Bioinformatics, 35(2), 2019, 243–250
doi: 10.1093/bioinformatics/bty583
Advance Access Publication Date: 6 July 2018
Original Paper

Structural bioinformatics

LigVoxel: inpainting binding pockets using 3D-convolutional neural networks

Miha Skalic[1], Alejandro Varela-Rial[2], José Jiménez[1], Gerard Martínez-Rosell[1] and Gianni De Fabritiis[1,3,*]

[1]Computational Science Laboratory, Universitat Pompeu Fabra, Barcelona Biomedical Research Park (PRBB), [2]Acellera, Barcelona Biomedical Research Park (PRBB), Doctor Aiguader 88, 08003 Barcelona, Spain and [3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluis Companys 23, 08010 Barcelona, Spain

*To whom correspondence should be addressed.

Associate Editor: Alfonso Valencia

Received on February 19, 2018; revised on May 23, 2018; editorial decision on July 3, 2018; accepted on July 4, 2018

Abstract

Motivation: Structure-based drug discovery methods exploit protein structural information to design small molecules binding to given protein pockets. This work proposes a purely data driven, structure-based approach for imaging ligands as spatial fields in target protein pockets. We use an end-to-end deep learning framework trained on experimental protein–ligand complexes with the intention of mimicking a chemist's intuition at manually placing atoms when designing a new compound. We show that these models can generate spatial images of ligand chemical properties like occupancy, aromaticity and donor–acceptor matching the protein pocket.

Results: The predicted fields considerably overlap with those of unseen ligands bound to the target pocket. Maximization of the overlap between the predicted fields and a given ligand on the Astex diverse set recovers the original ligand crystal poses in 70 out of 85 cases within a threshold of 2 Å RMSD. We expect that these models can be used for guiding structure-based drug discovery approaches.

Availability and implementation: LigVoxel is available as part of the PlayMolecule.org molecular web application suite.

Contact: gianni.defabritiis@upf.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Many methods in structure-based drug discovery focus on exploiting structural information of protein pockets in order to select and optimize possible ligand binders. A three dimensional spatial featurization of proteins and small molecules and a metric on such space is usually utilized to characterize the pocket or ligand. Shape similarity, for instance, is commonly used in several virtual screening pipelines as is the case of the volumetric aligned molecular shapes (VAMS) algorithm (Koes and Camacho, 2014) and the ROCS software (Hawkins *et al.*, 2007). Docking procedures such as PharmDock (Hu and Lill, 2014), use this type of information to refine poses. Another common application where three dimensional features are used is *de novo* drug-design, in order to build a binding site model with the intent of designing a novel ligand taking into account some constraints (Yuan *et al.*, 2011). Grid generation is also central in the elucidation of pharmacophores, such is the case of FLAPpharm (Cross *et al.*, 2012) and VolSite (Desaphy *et al.*, 2012), and their search (Ebalunode *et al.*, 2008).

Whether optimal features can be automatically learned from the data rather than imposed remains an open question. In the last few years, deep learning based approaches (Goodfellow *et al.*, 2016), have demonstrated state of the art performance on tasks such as 2-D image classification (Krizhevsky *et al.*, 2012), natural language processing (Goldberg, 2015), speech recognition (Graves *et al.*, 2013) and more. The promise of deep learning in computational biology, however, is still yet to be fully developed (Angermueller *et al.*, 2016), but is increasingly common. Proteins, for instance, can be seen from a computer vision point of view, as if they were three-dimensional images after featurization. Variants of this approach have already been applied in protein binding site prediction (Jiménez *et al.*, 2017), protein–ligand binding affinity prediction (Jiménez *et al.*, 2018), large-scale active-inactive classification of drug-like compounds (Ragoza *et al.*, 2017) and quality assessment of protein folds (Derevyanko *et al.*, 2018).

This work shows how end-to-end deep convolutional neural network (DCNNs) models (LeCun *et al.*, 1998) can generate ligand fields given the structure of a protein binding site. Unlike previous knowledge-based and hand-crafted rules approaches, the proposed model is purely data-driven with the intention to mimic a chemist's intuition at manually placing atoms when designing a new compound. Furthermore, one can adjust the properties of these generated fields by tuning several input parameters. We treat proteins and ligand structures from a computer vision perspective, by voxelizing them according to different physio-chemical properties, and show that the learned and predicted representation can be useful in several related applications by evaluating it against unseen co-crystal structures.

The proposed methods can be considered complementary to molecular interaction fields (MIF) methods such as GRID (Carosati *et al.*, 2004), where an interaction energy is calculated between a target molecule and a probe, which may be an atom or a group. However, there are key differences: MIFs create volumetric maps by iteratively placing the probe at each point and calculating the interaction energy with the macromolecule. In contrast, the method proposed here takes as input the whole area at once and produces all outputs simultaneously. This is done in an end-to-end fashion. Values representing protein shape are multiplied by model's weights generating latent feature maps. The operation is repeated sequentially, until an output feature map is generated. While MIFs predict interaction energies, the method described here gives probabilities of seeing atoms with a property at each point.

2 Approach

Both the protein cavity and the ligand are represented spatially via their voxelization into a discretized 1 Å cubic grid of side size 16 Å as in Jiménez *et al.* (2017). The influence of each atom to each voxel is determined by its atom type and their respective distance r:

$$n(r) = 1 - \exp\left(-(r_{\mathrm{vdw}}/r)^{12}\right), \qquad (1)$$

where r_{vdw} is the corresponding van der Waals radius of a particular atom. The pair correlation function in Eqn.1 assigns voxels in the featurization box that are close to atoms values close to 1, while voxels far away from the atoms will approach 0. Each of the voxels is featurized with information according to the physio-chemical nature of its neighbouring atoms depending on eight predefined rules defined in Table 1. Atom types were determined based on Autodock 4 (Morris *et al.*, 2009) definitions. Voxelization algorithm is available in Supplementary Material and all descriptor computations are available as part of the HTMD Python framework for molecular dynamics (Doerr *et al.*, 2016). Voxelization rules for defining these properties are equivalent for both protein and ligand. A schema on how our method works can be seen in Figure 1: Starting with a target protein pocket (A), and optionally setting up ligand design parameters such as the number of aromatic carbons or h-bond donors (B) we can then featurize the protein pocket using the detailed *channel* descriptors (C). This featurized pocket is then used as input for a trained 3d-convolutional neural network (D), which produces the desired spatial ligand properties (E). The work presented here focuses on generating four voxelized ligand properties: an approximation to the ligand shape (total excluded volume), aromatic rings and hydrogen-bond donors and acceptors.

3 Materials, methods and evaluation procedure

3.1 Datasets

Our model was both trained and evaluated on the 2013 release of the sc-PDB database (Desaphy *et al.*, 2015), which provides 9190 high-quality protein–ligand co-crystals extracted from the PDB (Berman *et al.*, 2000). For further evaluation, we also use the Astex diverse set, which includes 85 protein–ligand complexes considered to be of pharmaceutical interest (Hartshorn *et al.*, 2007) and has been previously used in several studies for docking software benchmarking (Davis and Baker, 2009; Ravindranath *et al.*, 2015).

In order to provide a fair benchmarking of the method, a cross-validation scheme based on protein sequence similarity was used. The goal was to test how well the proposed method generalizes to proteins that differ from the ones used in training. We extracted the latest available sequence similarity clusters from the PDB (Berman *et al.*, 2000) website using a 90% similarity cutoff and assigned each of the proteins available in the scPDB database to a particular cluster only if all of its chains belong to the same one, while proteins with chains belonging to different clusters are discarded. This procedure yielded a total of 8808 protein–ligand complexes belonging to 3305 different sequence similarity clusters that were then randomly grouped together into $k = 10$ splits with sizes ranging between 525 and 1230 complexes that we evaluate using k-fold cross-validation. In practice this procedure implies that k models are trained using all splits minus a single one, used for evaluation. All PDB IDs used in each split are available as Supplementary Material. In the case of the Astex diverse set, we make the sequence cluster difference between proteins in the scPDB database and the former for training, yielding a training set of size 7147.

Table 1. Rules for atom physical chemistry properties

Property	Rule
Hydrophobic	Aliphatic or aromatic C
Aromatic	Aromatic C
Hydrogen bond acceptor	Acceptor 1 H-bond or S Spherical N Acceptor 2 H-bonds or S Spherical O Acceptor 2 H-bonds S
Hydrogen bond donor	Donor 1 H-bond or Donor S Spherical H with either O or N partner
Positive ionizable	Gasteiger positive charge
Negative ionizable	Gasteiger negative charge
Metallic	Mg, Zn, Mn, Ca or Fe
Occupancy (Excluded volume)	All heavy atoms

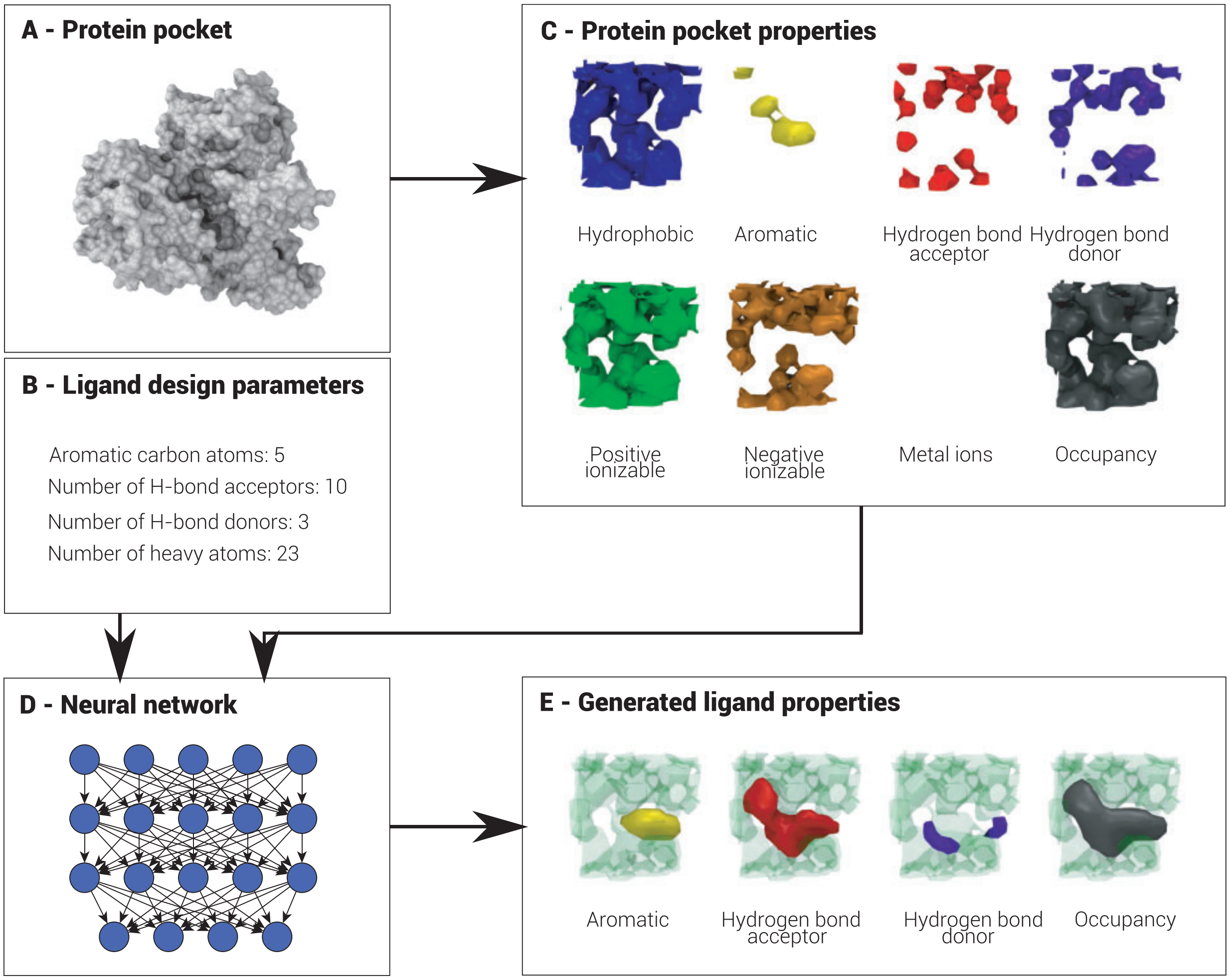


Fig. 1. Pipeline for generating ligand properties. We first voxelize the proteins (**A**), optionally providing several design parameters (**B**), according to eight different properties (**C**) then pass this input to a trained DCNN model (**D**), which will then predict the corresponding ligand properties (**E**)

3.2 Neural network architecture and learning

In order to train the proposed architecture, we aim to minimize the following two-part loss:

$$\mathcal{L}(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n} -y_i\log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i) + \mathcal{D}_{\mathrm{KL}}(y|\hat{y}) \qquad (2)$$

where

$$\mathcal{D}_{\mathrm{KL}}(y|\hat{y}) = \frac{\lambda}{n}\sum_{i=1}^{n}\left(\frac{1}{\mathrm{C}}\sum_{c=1}^{\mathrm{C}} s(y_{ic})\log\frac{s(y_{ic})}{s(\hat{y}_{ic})}\right), \qquad (3)$$

n corresponds to the total number of samples, $c \in \mathrm{C}$ to a particular and total number of channels, respectively, and s to a sorting operation. The first term in the expression corresponds to the binary cross-entropy (De Boer *et al.*, 2005) loss function commonly used in classification tasks, and whose goal is to minimize incorrect label predictions, while penalizing extremely confident wrong ones. The second term is a discrete approximation to the Kullback–Leibler (KL) divergence (Kullback and Leibler, 1951), which in turn aims to preserve a good approximation between predicted and actual output distributions. Finally, λ is a user-set parameter that controls the balance of each term to the overall loss. Intuitively, if using only binary cross-entropy, when the model is unsure where to put a property point it would give a low probability across the whole volume. By adding KL loss this low value distribution is incentivized to give points with higher probabilities, proportional to number of atoms specified in the input.

Two different types of model were trained in this work, named conditional and unconditional. The rationale behind the distinction is that in conditional models an additional four atom counts input was used (hydrogen donor and acceptor, aromatic and total number of atoms), while in the unconditional one we choose not to provide this information. The former uses both terms of the loss, with the second (corresponding to the KL-divergence) forcing the model and its additional input to place a fixed number of ligand fields in its predicted output. By adjusting or providing different additional channels to the same pocket channel inputs the conditional model can therefore be fine-tuned. In contrast, the unconditional model only uses the first part of the loss (corresponding to binary cross-entropy). Our proposed model architecture is presented in Figure 2, and code snippet is available in Supplementary Material. Model consists of seven sequential and independent layers of convolution, activated by a rectified linear unit (ReLU) non-linearity except the last one, which instead uses a sigmoid function, outputting values ranging between 0 and 1. The output of this last layer represents the

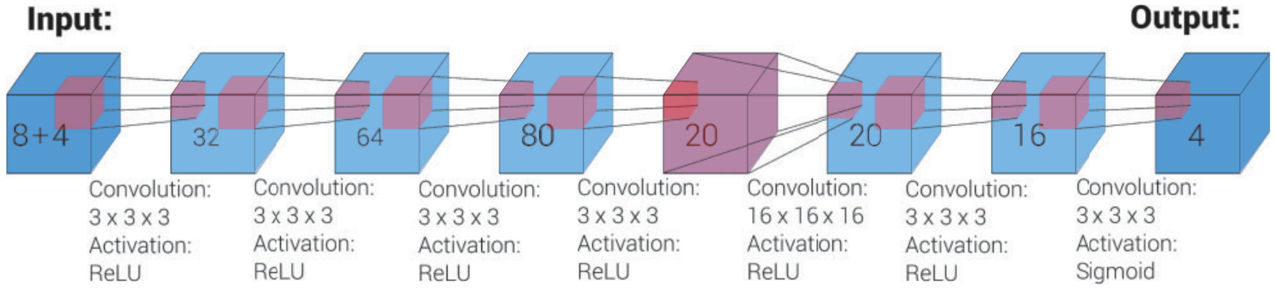

Input:
Output:
8+4
32
64
80
20
20
16
4
Convolution: 3 x 3 x 3 Activation: ReLU
Convolution: 3 x 3 x 3 Activation: ReLU
Convolution: 3 x 3 x 3 Activation: ReLU
Convolution: 3 x 3 x 3 Activation: ReLU
Convolution: 16 x 16 x 16 Activation: ReLU
Convolution: 3 x 3 x 3 Activation: ReLU
Convolution: 3 x 3 x 3 Activation: Sigmoid

Fig. 2. Architecture of the proposed DCNN model for ligand property generation. The number inside each layer, pictured as blue boxes, represents channel count, while the inner boxes represent the size of the input (3 or 16 units) and output (1 unit) of convolutional kernels (Color version of this figure is available at *Bioinformatics* online.)

predicted representation. All convolutions in the model use a filter size of 3 units, as commonly used in image processing tasks, except for the one in the fifth layer with uses an extensive 16 units one, so as to capture long-distance relationships in the input. Zero-padding is used in all feature maps and the number of latent feature maps in each layer has been chosen so that the model could be trained on a single GPU with 8GB of VRAM. It is common in the literature to use the back-propagation algorithm and some form of stochastic gradient descent to train deep neural networks. For this study, we chose Adam (Kingma and Ba, 2014) with default parameters for momentum scheduling ($\beta_1 = 0.99, \beta_2 = 0.999$) as model optimizer. The neural network was then trained during 300 epochs using randomized batch size of 128 samples, introducing molecule rotations and a shifting of 0.5 Å from the ligand center during training so as to prevent overfitting and set $\lambda = 10$ after visual inspection of the trained models. Model learning and predictions were done in the Keras (Chollet, 2015) deep learning library (keras.io) and the Theano (Theano Development Team, 2016) tensor backend.

4 Results

4.1 Ligand design

The predictive model is capable of generating useful property fields that reasonably fill pocket space, while avoiding protein clashes. One can further control the proposed conditional model by providing different input parameters, in practice affecting the properties of the generated fields. Figure 3 shows how these changes modify the behavior of the algorithm by generating bigger grids as input values increase. Additional examples of varying property counts are depicted in Supplementary Figure S3. Although the model itself learns how to place ligand hydrogen-bond acceptors next to protein hydrogen-bond donors and viceversa, one common shortcoming we observed is that for properties where variability can be high, the model is not efficient at generating distinguishable clouds, typically predicting grids around the hydrogen-bond partner. This phenomenon can be partially explained by the fact that donors and acceptors are located close in space, but can be positioned in a wide variety of angles.

4.2 Evaluation procedure

To evaluate the quality of the predicted property fields, we analyzed the volumetric overlaps between crystal ligand features and predicted features. To that end, three different and complimentary tests were carried out, where we either measure or maximize the overlap between the predicted fields and the actual ligand representations. The overlap scores were used to evaluate: (i) sensibility of predictions, comparing it to random ones and values generated by MIFs, (ii) ability to discriminate between random and crystal poses, taking into the account rotation and translation and (iii), ability to discriminate crystal ligand conformers from generated ones. The goal of the proposed tests is not to prove the method's superiority compared to existing approaches, but rather to show and motivate that the architecture can flexibly be used alongside similar representations in the same or related tasks.

Test (i) was carried out using the scPDB database, while for (ii-iv) the Astex diverse set was used instead. Overlap is measured using a discretized version of the Jaccard Index:

$$J = \frac{\#|V_p \cap V_r|}{\#|V_p \cup V_r|}, \tag{4}$$

where V_p and V_r correspond to arrays of voxels of a particular channel of the predicted and crystal representations, respectively, after using a threshold of $P = 0.75$ to discretize both arrays. This metric naturally balances both correctness and size of the prediction. In our first test, Equation (4) was used to demonstrate that our method produces sensible ligand representations, using random predictions as baseline. Results for test (i) can be checked in Figure 4 and Supplementary Figure S1 (varying thresholds values) and S2 (separating folds). For each of the splits, we measure the Jaccard index between the predicted and computed values of the crystal structures and compare it with a random baseline. The conditional model, using both parts of the loss and atom counts from the ligand is significantly better at accurately predicting defined volumetric clouds than the naive unconditional model. In particular it can be appreciated that the latter exhibits higher variance than the former, suggesting that the second term of the loss is in fact helping preserve a match between the original and predicted distribution. Furthermore, it can also be seen that the model accurately predicts the total excluded volume channel, with Jaccard indexes ranging between 0.5 and 0.6 for most cases and around 0.2 for randomized baseline. Results for the aromatic channel are also satisfactory, again with indexes ranging 0.4 and 0.5 and for some splits even surpassing it (randomized baseline of 0.1). Hydrogen-bond donor and acceptor results, while certainly more modest than the others are still significantly better than the randomized baseline. Finally, MIFs were generated using EasyMIFs (Ghersi and Sanchex, 2009) a freely available software. Aromatic, carbonyl and hydroxyl probes were used to generate interaction fields and compared to aromatic, H-bond acceptor and H-bond donor ground truth voxels, respectively. As seen in Figure 4 the MIFs give a poor estimation of the ligand

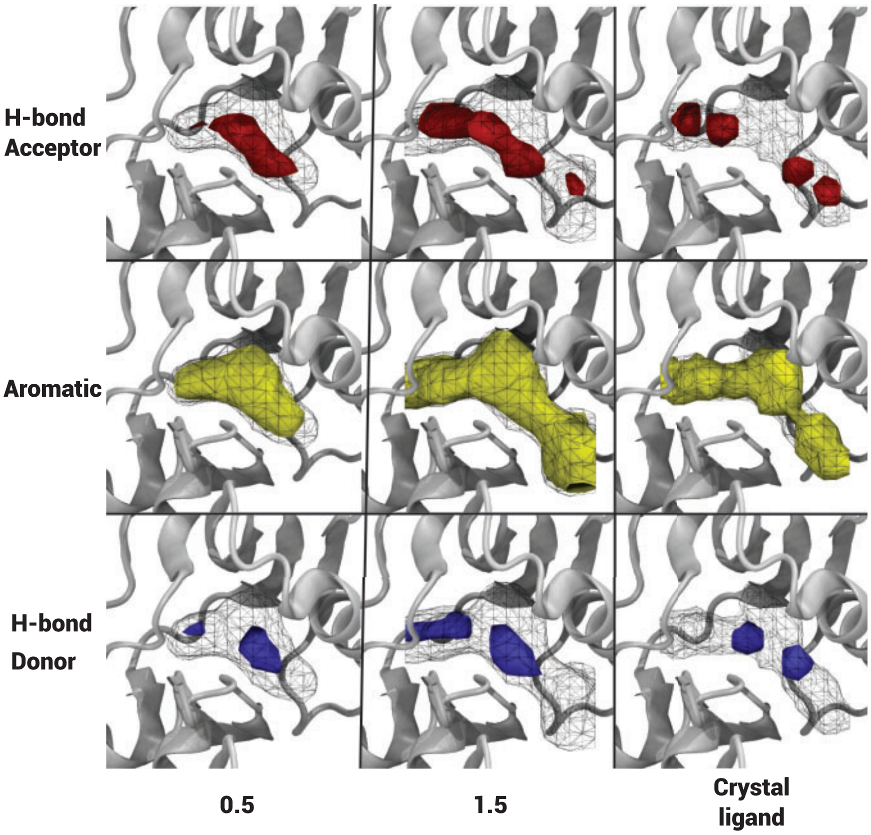


Fig. 3. Example of generated properties (PDB entry 1FPU). Ligand occupancy is displayed in wireframe while H-bond acceptors, aromatics and H-bond donors are in first, second and third row, respectively. The generated predictions shown in the column labeled 0.5 used half the atom counts of the cocrystallized ligand as the input. Column 1.5 follows the same logic, while the third shows actual cocrystallized ligand. As the atom count grows, the generated fields expand and are able to match more peripheral groups

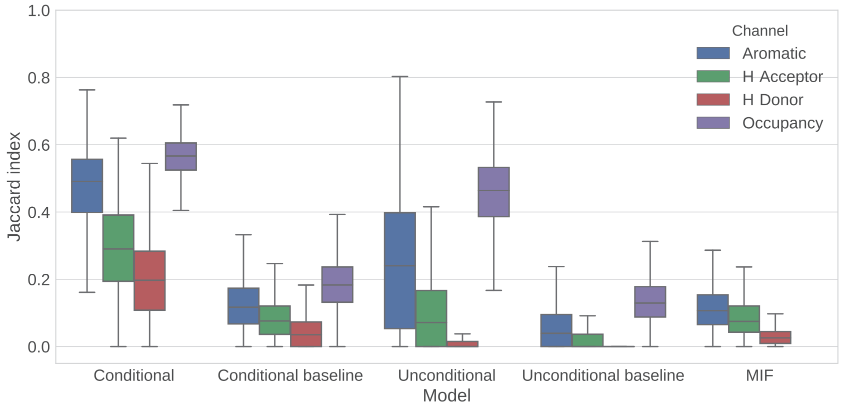


Fig. 4. Per complex results of predicted and actual physical chemistry values in the scPDB database using a Jaccard index, for the four channels of interest: aromatic, hydrogen-bond acceptor, hydrogen-bond donor and total excluded volume. The values are thresholded at 0.75 and the folds are pooled together. Per sample values in baseline groups are randomly shuffled within the sample and channel. Differences of all channels between groups conditional–conditional baseline, conditional–unconditional and unconditional–unconditional baseline are highly significant (*P*-value ≈0). Column MIF shows overlaps of the molecular interaction fields with the voxelization ground truth, at an energetic cut-off that maximizes the Jaccard index (Aromatic: $-5.10\,\text{kJmol}^{-1}$, H-acceptor: $-6.12\,\text{kJmol}^{-1}$, H-donor: $-6.33\,\text{kJmol}^{-1}$)

fields. The Jaccard indices are comparable with the conditional model baseline.

In our second test, we measure the model's ability to recover the crystal pose by rotating the crystallized ligand in the pocket around its own geometrical center with a $30^{\circ}\Delta$ and measure the overlap between each rotated state and the predicted fields. All rotated states were generated starting from a random orientation of the ligand. Ideally, the intention is to see that rotated states close to the crystal pose are ranked significantly higher than those far apart when considering volumetric overlaps. Analysis were further expanded by

Table 2. Results for pose and conformer enrichment analyses (tests ii–iii)

	Top 1	Top 5
Rotations	70 (82%)	75 (88%)
Rotations and translations	64 (75%)	80 (94%)
Conformer enrichment	74 (93%)	76 (95%)

Note: For test ii we first only took rotations into account and then translations.

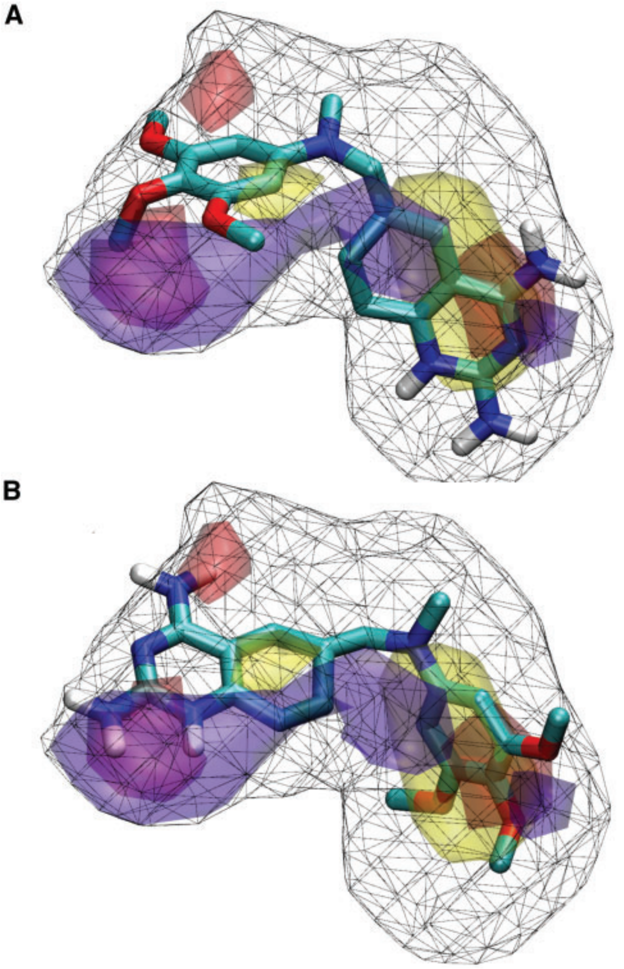


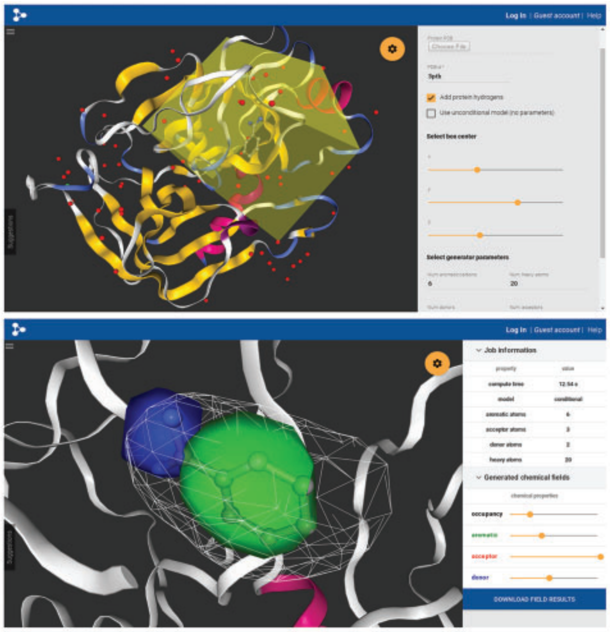
Fig. 5. Example of one wrong pose prediction. (**A**) Best ranked result using LigVoxel. (**B**) Native ligand pose. The excluded volume prediction is displayed in black wireframe, while aromatic and hydrogen bond acceptor and donor channels are displayed as clouds (Color version of this figure is available at *Bioinformatics* online.)

adding translations. In this last experiment the featurization window was slided ±3 Å in all three dimensions with a stride of 1 Å around the binding site.

In order to provide a single pose scoring function that does not depend on a threshold parameter, a second Jaccard-like index was defined using raw probabilities:

$$J_{\mathrm{raw}} = \sum_{v \in V} \frac{V_p \otimes V_r}{V_p + V_r - (V_p \otimes V_r)}, \tag{5}$$

where V_p and V_r now represent the raw three dimensional subgrid with probabilities and property values, respectively, $\otimes$ is an element-wise multiplication and $v \in V$ is a voxel in a subarray. J_{raw} indexes are then summed up: $J_s = \sum_i w_i J_{\mathrm{raw},i}$, previously weighting each one by the prevalence of their corresponding property i:



Protein PDB
Choose File
PDB id
3pth
☑ Add protein hydrogens
☐ Use unconditional model (no parameters)
Select box center
x
y
z
Select generator parameters
Num aromatic carbons: 6
Num heavy atoms: 20
Num donors
Num acceptors

Job information

property	value
compute time	12.54 s
model	conditional
aromatic atoms	6
acceptor atoms	3
donor atoms	2
heavy atoms	20

Generated chemical fields
chemical properties
occupancy
aromatic
acceptor
donor
DOWNLOAD FIELD RESULTS

Fig. 6. Web application user interface. (Top) Pocket and parameter selection phase for Trypsin (PDB Id. 3PTB). (Bottom) Aromatic and H-bond donor isosurface predictons with cocrystallized ligand superimposed

$$w_i = \frac{\#\text{atoms with property } i}{\#\text{total atoms}}. \tag{6}$$

In translation tests, a clash penalty was introduced by multiplying J_s by a discount factor P function, representing the proportion of the ligand in clash with the protein:

$$J_c = J_s\Big(1 - \tanh(\phi P)\Big), \tag{7}$$

where we set $\phi = 2$ and

$$P = \frac{\#|V_{\mathrm{prot,occ}} \cap V_{\mathrm{lig,occ}}|}{\#|V_{\mathrm{lig,occ}}|}, \tag{8}$$

with $V_{\mathrm{prot,occ}}$ and $V_{\mathrm{lig,occ}}$ representing protein and ligand occupancy voxels values, respectively, thresholded at $P = 0.75$. J_c can be considered as a discounted score of J_s, where the score is multiplied by a reward factor. The factor is 1 in case of no clashes and converges to 0 with bigger portion of the ligand volume clashing with the protein's.

We pick the best poses according to the detailed metric, taking into account possible ligand redundancies by discarding those poses with an RMSD below 2 Å with respect to any better scored pose. A pose was considered correct if the RMSD from its corresponding crystal was less than 2 Å. Results for test (ii) are presented in Table 2. After evaluation, an enrichment of scored poses below RMSD of 2 Å with respect to the crystallized ligand is observed: for the 85 protein–ligand pairs, and the 1728 total generated poses evaluated for each considering only rotations, those ranked first by the algorithm were under the 2 Å threshold in 70 pairs, while if we consider the 5 top ranked poses, the number of correct poses increases to 75. If translations are included, 64 and 81 are under the defined threshold for top 1 and top 5, respectively. This enrichment suggests that the proposed model and scoring function can effectively discriminate crystal poses. However, some ligand topologies represent a harder challenge; Figure 5 shows one case (PDB Id. 1S3V, ligand TQD) of unsuccessful pose recovery. In this example, the shape of the molecule and its almost symmetrical distribution of aromatic and h-bond groups along the ligand allows it to accommodate inside the predicted excluded volume with the two orientations shown, while fulfilling the aromatic and h-bond prediction restraints. The placement of two consecutive aromatic rings near the bigger aromatic patch prediction suggests why this pose was scored higher than those closer to the crystal. Similar scenarios were found during our testing, suggesting that molecules displaying functional groups symmetry are harder to recover - but it remains unclear if these alternative poses are simply less likely, but still viable.

For our third test, we design a conformer enrichment procedure. To that end, 10 ligand conformers were generated using RDKit (Landrum, 2012). Only poses with at least 2 Å difference between them were kept. For each generated conformer including the crystal conformer the best scoring 30° Δ pose was selected. Ranks of these poses were compared. Results for tests (iii) can also be found in Table 2, where 5 ligands had to be discarded from the analyses as RDKit failed to successfully provide 10 conformers. The proposed model is able to recover the crystal ligand out of the pool of generated conformers in 74 and 76 out of 80 cases for top 1 and 5 ranks, respectively. When considering only 10 generated poses, however, we did not observe successful pose prediction. Only in two cases the method was able to generate poses with an RMSD <2. This low performance could be attributed to non-exhaustive conformer generation, large Δ angle or translation.

5 Implementation and availability

An implementation of LigVoxel can be further explored by the scientific community through a web-application, part of the playmolecule.org suite. An example sesssion is shown in Figure 6. Users can submit their own protein binding site in.pdb format, choose either the conditional or unconditional model for field predictions and retrieve results, that can either be visualized using NGL (Rose and Hildebrand, 2015) or downloaded locally for further analysis.

6 Discussion

In this work, we have presented a novel approach for the generation of ligand images filling protein pockets based on deep neural networks. Results shown here suggest that these predictions are responsive to the number of atoms selected as input, they significantly overlap with ligand features of previously unseen ligands, and they can be used to select poses and conformers close to the native ligand orientation and geometry. We believe that approaches similar to the one proposed here will be important for structure based design of protein binders. For instance, in *de novo* drug-design, these generated fields could be used as reference for incremental atom placement. Furthermore, in computational docking, the proposed fields can assist classical scoring functions by overlap or proximity optimization. The extent to which these models can be applied in drug discovery and design pipelines has yet to be demonstrated fully, but we believe that the initial validation performed in this work is promising for future applications.

Funding

The authors thank Acellera Ltd. for funding. G.D.F. acknowledges support from MINECO (BIO2017-82628-P) and FEDER. This project has received funding from the European Union's Horizon 2020 Research and innovation programme under Grant Agreement No. 675451 (CompBioMed project).

Conflict of Interest: none declared.

References

Angermueller,C. *et al.* (2016) Deep learning for computational biology. *Mol. Syst. Biol.*, **12**, 878–816.

Berman,H.M. *et al.* (2000) The protein data bank. *Nucleic Acids Res.*, **28**, 235–242.

Carosati,E. *et al.* (2004) Hydrogen bonding interactions of covalently bonded fluorine atoms: from crystallographic data to a new angular function in the GRID force field. *J. Med. Chem.*, **47**, 5114–5125.

Chollet,F. (2015) Keras. https://github.com/fchollet/keras.

Cross,S. *et al.* (2012) GRID-based three-dimensional pharmacophores I: fLAPpharm, a novel approach for pharmacophore elucidation. *J. Chem. Inf. Model.*, **52**, 2587–2598.

Davis,I.W. and Baker,D. (2009) Rosettaligand docking with full ligand and receptor flexibility. *J. Mol. Biol.*, **385**, 381–392.

De Boer,P.T. *et al.* (2005) A tutorial on the cross-entropy method. *Ann. Oper. Res.*, **134**, 19–67.

Derevyanko,G. *et al.* (2018) Deep convolutional networks for quality assessment of protein folds. *ArXiv:1801.06252v1*.

Desaphy,J. *et al.* (2012) Comparison and druggability prediction of protein–ligand binding sites from pharmacophore-annotated cavity shapes. *J. Chem. Inf. Model.*, **52**, 2287–2299.

Desaphy,J. *et al.* (2015) Sc-PDB: a 3D-database of ligandable binding sites-10 years on. *Nucleic Acids Res.*, **43**, D399–D404.

Doerr,S. *et al.* (2016) HTMD: high-throughput molecular dynamics for molecular discovery. *J. Chem. Theory Comput.*, **12**, 1845–1852.

Ebalunode,J.O. *et al.* (2008) Novel approach to structure-based pharmacophore search using computational geometry and shape matching techniques. *J. Chem. Inf. Model.*, **48**, 889–901.

Ghersi,D. and Sanchez,R. *et al.* (2009) EasyMIFs and SiteHound: a toolkit for the identification of ligand-binding sites in protein structures. *Bioinformatics*, **25**, 3185–3186.

Goldberg,Y. (2015) A primer on neural network models for natural language processing. *arXiv Preprint arXiv: 1510.00726*, pp. 1–76.

Goodfellow. *et al.* (2016) *Deep learning*. MIT Press, Cambridge, MA. http://www.deeplearningbook.org.

Graves,A. *et al.* (2013) Speech recognition with deep recurrent neural networks. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 6645–6649.

Hartshorn,M.J. *et al.* (2007) Diverse, high-quality test set for the validation of protein–ligand docking performance. *J. Med. Chem.*, **50**, 726–741.

Hawkins,P.C. *et al.* (2007) Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.*, **50**, 74–82.

Hu,B., and Lill,M.A. (2014) PharmDock: a pharmacophore-based docking program. *J. Cheminform.*, **6**, 14.

Jiménez,J. *et al.* (2017) DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics*, **33**, 3036–3042.

Jiménez,J. *et al.* (2018) Kdeep: protein–ligand absolute binding affinity prediction via 3d-convolutional neural networks. *J. Chem. Inf. Model.*, **58**, 287–296.

Kingma,D. and Ba,J. (2014) Adam: a method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, pp. 1–13.

Koes,D.R. and Camacho,C.J. (2014) Shape-based virtual screening with volumetric aligned molecular shapes. *J. Comput. Chem.*, **35**, 1824–1834.

Krizhevsky,A. *et al.* (2012) ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.*, **25**, 1097–1105.

Kullback,S. and Leibler,R.A. (1951) On information and sufficiency. *Ann. Math. Statist.*, **22**, 79–86.

Landrum,G. (2012) RDKit: Open-source cheminformatics. http://www.rdkit.org. [Online; accessed December-2017].

LeCun,Y. *et al*. (1998) Gradient-based learning applied to document recognition. *Proc. IEEE*, **86**, 2278–2323.

Morris,G.M. *et al*. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem*., **30**, 2785–2791.

Ragoza,M. *et al*. (2017) Protein–ligand scoring with convolutional neural networks. *J. Chem. Inf. Model*., **57**, 942–957.

Ravindranath,P.A. *et al*. (2015) AutoDockFR: advances in protein–ligand docking with explicitly specified binding site flexibility. *PLoS Comput. Biol*., **11**, e1004586.

Rose,A.S. and Hildebrand,P.W. (2015) NGL viewer: a web application for molecular visualization. *Nucleic Acids Res*., **43**, W576–W579.

Theano Development Team (2016) Theano: a Python framework for fast computation of mathematical expressions. *arXiv e-Prints*, *arXiv:1605.02688v1*.

Yuan,Y. *et al*. (2011) LigBuilder 2: a practical de novo drug design approach. *J. Chem. Inf. Model*., **51**, 1083–1091.