Artificial Intelligence in the Life Sciences 3 (2023) 100072

Contents lists available at ScienceDirect

Artificial Intelligence in the Life Sciences

journal homepage: www.elsevier.com/locate/ailsci

Bayesian optimization for ternary complex prediction (BOTCP)

Arjun Rao[1,*], Tin M. Tunjic[1], Michael Brunsteiner[1], Michael Müller, Hosein Fooladi, Chiara Gasbarri, Noah Weber[2]

Celeris Therapeutics GmbH, Schmiedlstraße 3, Graz 8042, Austria

ARTICLE INFO

Keywords:
Bayesian optimization
Active learning
Machine learning
Ternary complex prediction
Targeted protein degradation

ABSTRACT

Proximity-inducing compounds (PICs) are an emergent drug technology through which a protein of interest (POI), often a drug target, is brought into the vicinity of a second protein which modifies the POI's function, abundance or localisation, giving rise to a therapeutic effect. One of the best-known examples for such compounds are heterobifunctional molecules known as proteolysis targeting chimeras (PROTACs). PROTACs reduce the abundance of the target protein by establishing proximity to an E3 ligase which labels the protein for degradation via the ubiquitin-proteasomal pathway. Design of PROTACs in silico requires the computational prediction of the ternary complex consisting of POI, PROTAC molecule, and the E3 ligase.

We present a novel machine learning-based method for predicting PROTAC-mediated ternary complex structures using Bayesian optimization. We show how a fitness score combining an estimation of protein-protein interactions with PROTAC conformation energy calculations enables the sample-efficient exploration of candidate structures. Furthermore, our method presents two novel scores for filtering and reranking which take PROTAC stability (Autodock-Vina based PROTAC stability score) and protein interaction restraints (the TCP-AIR score) into account. We evaluate our method using DockQ scores on a number of available ternary complex structures (including previously unevaluated cases) and demonstrate that even with a clustering that requires members to have a high similarity, i.e., with smaller clusters, we can assign high ranks to those clusters that contain poses close to the experimentally determined native structure of the ternary complexes. We also demonstrate the resultant improved yield of near-native poses[3] in these clusters.

1. Introduction

Targeted protein degradation is an emerging therapeutic modality which, instead of inhibiting the activity of a drug target, acts by inducing degradation of the target protein itself [1,2]. It employs so-called monofunctional degraders such as molecular glues or heterobifunctional degraders such as Proteolysis Targeting Chimeras (PROTACs) which act as proximity-inducing compounds. PICs bring together the protein of interest (POI) and an E3 ubiquitin ligase to induce subsequent ubiquitination and degradation of the target protein (See Fig. 1) [3–6].

Targeted protein degradation using PROTACs offers several advantages over traditional occupancy-based inhibitors such as Ciulli and Trainor [7]:

- Expanding the druggable space: there is no need for tight binding to the functional site of the POI, and as a result, PROTACs can target proteins more shallow binding pockets (e.g., transcription factors and scaffolding proteins).
- Working at a substoichiometric concentration: A single PROTAC molecule can induce degradation of multiple POI molecules over time.
- Improved efficacy: Recovery of target protein abundance (and therefore overall activity) is linked to protein synthesis in cells and not just to the absence of the PROTAC.
- Increased number of binding sites: PROTACs are typically larger than traditional inhibitors. This means that, albeit with a slight risk of side-effects due to less selective binding, binding sites that are traditionally inaccessable to other small molecules can be utilized. The promiscous binding of PROTAC fragments is typically compensated for by the fact that PROTAC efficiency is very specific to the linker designed.

* Corresponding author.
E-mail address: a.rao@celeristx.com (A. Rao).
[1] Equal contribution.
[2] Project Supervisor.
[3] The near-native pose is defined as a pose that has DockQ score ≥ 0.23.

https://doi.org/10.1016/j.ailsci.2023.100072
Received 21 November 2022; Received in revised form 8 March 2023; Accepted 5 April 2023
Available online 19 April 2023

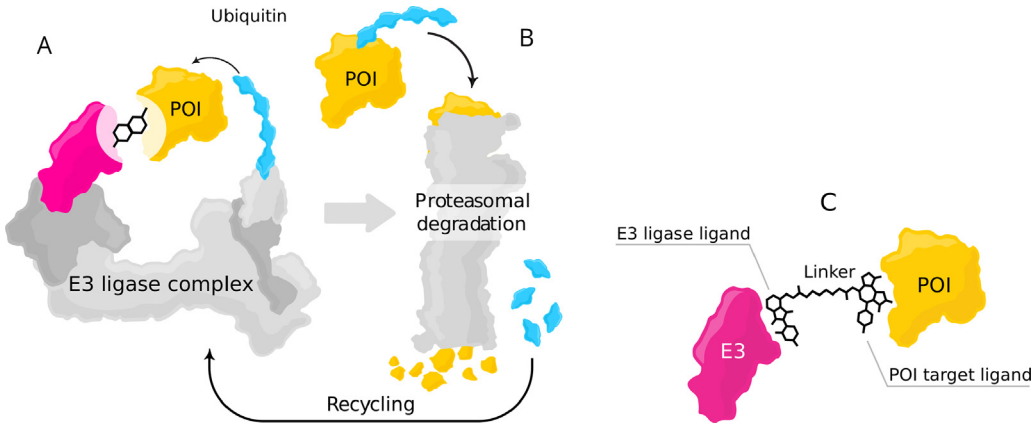

Fig. 1. Mode of action for targeted protein degradation: (A) A POI is shown in a proximity with an E3 ligase complex – the receptor part of E3 ligase is highlighted in magenta – a heterobifunctional degrader such as PROTAC brings together the POI and E3 ligase, which initiate the degradation process. (B) After formation of a ternary complex, the POI is marked with a polyubiquitin chain (blue), which subsequently leads to protein degradation through UPS. After degradation, the ubiquitin molecules are detached and recycled by the UPS. Moreover, the PROTAC molecule is detached and can induce degradation of multiple POIs (C) Ternary complex formed from the interaction of the POI, receptor part of the E3 ligase and heterobifunctional degrader. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Structurally, a PROTAC consists of a warhead that binds to the POI, an E3 ligand that binds into the E3 ligase, and a linker that connects these two parts. Current PROTACs utilize a limited number of E3 ligases. The most prominent E3 ligases are cereblon (CRBN) and Von-Hippel-Lindau Tumor Suppressor (VHL) due to the availability of high-affinity binders and favourable properties for both proteins. For CRBN binders, immunomodulatory drugs such as thalidomide [8], lenalidomide [9], and pomalidomide [10] are widely used. For VHL, binders were initially derived from a natural degron peptide and evolved to highly potent small molecules [11,12]. In addition, some other E3 ligases such as cell inhibitors of apoptosis protein (cIAP) [13], murine double minute 2 (MDM2) [14], and KEAP1 [15] have been harnessed for targeted protein degradation. There is an ongoing effort on expanding the library of E3 ligases and finding suitable potent ligands against other E3 ligases.

While the set of E3 ligases used for targeted protein degradation with PROTACs has so far been limited, they have been used to target a diverse range of POIs from different protein families. There are prominent examples with good evidence of safety and efficacy in initial clinical trials, including degraders targeting the androgen receptor (AR) [16,17], estrogen receptor (ER) [18], IRAK4, BCL-XL [19], Helios (IKZF2) [20], and GSPT1. Furthermore, in vivo studies for PROTAC targeting BRD4 [21], BTK [22,23], RIPK2 [24], and SMARCA2 [25] have shown promising results. These findings highlight the potential of PROTAC for the degradation of diverse protein families.

Judging the efficacy of a candidate PROTAC molecule in-silico is a daunting task owing to a plethora of interactions that are difficult to evaluate. For a PROTAC to be effective, it must at-least satisfy the following requirements. Firstly, it must be able to penetrate the cell membrane and stay stable inside the cell. Secondly, its interactions with the E3 ligase and the POI must result in the formation of a stable ternary complex where the POI is held near the E3 ligase. Finally, this complex must be one that lends itself to being tagged by ubiquitin so that the POI may be degraded.

In light of the above steps, it is clear that accurate and efficient prediction of the structure of the ternary complex is imperative for improvements at many stages of the PROTAC development process. For instance, computational modeling of ternary complexes allows the screening of multiple different E3 ligases for a candidate target, maximizing the stability of the ternary complex [26]. Knowing the structure of the ternary complex is the first step to calculating its stability, and experimental evidence shows that the formation of a stable ternary complex can lead to more efficient degradation [26].

Accurate and efficient ternary complex prediction is imperative for improvements at many stages of the PROTAC development process. For instance, computational modeling of ternary complexes allows the screening of multiple different E3 ligases for a candidate target, maximizing the stability of the ternary complex [26]. It can further facilitate linker design by investigating the impact of the linker on the ternary complex. These steps can incorporate different warheads, finding an ideal combination with available E3 ligase and binders and a linker.

However ternary complex prediction is in itself a challenging task, as the space of possible interactions between the two proteins and the PROTAC is astronomically large. The interactions include those between the warhead and the E3-binding fragment, and their respective proteins, the protein-protein interaction between the E3-ligase and the target protein, as well as the interactions and physical constraints imposed by the linker. Even if we know the fragments that bind to the E3 ligase and the POI, what constitutes an optimal linker is not yet fully understood [27]. The linker length and composition can have a significant impact on cooperativity and stability of the ternary complex [22,28], and consequently, on the efficiency of the degradation [29].

A method that addresses this problem in its entirety would need to predict the structure of the target protein, take into account it's flexibility, all of the various orientations of interacting side-chains, conformational changes that occur when binding to the PROTAC fragments, as well as the interaction of the linker to both proteins. This typically means computationally expensive and infeasibly long MD simulations. However, one arrives at a tractable subset of this problem using a few simplifying assumptions. Due to the abundance of data on binary interactions of proteins and ligands, it is often the case that we can find structures of the proteins co-crystallized with either the exact binder in the PROTAC, or a binder that structurally resembles it. An additional simplifying assumption is that proteins undergo only minor conformational changes to their backbone structure as a result of the interactions in the ternary complex. Prior works have, working with these assumptions, achieved success in estimating the structure of the ternary complex formed, and it is with these assumptions that we propose our method in this work.

In one of the first attempts for computationally predicting ternary complexes, Drummond et al. [30] proposed four different methods for the ensemble of PROTAC-mediated ternary complexes. With the most promising method of combining docking of the two protein-ligand structures with separate PROTAC conformer generation, they showed that it is possible to recover near-native (DockQ ≥ 0.23, see Section 2.7) ternary

complexes for a few available PDBs. In follow-up work [31], Drummond et al. added two additional methods which improved the previous performance. This work introduces clustering of the generated structures, an essential step in obtaining good results. Zaidman et al. [32] proposed another method, PRosettaC, which adds consecutive steps such as global protein-protein docking and local docking refinement, generating conformation of PROTACs that can fit in the docking solutions and final clustering. Although they have been able to reproduce a good amount of crystal structures of ternary complexes, they are initiating their prediction from bound structures, which does not reflect real-world conditions. Bound structures come from the decomposition of already available ternary complexes, but a priori we do not have access to this information. In one of the recent works, Weng et al. [33] proposed a method again based on global protein-protein docking, local refinement with RosettaDock [34], generating PROTACs using RDkit, and applying filtering and re-ranking criteria for predicting near-native ternary complexes. They started from unbound structures and showed that good results could be achieved with these methods. Furthermore, Li et al. [35], and Liao et al. [36] offer different approaches to rank ternary complexes where they use longer MD simulations. Their approaches result in good outcomes but are unsuitable for a real-world scenario, where we might try a large number of PROTACs, since they require a lot of computing resources and extended simulation time.

The first contribution of our current work is the demonstration of the successful use of a Bayesian optimization (BO) [37] framework to perform a sample-efficient exploration of the space of ternary complex candidates. BO is an established method that allows optimizing functions that are costly to evaluate by proposing promising evaluation candidates (here: candidate ternary complex structures). This leads to sample efficiency, i.e., fewer candidates need to be evaluated (e.g., by a docking software) compared to (e.g., Monte Carlo-) search methods, and allows the use of potentially complex fitness functions which are needed to characterize the multiple interactions within a ternary complex. Using this approach, we can efficiently sample poses that are viable candidates for the structure of the ternary complex, with the viability scored by a weighted sum of protein-protein interaction (PPI) score and a PROTAC score. The PPI score quantifies the interactions between the proteins in a particular candidate. In contrast, the PROTAC score considers the feasibility of embedding a conformation of the PROTAC molecule, given the position of the binding pockets in this configuration of proteins.

The second contribution of our work lies in designing of two novel scores specialized for filtering, and ranking. The first score is used for filtering favourable complexes by imposing constraints on the distances of the protein surfaces from one another – a version of the ambiguous interaction restraints energy (here abbreviated as “TCP-AIR Energy”) [38] that has been modified specifically for ternary complexes (see Fig. 5). The second score is the PROTAC stability score, which is calculated using a modified version of Autodock-Vina. Given a particular configuration of proteins, it packs the PROTAC between the proteins and indicates its stability and energetic favourability in this position. This score is useful for ranking.

We show in our current work (See Table 2) that, with efficient sampling with the BO loop, combined with filtering, clustering, and ranking using the above scores, we are able to effectively find highly ranked clusters with near-native poses. Specifically, our clustering is more confined compared to previous works, and each cluster contains only a small number of conformers. Our ability to rank these effectively gives us clusters with a high yield of near-native scores.

2. Methods

In this paper, we propose the Bayesian Optimization for Ternary Complex Prediction (BOTCP) method. At its core lies a Bayesian optimization (BO) loop capable of finding appropriate ensembles of ternary complexes by optimizing a combined fitness function which includes a fitness describing the strength of the interaction of the two proteins (“PPI fitness”) and a fitness modeling the possibility of connecting the binding pockets with the PROTAC (“PROTAC fitness”). We then perform a local optimization around the most promising candidates using a score which takes into account the stability of the PROTAC between the two proteins (“PROTAC stability score”, calculated by our own extension to Autodock-Vina). Afterwards, we cluster the resulting candidates before performing filtering using an extension to the ambiguous interaction restraint (AIR) energy [38], and re-ranking the final selections using the PROTAC stability score. In a nutshell, BOTCP consists of the following steps/stages:

1. Input initialization (Section 2.1)
2. Rigid pose sampling via Bayesian optimization (Section 2.2)
3. Local optimization with simulated annealing of the PROTAC stability score (Section 2.3)
4. Clustering, filtering using TCP-AIR energy, and re-ranking (Section 2.4)
5. Structural refinement (Section 2.5)

2.1. Initialization

At the time of writing, the protein database (PDB [39]) contained eighteen experimentally (X-ray) determined ternary complex structures, which include a PROTAC molecule, a protein of interest, and an E3 ligase. In comparison to the previous papers, we excluded (6BN8, 6BN9, and 6BNB) since the coordinates of PROTAC atoms have not been resolved and added three additional ternary complex structures (7JTO, 7Q2J, and 7JTP). In addition, we excluded (7PI4) since the POI is a kinase with a large flexible loop close to the protein-protein interface and the PROTAC molecule, which has not been resolved in the experimental structure. The remaining 14 structures are listed in Table 1. Each structure was downloaded from the PDB in its native format and preprocessed as follows. All (non-PROTAC) hetero atoms were removed, and the chains corresponding to the POI and the E3 ligase as well as the PROTAC molecule were saved in a single PDB file for comparison with predictions. In several cases, the asymmetric unit contained two or more ternary complexes, which were saved individually. The protein structures used as input for modeling were taken from other PDB entries (unbound structures) so as to emulate the practical use case where the experimental structure of the ternary complex is not available. The structures used here include three different E3 ligases and seven POIs. For each of these ten proteins, a PDB entry was selected according to the following criteria :

1. The protein has a bound ligand that is similar to the warhead or binder of the corresponding PROTAC molecule.
2. The structure comes with few and/or short gaps/unresolved loops and the resolution (*R*-factor) is good.

For one structure (7KHH), no PDB entry fulfilling the first criterion was available. Here we estimated the structure of the bound warhead by docking the warhead into the binding pocket, using the GNINA docking tool [40] with default parameters. The quality of the resulting complex structure was ascertained by visual inspection and comparison to the corresponding portion of the ternary complex structure. The resulting warhead and binder molecules (only non-hydrogen atoms) were saved in SDF format for usage as inputs to the BO method. Each of the ten individual single structures of E3 ligase or POI were then preprocessed using the following steps:

1. All heteroatoms (including water and small molecules) were removed while keeping ions (in four cases, the structure contained a Zinc ion complexed by four cysteine or histidine residues).
2. For those structures that had gaps, we downloaded the corresponding AlphaFold structures [41,42], aligned the experimental and the AlphaFold structures using Chimera [43], and used the coordinates of the AlphaFold structures to fill the gaps.

Table 1
Details of the protein structures used in this work. Columns show the PDB ID for each ternary complex as well as the chain identifiers (complex), the gene name, the PDB ID of the structure used as input for prediction (template), the chain IDs of E3 ligase and POI, the first and last residue number included into the used models (N-term, C-term). For the PROTAC, the colums show the residue ID, the nature of any ionizable group, and the net charge used in molecular models.

Complex	E3 ligase					POI					PROTAC		
PDB ID	Name	Template	Chain	N-term	C-term	Name	Template	Chain	N-term	C-term	Residue	Ionizable	Net charge
5T35_1	VHL	4W9H-I	D	62	204	BRD4-2	5UEU-A	A	349	457	759	no	0
5T35_2	VHL	4W9H-I	H	62	204	BRD4-2	5UEU-A	E	349	457	759	no	0
6BN7	CRBN	4TZ4-C	B	48	426	BRD4-1	3MXF-A	C	44	168	RN3	imine	0
6BOY	CRBN	4TZ4-C	B	48	426	BRD4-1	3MXF-A	C	44	168	RN6	imine	0
6HAX_1	VHL	4W9H-I	B	62	204	SMARCA2	6HAZ-A	A	1378	1490	FWZ	piperazine	1
6HAX_2	VHL	4W9H-I	F	62	204	SMARCA2	6HAZ-A	E	1378	1490	FWZ	piperazine	1
6HAY_1	VHL	4W9H-I	B	62	204	SMARCA2	6HAZ-A	A	1378	1490	FX8	piperazine	1
6HAY_2	VHL	4W9H-I	F	62	204	SMARCA2	6HAZ-A	E	1378	1490	FX8	piperazine	1
6HR2_1	VHL	4W9H-I	B	62	204	SMARCA4	6ZS2-A	A	1449	1569	FWZ	piperazine	1
6HR2_2	VHL	4W9H-I	F	62	204	SMARCA4	6ZS2-A	E	1449	1569	FWZ	piperazine	1
6SIS_1	VHL	4W9H-I	D	62	204	BRD4-2	5UEU-A	A	349	457	LFE	sec amine	1
6SIS_2	VHL	4W9H-I	H	62	204	BRD4-2	5UEU-A	E	349	457	LFE	sec amine	1
6W7O_1	BIRC2	6W74-A	C	266	349	BTK	5P9J-A	A	396	656	TL7	sec amine	1
6W7O_2	BIRC2	6W74-A	D	266	349	BTK	5P9J-A	B	396	656	TL7	sec amine	1
6W8I_1	BIRC2	6W74-A	D	266	349	BTK	5P9J-A	A	396	656	TKY	sec amine	1
6W8I_2	BIRC2	6W74-A	E	266	349	BTK	5P9J-A	B	396	656	TKY	sec amine	1
6W8I_3	BIRC2	6W74-A	F	266	349	BTK	5P9J-A	C	396	656	TKY	sec amine	1
6ZHC	VHL	4W9H-I	A	62	204	BCL2L1	4QVX-A	D	2	197	QL8	carboxylic acid	-1
7JTO	VHL	4W9H-I	L	62	204	WDR5	4QL1-A	B	32	333	VKA	2 piperazines	2
7JTP	VHL	4W9H-I	L	62	204	WDR5	4QL1-A	A	32	333	X6M	piperazine	1
7KHH	VHL	4W9H-I	C	62	204	BRD4-1	3MXF-A	D	44	168	WEP	no	0
7Q2J	VHL	4W9H-I	C	62	204	WDR5	4QL1-A	D	32	333	8KH	piperazine	1

Table 2
Results of the BOTCP method for ternary complex prediction on unbound structures before the structural refinement (using greedy clustering 2.4). [1]The best rank containing at least one model with DockQ≥0.23. [2]The total number of clusters. [3]The percentage of the near-native poses in that specific ranked cluster. For each ternary complex, we selected a cluster which includes an RRT with a high DockQ score. Columns also show the PDB ID and characterization of the best RRT in terms of f_{nat}, I_{RMSD}, L_{RMSD} and DockQ score (see text).

PDB ID	Rank of cluster[1]	Number of clusters[2]	% Near-native[3]	f_{nat}	I_{RMSD}	L_{RMSD}	DockQ
6HAY_1	2	914	100	0.61	1.07	3.19	**0.72**
6HAY_2	5	946	100	0.61	1.77	3.03	**0.64**
7JTO	15	403	50	0.5	3.05	14.82	**0.31**
6BN7	15	917	6.7	0.14	3.33	10.28	**0.24**
7Q2J	7	322	79.1	1.0	2.32	5.86	**0.66**
6W7O_1	11	269	9.4	0.17	3.77	7.31	**0.29**
6W7O_2	4	271	29.2	0.35	2.92	7.99	**0.36**
5T35_1	76	765	94.1	0.73	2.09	4.92	**0.61**
5T35_2	51	763	88.9	0.67	1.81	8.36	**0.53**
6W8I_1	13	362	96.6	1.0	2.248	4.41	**0.7**
6W8I_2	79	364	25	0.52	2.61	11.35	**0.38**
6W8I_3	6	365	84.6	0.81	1.85	3.7	**0.68**
6HAX_1	8	816	100	0.57	2.61	5.34	**0.51**
6HAX_2	14	795	100	0.83	1.39	8.95	**0.62**
7KHH	None	None	None	None	None	None	**None**
6BOY	19	720	80	0.61	2.27	10.04	**0.44**
6SIS_1	1	339	100	0.53	1.72	5.62	**0.55**
6SIS_2	8	365	100	0.65	2.14	4.56	**0.59**
6ZHC	142	670	75	0.82	3	19.15	**0.4**
7JTP	8	170	88.1	0.68	2.52	5.54	**0.55**
6HR2_1	26	727	41.7	0.44	2.27	12.95	**0.35**
6HR2_2	54	679	25	0.14	3.51	8.93	**0.26**

3. We cropped flexible N- and C-terminal tails (usually with poor confidence levels in AlphaFold predictions) as specified in 1.
4. Missing atoms were replaced using the PDBfixer tool [44].
5. We used Reduce [45] to protonate/add hydrogens, and, where appropriate, flip asparagine, glutamine, or histidine side chains.

The PROTAC molecules were extracted from the original PDB file and converted to SDF format using Open Babel [46]. If required, bond orders and protonation were corrected manually. Here we assumed that all weak bases (secondary and tertiary amines) and acids (carboxylic acid) are (de)protonated to form anions or cations. Since all models and force fields we use work with fixed protonation states only, we had to make a decision here. The most obvious choice was the protonation as expected in solution at physiological pH. As the pKa value of ionizable groups can change with the environment, e.g., when transferring from solution to a binding site, this protonation solution is not necessarily optimal. However, accounting for this would require the determination of pKa values on-the-fly, which is not straightforward. In a few cases, the presence of piperazines required a choice to protonate one of the two nitrogens. This choice was made based on chemical intuition and the corresponding expected pKA variations. All SDF files are provided in the SI. The protein PDB and PROTAC SDF files resulting from this procedure were used as input for the BO loop.

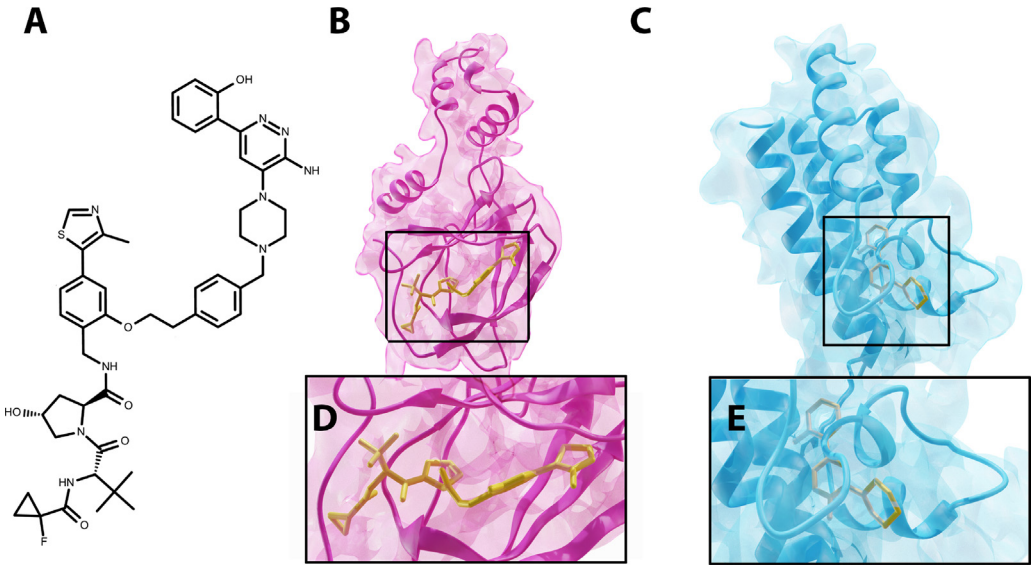

Fig. 2. Example of the required inputs for the ternary complex 6HAX (PDB ID). (A) PROTAC molecule. (B) and (D) E3 ligase docked to warhead. (C) and (E) POI docked to the warhead.

Our method uses three inputs: the POI with the docked warhead, the E3 ligase with the docked E3 binder, and the PROTAC SMILES (see Fig. 2 for an example), which are provided to the Bayesian optimization loop. It is more suitable to access the co-crystal structure of POI + warhead and E3 ligase + binder, but if this information is not available, it is required first to dock the corresponding ligands into their proteins and use these structures after molecular docking for the next step in out BOTCP pipeline.

We apply our method to the structures of proteins and docked ligands extracted from the co-crystal structure of known PROTAC ternary complexes (so-called bound protein data) and the much more challenging case of unbound data, where the co-crystal structure of known PROTAC ternary complexes is not available.

2.2. Rigid pose sampling via Bayesian optimization

In order to determine the structure of the ternary complex, we first recognize that the two proteins can interact in a variety of poses. The first step in our pipeline is to sample a set of such poses that are viable candidates for the ternary complex. To do so we consider the input POI and E3 ligase as rigid structures. In this case, each pose is characterised by the relative position and orientation of the POI with respect to the E3 Ligase. We then search through the space of poses, and sample poses that are viable structures for the ternary complex.

2.2.1. Representation of poses

In order to encode a pose, we consider a baseline pose where both the POI and E3 Ligase are shifted so that their center of masses line up at the origin. Keeping the position of the E3-ligase fixed, each pose then corresponds to a translation and rotation of the POI. This is denoted by a 7-D vector $\mathbf{x} \in \mathbb{R}^7$ that represents the relative rotation and translation (RRT) of the POI with respect to its position in the baseline pose, where $\mathbf{x}_{1\ldots3}$ is the relative translation, and $\mathbf{x}_{4\ldots7}$ is a quaternion representing a relative rotation. Due to the fact that a protein-protein configuration can be represented as such, for the remainder of this paper, we will refer to a specific protein-protein configuration using the term RRT. Our objective above then becomes, to sample a set of RRT's that are viable candidates for the structure of the ternary complex.

2.2.2. Fitness of a pose (RRT)

Here we make clear what we mean by a particular RRT being a 'viable candidate' for the structure of the ternary complex. For a particular RRT $\mathbf{x} \in \mathbb{R}^7$, we define a fitness function $y(x)$ that is a weighted combination of two fitness components:

$$y(\mathbf{x}) = y_{\text{PPI}}(\mathbf{x}) + \lambda \cdot y_{\text{PROTAC}}(\mathbf{x}), \tag{1}$$

where λ determines the relative weighting of the PPI fitness y_{PPI} vs. the PROTAC fitness y_{PROTAC} (we used $\lambda = 1$). The PPI fitness represents the strength of interaction between the two proteins. While strong PPIs are important for a stable ternary complex, there are cases in which the proteins are oriented such that the respective binding pockets for the E3-binder and the warhead are too far apart for the length of the linker. On the other hand, the proteins' relative positions might result in binder and warhead being too close, leading to steric clashes within the linker itself. The PROTAC fitness $y_{\text{PROTAC}}(\mathbf{x})$ ensures these cases have a low overall fitness values.

PPI fitness A number of protein-protein docking scoring schemes have been proposed in the literature. In this work, we use the distance-scale finite ideal-gas reference (DFIRE) energy score [47] as y_{PPI}. This score is a relatively unbiased estimate of the validity of the protein-protein interaction while allowing rapid evaluation. The PPI fitness serves to filter out infeasible protein-protein orientations by assigning a low score to RRTs that result in steric clashes or the proteins being too far apart.

However, manual inspection of the native ternary complex structures show that the poses corresponding to the crystal structure do not correspond to high surface contact between the proteins. This means that optimizing the DFIRE score is likely to bias us away from sampling near-native poses. Hence we choose to cutoff the value of DFIRE score at 6 as below

$$y_{\text{PPI}} = \min(6.0, \text{DFIRE}(\mathbf{x})) \tag{2}$$

PROTAC fitness

Our input consists of the POI + docked warhead and the E3 ligase + docked binder. Since each RRT corresponds to a translation and rotation of the POI, this means that the docked coordinates of the warhead is also transformed accordingly. The PROTAC fitness $y_{\text{PROTAC}}(\mathbf{x})$ aims to quantify the feasibility of finding a valid linker conformer given the docked coordinates of the binder and warhead corresponding to a particular RRT $\mathbf{x}$. We generate conformations by minimizing the energy of the UFF molecular forcefield [48] (using RDKit [49]). However, we additionally constrain the atoms of the binder and warhead at their respective docked coordinates. We then sample linker conformations by minimizing the UFF energy [48] starting from a random (unconstrained) initial conformer generated using distance geometry. The choice of the

simple UFF forcefield is because it allows us to constrain atoms to atoms to specific positions by using a quadratic potential that penalizes deviation from these target positions. This quadratic potential proves useful when approximating the PROTAC score (see following section)

After generating 4 candidate conformers, we choose the output to be the conformer with the lowest UFF energy. The choice of 4 conformers is because it is the smallest number of conformers we find is needed to reliably report the minimum energy constrained conformer. The PROTAC fitness is defined as the negative square-root of the minimized UFF energy of this conformer.

Note that when calculating the PROTAC score above, we do not consider the interactions of the PROTAC with the proteins, thus, it cannot be used to filter out cases with steric clashes between the linker and the proteins. These cases are taken into account by the PROTAC stability score (see Section 2.3).

Approximating PROTAC fitness with a neural-network

The PROTAC fitness for a particular RRT is a function of the energy obtained after minimization of the UFF force-field energy. This minimization is computationally expensive, and the result of the minimization is not an analytic function of the RRT, which means that it is expensive to optimize the RRT with respect to this score. As a solution, we have found that we can train a neural network model that can effectively approximate the PROTAC fitness for a specific RRT.

For a particular POI, E3 ligase, and PROTAC, we train a neural network model which maps the 7D RRT vector $\mathbf{x}$ to the PROTAC fitness values generated using RDKit (as described above). The model consists of two components: a simple model which learns to predict asymptotic component of the energy, which is asymptotically linear w.r.t the distance between the binding pockets of the PROTAC fragments for the particular RRT, and a 3-hidden layer neural network which learns to predict the residual component, capturing the more complex aspects of the score such as self-collisions that may arise due to the RRTs bringing the binding pockets too close.

Prior to the BO loop, we pre-train this model using a dataset of RDKit-generated conformers for different RRT, and use this model to calculate an approximation to the PROTAC fitness for the $\mathbf{x}$ values proposed by the Bayesian optimization loop. The advantage is that, after being pre-trained, computing the fitness for a new RRT using the model requires just one forward pass, making the fitness computation very fast. We find that even though we require additional computation to train the model, the resultant speed gain makes up for it when optimizing the PROTAC score in the BO loop.

Local fitness ascent

As is often seen in biological systems, small changes in position can drastically affect fitness values for example, as a result of steric clashes. This makes the fitness landscape highly irregular with regions of low fitness directly adjacent to regions of the highest fitnesses. In order to effectively learn this fitness landscape via the Bayesian optimization, we modify the fitness function as follows. For a particular RRT $\mathbf{x}$, we perform a local ascent in the space of RRTs using a black box optimization algorithm (Nelder-mead simplex ascent [50] for 15 steps), and take the final optimized fitness $y_{\text{opt}}(\mathbf{x})$ to be the fitness corresponding to $\mathbf{x}$. This optimization means that even RRTs with steric clashes will be mapped to nearby RRTs that have high fitnesses, making the landscape a lot more regular. Additionally, the maxima of $y_{\text{opt}}(\mathbf{x})$ correspond to the maxima of $y(\mathbf{x})$ meaning that we can maximize one instead of the other. This approach is described in Algorithm 1.

Algorithm 1 Evaluation of fitness.

function EVALUATEFITNESS(X)
  for $\mathbf{x}_i$ **in** X **do**
    maximize $y(\mathbf{x})$ from initial point $\mathbf{x}_i$ using black box optimization ▷ Equation∼1
    $y_{\text{opt},i} \leftarrow$ final $y(\mathbf{x})$ after k steps of ascent
  end for
  return $(y_{\text{opt},1}, y_{\text{opt},2}, \cdots y_{\text{opt},|X|})$ ▷ concatenate
end function

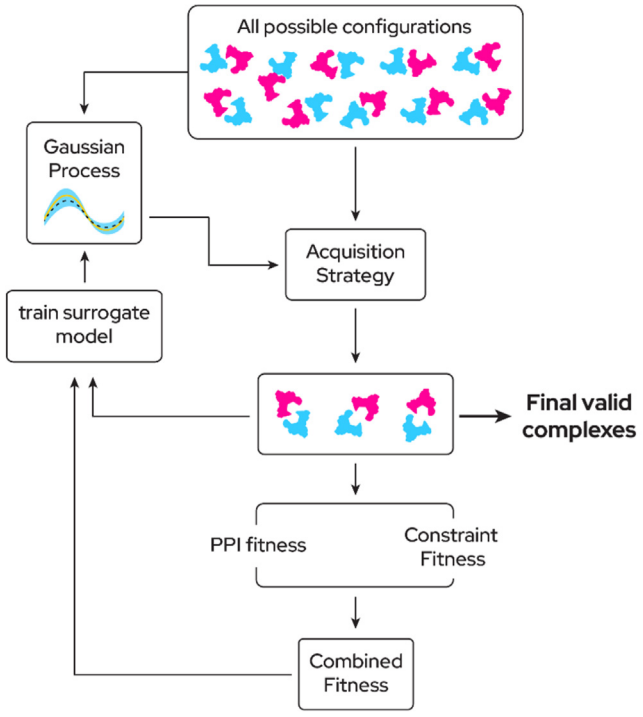

Fig. 3. Overview of Bayesian optimization pipeline. We are sampling from the roto-translational space of POI. Then we assign a score for each sample $D = \{(\mathbf{x}_i, y_i)\}$. Then, the dataset D is used for updating the GP into posterior GP. For sampling the next batch of points, the acquisition function is used. In our proposed method, we used the UCB acquisition function to determine the next sampling points. This iterative process continues until the end of some Bayesian optimization steps. The number of iterations is a hyperparameter that controls when do we get the list of final valid complexes.

2.2.3. Bayesian optimization

When sampling viable RRTs, We wish to select each RRT $\mathbf{x}$ by optimizing the black box function $y(\mathbf{x})$ described above. However, while we can evaluate $y(\mathbf{x})$ at any RRT $\mathbf{x}$, the evaluations are costly, and the function has no analytical form, i.e. we do not have any information about the gradient, thus becoming a problem of gradient-free or black-box optimization. In this work, we choose to approach this problem via the technique of Bayesian Optimization [51,52] as it is known to be very sample efficient, which is useful in the current context owing to the computational complexity of the fitness function we are optimizing. The way we achieve this is through the use of a surrogate model $f(\mathbf{x})$. Assuming that we have a set of M RRTs $\mathbf{x}_i \;::\; i \in \{1 \dots M\}$, for which we have calculated the corresponding fitness values y_i, the surrogate model is trained so that $f(\mathbf{x}_i)$ approximates the true value of the function y_i. However the speciality of a surrogate model is its ability to extend its knowledge of the known RRTs $\mathbf{x}_i, y_i$ onto unknown RRTs. More specifically, we can evaluate $f(\mathbf{x})$ on a previously unseen RRT, for which the model returns it's prediction, and more importantly, its uncertainty in its prediction. $f(\mathbf{x})$ is typically chosen to be computationally cheap to calculate as well as analytical, so that it can be optimized using gradient information.

This allows us to use the following loop to sample RRTs with high fitnesses (see Fig. 3 and Algorithm 2). We start with a random set of initial RRTs, evaluate the fitnesses for these RRTs, and train the surrogate model. Using the surrogate model, we use an acquisition strategy to sample a new set of RRTs with a high predicted fitness (exploitation)

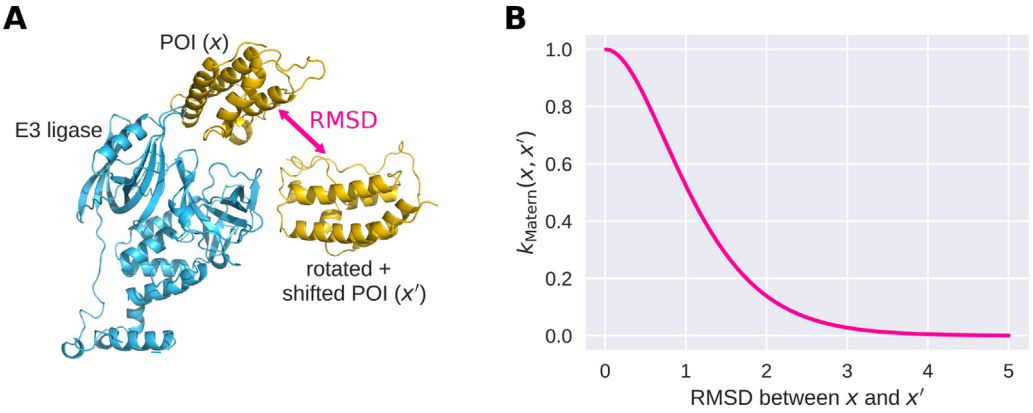

Fig. 4. Illustration of the Matern kernel, used as covariance kernel for the Gaussian process. (A) RMSD between different POI positions relative to the E3 ligase (i.e., different RRTs **x**). (B) Matern kernel (see Eq. (4)) as a function of the RMSD.

Algorithm 2 The BO loop.

$X_{\text{train}} \leftarrow N_{\text{init}}$ randomly sampled RRTs
$\mathbf{y}_{\text{train}} \leftarrow$ corresponding fitness values
repeat
  Fit surrogate model (GP) $\mathcal{M}$ using $X_{\text{train}}, y_{\text{train}}$
  $X_{\text{new}} \leftarrow$ AcquirePoints($\mathcal{M}$)
  $\mathbf{y}_{\text{new}} \leftarrow$ EvaluateFitness(X_{new})
  $X_{\text{train}} \leftarrow$ Concatenate(X_{train}, X_{new})
  $\mathbf{y}_{\text{train}} \leftarrow$ Concatenate($\mathbf{y}_{\text{train}}$, $\mathbf{y}_{\text{new}}$)
until stopping condition is met

and/or high uncertainty (exploration). We then evaluate the fitnesses for these points and retrain the surrogate model and repeat the procedure. The details of the surrogate model and acquisition strategy are described below.

Surrogate model

While in principle Bayesian optimization can utilize different surrogate models, we employed the widely used Gaussian process (GP) [53,54]. A GP is a model that places a prior on the space of functions that can be used to describe the data, by specifying a prior on the mean of the function value $m(\mathbf{x})$ and via a covariance kernel $k(\mathbf{x}, \mathbf{x}')$ which describes the covariance between the function values at any two points **x** and **x**′, i.e.,

$$f(\mathbf{x}) \sim \mathcal{GP}\Big(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\Big) \,. \tag{3}$$

We use a constant learned value $m(\mathbf{x}) \equiv m_0$ as prior mean. As covariance kernel, we use the a well-known Matern kernel [54] (see Fig. 4), i.e.,

$$k(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)}\Big(\sqrt{2\nu}d\Big)^{\nu} K_{\nu}\Big(\sqrt{2\nu}d\Big) \tag{4}$$

where d is the distance between **x** and **x**′, ν is a smoothness parameter (taking on values $\frac{1}{2}$, $\frac{3}{2}$, or $\frac{5}{2}$, with larger values increasing smoothness), and K_{ν} is a modified Bessel function.

In our optimization, We use $\nu = \frac{5}{2}$, and as distance d, we use the RMSD between the atomic coordinates of the atoms of the POI scaled by a learned parameter Θ. We make use of the fact that the RMSD between two rigid transformations of the POI can be efficiently calculated by precalculating the moment-of-inertia tensor for the POI, thus not requiring us to transform all the coordinates for each RRT. This makes this kernel very computationally efficient to compute.

Acquisition strategy

Bayesian optimization requires a strategy by which we select new candidate RRTs for which we evaluate the fitness. Intuitively, we want to perform exploration by sampling points for which the surrogate model has no information about their potential value, while also performing exploitation by sampling points for which the surrogate model estimates a high fitness value. At the same time, we should avoid sampling points where the surrogate model already possesses enough information to predict that the fitness will be low. Formally, we define an acquisition function $\alpha(\mathbf{x})$, and evaluate the fitness on values **x** which maximize $\alpha(\mathbf{x})$. There are different acquisition strategies allowing to balance exploration and exploitation [37]. We found that the popular upper confidence bound (UCB) acquisition function [37,55] is effective in our scenario. It is defined as

$$\alpha(\mathbf{x}; \beta) = \mu(\mathbf{x}) + \sqrt{\beta}\sigma(\mathbf{x}) \,, \tag{5}$$

where $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$ are the predicted mean and standard deviation of $f(\mathbf{x})$ (which can be calculated from the existing data points using the mean function and covariance kernel). The parameter $\beta > 0$ defines the balance of exploration versus exploitation (a large β leads to large α at points where the standard deviation is high, thus increasing the exploration, while a small β leads to large α at points with a high predicted mean, thus increasing exploitation).

Our full acquisition strategy is described in Algorithms 3 and 4. In each iteration of the BO loop, we sample two batches of new points, one for exploitation and for exploration, by optimizing the above acquisition function, with low and high β values respectively. This optimization is

Algorithm 3 Acquiring new points for evaluation.

function AcquirePoints($\mathcal{M}$)
  $X_{\text{rand}} \leftarrow N_{\text{rand}}$ randomly sampled RRTs
  $X_{\text{optimized}}, \alpha_{\text{optimized}} \leftarrow$ LocalOptimizeUCBAcq($\mathcal{M}$, X_{rand}, $\beta := 10.0$)
  $X_{\text{exploit}} \leftarrow$ filter N_{exploit} values from $X_{\text{optimized}}$ with the highest $\alpha_{\text{optimized}}$
  $X_{\text{perturbed}} \leftarrow$ random perturbations of X_{exploit}
  $X_{\text{explore}} \leftarrow$ LocalOptimizeUCBAcq($\mathcal{M}$, $X_{\text{perturbed}}$, $\beta := 100.0$)
  $X_{\text{new}} \leftarrow$ Concatenate(X_{exploit}, X_{explore})
  return X_{new}
end function

Algorithm 4 Acquisition optimization algorithms.

function LocalOptimizeUCBAcq($\mathcal{M}$, X, β)
  for $\mathbf{x}_i$ **in** X **do**
    maximize $\alpha(\mathbf{x}; \beta)$ from initial point $\mathbf{x}_i$ using gradient ascent ▷ Equation~5
    $\mathbf{x}_i^* \leftarrow$ final **x** after k steps of ascent
  end for
  return $(\mathbf{x}_1^*, \mathbf{x}_2^*, \ldots \mathbf{x}_{|X|}^*)$, $(\alpha(\mathbf{x}_1^*), \alpha(\mathbf{x}_2^*), \ldots \alpha(\mathbf{x}_{|X|}^*))$ ▷ concatenate
end function

performed using gradient ascent (using ADAM [56]). We use the open-source BoTorch package [57] to solve this optimization problem and implement the BO loop.

2.3. Local optimization with simulated annealing of the PROTAC stability score

From the above points, we filter out RRTs with infeasible PROTAC scores or PPI scores, i.e. we select only complexes that have PPI scores that are at-least 6.0 (the clipped value), and PROTAC scores that correspond to a UFF energy lesser than 1.5 times the minimum UFF energy. We next locally optimized the remaining RRTs using a combination of the PPI score and a PROTAC stability score. We performed the optimization with the simulated annealing.

These thresholds above were chosen heuristically to optimize the number of near-native structures that pass the filter for the compounds 6HAX_1, 6HAX_2, 6HAY_1, 6HAY_2. These were not optimized subsequently based on the final ranking in order to avoid overfitting. Any parameter that is mentioned in the following sections is set based on the data generated when running the BOTCP protocol for the above four compounds.

Simulated annealing We used a sum of the PROTAC stability score, described underneath, and the PPI score to perform local optimization of the remaining candidates using simulated annealing [58]. For any RRT, simulated annealing proceeds via a sequence of 10 steps, where in each step we randomly sample an RRT near the current RRT. We evaluate the weighted sum of the PROTAC stability and PPI scores and stochastically accept and reject it based on that. Note, that in order to not waste computation on infeasible RRTs, the randomly generated samples above are pre-optimized to have a reasonable PPI and PROTAC score.

As this process allows the PROTAC conformation to change (both explicitly through simulated annealing and implicitly within Autodock-Vina), it represents a structural refinement / packing of the PROTAC conformer.

PROTAC stability score

We calculated a PROTAC stability score using an extension of AutoDock-Vina [59]. This score takes into account the stability of the PROTAC within the context of the protein configuration corresponding to a particular RRT (which is ignored in the calculation of the PROTAC fitness, see above).

In order to do this, we extended Autodock-Vina by implementing an additional restoring energy, that allows one to constrain the position of a subset of atoms of the ligand (which is being docked) to specific coordinates. Consider for example, a particular ligand where a subset of atoms S_c are constrained to specific coordinates. Now let's denote for a particular conformation of the ligand, the coordinates of atom j by $r_j \in \mathbb{R}^3$, and its constrained position by $r_{0,j} \in \mathbb{R}^3$, then the restoring energy E^R is given by

$$E^R = k^R \sum_{j \in S_c} \left\| r_j - r_{0,j} \right\|^2 \tag{6}$$

This energy grows quadratically, and imposes a quadratic penalty on the deviation of an atoms position from it's constrained position. The gradient of this restoring energy is termed the restoring force. We used this force to constrain the E3 binder and the warhead to have the coordinates of the respective binding pockets given the current protein-protein RRT. To calculate the PROTAC stability score for a particular RRT, we start from 10 randomly generated initial PROTAC conformations, and for each of them perform the following:

1. We optimize the PROTAC molecule's conformation under only the restoring force and the intra-molecular part of the Vina force field, creating a viable conformer with the binding fragments held at their positions.
2. We then turned off the restoring force (i.e. set $k^R = 0$) and perform an optimization using the full (intra- and inter-molecular) Vina force field.
3. For the conformation that it ends up in, we calculate the energy according to the vina score (called the Vina energy E^V), as well as the restoring energy E^R using the original value of k^R

To understand the PROTAC stability score, We describe the heuristic interpretations of the Vina energy and the restoring energy. First, the Vina energy E_i^V (i indexes a particular PROTAC conformer from among the 10 sampled), which is the combination of the intermolecular and internal energy of the conformer. Second, the restoring energy, E_i^R which measures the deviation of the E3 binder and warhead from their constrained positions in the particular RRT (as in Eq. (6)).

To generate a score corresponding to the stability of the PROTAC in a particular protein-protein RRT, we look at the following cases:

1. E_i^V is low and E_i^R is low: the conformation is favourable (has low vina energy) and the warhead and E3 ligase are stable in their respective binding pockets
2. E_i^V is low and E_i^R is high: the conformation is favourable (has low vina energy), but the warhead and E3 ligase need to be displaced to achieve this conformation.
3. E_i^V is high: the conformation has a high vina energy and is unfavourable.

We thus define the following weighted average of the Vina energy as the PROTAC stability score:

$$y_{\text{stability}} = \log \left(\frac{\sum_{i=1}^{n} e^{\frac{-E_i^V}{\eta^V}} e^{\frac{-E_i^R}{\eta^R}}}{\sum_{i=1}^{n} e^{\frac{-E_i^R}{\eta^R}}} \right) \tag{7}$$

where η^V and η^R are scaling parameters set to 2 and 50 respectively. These values were again picked heuristically to bring the vina energy and constraint energy within the same range. This score takes an average that assigns a higher weightage to conformations that are stable (i.e. higher E_i^R), while not penalizing the score too heavily in the presence of outliers.

2.4. Clustering and filtering using TCP-AIR energy

After generating this refined set of candidate RRTs, we perform a clustering step. These clusters are first filtered using the TCP-AIR energy, then, we re-rank them to obtain the final output set of RRTs.

Clustering

We cluster the set of candidates using the PROTAC stability score. Clusters are greedily created by selecting the RRT with the highest score, choosing that as the center of the new cluster, and assigning all sampled RRTs with an RMSD (vs. the center of the cluster) of below 5 Å to this cluster. This step was repeated (starting again with the RRT with the highest fitness which was not yet part of any cluster) until all points were assigned to some cluster. It is to be noted here that clusters with a single element are also considered.

Filtering using TCP-AIR energy

We use a filtering technique based on the ambiguous interaction restraints (AIRs) energy [60] that has been used extensively in Dominguez et al. [38] for energy calculation. The intuition behind using the TCP-AIR as a filtering step is that we restrict the distance between two proteins concerning each other's interaction residues (active and passive). It allows us to specify a distance measure between the two proteins that considers the distance between the surfaces of the proteins rather than just the distance between binding pockets. In order to calculate the TCP-AIR energy terms, active and passive residues need to be defined for both proteins. Active and passive residues are chosen based on two criteria: first, they have some atoms on the protein's surface, and second, they lie within some distance of the atoms of the fragments binding to this particular protein. We used distances of 5 Å and 7.5 Å for active and passive distances, respectively. For instance, the residues in the POI within 5 Å of the warhead were defined as active residues for POI.

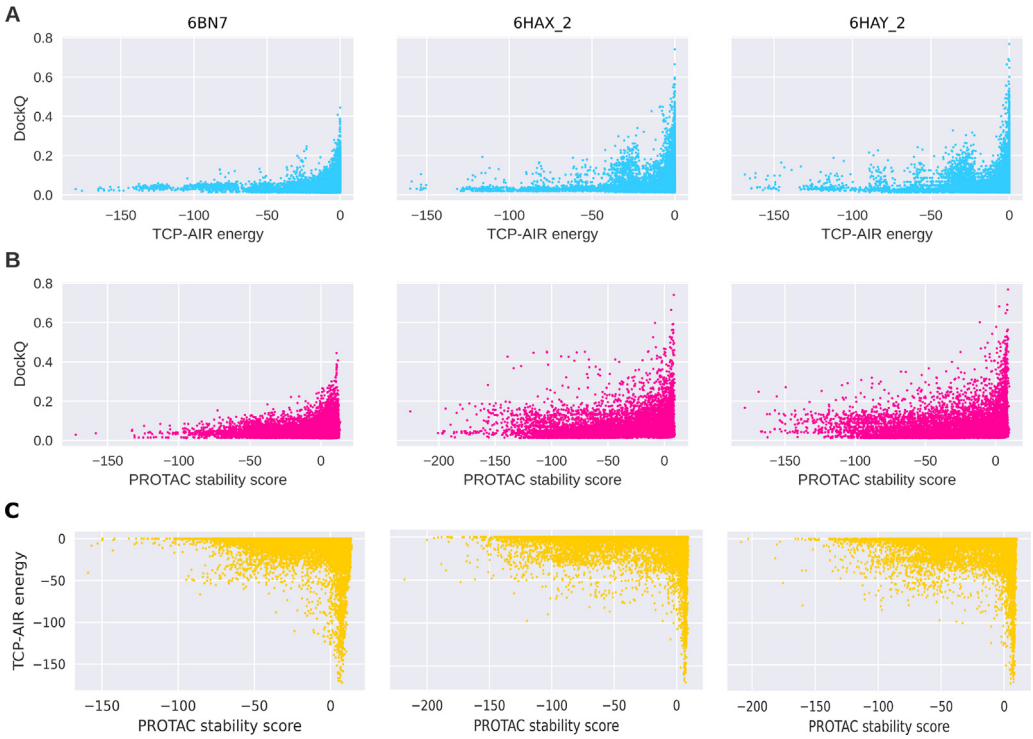

Fig. 5. Usefulness of TCP-AIR energy and the PROTAC stability score for filtering. A) Correlation between the TCP-AIR energy and the DockQ score for different example ternary complexes. As high DockQ scores are only achieved when the TCP-AIR energy is 0 or close to 0, the later can be used as a filtering criterion. B) Same as A, but for the PROTAC stability score. C) Correlation between the PROTAC stability score and the TCP-AIR energy for 6BN7 ($r = -0.01$, $\rho = 0.09$), 6HAX_2 ($r = -0.07$, $\rho = -0.08$) and 6HAY_2 ($r = -0.05$, $\rho = -0.07$), with r being the pearson correlation coefficient and ρ the spearman's rank correlation coefficient. The low correlation indicates that the two scores capture independent aspects of the near-native pose, thus justifying the use of both scores.

For these residues, we calculate the effective distances (as done in previous work [38]). These are used to calculate the TCP-AIR energy (using the NOE energy formula, [60]) using three threshold values L, U, and S. L and U are the lower and upper bounds on the effective distance, while S is the value from where the computation of the energy changes from quadratic to linear. These parameters need to be specified for each ternary complex separately since each structure has different distances between two active sites of the proteins (due to the space occupied by the PROTAC molecule). To set the values, we calculated the average distances of each active residue of one protein with the active and passive residues of the other protein, resulting in two distances for each candidate. We considered all of our samples with PROTAC stability score > 5. For each of them, we took an average of the two proteins' distances and defined L as the 25th quantile. We then defined $U = L + 3$ Å and $S = U + 2$ Å. When filtering, we preserved all clusters with at least one candidate with TCP-AIR energy larger than the 90th quantile of all AIR energies (see Fig. 5).

All of the parameters defined above were defined and optimized using four ternary structures 6HAX_1, 6HAX_2, 6HAY_1, and 6HAY_2. The active and passive distances were decided by inspecting the structures of these ternary complex, which showed it to be a reasonable assumption that the native complex will contain a sizeable fraction of interacting residues within 5 Angstroms from the binding site, and almost certainly 7.5 A from it. The parameters L, U, S, were manually tuned to maximize the number of near-native complexes filtered for these 4 cases.

Ranking of clusters

After filtering the clusters, we rank the remaining clusters using the PROTAC stability score (see above, Section 2.3). Clusters are ranked according to the largest PROTAC stability score of their members. We then calculate the rank of the near-native clusters (Table 2). For the structural refinement, we generate a diverse set of complexes by selecting the structure with the highest PROTAC stability score out of each of the 100 top-ranked clusters. For each of these, we use the 10 PROTAC conformers generated by AutoDock-Vina (see Section 2.3) as input to the structural refinement (i.e., the structural refinement is performed for 1000 structures).

2.5. Structural refinement

Starting with a given number of ternary complex structures from re-ranked and clustered TCP output, we performed further structural refinement using classical force field-based molecular simulation with an implicit solvent model. Force field parameters for each molecule were taken from the Amber ffSB14 force field [61] for proteins, and GAFF2 [62] for PROTAC molecules. Partial charges for the PROTAC molecules were determined as AM1BCC charges [63] using antechamber/Ambertools version 21 [64]. MD simulations were performed using OpenMM [65] with a Langevin type integrator and a time step of two femtoseconds. The implicit solvent model used was the second GBneck model (GBN2 [66]). Each structure was first optimized through energy minimization for 100 steps to remove close contacts, followed by a simulated annealing run, with the temperature varying from 300 to 0 K in a linear fashion, for a total simulation time of 200 ps. This procedure was repeated four times to approximate the average total free energy of solvation and standard deviations for each input structure. The resulting energies are essentially a simplified version of MM-GB/SA energies, where a calculation of entropic contributions is omitted. Including entropy terms would require much longer simulation times, and, given the approximate nature of current approaches for entropy estimation [67], we consider the benefit of including this contribution limited. The elevation of temperature above 300 K may lead to the partial unfolding of proteins, which can not be compensated for due to the limited duration of the MD simulation (200 ps). As a consequence, it is important to acknowledge that short MD runs serve primarily as a means for relaxing the protein structure, removing any clashes, and exploring the local

conformational space to identify a near-local minimum and not for exhaustive sampling. We assume that the previous steps, namely Bayesian optimization sampling and local optimization of the PROTAC stability score, enable us to attain larger-scale sampling.

2.6. Re-clustering and re-ranking

The refined structures (1000 of them) were clustered by the fraction of common contact (FCC) clustering method [68]. Afterwards, the clusters were ranked based on size (the cluster with the highest number of elements being the first-ranked). The clustering threshold was set to 0.5 and minimum number of elements in a cluster was set to 2. The rest of non-clustered structures were not taken into consideration.

2.7. Evaluation metric

For evaluating the quality of the candidates generated by the BO loop (Table 2), we used the DockQ score [69], which is based on the metrics underlying the well-known critical assessment of predicted interactions (CAPRI) score [70]: f_{nat}, I_{RMSD}, and L_{RMSD}. The DockQ score for a particular candidate provides a quantitative measure of how closely the interaction of proteins in this candidate structure resembles that of proteins in a reference native structure. It is a function of the following metrics

f_{nat} is the fraction of protein's native contacts that a candidate recovers. I_{RMSD} is calculated over all backbone atoms for those residues found in the interface of the reference structure (residues within a radius of 5 Å). L_{RMSD} is calculated based on the backbone atoms of the target protein. Both RMSD values are calculated between the candidate and native model (after aligning the receptor structures, i.e., the E3 ligase).

The DockQ score combines these measures into a single scalar value between 0 and 1 (which is the score of the reference structure), with higher values corresponding to ternary complexes which resemble the original structure more closely. The distribution of the DockQ scores has been defined as follows [69]:

- $0 <$ DockQ < 0.23 - Incorrect complex
- $0.23 \leq$ DockQ < 0.49 - Acceptable quality complex
- $0.49 \leq$ DockQ < 0.80 - Medium quality complex
- $0.80 \leq$ DockQ $<= 1.0$ - High quality complex

All candidate poses that achieve a DockQ > 0.23 compared to the native crystal structure are termed as near-native poses

2.8. Fitness landscape analysis

Using 5 exemplary ternary complexes from PDB (bound structures), we extracted the RRT corresponding to the native pose and modified them to investigate the resulting change of the fitness $y(\mathbf{x})$. For each ternary complex, we generated 4000 random near-native samples by first randomly shifting the POI before rotating them. The shift directions were drawn from a von Mises–Fisher distribution with the mean set to the offset of the POI (i.e., pointing away from the E3 ligase) and the directional parameter $\kappa = 1$ (i.e., shifts moderately concentrated in this direction). The shift distances were drawn from an exponential distribution with $\lambda = 1$ Å. The rotations were generated by drawing a random rotation direction from a uniform distribution (no directions favored) and a rotation angle drawn from a von Mises distribution with mean 0 and $\kappa = 10$ (i.e., small rotations favored).

3. Results

3.1. Neural-network based approximation of PROTAC score

The PROTAC score is the negative square root of the energy achieved after minimizing the UFF force-field. The value of this PROTAC score achieves for an un-strained conformation of the PROTAC ranges from -40 to -15 over the set of PROTACs in the database. The root mean square deviation for the same set of PROTACS ranges from 1.2 to 2.4, with a mean RMS deviation of 1.18. The small error of prediction compared to the magnitude of the score value suggests that the learned model is a sufficiently accurate approximation of the computationally expensive PROTAC score.

3.2. Evaluation of fitness functions

To check whether our proposed fitness functions are a meaningful objective for the BO loop, we analyzed the change of the fitness values around the experimentally determined ternary complexes. Using bound structures, we created random near-native RRTs by altering the position of the POI using small, random translations and rotations (see Section 2.8 for details). We then compared the values of the PPI and PROTAC fitnesses of these random RRTs to the fitness of the actual native complex (Fig. 6). The results show that these fitnesses are higher at the native RRT compared to other RRTs in its vicinity, thus indicating the their usefulness as objective for the BO loop.

3.3. Evaluation of the BO loop

We futhermore evaluated the BO method to test whether it performs adequate exploration and exploitation. Ideally, the distribution of candidates acquired by the BO loop is clustered in regions where the fitness $y(\mathbf{x})$ is high. We compared the samples generated by the BO loop to a set of randomly sampled candidates, and found that this is indeed the case (Fig. 7). We recap here that each orientation is sampled as a 7-D vector containing a 3-D translation (in Angstroms) and a 4-D rotation quaternion. Fig. 7A shows us that the samples drawn by the BO Loop are concentrated in a specific region where the PPI interactions and PROTAC score are high. Fig. 7C shows us that while random sampling occasionally finds RRTs with high fitness values, most samples are sub-optimal leading to a lower mean fitness of 2.8. In comparison, the concentrated sampling by the BO loop results in a more consistently high fitnesses with mean 17.3 and a higher maximum value. Additionally, the spread of points observed in both Fig. 7A and B show that we do not converge prematurely to the optimum but also sufficiently explore the space of possible RRTs.

3.4. Prediction of near-native poses

We run our pipeline on 14 ternary complexes, some of which have more than one TC structure in the asymmetric unit of the experimental crystal structure leading to a total of 22 cases. By running our pipeline and comparing our results to these 22 native ternary complex structures, we have found that the BO loop can effectively sample the near-native poses. In addition, the filtering and ranking procedures can assign them a relevant rank. For 21 out of 22 cases, we could sample at least acceptable quality complexes. For 15 out of 22 (Table 2), we were able to rank them within the top 20 clusters, and for 12 out of 22 complexes, we were able to recover a medium quality ternary structure. Additionally, 15 out of 22 clusters have at least 50% of the near-native poses in the reported cluster (see Table 2).

This method proves that even with using a relatively simple PPI interaction and unbound structures, taking into account the constraints imposed by the linker, with the incorporation of the PROTAC score in the BO loop, as well as the TCP-AIR energy for filtering and PROTAC stability score for ranking, we can recover near-native predictions of the ternary complex structures. Some of the apparent issues with certain cases are discussed below.

In the 7KHH structure, the binding mode of the binder to the E3 ligase differs from the binding mode found in the input structure of the receptor, derived by docking from 4W9H (See Supplement Fig. S1). The PROTAC score, which is calculated by constraining the PROTAC fragments to the positions in the input, thus returns very poor scores for the

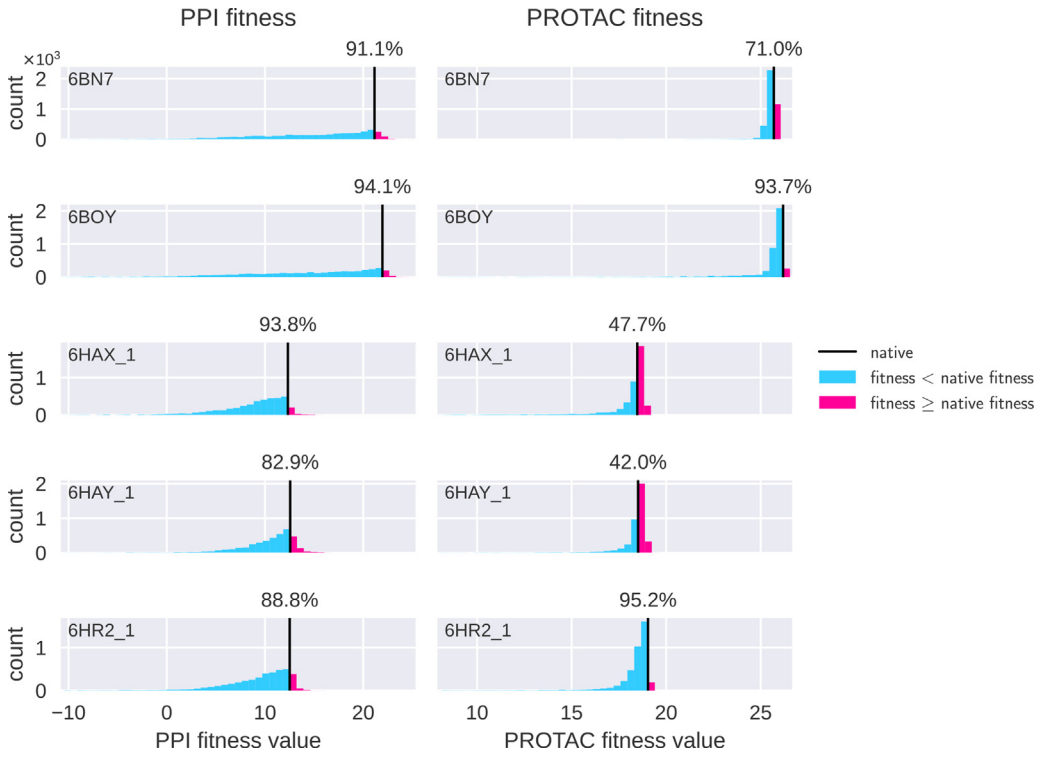

Fig. 6. Native complexes have high fitness values. The columns show (for five sample ternary complexes) histograms of PPI fitness (left) and PROTAC fitness (right) for random near-native RRTs, created by randomly shifting and rotating the POI (see text for details). Each plot contains the histogram of 4096 such RRTs, with the percentile of the fitness of the native RRT displayed above the black line. The actual native complex (black line) achieves a comparitively high fitness in most cases. In a few cases where the constraint-fitness has a lower percentile, we find that this is compensated for by the higher percentile of the PPI score.

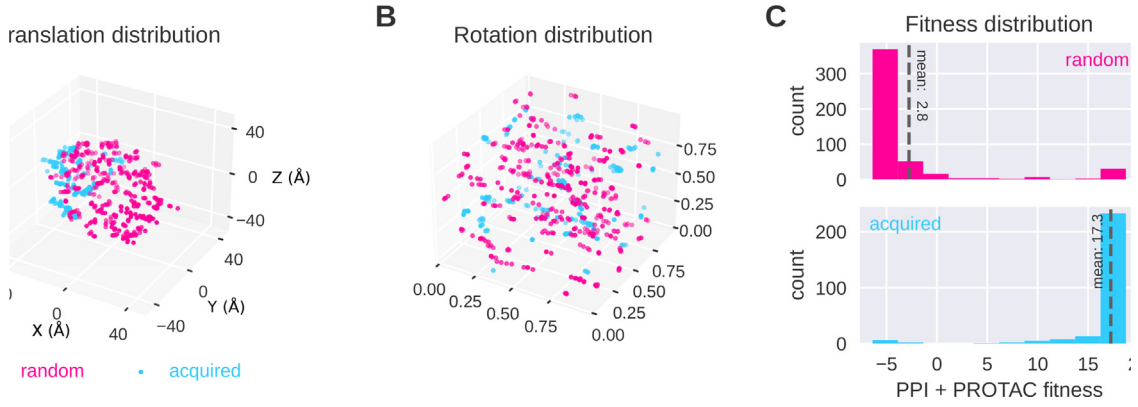

Fig. 7. Bayesian optimization results in higher fitness values than random search. Shown here are samples from the acquisition function at iteration 15 of the BO loop for the 6HAX complex, compared to random samples picked with proteins close to each other. (A) The Translation coordinates of the RRTs (in Å) from random sampling (pink) and proposed by Bayesian optimization (blue). The latter leads to a candidate concentration in a promising region. The spread of points also suggests that the BO loop sufficiently explores the space without converging prematurely to any specific minima. Points sampled via both random sampling and BO are subject to a local optimization (B) The Rotation quaternions from the RRTs, mapped uniformly onto a 3-D unit cube from random sampling and Bayesian optimization. This mapping is described in the Supplement Section S1. The diversity of rotation quaternions shows the extent of exploration. (C) Resulting fitness histograms. While random samples find some conformations with high fitness, the best fitness acquired by the Bayesian optimization is higher, and the mean fitness is substantially better, confirming that our samples efficiently search through the space of higher fitness values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

near-native RRTs. Thus in Table 2, we see no near-native poses were sampled for 7KHH. This issue however is partially addressed by structural refinement which is free to refine the PROTAC position (Table 4).

For the case of 6W8I, the variance in performance can be attributed to the fact that it consists of two sets of very different poses. 6W8I_1 and 6W8I_2 are qualitatively similar (RMSD of 7 Å) compared to 6W8I_3 (RMSD of 33 Å w.r.t 6W8I_1). The reason for the lower performance in 6W8I_2 is due to the fact that the orientation of the LYS residue in the input warhead gets in the way of the PROTAC for all RRTs that are near-native (See Figure S2 in Supplement).

After the use of structural refinement, we can see clear improvements. In 6 cases, we were able to rank near-native cluster as first. For 21/22 cases, we obtained complexes with the DockQ≥ 0.23. For 16 out of 22 (Table 4), we were able to rank them within the top 15 clusters. For 10 out of 22 cases, we could recover at least a medium-quality ternary structure. Additionally, in 17/22 cases, in the reported highly-ranked

Table 3
Results and the comparison between BOTCP and Weng et al. (FRODOCK-Based Protocol) methods for ternary complex prediction on unbound structures before the structural refinement. The models quality is marked with * (**: high, *: medium and no *: acceptable). [1]The best rank containing at least one model with DockQ≥0.23. [2]The total number of clusters. [3]The percentage of the near-native poses in that specific ranked cluster (available only for BOTCP). [4]The best possible dockq score with rigid docking (we align an unbound target to a native complex and calculate the DockQ score). For each ternary complex, we selected a cluster which includes an RRT with a high DockQ score. None indicates there were no near-native complexes found and an empty field that a method was not run for that complex. The columns also show the PDB ID and characterization of the best RRT in terms of DockQ score (see text).

	Weng et al.			BOTCP				
PDB ID	Cluster rank[1]	#Clusters[2]	DockQ	Cluster rank[1]	#Clusters[2]	DockQ	%Near-native[3]	Best DockQ[4]
6HAY_1	15	24	0.46	**3***	374	**0.72**	98.94	0.89
6HAY_2				**1***	377	**0.55**	30.47	0.87
6HAX_1	8	16	0.44	**7***	300	**0.63**	1.04	0.85
6HAX_2				**6**/21*	283	**0.39**	0.14	0.85
6W7O_1	2*	15	**0.68**	**1**/9*	102	0.48	35.79	0.84
6W7O_2				**1**/4*	106	**0.32**	17.24	0.83
5T35_1	**4****	35	**0.83**	5/15*	370	0.39	0.84	0.85
5T35_2				**2**/27*	315	**0.39**	31.45	0.88
6BN7	**3**	25	0.26	30	606	0.31	3.33	0.8
6BN8	2	31	**0.26**					
6BN9	3	17	**0.27**					
6BNB	1	1	**0.47**					
6BOY	**3**	30	0.39	12*	423	**0.69**	77.78	0.81
6SIS_1	2*	14	**0.75**	**1***	161	0.59	17.76	0.86
6SIS_2				**1**	189	**0.36**	2.35	0.83
7KHH	5	29	**0.43**	**3**	162	0.23	0.44	0.92
7KHH_2	3*	15	**0.65**					
6HR2_1	21	26	**0.48**	**5**/18*	302	0.28	1.11	0.81
6HR2_2				**4**/21*	327	**0.27**	0.60	0.84
6ZHC	**2***	22	**0.58**	8	194	0.39	5.9	0.91
6W8I_1	None	None	None	**5***	191	**0.57**	16.67	0.84
6W8I_2	None	None	None	**93**	196	**0.38**	4	0.79
6W8I_3	None	None	None	**1**/16*	188	**0.27**	8.7	0.86
7JTO				**4**	199	**0.31**	4.49	0.76
7Q2J				**8****	69	**0.8**	1.1	0.88
7JTP				**2**/5*	49	**0.39**	0.43	0.84

near-native cluster, at least 50% of the poses in it are near-native poses (see Table 4).

3.5. Computational performance

Our results do not require the use of GPUs to compute the scores and, using 128 standard x64 CPU cores (16 AWS EC2 c5.4xlarge instances), for each complex, our results including filtering and re-ranking excluding structural refinement, typically take less than 2 h. This represents a significant improvement over conventional techniques, that attempt to handle many interactions in a ternary complex via GPU intensive molecular dynamics simulations which typically take days or weeks to report a near-native pose. Note that this code is a prototype with significant room to optimize the various components further still.

3.6. Comparison to previous work

Ternary complex prediction models have frequently been evaluated on bound structures (e.g., PRosettaC [32]). We have demonstrated in this work that effective ranking of predicted structures in the more challenging scenario of unbound structures is possible with our method. We are aware of one other approach based on unbound structures has been presented by Weng et al. [33], who have shown successful ranking of the near-native clusters within the top-15 clusters (Tables 3 and 4). However, Weng et al. used clustering with FCC using a threshold of 0.5, which is relatively low, leading to a large spread of poses within the resulting clusters. The same threshold was used in our case, but only after refinement, resulting in much bigger near-native percentage of the ranked clusters. When we used fcc with threshold 0.5 before refinement, we got much worse results with some near-native percentages being lower than 1% (Table 3). Even though there may be a near-native pose in such a cluster, a threshold of 0.5 implies that the fraction of near-native poses in the cluster would be low due to the diversity of poses within the cluster. Moreover, there is a higher likelihood that the best-scored pose in this cluster is, in fact, not a near-native pose.

Accordingly, a cluster ranked high in our case has lower entropy, i.e., the yield of near-native structures in it is much higher. With this form of constrained clustering, we have near-native clusters ranked within the top-15 clusters in 15 out of 22 cases before structural refinement and 16 out of 22 cases after structural refinement. These results demonstrate the efficacy of the Vina and TCP-AIR energies. It should be noted that our work contains evaluation on the most of the known ternary complexes when compared with other works. Adjusting the parameters for our method and ensuring it works across all of the 22 and not 8 or 10 ternary structures, underscores the generalization capabilities of our method.

4. Conclusion and outlook

We have demonstrated our BOTCP module for predicting unbound rigid ternary complex structures (Table 2). Our method can sample and appropriately rank the near-native poses for most of them.

At the heart of this, is the fact that Bayesian optimization learns to sample regions of high scores in a few iterations. In all of our simulations, we sample only around 2500 points, where the BO loop converges rougly after sampling around 700 points. This sample efficiency suggests modifying the fitness function to use more computationally expensive but informative data (e.g., the PROTAC stability score) for the BO loop. Such a fitness could take into account the interactions of the linker with the proteins, which has the potential to significantly improve the yield of near-native poses.

Currently, there are several avenues of improvement for both the quality of results and computational efficiency. The subsequent struc-

Table 4
Results and the comparison between BOTCP and Weng et al. (RosettaDock-Based Refinement) methods for ternary complex prediction on unbound structures after the structural refinement. The models quality is marked with * (**: high, *: medium and no *: acceptable). [1]The best rank containing at least one model with DockQ≥0.23. [2]The total number of clusters. [3]The percentage of the near-native poses in that specific ranked cluster (BOTCP). For each ternary complex, we selected a cluster which includes an RRT with a high DockQ score. None indicates there were no near-native complexes found and an empty field that a method was not run for that complex. The columns also show the PDB ID and characterization of the best RRT in terms of DockQ score (see text).

	Weng et al.			BOTCP			
PDB ID	Cluster rank[1]	#Clusters[2]	DockQ	Cluster rank[1]	#Clusters[2]	DockQ	%Near-native[3]
6HAY_1	7	88	0.48	**1***	117	**0.6**	100
6HAY_2				**1***	117	**0.59**	100
6HAX_1	2	51	**0.45**	**1**/10*	129	0.35	100
6HAX_2				**1**/10*	129	**0.36**	100
6W7O_1	**9***	37	**0.56**	42	99	0.44	40
6W7O_2				**42**	99	**0.45**	40
5T35_1	3*	102	**0.72**	**1***	89	0.54	100
5T35_2				**1***	89	**0.51**	100
6BN7	**1***	69	**0.66**	8	130	0.27	30
6BN8	**1***	80	**0.7**				
6BN9	**1***	48	**0.67**				
6BNB	**1***	1	**0.5**				
6BOY	**4***	78	**0.68**	11	133	0.40	83
6SIS_1	**1***	15	**0.72**	2*	70	0.56	94
6SIS_2				**2***	70	**0.56**	96.5
7KHH	**24***	99	**0.5**	28	121	0.34	66.7
7KHH_2	**8***	52	**0.65**				
6HR2_1	**6**	75	0.43	8*	92	**0.56**	100
6HR2_2				**8**	92	**0.48**	100
6ZHC	**15***	75	**0.58**	None	None	None	None
6W8I_1	13	35	0.31	**10***	113	**0.51**	100
6W8I_2	None	None	None	**86**	128	**0.35**	100
6W8I_3	**13**	35	0.28	21*	113	**0.51**	100
7JTO				**13**	98	**0.48**	82
7Q2J				**6***/34**	82	**0.51**	100
7JTP				**7**/21*	46	**0.40**	15.8

tural refinement of the structures provided from the BOTCP module currently takes about 25 h but could be further optimized. For instance by running the structural refinement for less than 1000 poses or by using a smaller number of PROTAC conformations per structure. While the structural refinement is still computationally intensive, we see that the results without structural refinement suggest that the current scores and sampling technique can be beneficial in assessing the interactions for a given ternary complex.

As mentioned in the introduction, the current work predicts the ternary complex structure most effectively under a set of specific assumptions, i.e. that the proteins do not undergo significant backbone conformational changes, and that we have information regarding the E3-ligase binder, the warhead and the respective binding pockets. Currently, the PROTAC stability score and the PPI score are affected because the side-chains are misaligned with the side chains as in the original crystal structure. These limitations affect us for instance in the case of 7KHH and 6W8I_2, where misaligned binders and side-chains in the input impact performance.

Future work can incorporate more informative PPI and constraint scores in the BO loop which will improve the filteration of non-native poses. Our simulations show that the native structure consistently has a PROTAC stability fitness that is among the highest observed for that complex, which promises significant improvements if this score includes side-chain flexibility (e.g., using Autodock-Vina).

The current approach to validate ternary complex prediction is hindered by the severe lack of availability of structures of ternary complexes. A promising direction of research is thus to develop a score that is derived from the poses sampled by BOTCP, which can be validated against more abundantly available data on the dissociation constant of a ternary complex.

To summarize, we demonstrate in this work the successful application of Bayesian optimization as well as the design of two specialized scores that can effectively sample and rank highly clusters of near-native poses for Ternary Complexes.

Data availability

All the processed input structures as well as the poses with the best DockQ score from each of the 100 best ranked clusters are available at https://storage.googleapis.com/public_botcp_data/Publication_Data_Final.zip

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Arjun Rao, Tin Tunjic, Michael Brunsteiner, Chiara Gasbarri, and Noah Weber are currently employed by Celeris Therapeutics GmbH, Graz Austria. The methodologies discussed in this paper are currently actively being used by Celeris Therapeutics GmbH in order to design Proximity Inducing Compounds for the purpose of Targeted Protein Degradation. Michael Müller and Hosein Fooladi were employed by Celeris Therapeutics during their contribution to this work

CRediT authorship contribution statement

Arjun Rao: Conceptualization, Investigation, Methodology, Software, Validation, Writing – review & editing. **Tin M. Tunjic:** Conceptualization, Investigation, Methodology, Software, Writing – review & editing, Validation. **Michael Brunsteiner:** Data curation, Methodology, Investigation, Software, Writing – review & editing. **Michael Müller:** Software, Validation, Visualization. **Hosein Fooladi:** Software. **Chiara Gasbarri:** Data curation, Validation. **Noah Weber:** Conceptualization, Supervision, Writing – review & editing.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.ailsci.2023.100072.

References

[1] Schapira M, Calabrese MF, Bullock AN, Crews CM. Targeted protein degradation: expanding the toolbox. Nat Rev Drug Discov 2019;18(12):949–63.
[2] Bai N, Riching KM, Makaju A, Wu H, Acker TM, Ou S-C, et al. Modeling the CRL4A ligase complex to predict target protein ubiquitination induced by cereblon-recruiting PROTACs. J Biol Chem 2022;298:101653.
[3] Schreiber SL. The rise of molecular glues. Cell 2021;184(1):3–9.
[4] Dong G, Ding Y, He S, Sheng C. Molecular glues for targeted protein degradation: from serendipity to rational discovery. J Med Chem 2021;64(15):10606–20.
[5] Neklesa TK, Winkler JD, Crews CM. Targeted protein degradation by PROTACs. Pharmacol Ther 2017;174:138–44.
[6] Pettersson M, Crews CM. Proteolysis targeting chimeras (PROTACs)—Past, present and future. Drug Discov Today 2019;31:15–27.
[7] Ciulli A, Trainor N. A Beginner's guide to PROTACs and targeted protein degradation. Biochemist 2021;43(5):74–9.
[8] Fischer ES, Böhm K, Lydeard JR, Yang H, Stadler MB, Cavadini S, Nagel J, Serluca F, Acker V, Lingaraju GM, et al. Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide. Nature 2014;512(7512):49–53.
[9] Fink EC, Ebert BL. The novel mechanism of lenalidomide activity. Blood, J Am Soc Hematol 2015;126(21):2366–9.
[10] Gao S, Wang S, Song Y. Novel immunomodulatory drugs and neo-substrates. Biomark Res 2020;8(1):1–8.
[11] Buckley DL, Van Molle I, Gareiss PC, Tae HS, Michel J, Noblin DJ, Jorgensen WL, Ciulli A, Crews CM. Targeting the von Hippel–Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1α interaction. J Am Chem Soc 2012;134(10):4465–8.
[12] Galdeano C, Gadd MS, Soares P, Scaffidi S, Van Molle I, Birced I, et al. Structure-guided design and optimization of small molecules targeting the protein–protein interaction between the von Hippel–Lindau (VHL) E3 ubiquitin ligase and the hypoxia inducible factor (HIF) alpha subunit with in vitro nanomolar affinities. J Med Chem 2014;57(20):8657–63.
[13] Schiemer J, Horst R, Meng Y, Montgomery JI, Xu Y, Feng X, et al. Snapshots and ensembles of BTK and cIAP1 protein degrader ternary complexes. Nat Chem Biol 2021;17(2):152–60.
[14] Hines J, Lartigue S, Dong H, Qian Y, Crews CM. MDM2-recruiting PROTAC offers superior, synergistic antiproliferative activity via simultaneous degradation of BRD4 and stabilization of P53. Cancer Res 2019;79(1):251–62.
[15] Wei J, Meng F, Park K-S, Yim H, Velez J, Kumar P, et al. Harnessing the E3 ligase KEAP1 for targeted protein degradation. J Am Chem Soc 2021;143(37):15073–83.
[16] Petrylak D.P., Gao X., Vogelzang N.J., Garfield M.H., Taylor I., Dougan Moore M., Peck R.A., Burris H.A. III. First-in-human phase I study of ARV-110, an androgen receptor (AR) PROTAC degrader in patients (PTS) with metastatic castrate-resistant prostate cancer (MCRPC) following enzalutamide (ENZ) and/or abiraterone (ABI)2020.
[17] Neklesa T, Snyder LB, Willard RR, Vitale N, Pizzano J, Gordon DA, et al. ARV-110: an oral androgen receptor PROTAC degrader for prostate cancer. J Clin Oncol 2019;37(7):14–16.
[18] Flanagan J., Qian Y., Gough S., Andreoli M., Bookbinder M., Cadelina G., Bradley J., Rousseau E., Willard R., Pizzano J., et al. Abstract P5-04-18: ARV-471, an oral estrogen receptor PROTAC degrader for breast cancer. 2019.
[19] He Y, Koch R, Budamagunta V, Zhang P, Zhang X, Khan S, Thummuri D, Ortiz YT, Zhang X, Lv D, et al. DT2216–a Bcl-xL-specific degrader is highly active against Bcl-xL-dependent T cell lymphomas. J Hematol Oncol 2020;13(1):1–13.
[20] Hansen JD, Correa M, Nagy MA, Alexander M, Plantevin V, Grant V, et al. Discovery of CRBN E3 ligase modulator CC-92480 for the treatment of relapsed and refractory multiple myeloma. J Med Chem 2020;63(13):6648–76.
[21] Zhou B, Hu J, Xu F, Chen Z, Bai L, Fernandez-Salas E, et al. Discovery of a small-molecule degrader of bromodomain and extra-terminal (BET) proteins with picomolar cellular potencies and capable of achieving tumor regression. J Med Chem 2018;61(2):462–81.
[22] Zorba A, Nguyen C, Xu Y, Starr J, Borzilleri K, Smith J, Zhu H, Farley KA, Ding W, Schiemer J, et al. Delineating the role of cooperativity in the design of potent PROTACs for BTK. Proc Natl Acad Sci 2018;115(31):E7285–92.
[23] Dobrovolsky D, Wang ES, Morrow S, Leahy C, Faust T, Nowak RP, Donovan KA, Yang G, Li Z, Fischer ES, et al. Bruton tyrosine kinase degradation as a therapeutic strategy for cancer. Blood, J Am Soc Hematol 2019;133(9):952–61.
[24] Mares A, Miah AH, Smith IE, Rackham M, Thawani AR, Cryan J, Haile PA, Votta BJ, Beal AM, Capriotti C, et al. Extended pharmacodynamic responses observed upon PROTAC-mediated degradation of RIPK2. Commun Biol 2020;3(1):1–13.
[25] Kofink C, Trainor N, Mair B, Wöhrle S, Wurm M, Mischerikow N, Bader G, Rumpel K, Gerstberger T, Cui Y, et al. A selective and orally bioavailable VHL-recruiting PROTAC achieves SMARCA2 degradation in vivo. ChemRxiv 2022. doi:10.26434/chemrxiv-2022-q63s3.
[26] Hughes SJ, Ciulli A. Molecular recognition of ternary complexes: a new dimension in the structure-guided design of chemical degraders. Essays Biochem 2017;61(5):505–16.
[27] Bemis TA, La Clair JJ, Burkart MD. Unraveling the role of linker design in proteolysis targeting chimeras: miniperspective. J Med Chem 2021;64(12):8042–52.
[28] Gadd MS, Testa A, Lucas X, Chan K-H, Chen W, Lamont DJ, et al. Structural basis of PROTAC cooperative recognition for selective protein degradation. Nat Chem Biol 2017;13(5):514–21.
[29] Troup RI, Fallan C, Baud MG. Current strategies for the design of PROTAC linkers: a critical review. Explor Target Anti-Tumor Ther 2020;1:273–312.
[30] Drummond ML, Williams CI. In silico modeling of PROTAC-mediated ternary complexes: validation and application. J Chem Inf Model 2019;59(4):1634–44.
[31] Drummond ML, Henry A, Li H, Williams CI. Improved accuracy for modeling PROTAC-mediated ternary complex formation and targeted protein degradation via new in silico methodologies. J Chem Inf Model 2020;60(10):5234–54.
[32] Zaidman D, Prilusky J, London N. Prosettac: rosetta based modeling of PROTAC mediated ternary complexes. J Chem Inf Model 2020;60(10):4894–903.
[33] Weng G, Li D, Kang Y, Hou T. Integrative modeling of PROTAC-mediated ternary complexes. J Med Chem 2021;64(21):16271–81.
[34] Marze NA, Roy Burman SS, Sheffler W, Gray JJ. Efficient flexible backbone protein–protein docking for challenging targets. Bioinformatics 2018;34(20):3461–9.
[35] Li W, Zhang J, Guo L, Wang Q. Importance of three–body problems and protein–protein interactions in proteolysis-targeting chimera modeling: insights from molecular dynamics simulations. J Chem Inf Model 2022;25(3):523–32.
[36] Liao J, Nie X, Unarta IC, Ericksen SS, Tang W. In silico modeling and scoring of PROTAC-mediated ternary complex poses. J Med Chem 2022;65(8):6116–32.
[37] Brochu E., Cora V.M., De Freitas N.. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:101225992010.
[38] Dominguez C, Boelens R, Bonvin AM. Haddock: a protein- protein docking approach based on biochemical or biophysical information. J Am Chem Soc 2003;125(7):1731–7.
[39] Berman HM. The protein data bank. Nucleic Acids Res 2000;28(1):235–42. doi:10.1093/nar/28.1.235.
[40] McNutt AT, Francoeur P, Aggarwal R, Masuda T, Meli R, Ragoza M, et al. GNINA 1.0: molecular docking with deep learning. J Cheminform 2021;13(1):1–20. doi:10.1186/s13321-021-00522-2.
[41] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9. doi:10.1038/s41586-021-03819-2.
[42] Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 2022;50(D1):D439–44. doi:10.1093/nar/gkab1061.
[43] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem 2004;25(13):1605–12. doi:10.1002/jcc.20084.
[44] Eastman P.. PDBFixer. 2019. https://github.com/pandegroup/pdbfixer.
[45] Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci 2018;27(1):293–315. doi:10.1002/pro.3330.
[46] O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open babel: an open chemical toolbox. J Cheminform 2011;3(1):33. doi:10.1186/1758-2946-3-33.
[47] Zhang C, Liu S, Zhu Q, Zhou Y. A knowledge-based energy function for protein- ligand, protein- protein, and protein- dna complexes. J Med Chem 2005;48(7):2325–35.
[48] Rappé AK, Casewit CJ, Colwell K, Goddard WA III, Skiff WM. Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc 1992;114(25):10024–35.
[49] Landrum G.. Rdkit: open-source cheminformatics software2016;https://github.com/rdkit/rdkit/releases/tag/Release_2016_09_4.
[50] Nelder JA, Mead R. A simplex method for function minimization. Comput J 1965;7(4):308–13. doi:10.1093/comjnl/7.4.308.
[51] Mockus J, Tiesis V, Zilinskas A. The application of Bayesian methods for seeking the extremum. Towards Glob Optim 1978;2:117–29.
[52] Agnihotri A, Batra N. Exploring Bayesian optimization. Distill 2020;5(5):e26.
[53] Rasmussen CE. Gaussian processes in machine learning. In: Summer school on machine learning. Springer; 2003. p. 63–71.
[54] Kanagawa M., Hennig P., Sejdinovic D., Sriperumbudur B.K.. Gaussian processes and kernel methods: a review on connections and equivalences. arXiv preprint arXiv:1807025822018.
[55] Auer P. Using confidence bounds for exploitation-exploration trade-offs. J Mach Learn Res 2002;3(Nov):397–422.
[56] Kingma D.P., Ba J.. Adam: a method for stochastic optimization. arXiv preprint arXiv:141269802014.
[57] Balandat M, Karrer B, Jiang DR, Daulton S, Letham B, Wilson AG, Bakshy E. BoTorch: a framework for efficient Monte–Carlo Bayesian optimization. Advances in neural information processing systems, vol 33; 2020. http://arxiv.org/abs/1910.06403
[58] Kirkpatrick S, Gelatt Jr CD, Vecchi MP. Optimization by simulated annealing. Science 1983;220(4598):671–80.
[59] Eberhardt J, Santos-Martins D, Tillack AF, Forli S. Autodock vina 1.2. 0: new docking methods, expanded force field, and python bindings. J Chem Inf Model 2021;61(8):3891–8.
[60] Nilges M. A calculation strategy for the structure determination of symmetric demers by 1H NMR. Proteins 1993;17(3):297–309.
[61] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11(8):3696–713. doi:10.1021/acs.jctc.5b00255.
[62] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25(9):1157–74. doi:10.1002/jcc.20035.
[63] Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. J Comput Chem 2002;23(16):1623–41. doi:10.1002/jcc.10128.

[64] Case D., Aktulga H., Belfon K., Ben-Shalom I., Brozell S., Cerutti D.S., Cheatham T.E. III, Cisneros G., Cruzeiro V., Darden T., Duke R., Giambasu G., Gilson M., Gohlke H., Goetz A., Harris R., Izadi S., Izmailov S., Jin C., Ka K.. Amber. 2021.
[65] Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. PLoS Comput Biol 2017;13(7):e1005659. doi:10.1371/journal.pcbi.1005659.
[66] Nguyen H, Roe DR, Simmerling C. Improved generalized born solvent model parameters for protein simulations. J Chem Theory Comput 2013;9(4):2020–34. doi:10.1021/ct3010485.
[67] Ekberg V, Ryde U. On the use of interaction entropy and related methods to estimate binding entropies. J Chem Theory Comput 2021;17(8):5379–91. doi:10.1021/acs.jctc.1c00374.
[68] Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis P, Karaca E, Melquiond ASJ, et al. Clustering biomolecular complexes by residue contacts similarity. Proteins 2012;80(7):1810–17.
[69] Basu S, Wallner B. Dockq: a quality measure for protein-protein docking models. PLoS One 2016;11(8):e0161879.
[70] Lensink MF, Nadzirin N, Velankar S, Wodak SJ. Modeling protein-protein, protein-peptide, and protein-oligosaccharide complexes: CAPRI 7th edition. Proteins 2020;88(8):916–38. doi:10.1002/prot.25870.