Authors: Vilar, Santiago; Byler, Kendall; Ingman, Victoria; Benson, Maurice; Petroff, Anna B.; Hendrix, Paul; Goldhagen, Guy; Shipman, William J.; Keinan, Shahar.

NNMT drug discovery using Quantum-Aided Drug Design (QuADD)

Introduction

At Polaris Quantum Biotech, we're leveraging quantum computing to accelerate drug discovery [1]. Once a druggable protein target is identified, it becomes a race to discover a preclinical drug candidate for the project. One of the slowest steps in drug discovery is generating a diverse library of starting candidates. In the past, these libraries have been limited to generic options that are not specific to the protein pocket in question, searches of published literature, and unpatentable small molecules. To accommodate for these limitations, companies resort to expensive, iterative wet lab tests that take years to complete. These options generally result in a starting library with a small chemical space, low diversity, and high bias.

To solve this problem, Polaris Quantum Biotech treats drug design as a multi-objective optimization problem that a quantum annealing computer can solve. Because quantum annealing technology can search unprecedented chemical spaces at high speed, it is possible to design a combinatorial solution space of 10^{30} and then search it for the combinations of fragments that satisfy a given set of criteria. For example, the output libraries of selected molecules will have structures that are most complementary to the specific binding pocket in question, maximize binding, metabolic stability, minimize toxicity, and ensure that all are readily synthesizable.

We developed Quantum-Aided Drug Design, QuADD, which is a Software as a Service (SaaS) platform for library generation. With the input of a 3D structure of the protein binding pocket and bound ligand, QuADD generates a multi-billion search space to find lead-like hits that are novel, bioavailable, and synthesizable. This output, which includes candidates for best-in-class, first-in-class, and scaffold-jumping structures, provides an enriched starting library of 1K-10K molecules (depending on binding-pocket size) for your drug discovery project. This number of molecules is great enough in size to include a range of options, but small enough that it can be handled with traditional computer-aided drug design programs or high-throughput screening. QuADD includes key parts of our established pipeline that have been discussed in Nature Biopharma Dealmakers [1], Ars Technica [2], and Quantum Insider [3].

This paper details a case study for the target Nicotinamide N-Methyltransferase target (NNMT), which is an important and emergent target for obesity, insulin-related diseases, and different types of cancer [4]. As the starting point for our QuADD calculations, we selected the human crystallized NNMT protein structure in complex with the bisubstrate inhibitor MS2756 from the Protein Data Bank (PDB code: 6CHH) [5]. Results demonstrate that QuADD recovers clinically-proven fragments, discriminates between control and decoy molecules, identifies molecules that are energetically favorable in the pocket, have a rational binding profile, and present drug-like properties. Taken together, QuADD can facilitate drug discovery by providing enriched small molecule libraries for specific protein pockets, an advantage that lowers the risk and

shortens the time to clinical testing.

Retrieval of molecular fragments in QuADD

First, we checked QuADD's "backward compatibility", i.e. for QuaDD to identify the template molecule's fragments, as well as other fragments that have been identified in the literature as relevant to NNMT. Our results clearly show that QuADD identifies both the template fragments and fragments of other known binders, including those of medicines that are currently on the market.

In this test, QuADD fragmented the template ligand into five subregions and searched for molecular fragments with optimal properties in the BRICS library [6]. Results showed that QuADD recovered the fragments from the template in excellent ranking positions (see Figure 1). QuADD also retrieved fragments of other known NNMT binders. This control set included 17 binders with scaffolds similar to the template ligand [7] and two bisubstrate inhibitors with different scaffolds from the Protein Data Bank (PDB codes: 7NBM, 7NBJ) [5]. Control ligands were divided into 24 fragments according to the predefined subregions. QuADD then searched a large chemical space of thousands of fragments and retrieved 15 test fragments: 100% of fragments in R1 (adenine and quinoline), 100% in R2 (sugar/ribose and piperazine), 50% in R3 (Nitrogen atom), 60% in R4 (3 out of 5 aliphatic polar analogs), and 54% of fragments in R5 (7 out of 13 benzamide analogs). QuADD showed excellent recognition of the template and control fragments.

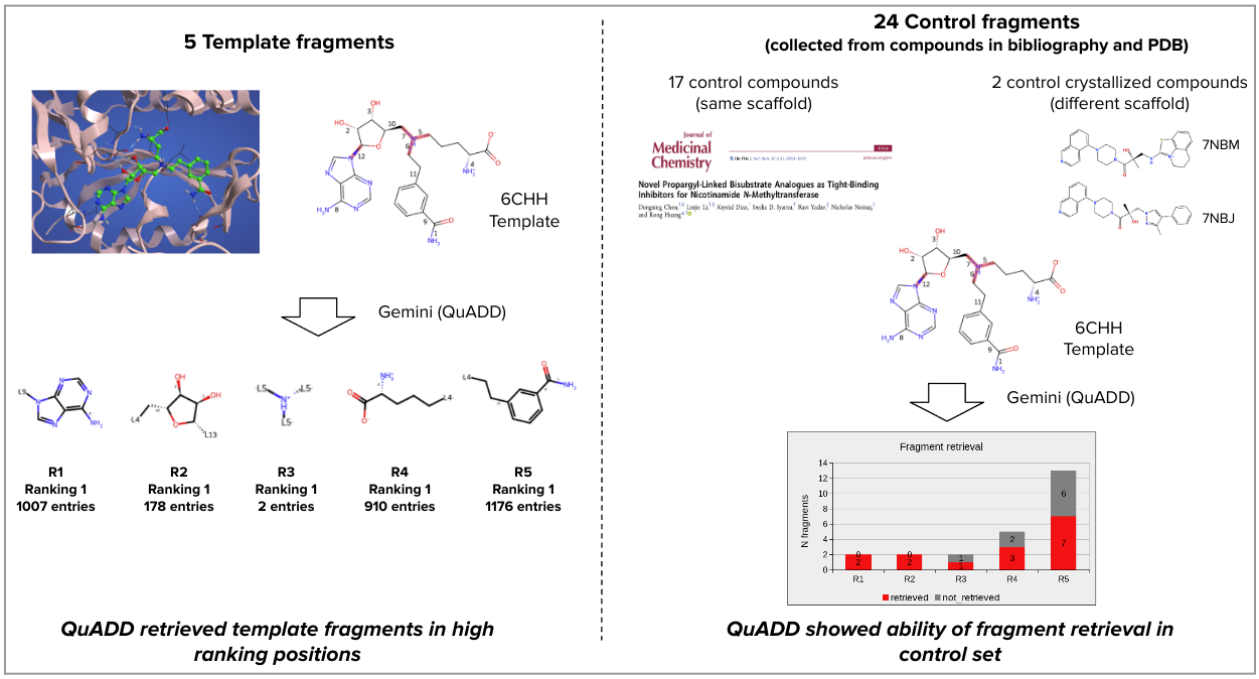

Figure 1. Fragment retrieval by QuADD: template and control sets.

QuADD library has better NNMT binders

Next, we investigated the quality of NNMT binders in the molecules generated by

QuADD for the NNMT binding pocket using docking. In order to test if QuADD produced molecules with favorable binding, we made a database of molecules from three sources: 1) the top 2,000 candidates from the QuADD library, 2) 2,000 decoys with similar drug-like descriptors (including logP, molecular weight, and topological polar surface area, among others), and 3) 17 control molecules with known binding in the NNMT pocket. The database was docked to the NNMT protein and docking scores were compared for the three sets. Results (Figure 2, left side) showed a clear prioritization of the control over decoys. However, no prioritization of control over QuADD was observed, indicating that QuADD returned molecules with equivalent docking scores as known binders. QuADD molecules yielded excellent docking rankings, reflecting that they had favorable binding in the pocket. Docking ranking also yielded a clear prioritization of QuADD over decoys. Figure 2 shows the docking results sorted by ranking.

Without QuADD, it is common practice for a computational or medicinal chemist to spend months designing candidate molecules for a protein pocket. How do the QuADD results, which are automatic and are returned in 1-2 days, compare to those hand-curated over the course of two months by an expert? To answer this question, we tested the QuADD results against a Chemist-Designed Library. As shown in Figure 2, QuADD-designed molecules had a more favorable docking rank than Chemist-designed molecules and outperformed the chemist timeframe. QuADD saves time and effort in drug discovery research.

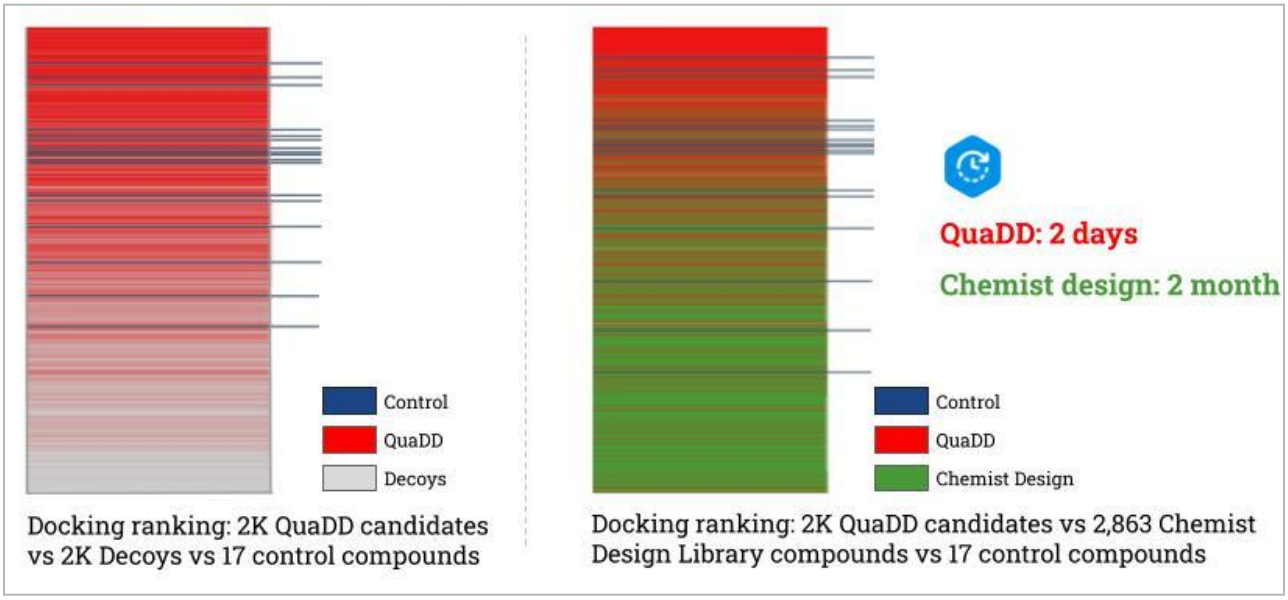

Figure 2. Docking ranking for QuADD, control, decoys, and Chemist Design Library.

QuADD compounds bind in a similar way as the template

We studied the protein-ligand interaction profile of the QuADD candidates and found that QuADD molecules bind in a similar mode as the template molecule

(co-crystallized ligand). The results were compared to the binding profiles generated by the random decoys. In this test, we docked the set of QuADD molecules and decoys to the protein, generating 25 poses per ligand. Then, we calculated Protein-Ligand Interaction Fingerprints (PLIFs) [8] for both sets and for the template molecule. We computed the Tanimoto coefficient (TC) between all poses for each ligand (including QuADD and decoys) and the template and selected the pose with the maximum TC. TC can span values from 1 (total similarity) to 0 (total dissimilarity). When we compared both groups, QuADD, and decoys, we observed that the TC values were higher for the QuADD group (see Figure 3 with the TC for both groups). The protein-ligand interaction profiles of the QuADD molecules were more similar to the template binding profile than the decoys. The QuADD library is composed of molecules that bind the protein in a similar mode as the original co-crystallized template. Figure 4 shows the binding mode of a QuADD candidate along with the crystallized ligand.

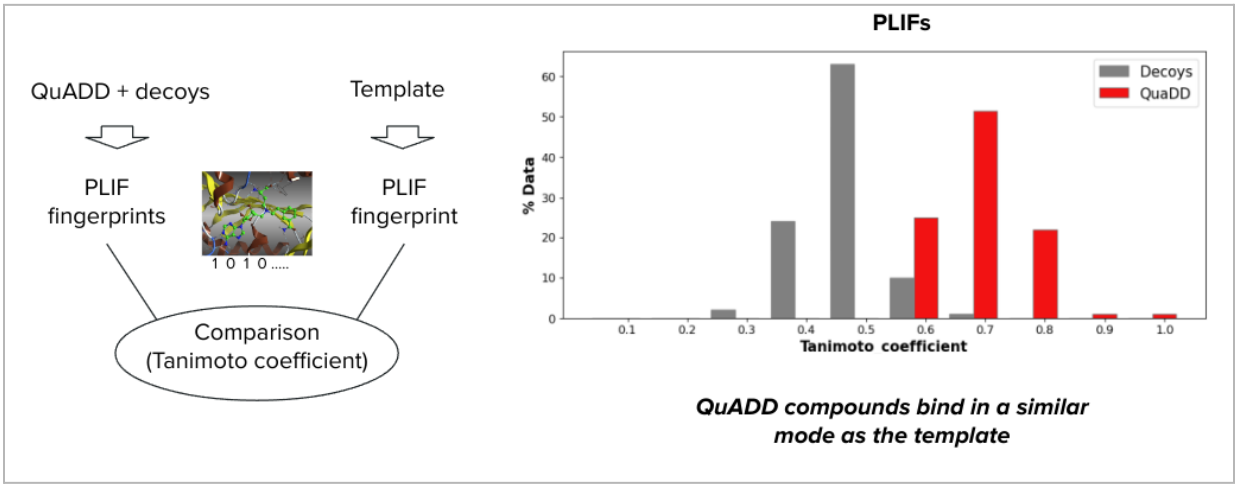

Figure 3. Binding profiles comparison between QuADD-template and decoys-template.

QuADD compounds are novel and diverse

We also measured the similarity/diversity between the QuADD library and the template molecule. We calculated the MACCS fingerprint, composed of 166 structural keys and commonly used to compare molecular similarity [9]. Then, we computed the Tanimoto coefficient (TC) as a measure of similarity between pairs of fingerprints, taking into account each QuADD molecule and the template. The QuADD library showed structural diversity against the template molecule. Most of the QuADD molecules (~80% of the data) yielded TC values against the template between 0.55 and 0.75. Only 0.38% of the QuADD library was similar to the template with a TC value greater or equal to 0.85. QuADD molecules present a different scaffold than the template but still with similar 3D pharmacophoric features that allow an analogous binding interaction.

Additionally, we computed the TC for all the pairs of compounds within the QuaDD set and plotted the results in a cluster heatmap (see Figure 5). The structural cluster analysis showed multiple subsets within the library with high diversity. The QuADD library is not similar to the template and is structurally diverse.

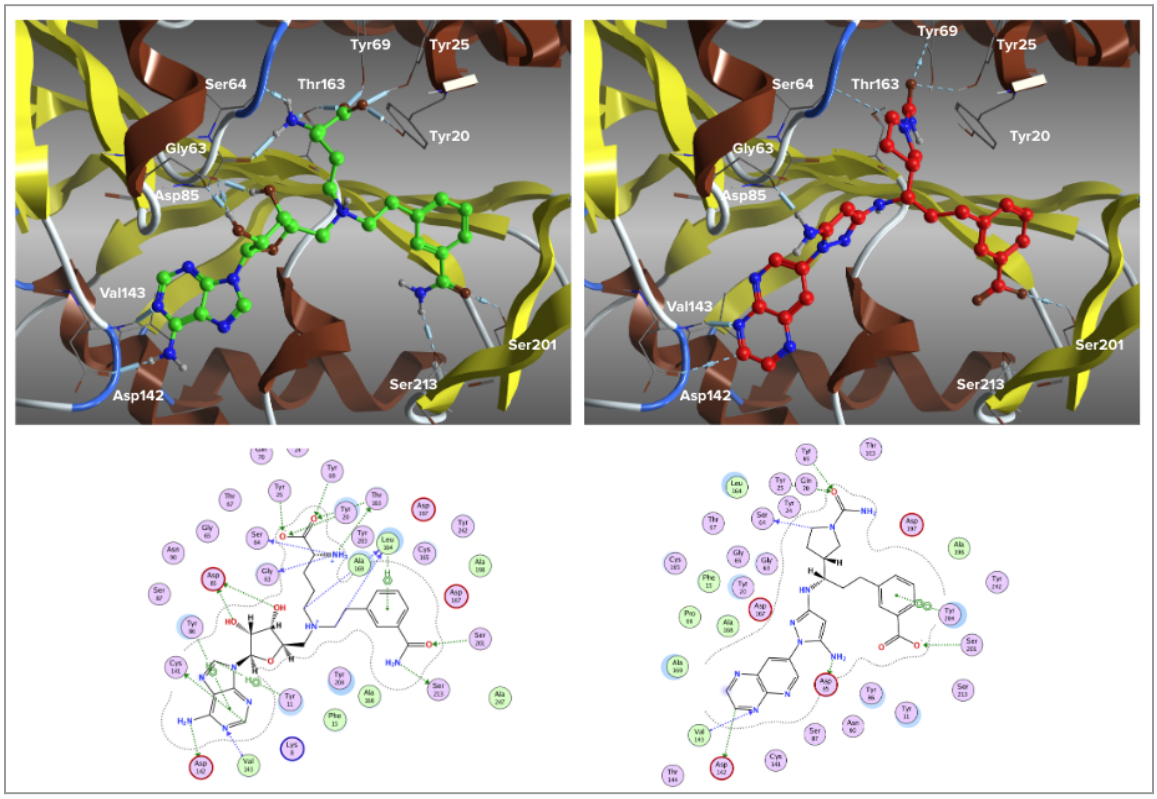

Figure 4. Binding modes for the crystallized ligand (green color) and one of the QuADD candidates (red color).

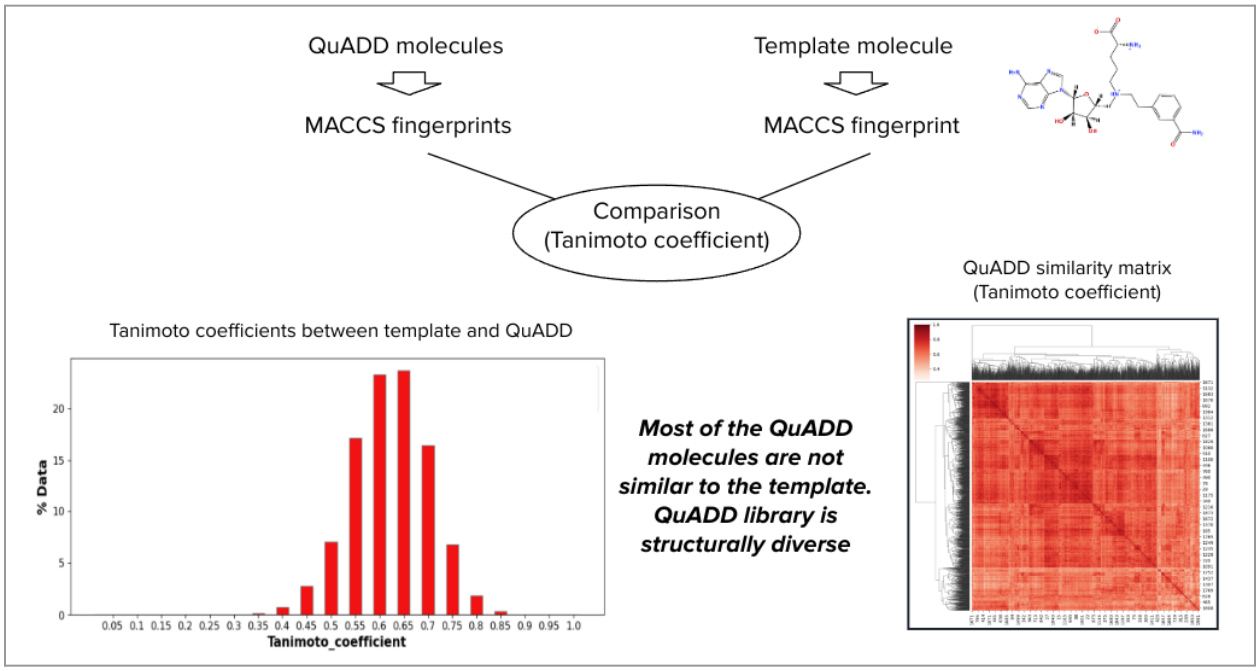

Figure 5. Structural similarity comparison between QuADD and the template molecule, and similarity/diversity within the QuADD library.

QuADD compounds are Drug-like

We tested the drug-like (drug-like as defined by Lipinski, other definitions are easy to implement) quality of the molecules in the QuADD library and compared the results against a representative set of NNMT binders extracted from ChEMBL [10]. We collected the NNMT ChEMBL ligands with $IC_{50}/K_i \leq 10$ µM. Repeated molecules and peptidic drugs were eliminated (all the described NNMT peptides have a high molecular weight, more than 1,000 Da). The final ChEMBL set contained 62 ligands. We calculated drug-like properties in both ChEMBL ligands and the QuADD library and compared the results. The drug-like properties analyzed in this test included topological polar surface area (TPSA), number of rotatable bonds (NRot), number of hydrogen bond donors (HBD), number of hydrogen bond acceptors (HBA), molecular weight (MW), partition coefficient octanol/water (logP), and solubility (logS). The QuADD library showed better drug-like properties because we used drug-like filters as part of our QuADD protocol. QuADD is designed to facilitate drug development and generates great starting point molecules with optimal values in the drug-likeness space. This custom library generation platform heightens the probability of quickly detecting preclinical leads while also reducing the risk of bias in conventional approaches.

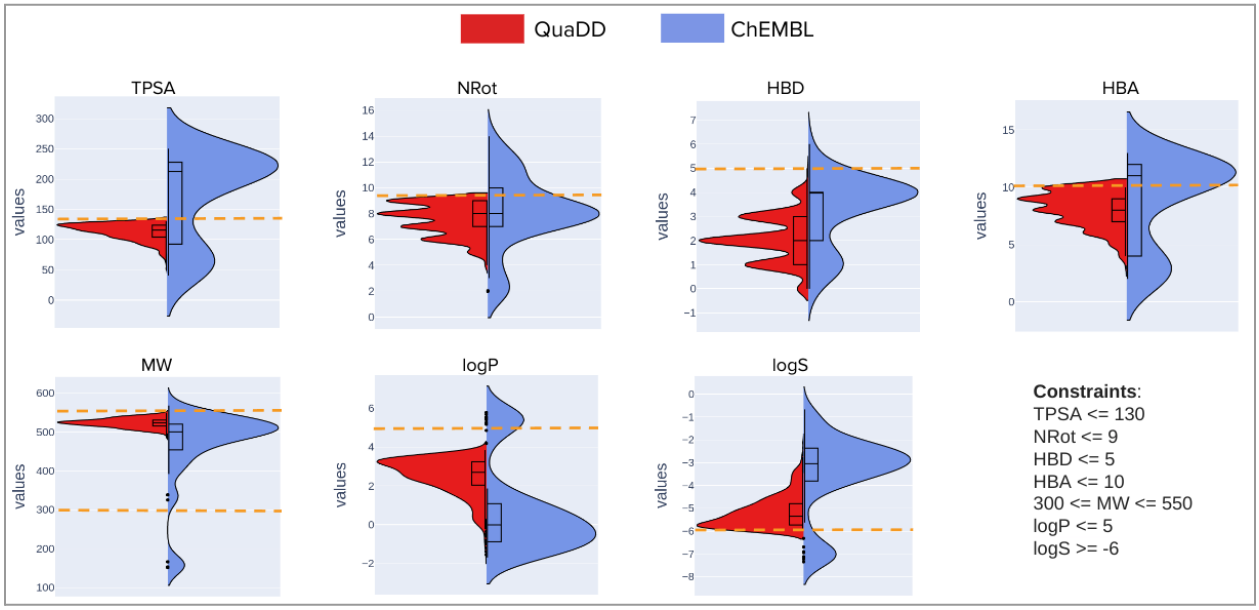

Figure 6. Drug-like properties comparison between 62 ChEMBL ligands and QuADD library. The dotted red line shows the value of the applied constraint for this property.

Conclusion

This test case demonstrates the accuracy and quality of the QuADD library design platform. QuADD yielded excellent results in the generation of potential drug-like NNMT binders. The performance of QuADD was assessed from multiple perspectives, including fragment recovery, docking score, geometrical binding profile, diversity, and drug-like

properties. QuADD recovered the fragments from the template molecule with a high ranking. Our methodology also showed great results in the recovery of fragments from a control set, including multiple structural scaffolds. The library generated by QuADD was evaluated by molecular docking and results were compared against different sets, including control, decoys and a library manually designed by an expert. QuADD clearly yielded the best set of binders in a much shorter time frame. QuADD generated a library composed of molecules with similar binding profiles as the template but at the same time with different structural characteristics and a high degree of diversity. QuADD also has more suitable drug-like properties than previously discovered ChEMBL ligands. In summary, the QuADD library is composed of potential binders structurally diverse than the template but with a similar binding profile and with optimal drug-like properties. These characteristics make QuADD an excellent tool to facilitate drug discovery and development, de-risk preclinical stages and shorten the time to reach clinical programs.

References

[1] POLARISqb. Harnessing the power of quantum computing for drug discovery. https://www.nature.com/articles/d43747-023-00021-3.
[2] Timmer, J. What are companies doing with D-Wave's quantum hardware? https://arstechnica.com/science/2023/01/companies-are-relying-on-quantum-annealers-for-useful-computations/
[3] SWAYNE, M. POLARISqb, Auransa Report Using Quantum Computers to Identify Promising Treatment Targets for Triple Negative Breast Cancer https://thequantuminsider.com/2022/08/12/polarisqb-auransa-report-using-quantum-computers-to-identify-promising-treatment-targets-for-triple-negative-breast-cancer/
[4] Gao Y, Martin NI, van Haren MJ. Nicotinamide N-methyl transferase (NNMT): An emerging therapeutic target. *Drug Discov. Today* 2021, 26(11), 2699-2706.
[5] RCSB PDB Protein Data Bank. https://www.rcsb.org/
[6] Degen J, Wegscheid-Gerlach C, Zaliani A, Rarey M. On the art of compiling and using 'drug-like' chemical fragment spaces. *ChemMedChem* 2008, 3(10), 1503-1507.
[7] Chen D, Li L, Diaz K, Iyamu ID, Yadav R, Noinaj N, Huang R. Novel Propargyl-Linked Bisubstrate Analogues as Tight-Binding Inhibitors for Nicotinamide *N*-Methyltransferase. *J. Med. Chem.* 2019, 62(23), 10783-10797.
[8] Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res.* 2015, 43(W1), W443-447.
[9] Daylight Chemical Information Systems, Inc. https://www.daylight.com/dayhtml/doc/theory/theory.finger.html
[10] ChEMBL Database - EMBL-EBI. European Bioinformatics Institute. https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL2346486/