SurfBind: Surface Distance aided Geometric Deep Learning for Binding Conformations

Jiamin WU [*1], **Wenqi ZENG** [*1 †], **Song LIU** [2, 4], **Yuan YAO** [1, 3 ‡]

[1]Department of Mathematics, Hong Kong University of Science and Technology
[2]Department of Statistics and Data Science, Southern University of Science and Technology
[3]Department of Chemical and Biological Engineering, Hong Kong University of Science and Technology
[4]International Digital Economy Academy (IDEA)
{ jwubz, wzengad}@connect.ust.hk, liusong@idea.edu.cn, yuany@ust.hk

Abstract

A core task in computer-aided drug discovery is the optimization of lead compounds with high binding affinity to the target proteins. The binding process is desired to find the proper position and the correct relative orientation of the "key" (the ligand), which will open up the "lock" (the protein). During the process, existing deep learning methods usually overlook surface intersection between ligands and targets, i.e., part of the ligands goes into the protein interior. In this paper we present our SurfBind model, a two-stage deep learning method aided by the surface distance function (SDF). Our model will produce pairwise distance distribution to encode the multi-valued possible relative positions. By exerting effective SDF constraint to distance likelihood potential, SurfBind derives rational conformations lessening clash to the protein, and also reduces root mean square deviation (RMSD) for ultra-large ligands. To the best of our knowledge, the performance of SurfBind working as a score function on docking and screening power achieves SOTA on CASF-2016 benchmark.

1 Introduction

When it comes to protein-ligand binding, the target protein is postulated as rigid and encoded into high dimensional feature space. For encoded protein, finding the rational ligand pose is an significant but challenging task. One direction to settle protein-ligand binding is to predict intermediate values enlightening binding interactions like binding affinity (Jones et al. 2021), score functions (SF) (Sánchez-Cruz et al. 2021; Zheng et al. 2022), RMSD between given binding pose and target protein (Bao, He, and Zhang 2021). These intermediate values can be adopted as measures to binding ability, however, challenges to generate suitable conformations are regular visitors in these methods.

Another direction is end-to-end models such as EQUIDOCK (Ganea et al. 2021), EquiBind (Stärk et al. 2022) and TANKBind (Lu et al. 2022). They ensure SE(3)-equivariance for 3D inputs and consider physical constraints like dihedral angles. However, aforementioned methods impose insufficient restriction on orientation of ligands against protein surfaces. Though equipped with loss to prevent steric clashes (Ganea et al. 2021; Stärk et al. 2022), they still suffer from the surface intersection between ligands and targets, i.e., part of the ligands go "inside" the protein surface. Such unrealistic molecules will have extremely low binding affinity although their RMSDs seem good.

Recently, a two-stage deep learning method, DeepDock (Méndez-Lucio et al. 2021) is proposed to predict pairwise distance distribution encoding flexible binding outcomes. Distinct from determinate values, the distance likelihood can be aggregated into a statistical potential to build either a score function (SF) or an energy function. SF gives out binding ratings to screen out unreasonable ligand conformations and get reasonable ranks. While energy function guides the optimization in differential evolution (Storn and Price 1997) to generate binding conformations for ligands. However, the mesh representation (Gainza et al. 2019) for protein in DeepDock requires pre-computed geometrical and chemical features, which is time consuming and takes more than 2 days on the PDBBind database v2019 (17679 protein-ligand complexes). Furthermore, the simple feature extractor GCN does not guarantee the SE(3)-equivariance and thus may cause message passing changes with Euclidean 3D transformations. Without any penalty on clash, DeepDock also suffers from surface penetration issue like its deep learning counterparts.

Therefore, we raise SurfBind, a two-stage deep learning method aided by the surface distance function (SDF) (Zhu et al. 2010; Park et al. 2019; Venkatraman et al. 2009; Bordner and Gorin 2007) on the basis of DeepDock. Fig. 1 provides workflow for SurfBind. For protein surface, in view of training efficiency, SurfBind uses point clouds sampled from a level set of SDF as the representation. A modified protein structure extractor based on Dmasif (Sverrisson et al. 2021) provides the SE(3) equivariant transformation and trains the chemical and geometrical features from raw, without any beforehand experiment-basis calculation. As for ligand graph representation, SurfBind adds pseudo-edges for the geometrically close but non-bonded atoms to avoid the under-determination of long-range information. To address the surface intersection issue between ligand and protein,

[*]These authors contributed equally.
[†]This work is done during an internship in the International Digital Economy Academy (IDEA).
[‡]Corresponding authors: Song LIU, Yuan YAO

SDF constraint assists the statistical potential trained at the first stage to screen out irrational conformations in the second stage. With SDF constraints and pseudo-edges, the optimized conformers are more likely to fit in the pocket shapes and reduce the RMSD to reference conformations for large molecules. We compare SurfBind to DeepDock and physical based state-of-the-art methods on the CASF-2016 benchmark (Su et al. 2019) and obtain better results on docking and screening powers.

Compared to DeepDock, our innovations can be listed threefold:

- For protein surface, by sampling the level set of surface distance function (SDF), point clouds are used as the representation and provides abundant information by dense coverage of the surface. For features attached to each point cloud, chemical information and geometrical curvatures are directly learned by the feature extractor with SE(3)-equivariance, averting time-consuming hand-curated items.
- For ligand graph, novel pseudo-edges for close non-bonded atoms in the 3D space are added to take non-covalent interactions into consideration. The performance on ultra-large ligand who has plentiful rotatable bonds or flexible rings is improved.
- In the second stage, a new application of SDF constraint comes into play. SDF constraint is directly combined into outputs from the first stage and penalize the statistical potential for undesired clashed position.

2 Methodology

SurfBind can be divided into four main blocks: protein structure extractor, ligand structure extractor, mixture density layer, optimization and ranking under SDF constraint. Fig. 2 shows details for processing the target protein and ligand in the first stage.

In the protein structure extractor, the target protein surface is represented as point clouds. As point clouds are uninformative, extra chemical and geometrical curvature features are added. Chemical features are learned by GCN from chemistry properties while geometrical curvatures are calculated from point clouds. A subsequent GCN handles point clouds and features to attain the protein embedding. In the ligand structure extractor, the ligand is represented as a 2D molecular graph with pseudo edges. Again GCN execute the embedding for the ligand graph. Then, a mixed density layer converts the concatenation of protein and ligand feature embeddings to the final output, pairwise distance distribution. Eventually, SDF constraint accompanies a distance potential and screens out unreasonable binding outcomes during optimization or ranking.

2.1 Protein Structure Extractor

Point Cloud Representation from Surface Distance Our model represents the molecular surface of the rigid binding site as oriented point cloud following (Sverrisson et al. 2021), where each point cloud is associated with a unit normal vector. The point cloud is sampled close to the surface derived from the SDF in Eq. 1.

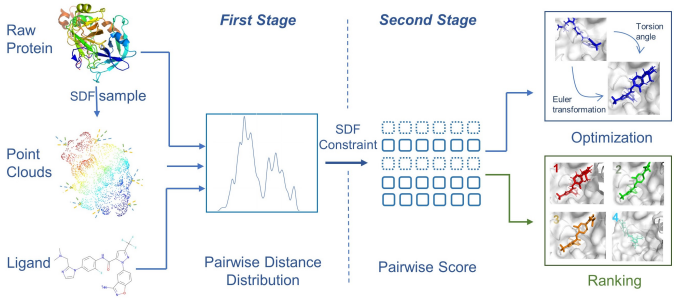

Figure 1: Workflow of SurfBind. The first stage is a deep learning model to utilize raw protein, point clouds from SDF sampling, and raw ligand to predict pairwise distance distribution. The second stage is to apply SDF constraint on distance likelihood potential to yield binding scores, which can be applied on two downstream tasks, optimization and ranking.

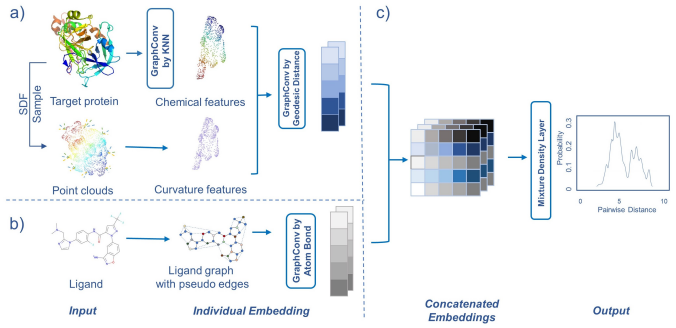

Figure 2: First stage for SurfBind. The procedure is SE(3) equivariant and consists of three major modules, a) protein structure extractor for point clouds with trainable chemical and geometrical curvature features. b) ligand structure extractor for 2D ligand graph, with atom type and bond type as features. c) a mixture density layer to produce pairwise distance distribution.

With input atom clouds $\{\mathbf{a}_j\}_{j=1}^N \subset \mathbb{R}^3$, the point cloud $\{\mathbf{p}_i\}_{i=1}^n \subset \mathbb{R}^3$ is sampled near the protein surface as level set $\mathrm{SDF}(\mathbf{p}) = \gamma_1$, where γ_1 is a hyperparameter. σ_j is experimental atom radius for $\mathbf{a}_j$, $w(\mathbf{p})$ is the averaged atom radius for $\mathbf{p}$. Normal vectors $\{\mathbf{n}_i\}_{i=1}^n \subset \mathbb{R}^3$ are taken as the gradient of the level set.

$$\mathrm{SDF}(\mathbf{p}) = -w(\mathbf{p}) \cdot \log \sum_{j=1}^{N} \exp(-\|\mathbf{p} - \mathbf{a}_j\|/\sigma_j) \quad (1)$$

Unlike mesh representation, point clouds benefit from computation efficiency because they avoid connectivity calculation or structure partition. As a result, SurfBind can have a speedup to obtain finer representation for protein surface. However, point clouds are random discrete points so extra information needs to be added. Different from using time-consuming pre-computed chemical attributes in DeepDock, SurfBind handles protein from raw and assigns trainable

chemical and geometrical curvature features to each point cloud, as shown in the next part.

Trainable Chemical and Geometrical Features The chemical features are produced by applying graph convolution neural network (GCN) on $\{\mathbf{a}_j\}_{j=1}^N \subset \mathbb{R}^3$. The chemical properties, such as 32 atom types (C, H, O, N, S, Se, Be, B, F, Mg, Si, P, Cl, V, Fe, Co, CU, Zn, As, Br, Ru, Rh, Sb, I, Re, Os, Ir, Pt, Hg, Ca, Na, Ni), weighted average distances, and neighboring chemical information of nearest k atoms ($k = 16$), will be utilized. Such design assists the informative encoding of atom itself and surrounding environments.

For geometrical features, SurfBind aims to get the Gaussian curvature and mean curvature. Assume the surface shape of a protein comes from sampling of a 2-dimensional Riemannian manifolds $\mathcal{M}$ embedded in $\mathbb{R}^3$. Around each point cloud in $\mathcal{M}$, the manifold is homeomorphic to the tangent plane $T_p\mathcal{M}$. The differential of the Gauss map at $\mathbf{p}_i$ is a self-adjoint operator on $T_p\mathcal{M}$ and conventionally called as Weingarten map. Curvatures are determined by Weingarten map matrix, where Gaussian curvature is the determinant and mean curvature is the trace.

We follow (Cao et al. 2019) to estimate Weingarten map for oriented point clouds. For $\mathbf{p}_i$ sampled from $\mathcal{M}$, extend its corresponding normal vector $\mathbf{n}_i$ to a local coordinate system $(\mathbf{n}_i, \mathbf{u}_i, \mathbf{v}_i)$ where $(\mathbf{u}_i, \mathbf{v}_i)$ is the orthogonal basis in $T_p\mathcal{M}$. With k-nearest neighbors $[\mathbf{p}_{i1}, \mathbf{p}_{i1}, ..., \mathbf{p}_{ik}]$ around $\mathbf{p}_i$, we project the relative coordinates and relative normal vector to tangent space as:

$$\Delta\mathbf{p}_i^{\perp} = [(\mathbf{p}_{i1} - \mathbf{p}_i, \mathbf{p}_{i2} - \mathbf{p}_i, ..., \mathbf{p}_{ik} - \mathbf{p}_i)^T] \cdot [\mathbf{u}_i, \mathbf{v}_i]$$
$$\Delta\mathbf{n}_i^{\perp} = [(\mathbf{n}_{i1} - \mathbf{n}_i, \mathbf{n}_{i2} - \mathbf{n}_i, ..., \mathbf{n}_{ik} - \mathbf{n}_i)^T] \cdot [\mathbf{u}_i, \mathbf{v}_i]$$

The Weingarten map $G_i \in \mathbb{R}^{2\times 2}$ is given as $G_i = ((\Delta\mathbf{p}_i^{\perp})^T\Delta\mathbf{p}_i^{\perp} + \delta\mathbb{I}_{2\times 2})^{-1}((\Delta\mathbf{p}_i^{\perp})^T\Delta\mathbf{n}_i^{\perp})$ with $O(n^{-2/3})$ convergence rate (Cao et al. 2019). Now that Weingarten map is obtained, the Gaussian curvature is $K_i = \det(G_i)$, the mean curvature is $H_i = \text{trace}(G_i)/2$. In practice, we scale $\mathbf{n}_i$ with area-weighted average of neighboring points normals at different radius as $[1, 2, 3, 5, 10]$ and produce in total 5 Gaussian curvatures and 5 mean curvatures.

Up to now, the local system of Cartesian coordinates $(\mathbf{n}_i, \mathbf{u}_i, \mathbf{v}_i)$ at each point $\mathbf{p}_i$ has built an orthogonal coordinate frame up to one rotation in tangent space by numerical methods. It is necessary to further modify such local reference frame in order to be robust under point cloud noise and non-rigid deformation. Similar to (Melzi et al. 2019), we update our frame on the intrinsic gradient of a scalar field defined on the input shape. To incorporate chemical features and geometrical features, $\mathbf{f_i}$, define a scalar function $f_i = \text{MLP}(\mathbf{f_i})$. The gradient on f can be approximated by a derivative of geometric convolution containing Gaussian filter from (Sverrisson et al. 2021):

$$\cos\Delta\theta_i \leftarrow \frac{1}{n}\sum_{j=1}^{n} w(d_{ij}) \cdot (\mathbf{p}_i - \mathbf{p}_j)^T \cdot \mathbf{u}_i \cdot f_i$$
$$\cos\Delta\phi_i \leftarrow \frac{1}{n}\sum_{j=1}^{n} w(d_{ij}) \cdot (\mathbf{p}_i - \mathbf{p}_j)^T \cdot \mathbf{v}_i \cdot f_i$$

where $d_{ij} = \|\mathbf{p}_i - \mathbf{p}_j\| \cdot (2 - \langle\mathbf{n}_i, \mathbf{n}_j\rangle)$ stands for an approximation for geodesic distance, $w(d_{ij}) = \exp(-d_{ij}^2/2\sigma^2)$ stands for Gaussian window. Updated basis for tangent plane $\mathbf{u}_i^{\text{update}}, \mathbf{v}_i^{\text{update}}$ will be:

$$\mathbf{u}_i^{\text{update}} = \cos\Delta\theta_i \cdot \mathbf{u}_i + \sin\Delta\theta_i \cdot \mathbf{v}_i,$$
$$\mathbf{v}_i^{\text{update}} = \cos\Delta\phi_i \cdot \mathbf{u}_i + \sin\Delta\phi_i \cdot \mathbf{v}_i$$

Finally, with updated local reference frame $[\mathbf{n}_i, \mathbf{u}_i^{\text{update}}, \mathbf{v}_i^{\text{update}}]$, the embedding features in $(l+1)$-th layer $\mathbf{f}_i^{(l+1)}$ is computed by graph convolution network. Different from other work where convolution kernels are defined over Euclidean distance using K-NN or radius graph, here we consider the approximated geodesic distance, which can be beneficial to differentiate specific surfaces. To be specific, the neighborhood $\mathcal{N}_i$ is localized by the approximated geodesic distance d_{ij} by a Gaussian filter as $\mathcal{N}_i = \{j : \exp(-d_{ij}^2/2\sigma^2) \geq \gamma\}$ where γ is given radius threshold.

$$\mathbf{f}_i^{(l+1)} = \varphi_1(\sum_{j\in\mathcal{N}_i} \varphi_2(\mathbf{p}_i - \mathbf{p}_j)[\mathbf{n}_i, \mathbf{u}_i^{\text{update}}, \mathbf{v}_i^{\text{update}}]) \cdot \mathbf{f}_j^{(l)})$$

2.2 Ligand Structure Extractor

The ligand is represented as a 2D undirected graph, where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Vertices $\mathbf{v}_i \in \mathcal{V}$ stands for atoms in the ligand with atom type as attached features, while edges $\mathbf{e}_{i,j} \in \mathcal{E}$ stands for covalent bond between atoms with bond type as features. A ligand graph will be processed through GCN with residue block, extracting information in atom and its local environments.

Pseudo-edge for Long Range Interaction The input of ligands are 2D molecular graphs, whose edges are the covalent bonds of the molecules (single, double, triple or aromatic). This simple representation ignores the long-range interaction of atoms without covalent bonds. In these cases, the atoms not directly connected in the 2D molecular graph may have non-covalent interactions in the real 3D space, as the accessible conformation space of the molecules is vast. Therefore, the performance of ultra-large molecules with numerous rotatable bonds or flexible rings suffers underestimation. To deal with this issue, a pseudo-edge is added for close non-bonded atom if the Euclidean distance of two atoms is smaller than the distance threshold. A proper distance threshold is treated as hyperparameter that to be chosen by experiments.

2.3 Mixture Density Layer

After aforementioned protein and ligand extractor, the processed embeddings are combined to model the interactions between ligand and protein. In the final step, the concatenated embeddings will be processed by a mixture density network (MDN) (Bishop 1994) to encode pairwise distance between protein and ligand as mixed Gaussian distribution $P(d_{ij}|\mathbf{p}_i, \mathbf{v}_j)$. Compare to predicting a definite distance, the distance distribution is more adequate to model the multi-valued possible relative position between protein and ligand when binding.

The MDN layer allows the distribution of protein-ligand distance to be learned by loss function, where $w_{ij,k}, \mu_{ij,k}, \sigma_{ij,k}$ are outputs from MDN layer formed by K Gaussian distribution ($K = 10$):

$$\begin{aligned}\mathcal{L} &= -\frac{1}{nm}\sum_{i}^{n}\sum_{j}^{m}\log P(d_{ij}|\mathbf{p}_i, \mathbf{v}_j) \\ &= -\frac{1}{nm}\sum_{i}^{n}\sum_{j}^{m}\log\sum_{k=1}^{K} w_{ij,k}\mathcal{N}(\mu_{ij,k}, \sigma_{ij,k}) \end{aligned} \quad (2)$$

Euclidean Equivariant To summarize, aside from the invariance in group representation theory, SurfBind adopts scalarization, a generic way, to achieve invariance. Typically, geometric vectors $\mathbf{p}$ are transformed into intrinsic scalars, like curvatures and distance used in producing normal vectors. These scalars are invariant to Euclidean transformations hence ensures SE(3)-equivariance. Despite the simplicity, such design provides fundamental construction and is efficient enough and universal to achieve equivariance.

2.4 SDF Constraint on Statistical Potential

Learning from the 3D collision problem in computer vision fields, SurfBind imposes a hard SDF constraint in Eq.3 for ligand atoms inside the protein. The protein surface is defined as $\text{SDF}(\mathbf{v}) = \gamma_2$, where weights $w = 0.35$ and $\sigma = 0.35$ are referenced from (Venkatraman et al. 2009), $\gamma_2 = 0.5$ is chosen to approximate the van der Waal's radius (1.5 Å) for an isolated atom.

$$\text{SDF}(\mathbf{v}) = -w \cdot \log\sum_{i=1}^{n}\exp(-\|\mathbf{v} - \mathbf{p}_i\|/\sigma) \quad (3)$$

Statistical potential is the aggregation of protein-ligand pairwise distance function in Eq. 2. SDF is directly combined into this potential to decrease scores for undesired ligand positions shown as Eq. 4.

$$Surf_{score} = \frac{1}{nm}\sum_{i}^{n}\sum_{j}^{m}\log P(d_{ij}|\mathbf{p}_i, \mathbf{v}_j)\cdot\mathbb{1}(\text{SDF}(\mathbf{v}_j) > \gamma) \quad (4)$$

$Surf_{score}$ can be applied in multiple scenarios. Firstly in *optimization task*, for an optimization algorithm, e.g., differential evolution, Eq. 4 is also known as energy function which can be minimized to generate optimal conformation. Moreover, in the *ranking task*, the statistical potential based on the pairwise distance distribution in Eq. 4 can be treated as a score function. Given proteins and ligands, such score function can give out binding ratings which indicate relative ranks. The ranks can be applied to screen out most likely binding poses for ligands or binding complexes for protein-ligand pairs.

3 Experiment Settings

3.1 Datasets

Train data SurfBind is trained on PDBBind database v2019 (Liu et al. 2017), which provides binding data and processed structural files for the biomolecular complexes in the Protein Data Bank (PDB) (Berman, Henrick, and Nakamura 2003). The specific version, PDBBind v2019 contains 17679 protein-ligand complexes. After excluding the complexes in the test set or containing unrecognizable atom types, 14603 complexes are finally used for train and valid set with random 0.8/0.2 allocation.

Test data SurfBind is tested on an open-access benchmark, the conventional Comparative Assessment of Scoring Functions (CASF) dataset (Li et al. 2018). The latest version, CASF 2016 (Su et al. 2019), contains 285 protein-ligand complexes with reliable structures.

3.2 Evaluation Setup

Evaluation Metrics In the optimization task, the performance is evaluated on RMSD. In the ranking task, as SurfBind is not trained to predict experimental scores on purpose, we are more interested in *docking power* and *screening power* (Su et al. 2019). Docking power assesses the ability of a score function to identify native binding poses among randomly generated decoys. Screening power evaluates the ability to identify true protein-ligand binders for given target proteins, among cross docking conformations. Good performance on these two metrics indicates that conformations similar to native binding pose will be top-ranked.

Baselines For optimization task, to address surface intersection issue, baselines are mishandled cases by DeepDock (Méndez-Lucio et al. 2021). For ranking task, the baselines are DeepDock and the Top 4 score functions recorded in (Su et al. 2019), containing both traditional and machine learning driven ones.

SurfBind Versions We provide three versions for SurfBind models: 1) vanilla SurfBind without SDF constraint 2) SurfBind-S, for adding SDF constraint into SurfBind 3) SurfBind-V, the incremental model for using Autodock Vina score as a correction on the basis of SurfBind-S. The idea in SurfBind-V, to fine-tune with traditional score function, has witnessed appealing precedents such as OnionNet-SFCT (Zheng et al. 2022) and $\Delta_{\text{Vina}}\text{RF}_{20}$ (Li et al. 2015). For the combination:

$$\text{Score} = \alpha \cdot Surf_{score} + \beta \cdot Vina_{score}$$

where $Surf_{score}$ is the normalized score by SurfBind in Eq. 4. Weights $\alpha = 0.5$, $\beta = -0.5$ are used in experiments as Autodock Vina shows a negative correlation to SurfBind score.

3.3 Implementation Details

SurfBind is implemented with Pytorch (Paszke et al. 2019). For data processing, we use PyTorch Geometric (Fey and Lenssen 2019). SurfBind is trained on 1 Tesla A100 with Adam optimizer with learning rate $\text{lr} = 0.002$ equipped with cosine annealing. Reproducible code will be provided.

4 Results

4.1 Optimization Task

In the optimization, the differential evolution is used to get the optimal conformation. Adding SDF constraint helps to

Table 1: Forward Screening Power for SurfBind, DeepDock and Top 4 score functions.

Forward Screening Power	Success rate ↑			Enhancement Factor (EF) ↑		
	Top1%	Top5%	Top10%	Top1%	Top5%	Top10%
Autodock Vina (2012; 2010)	29.8%	40.4%	50.9%	7.70	4.01	2.87
ChemPLP@GOLD (2009)	35.1%	61.4%	64.9%	11.91	5.29	3.59
GlideScore-SP (2004; 2004; 2006)	36.8%	54.4%	63.2%	11.44	5.83	3.98
$\Delta_{\text{Vina}}\text{RF}_{20}$ (2015)	42.1%	49.1%	54.4%	11.73	4.43	3.10
DeepDock (2021)	43.9%	61.4%	82.5%	16.41	7.36	5.16
SurfBind-S	**54.4%**	71.9%	**86.0%**	**20.98**	8.13	5.12
SurfBind-V	**54.4%**	**73.7%**	84.2%	20.68	**8.23**	**5.20**
SurfBind	**54.4%**	**73.7%**	84.2%	20.43	8.10	5.08

derive rational conformations adapting to protein surface, reducing surface intersection issue. Fig. 3 shows multiple surface clash cases by DeepDock. Given identical input as DeepDock, SurfBind can elude such collision. As shown in Table 3, for the molecules with more than 30 rotatable bonds,

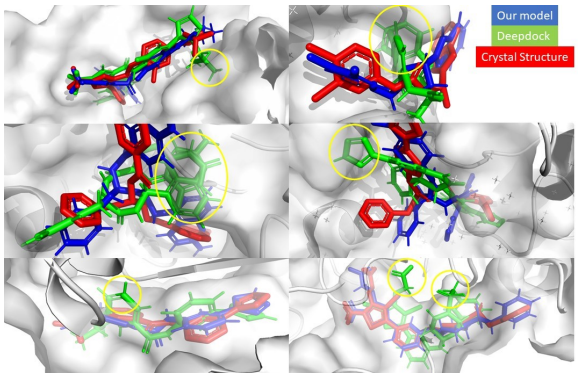

Figure 3: Surface intersection examples for ligand 1z6e, 3jvr, 4cra, 4crc, 4de2, 2qbp respectively (left to right, top to bottom). Green: DeepDock, Blue: SurfBind, Red: Crystal Structure. For DeepDock, the ligand may go inside the protein surfaces (see the yellow circle). While for SurfBind, the generated ligand can avoid surface collision with protein surface.

Table 2: Docking Power for SurfBind, DeepDock and Top 4 score functions. The correlation is based on 0-10Å.

Docking Power ↑	Top 1	Top 2	Top 3	Spearman's correlation
DeepDock (2021)	87.0%	92.6%	94.4%	0.830
$\text{DrugScore}^{\text{CSD}}$ (2005)	87.4%	93.3%	95.1%	0.609
GlideScore-SP (2004; 2004) (2006)	87.7%	91.9%	93.7%	0.512
$\Delta_{\text{Vina}}\text{RF}_{20}$ (2015)	89.1%	94.4%	96.5%	0.608
Autodock Vina (2012; 2010)	90.2%	95.8%	97.2 %	0.605
SurfBind	95.4%	**97.9%**	98.2%	**0.901**
SurfBind-S	**95.8%**	97.5%	98.2%	0.899
SurfBind-V	95.1%	**97.9%**	**99.3%**	0.713

Table 3: Ligand RMSD performance on ultra-large molecules with a large number of rotatable bonds and atoms.

Ligand RMSD ↓	rotatable bounds			atoms	
	> 10	> 30	Top 10%	> 40	Top 10%
DeepDock (2021)	3.96	6.47	3.95	4.48	3.95
SurfBind	3.62	4.59	3.63	3.90	3.49
SurfBind-S	**3.55**	**4.46**	**3.56**	**3.82**	**3.33**

the average RMSD is reduced from 6.47 to 4.46. For the molecules with more than 40 atoms, the average RMSD is reduced from 4.48 to 3.82.

4.2 Ranking Task

In docking power and screening power, SurfBind surpasses the performance of DeepDock and other 4 best score functions. As docking power results shown in Table. 2, all three versions of SurfBind outperform DeepDock and the other score functions evaluated in the same benchmark. Results on screening power in Table. 1 are consistent with docking power, showing our method possess advantage in identifying acceptable binding poses as top-ranked. The SurfBind-V, to finetune SurfBind score with Autodock Vina score, is a successful attempt to boost the performance.

5 Conclusion

We introduce a ligand-protein docking scheme to calculate pairwise distance distribution as statistical potential. The use of a trainable surface point cloud encoder and scalar features makes the model Euclidean invariant. With the SDF constraint and pseudo-edge construction, SurfBind can derive more rational conformations matching the pocket shape and reduce RMSDs for generated ligands. Working as a score function, SurfBind achieves surpassing results on the docking and screening power for CASF-2016 benchmark. Further exploration contains multiple directions, e.g., to incorporate SDF and ligand generation in an end-to-end manner, explore relaxed restriction on rigid protein position.

References

Bao, J.; He, X.; and Zhang, J. Z. H. 2021. DeepBSP - a Machine Learning Method for Accurate Prediction of Protein-

Ligand Docking Structures. *Journal of chemical information and modeling*.

Berman, H.; Henrick, K.; and Nakamura, H. 2003. Announcing the worldwide protein data bank. *Nature Structural & Molecular Biology*, 10(12): 980–980.

Bishop, C. M. 1994. Mixture density networks.

Bordner, A. J.; and Gorin, A. A. 2007. Protein docking using surface matching and supervised machine learning. *Proteins: Structure*, 68.

Cao, Y.; Li, D.; Sun, H.; Assadi, A. H.; and Zhang, S. 2019. Efficient curvature estimation for oriented point clouds. *stat*, 1050: 26.

Fey, M.; and Lenssen, J. E. 2019. Fast graph representation learning with PyTorch Geometric. *arXiv preprint arXiv:1903.02428*.

Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; and Shenkin, P. S. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry*, 47 7: 1739–49.

Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; and Mainz, D. T. 2006. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry*, 49 21: 6177–96.

Gainza, P.; Sverrisson, F.; Monti, F.; Rodolà, E.; Boscaini, D.; Bronstein, M. M.; and Correia, B. E. 2019. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods*, 17: 184–192.

Ganea, O.-E.; Huang, X.; Bunne, C.; Bian, Y.; Barzilay, R.; Jaakkola, T.; and Krause, A. 2021. Independent SE(3)-Equivariant Models for End-to-End Rigid Protein Docking. *ArXiv*, abs/2111.07786.

Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; and Banks, J. L. 2004. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry*, 47 7: 1750–9.

Handoko, S. D.; Ouyang, X.; Su, C. T.-T.; Kwoh, C. K.; and Ong, Y. 2012. QuickVina: Accelerating AutoDock Vina Using Gradient-Based Heuristics for Global Optimization. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 9: 1266–1272.

Jones, D.; Kim, H.; Zhang, X.; Zemla, A. T.; Stevenson, G.; Bennett, W. F. D.; Kirshner, D. A.; Wong, S. E.; Lightstone, F. C.; and Allen, J. E. 2021. Improved Protein-ligand Binding Affinity Prediction with Structure-Based Deep Fusion Inference. *Journal of chemical information and modeling*.

Korb, O.; Stützle, T.; and Exner, T. E. 2009. Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *Journal of chemical information and modeling*, 49 1: 84–96.

Li, H.; Leung, K.-S.; Wong, M. H.; and Ballester, P. J. 2015. Improving AutoDock Vina Using Random Forest: The Growing Accuracy of Binding Affinity Prediction by the Effective Exploitation of Larger Data Sets. *Molecular Informatics*, 34.

Li, Y.; Su, M.; Liu, Z.; Li, J.; Liu, J.; Han, L.; and Wang, R. 2018. Assessing protein-ligand interaction scoring functions with the CASF-2013 benchmark. *Nature protocols*, 13 4: 666–680.

Liu, Z.; Su, M.; Han, L.; Liu, J.; Yang, Q.; Li, Y.; and Wang, R. 2017. Forging the basis for developing protein–ligand interaction scoring functions. *Accounts of chemical research*, 50(2): 302–309.

Lu, W.; Wu, Q.; Zhang, J.; Rao, J.; Li, C.; and Zheng, S. 2022. TANKBind: Trigonometry-Aware Neural NetworKs for Drug-Protein Binding Structure Prediction. *bioRxiv*.

Melzi, S.; Spezialetti, R.; Tombari, F.; Bronstein, M. M.; Stefano, L. D.; and Rodola, E. 2019. Gframes: Gradient-based local reference frame for 3d shape matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4629–4638.

Méndez-Lucio, O.; Ahmad, M.; del Rio-Chanona, E. A.; and Wegner, J. K. 2021. A Geometric Deep Learning Approach to Predict Binding Conformations of Bioactive Molecules. *Nature Machine Intelligence*.

Park, J. J.; Florence, P. R.; Straub, J.; Newcombe, R. A.; and Lovegrove, S. 2019. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 165–174.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32.

Sánchez-Cruz, N.; Medina-Franco, J. L.; Mestres, J.; and Barril, X. 2021. Extended connectivity interaction features: improving binding affinity prediction through chemical description. *Bioinformatics*.

Stärk, H.; Ganea, O.-E.; Pattanaik, L.; Barzilay, R.; and Jaakkola, T. 2022. EquiBind: Geometric Deep Learning for Drug Binding Structure Prediction. In *ICML*.

Storn, R.; and Price, K. 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4): 341–359.

Su, M.; Yang, Q.; Du, Y.; Feng, G.; Liu, Z.; Li, Y.; and Wang, R. 2019. Comparative Assessment of Scoring Functions: The CASF-2016 Update. *Journal of chemical information and modeling*, 59 2: 895–913.

Sverrisson, F.; Feydy, J.; Correia, B. E.; and Bronstein, M. M. 2021. Fast end-to-end learning on protein surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15272–15281.

Trott, O.; and Olson, A. J. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2): 455–461.

Velec, H. F. G.; Gohlke, H.; and Klebe, G. 2005. DrugScore(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. *Journal of medicinal chemistry*, 48 20: 6296–303.

Venkatraman, V.; Yang, Y. D.; Sael, L.; and Kihara, D. 2009. Protein-protein docking using region-based 3D Zernike descriptors. *BMC Bioinformatics*, 10: 407 – 407.

Zheng, L.; Meng, J.; Jiang, K.; Lan, H.; Wang, Z.; Lin, M.; Li, W.; Guo, H.; Wei, Y.; and Mu, Y. 2022. Improving protein–ligand docking and screening accuracies by incorporating a scoring function correction term. *Briefings in Bioinformatics*, 23.

Zhu, L.; Zhang, X.; Ding, H.; and Xiong, Y. 2010. Geometry of Signed Point-to-Surface Distance Function and Its Application to Surface Approximation. *J. Comput. Inf. Sci. Eng.*, 10.