV = \frac{TN}{TN+FN}$$

The positive and negative predictive values (PPV and NPV respectively) are the proportions of positive and negative results in statistics and diagnostic tests that are true positive and true negative results, respectively.