Supplementary Material for MEMES: Machine learning framework for Enhanced MolEcular Screening

Sarvesh Mehta,[†] Siddhartha Laghuvarapu,[†,§] Yashaswi Pathak,[†,§] Aaftaab Sethi,[‡] Mallika Alvala,[¶] and U. Deva Priyakumar[*,†]

†*Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500 032, India*

‡*Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research, Hyderabad 500 037, India*

¶*School of Pharmacy and Technology Management, Narsee Monjee Institute of Management Sciences, Hyderabad*

§*Contributed equally to this work*

E-mail: deva@iiit.ac.in

Phone: +91 40 6653 1161. Fax: +91 40 6653 1413

Table S1: Overlap of top N molecules sampled by ExactMEMES framework(CDDD, Mol2Vec and ECFP) with actual top hits from Zinc250K against both target proteins. Reported results are for three runs.

Target Protein	Molecular Embedding	Run No.	N	No. of molecules matched
TTBK1	Mol2Vec	1	100	98
TTBK1	Mol2Vec	2	100	96
TTBK1	Mol2Vec	3	100	96
TTBK1	Mol2Vec	1	500	467
TTBK1	Mol2Vec	2	500	461
TTBK1	Mol2Vec	3	500	462
TTBK1	CDDD	1	100	100
TTBK1	CDDD	2	100	100
TTBK1	CDDD	3	100	100
TTBK1	CDDD	1	500	489
TTBK1	CDDD	2	500	487
TTBK1	CDDD	3	500	487
TTBK1	ECFP	1	100	92
TTBK1	ECFP	2	100	90
TTBK1	ECFP	3	100	92
TTBK1	ECFP	1	500	423
TTBK1	ECFP	2	500	419
TTBK1	ECFP	3	500	424
SARS-CoV-2 M^{pro}	Mol2Vec	1	100	99
SARS-CoV-2 M^{pro}	Mol2Vec	2	100	98
SARS-CoV-2 M^{pro}	Mol2Vec	3	100	98
SARS-CoV-2 M^{pro}	Mol2Vec	1	500	465
SARS-CoV-2 M^{pro}	Mol2Vec	2	500	469
SARS-CoV-2 M^{pro}	Mol2Vec	3	500	472
SARS-CoV-2 M^{pro}	CDDD	1	100	99
SARS-CoV-2 M^{pro}	CDDD	2	100	98
SARS-CoV-2 M^{pro}	CDDD	3	100	99
SARS-CoV-2 M^{pro}	CDDD	1	500	480
SARS-CoV-2 M^{pro}	CDDD	2	500	483
SARS-CoV-2 M^{pro}	CDDD	3	500	484
SARS-CoV-2 M^{pro}	ECFP	1	100	80
SARS-CoV-2 M^{pro}	ECFP	2	100	79
SARS-CoV-2 M^{pro}	ECFP	3	100	80
SARS-CoV-2 M^{pro}	ECFP	1	500	387
SARS-CoV-2 M^{pro}	ECFP	2	500	393
SARS-CoV-2 M^{pro}	ECFP	3	500	390

Table S2: Overlap of top N molecules sampled by DeepMEMES framework(Mol2Vec, CDDD, ECFP) with actual top hits from Enamine HTS Collection against target protein TTBK1. Reported results are for three runs.

Molecular Embedding	Run No.	N	No. of molecules matched
CDDD	1	100	89
CDDD	2	100	88
CDDD	3	100	89
CDDD	1	500	448
CDDD	2	500	448
CDDD	3	500	448
Mol2Vec	1	100	99
Mol2Vec	2	100	96
Mol2Vec	3	100	96
Mol2Vec	1	500	487
Mol2Vec	2	500	485
Mol2Vec	3	500	488
ECFP	1	100	78
ECFP	2	100	77
ECFP	3	100	76
ECFP	1	500	362
ECFP	2	500	361
ECFP	3	500	356

Table S3: Overlap of top N molecules sampled by DeepMEMES framework(Mol2Vec) with actual top hits from Ultra Large Docking Library against target protein AmpC. Reported results are for three runs.

Molecular Embedding	Run No.	N	No. of molecules matched
Mol2Vec	1	500	445
Mol2Vec	2	500	437
Mol2Vec	3	500	437
Mol2Vec	1	1000	900
Mol2Vec	2	1000	889
Mol2Vec	3	1000	880
Mol2Vec	1	5000	4396
Mol2Vec	2	5000	4354
Mol2Vec	3	5000	4342

Table S4: Overlap of top N molecules sampled by DeepMEMES framework(CDDD, Mol2Vec and ECFP) with actual top hits from Zinc250K against both target proteins. Reported results are for three runs.

Target Protein	Molecular Embedding	Run No.	N	No. of molecules matched
TTBK1	Mol2Vec	1	100	97
TTBK1	Mol2Vec	2	100	95
TTBK1	Mol2Vec	3	100	98
TTBK1	Mol2Vec	1	500	464
TTBK1	Mol2Vec	2	500	473
TTBK1	Mol2Vec	3	500	471
TTBK1	CDDD	1	100	98
TTBK1	CDDD	2	100	99
TTBK1	CDDD	3	100	98
TTBK1	CDDD	1	500	469
TTBK1	CDDD	2	500	474
TTBK1	CDDD	3	500	472
TTBK1	ECFP	1	100	74
TTBK1	ECFP	2	100	72
TTBK1	ECFP	3	100	73
TTBK1	ECFP	1	500	329
TTBK1	ECFP	2	500	318
TTBK1	ECFP	3	500	333
SARS-CoV-2 M^{pro}	Mol2Vec	1	100	98
SARS-CoV-2 M^{pro}	Mol2Vec	2	100	95
SARS-CoV-2 M^{pro}	Mol2Vec	3	100	97
SARS-CoV-2 M^{pro}	Mol2Vec	1	500	473
SARS-CoV-2 M^{pro}	Mol2Vec	2	500	458
SARS-CoV-2 M^{pro}	Mol2Vec	3	500	471
SARS-CoV-2 M^{pro}	CDDD	1	100	98
SARS-CoV-2 M^{pro}	CDDD	2	100	95
SARS-CoV-2 M^{pro}	CDDD	3	100	99
SARS-CoV-2 M^{pro}	CDDD	1	500	471
SARS-CoV-2 M^{pro}	CDDD	2	500	462
SARS-CoV-2 M^{pro}	CDDD	3	500	473
SARS-CoV-2 M^{pro}	ECFP	1	100	71
SARS-CoV-2 M^{pro}	ECFP	2	100	62
SARS-CoV-2 M^{pro}	ECFP	3	100	63
SARS-CoV-2 M^{pro}	ECFP	1	500	302
SARS-CoV-2 M^{pro}	ECFP	2	500	303
SARS-CoV-2 M^{pro}	ECFP	3	500	295

Table S5: Overlap of top N molecules sampled by DeepMEMES framework(CDDD) and Deep Docking with actual top hits from Zinc250K against both target proteins. Reported results are for three runs.

Target Protein	Method	Run No.	N	No. of molecules matched
TTBK1	Deep Docking	1	500	59
TTBK1	Deep Docking	2	500	68
TTBK1	Deep Docking	3	500	69
TTBK1	DeepMEMES	1	500	469
TTBK1	DeepMEMES	2	500	474
TTBK1	DeepMEMES	3	500	472
SARS-CoV-2 M^{pro}	Deep Docking	1	500	67
SARS-CoV-2 M^{pro}	Deep Docking	2	500	67
SARS-CoV-2 M^{pro}	Deep Docking	3	500	70
SARS-CoV-2 M^{pro}	DeepMEMES	1	500	471
SARS-CoV-2 M^{pro}	DeepMEMES	2	500	462
SARS-CoV-2 M^{pro}	DeepMEMES	3	500	473

Table S6: Time taken (in hrs) by different steps of proposed method and Deep Docking for identifying top hits against target protein TTBK1 from Zinc-250K drug library. Reported metrics are of a single run of DeepMEMES and Deep Docking.

	DeepMEMES	Deep Docking	Docking of all compounds
Clustering	0.0872	-	-
Embedding Calculation	0.00294	0.00252	-
Docking Calculation	80.375	74.590	1012.855
Training	0.955	0.904	-
Total Time	81.420	75.49652	1012.855

Table S7: Overlap of top N molecules sampled by DeepMEMES framework(CDDD) and Deep Docking with actual top hits from Enamine dataset against TTBK1 target protein. Reported results are for three runs.

Target Protein	Method	Run No.	N	No. of molecules matched
TTBK1	Deep Docking	1	500	60
TTBK1	Deep Docking	2	500	56
TTBK1	Deep Docking	3	500	69
TTBK1	DeepMEMES	1	500	469
TTBK1	DeepMEMES	2	500	474
TTBK1	DeepMEMES	3	500	472

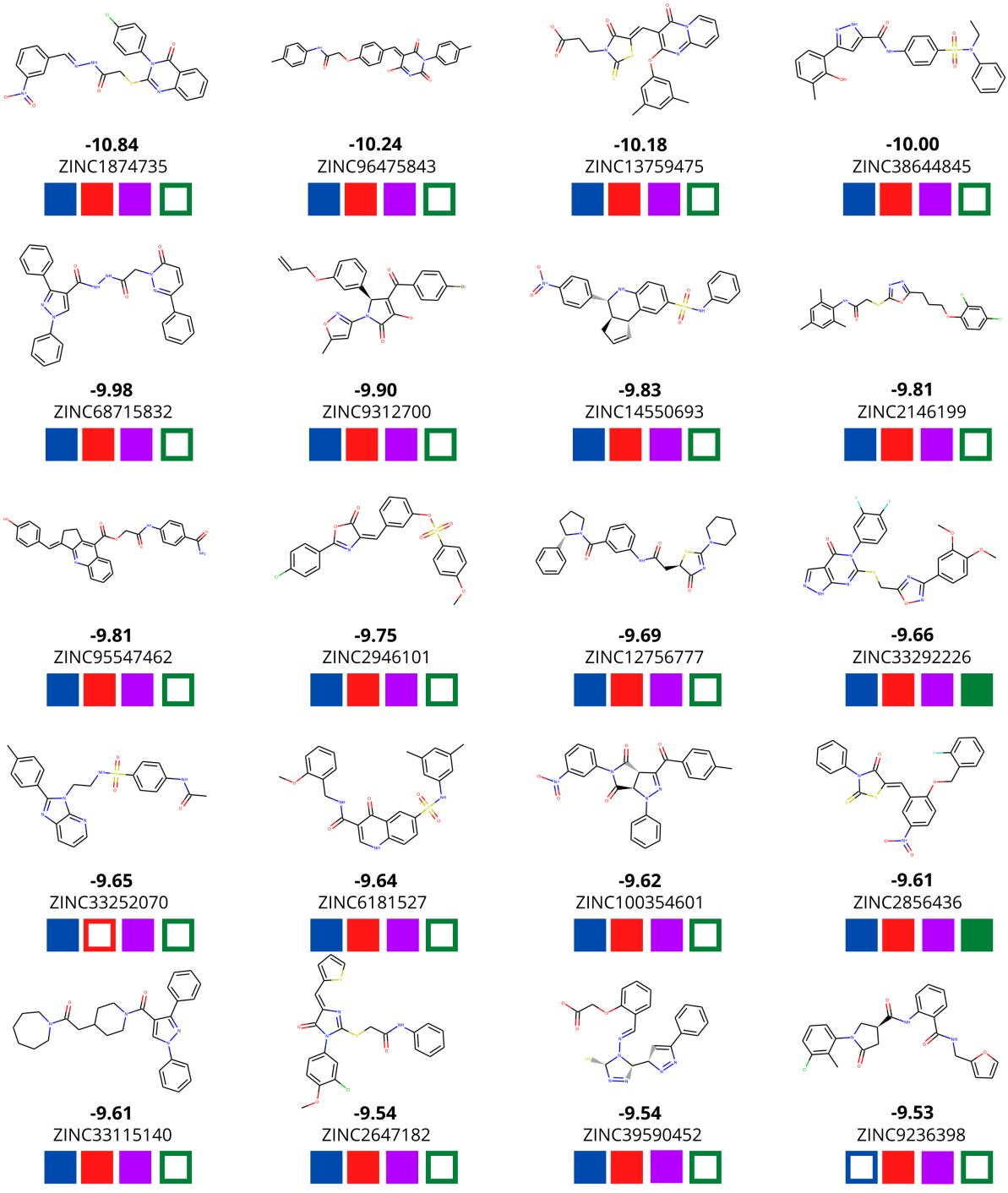


Figure S1: Top 20 hits for target protein TTBK1 found in Zinc-250K dataset. The color codes indicate if the molecules have also been found using ExactMEMES framework or random search.

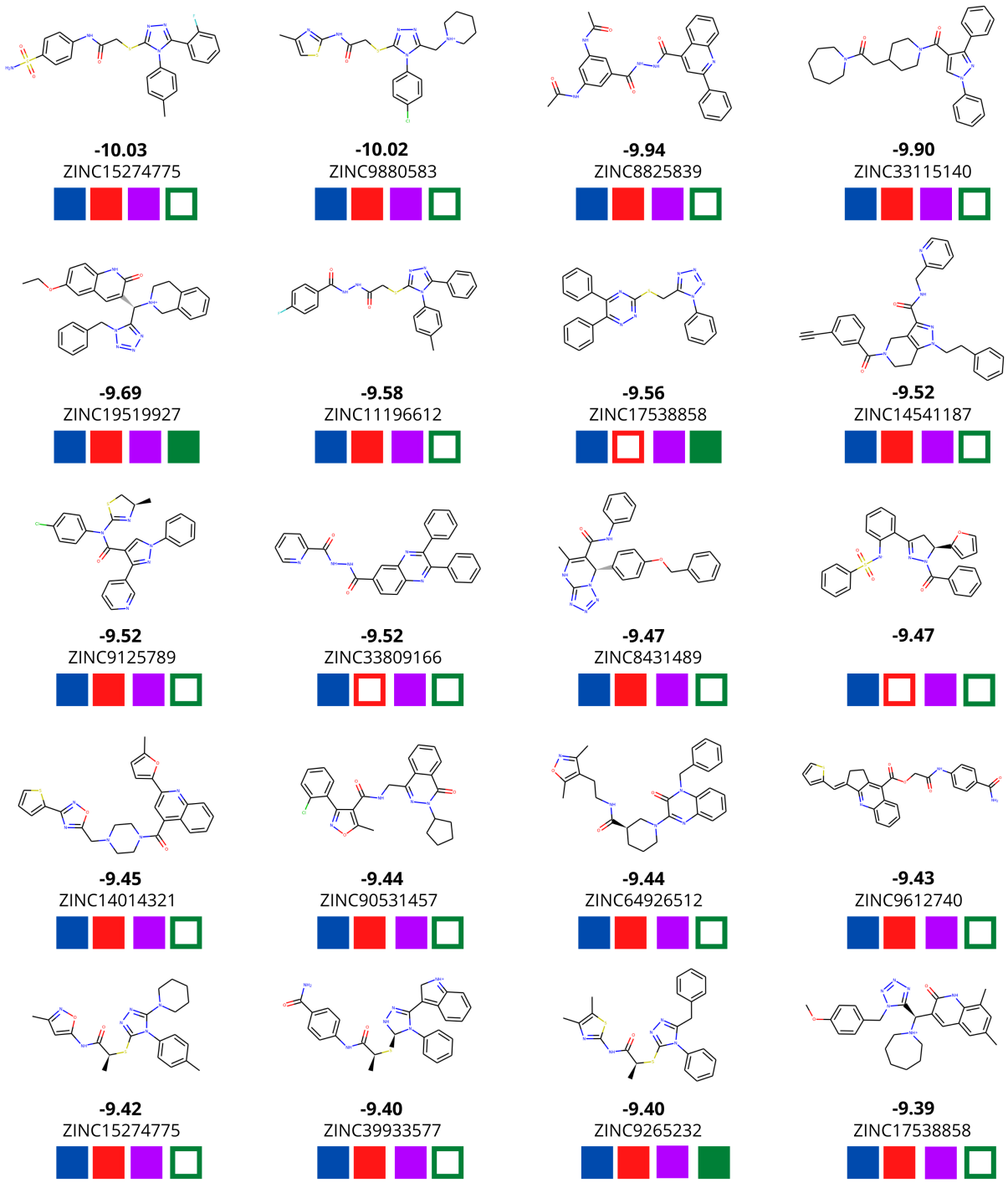
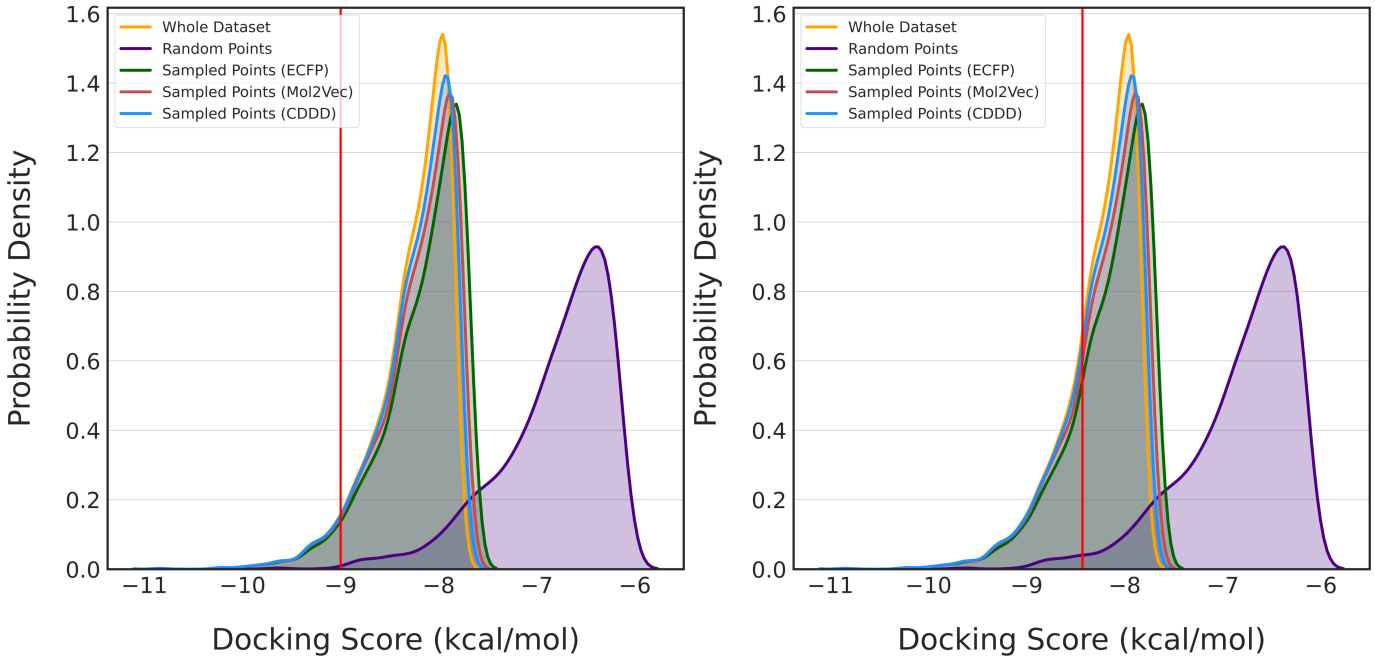
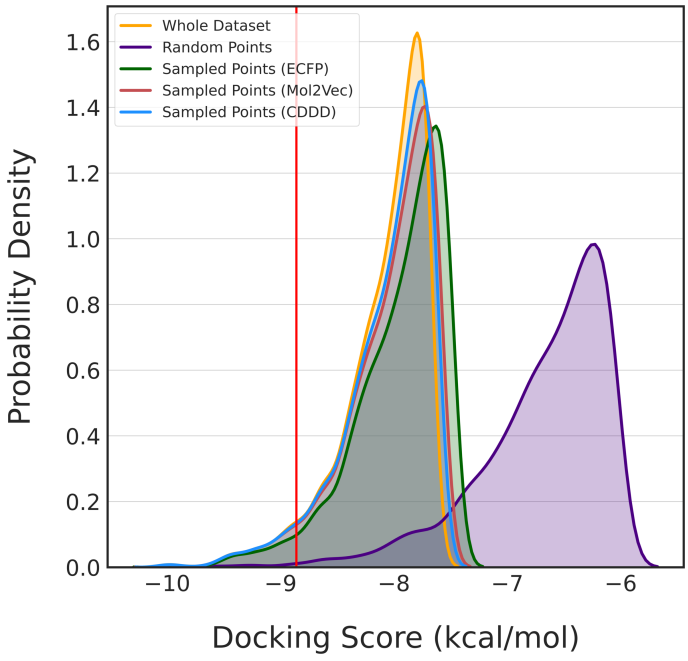


Figure S2: Top 20 hits for target protein SARS-CoV-2 M^{pro} found in Zinc-250K dataset. The color codes indicate if the molecules have also been found using ExactMEMES framework or random search.

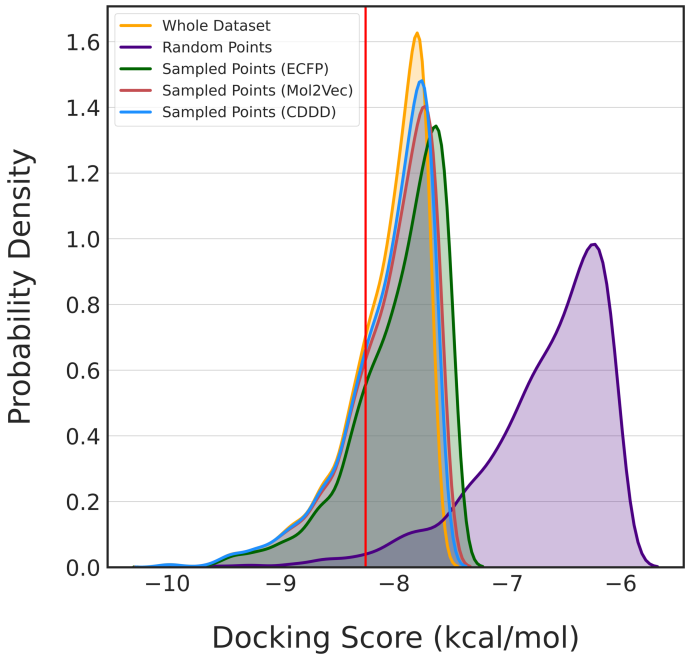


(a) Target Protein: TTBK1

(b) Target Protein: TTBK1

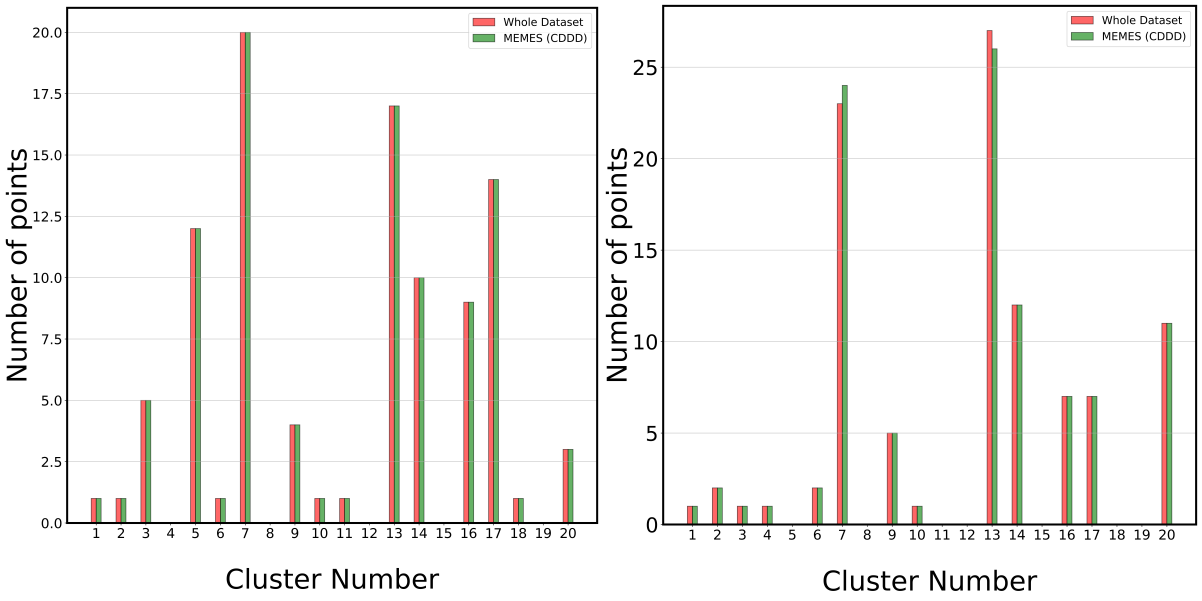


(c) Target Protein: SARS-CoV-2 M^{pro}



(d) Target Protein: SARS-CoV-2 M^{pro}

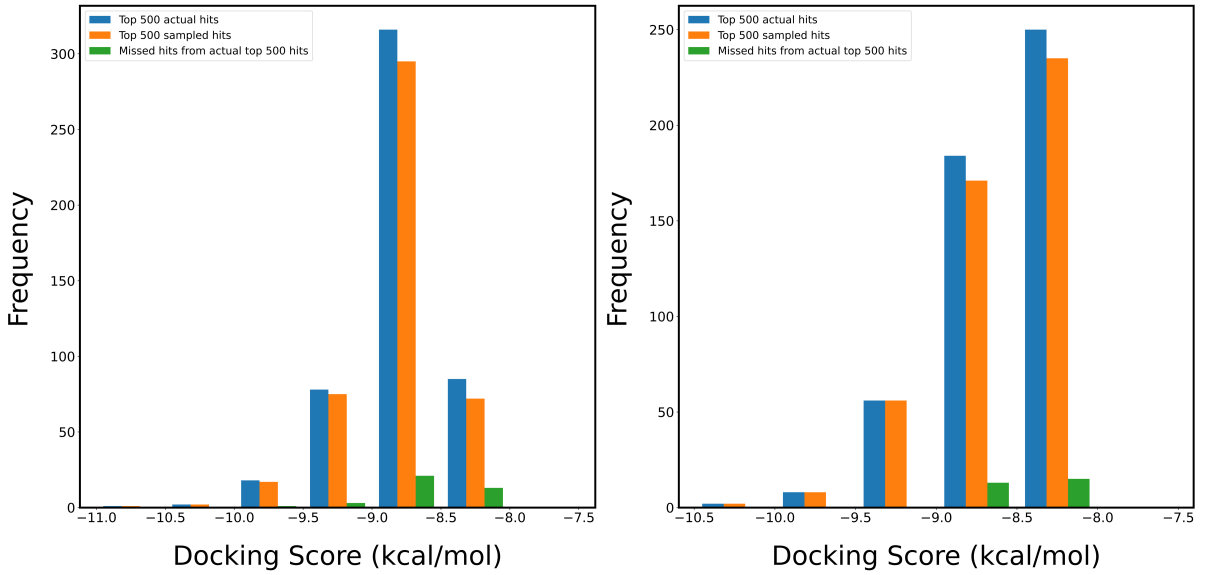
Figure S3: Figure shows the distribution of docking scores for top 2000 molecules sampled by ExactMEMES, random sampling, and actual top 2000 docking hits of the complete Zinc-250K molecular library. Vertical red line shows the cut-off docking score for top 100 molecules in (a) and (c), and for top 500 molecules in (b) and (d).



(a) Target Protein: TTBK1

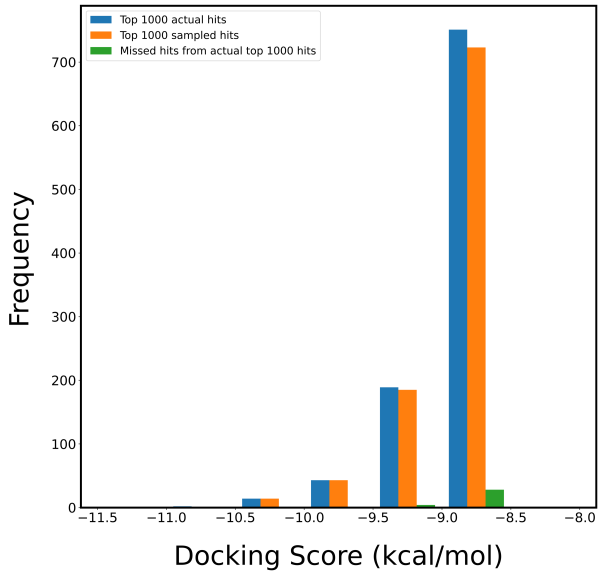
(b) Target Protein: SARS-CoV-2 M^{pro}

Figure S4: The figure shows the distribution of top molecules sampled by ExactMEMES (with CDDD as molecular embedding) from Zinc-250K across 20 different clusters.

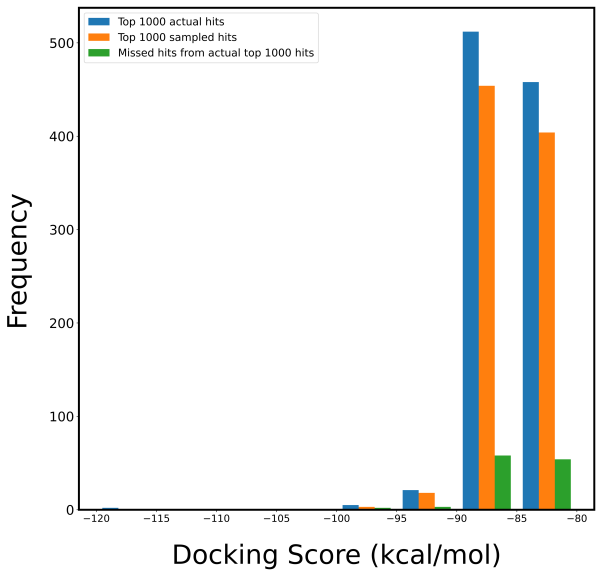


(a) Target Protein: TTBK1

(b) Target Protein: SARS-CoV-2 M^{pro}

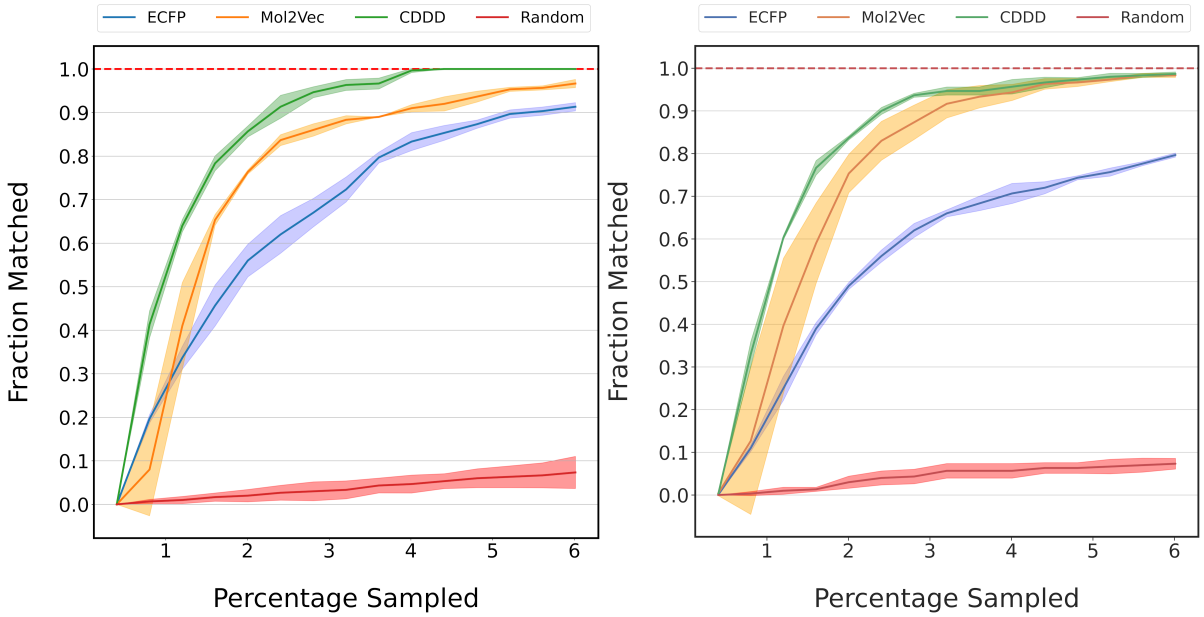


(c) Target Protein: TTBK1



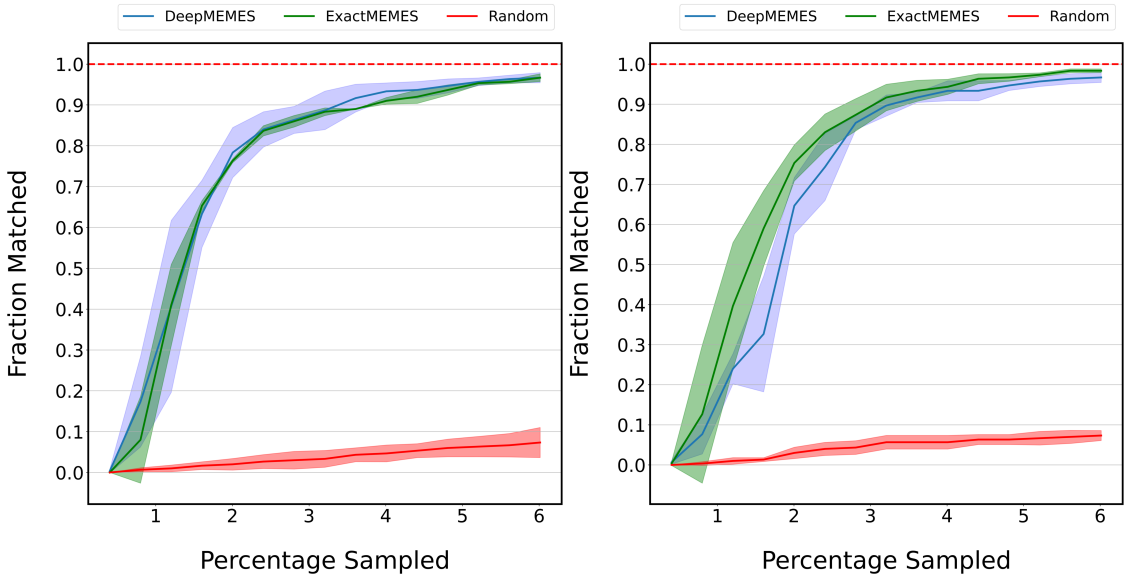
(d) Target Protein: AmpC

Figure S5: The histogram shows the distribution of top molecules identified by MEMES across different bins (according to docking score) as well as the spread of actual top hits missed by proposed framework. (a) and (b) shows the distribution of molecules sampled by ExactMEMES (with Mol2Vec as molecular embedding) from Zinc-250K dataset. (c), and (d) shows the distribution of molecules sampled by DeepMEMES (with Mol2Vec as molecular embedding) from Enamine dataset, and Ultra Large Docking library, respectively.



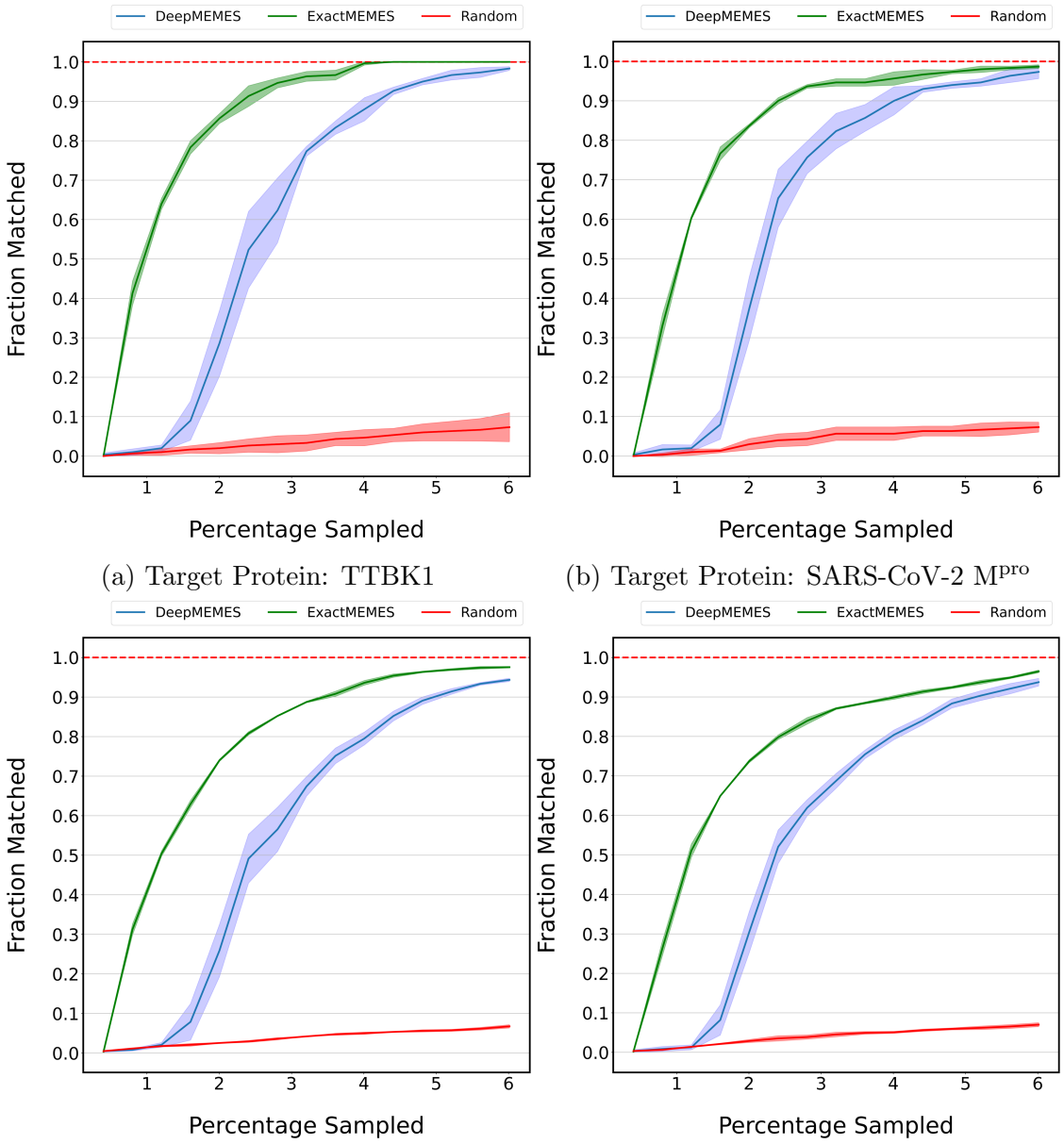
(a) Target Protein: TTBK1 (b) Target Protein: SARS-CoV-2 M^{pro}

Figure S6: The figure shows the fraction of top 100 molecules sampled by ExactMEMES that matches with actual top hits in the Zinc-250K docking library (for target receptor TTBK1 and SARS-CoV-2 M^{pro}) against the percentage of dataset sampled.



(a) Target Protein: TTBK1 (b) Target Protein: SARS-CoV-2 M^{pro}

Figure S7: To compare the performance of ExactMEMES and DeepMEMES, fraction of the top 100 molecules sampled by MEMES from Zinc-250K dataset that are actual top hits is plotted against the percentage of dataset sampled. Mol2Vec as featurization technique was used for this comparison. The reported trial results are average of 3 runs and the shaded region represent standard deviation across these runs.



(a) Target Protein: TTBK1

(b) Target Protein: SARS-CoV-2 M^{pro}

(c) Target Protein: TTBK1

(d) Target Protein: SARS-CoV-2 M^{pro}

Figure S8: To compare the performance of ExactMEMES and DeepMEMES, fraction of the top molecules sampled that are actual top hits from Zinc-250K drug library is plotted against the percentage of dataset sampled. CDDD as featurization technique was used for this comparison. The reported trial results are average of 3 runs and the shaded region represent standard deviation across these runs. (a) and (b) compares for top 100 sampled molecules. (c) and (d) compares for top 500 sampled molecules.

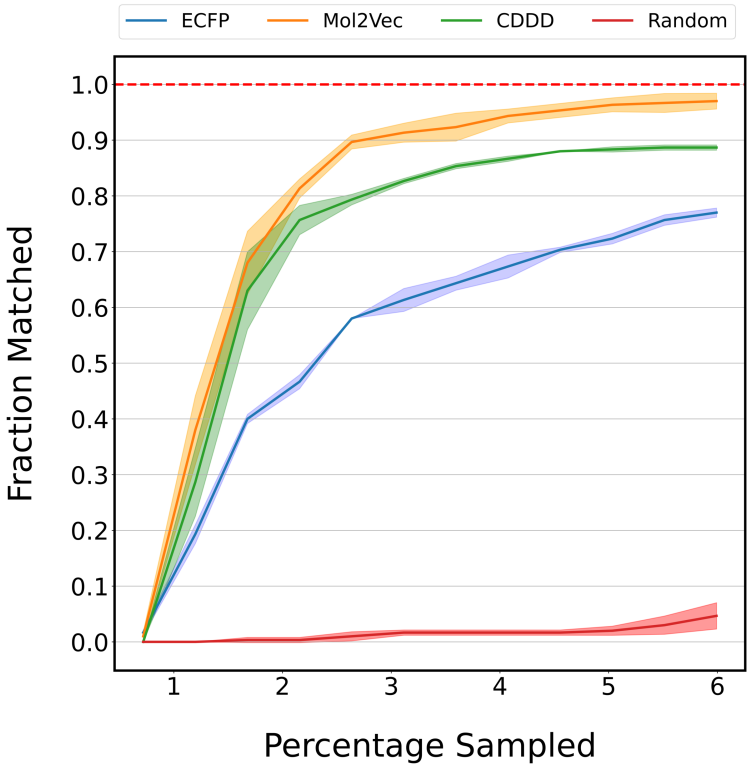
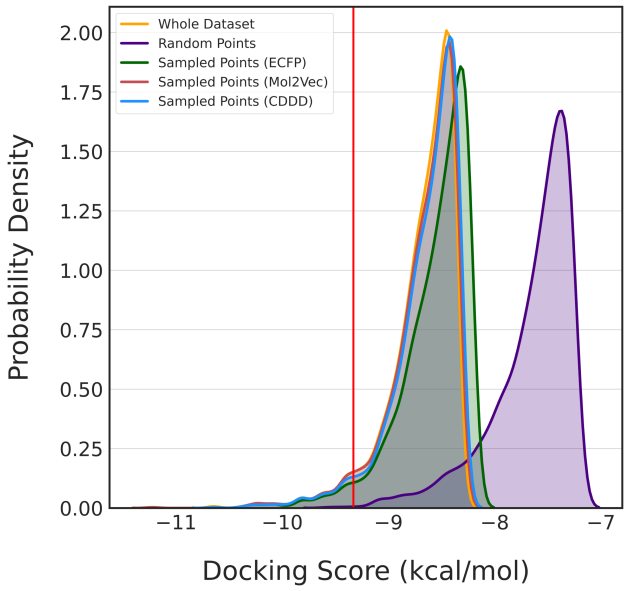
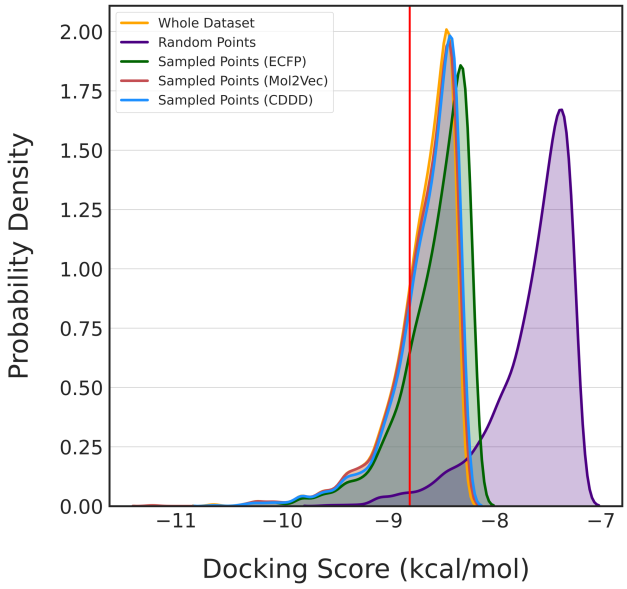


Figure S9: Figure shows the fraction of top 100 molecules that matches with actual top hits from Enamine dataset (for target receptor TTBK1) against the percentage of dataset sampled.

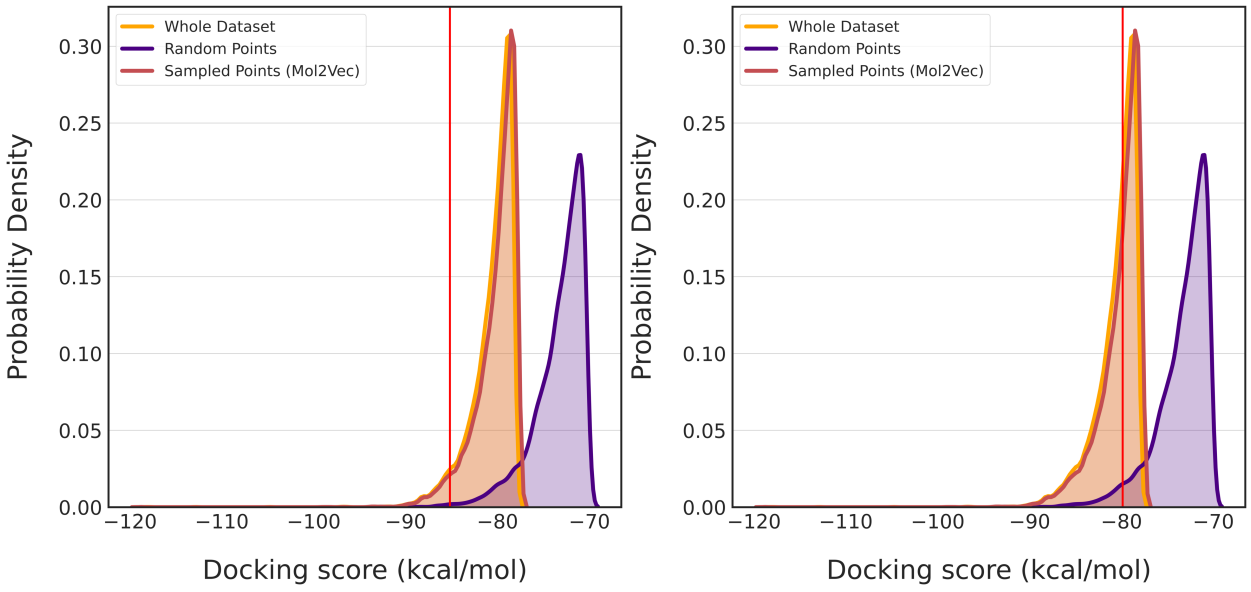


(a)



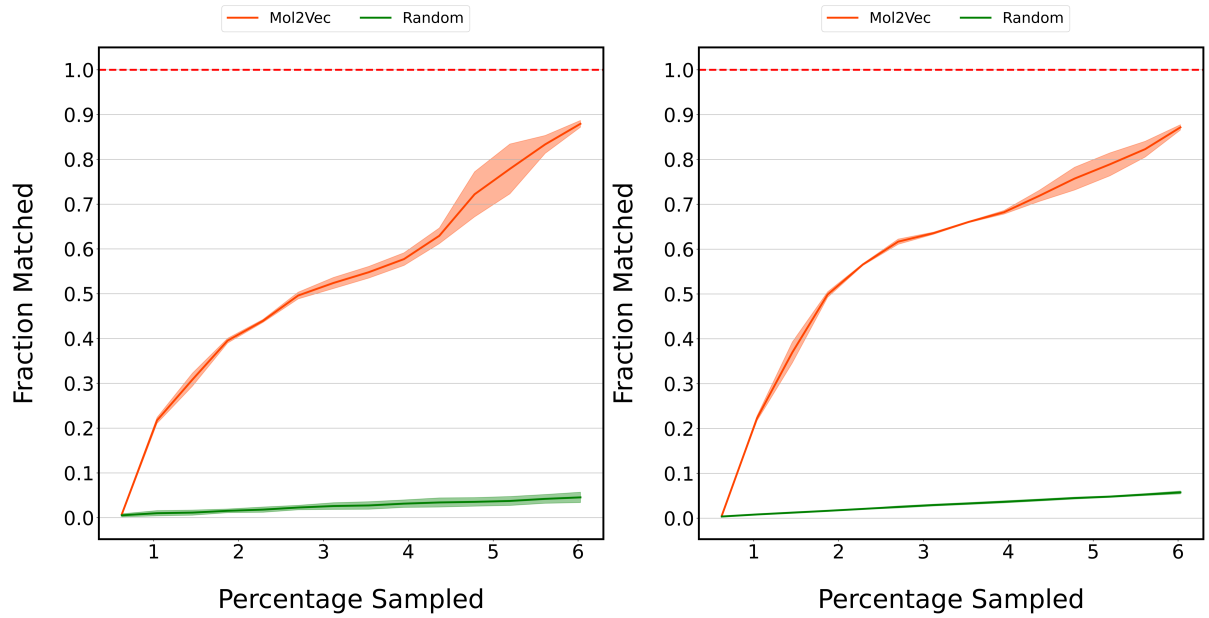
(b)

Figure S10: Performance of DeepMEMES on Enamine Dataset against target receptor TTBK1. Figure shows the distribution of docking scores of top 2000 molecules sampled by DeepMEMES, random sampling, and actual top 2000 docking hits of the complete Enamine Dataset. Vertical red line shows the cut-off docking score for top 100 molecules in (a) and top 500 molecules in (b).



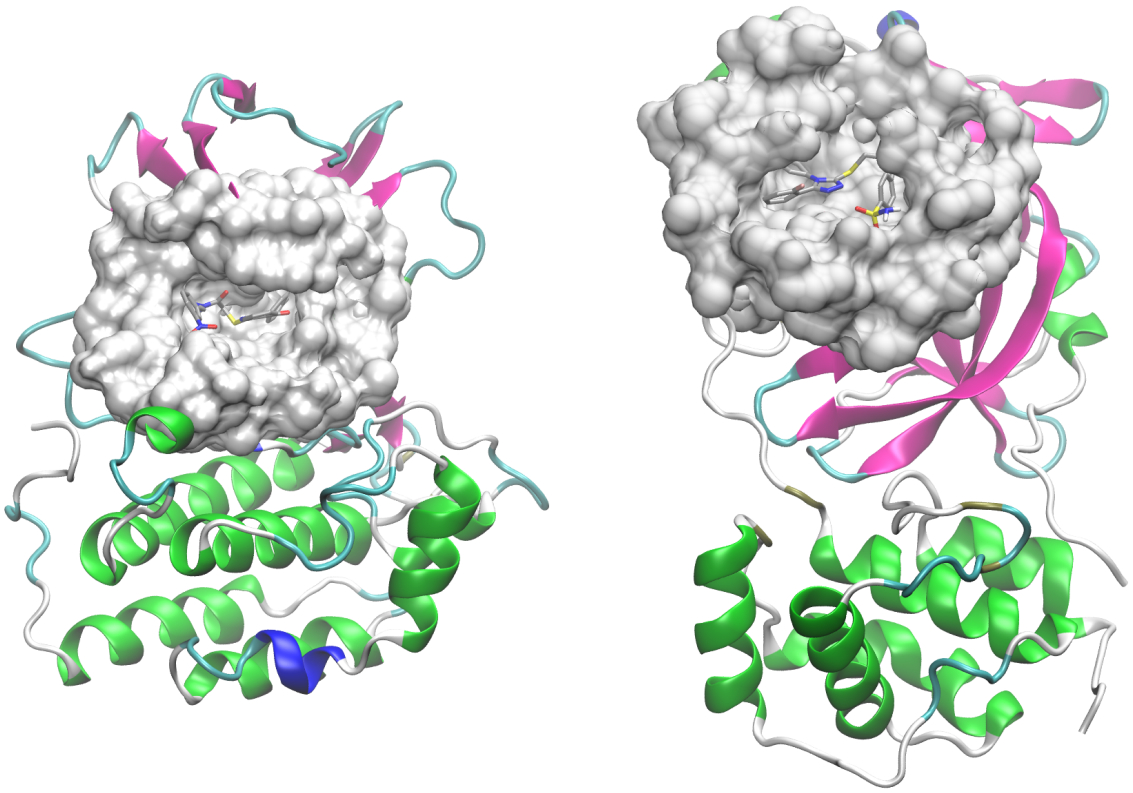
(a) (b)

Figure S11: Performance of DeepMEMES on Ultra Large Docking Library against target protein AmpC. Figure shows the distribution of docking scores of top 10000 molecules sampled by DeepMEMES, random sampling, and actual top 10000 docking hits of the complete dataset. Vertical red line shows the cut-off docking score for top 500 molecules in (a) and top 5000 molecules in (b).



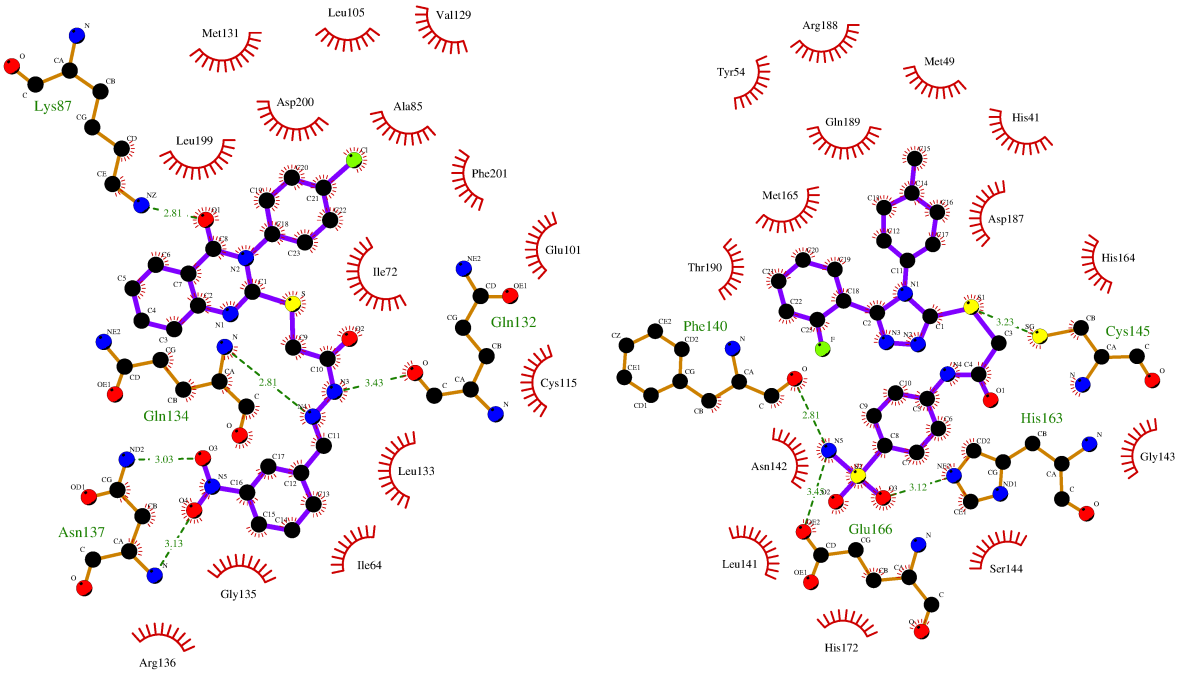
(a) Top 500 molecules (b) Top 5000 molecules

Figure S12: The figure shows the fraction of top molecules sampled by DeepMEMES that are actual top hits from Ultra Large Docking Library (for target protein AmpC) against the percentage of dataset sampled.



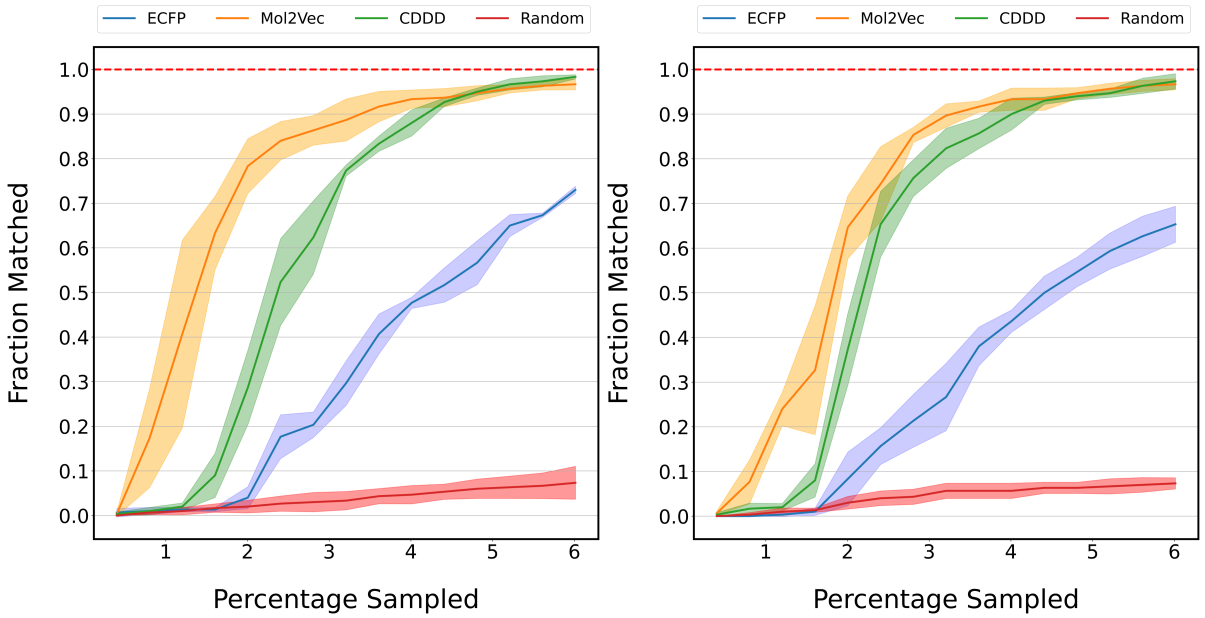
(a) Target Protein: TTBK1 (b) Target Protein: SARS-CoV-2 M^{pro}

Figure S13: 3D Protein-Ligand Complex of target protein and the best hit (molecule with most negative docking score) in complete Zinc-250K library.



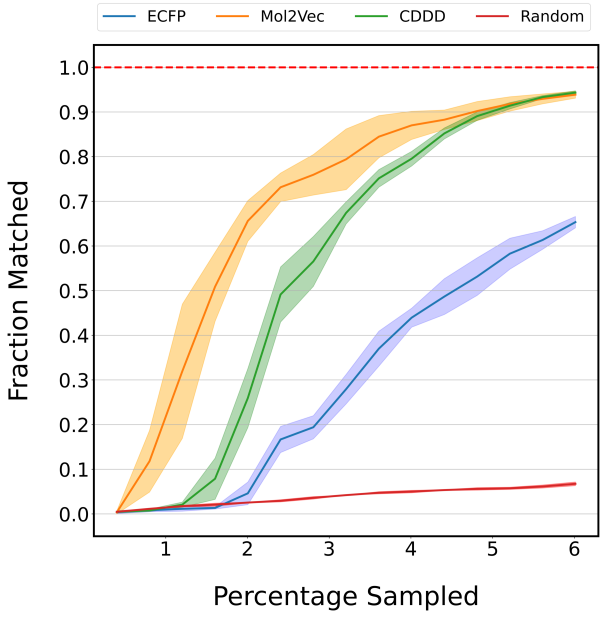
(a) Target Protein: TTBK1 (b) Target Protein: SARS-CoV-2 M^{pro}

Figure S14: Ligplot image showing the interaction between the target protein and best hit in complete Zinc-250K library.

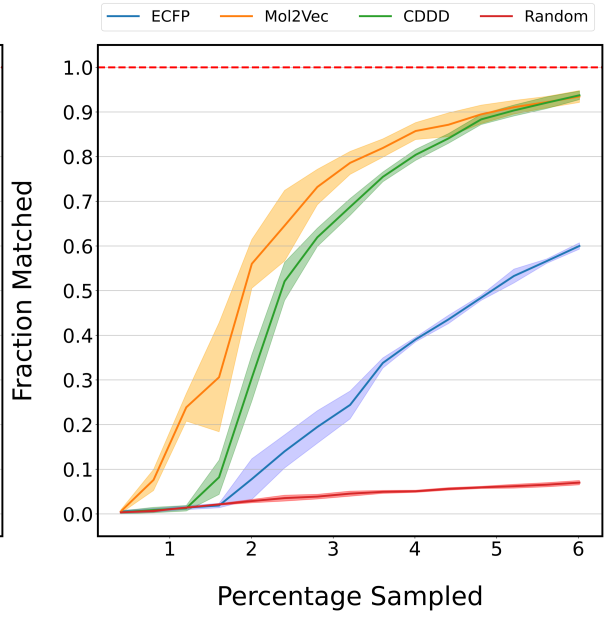


(a) Target Protein: TTBK1

(b) Target Protein: SARS-CoV-2 M^{pro}

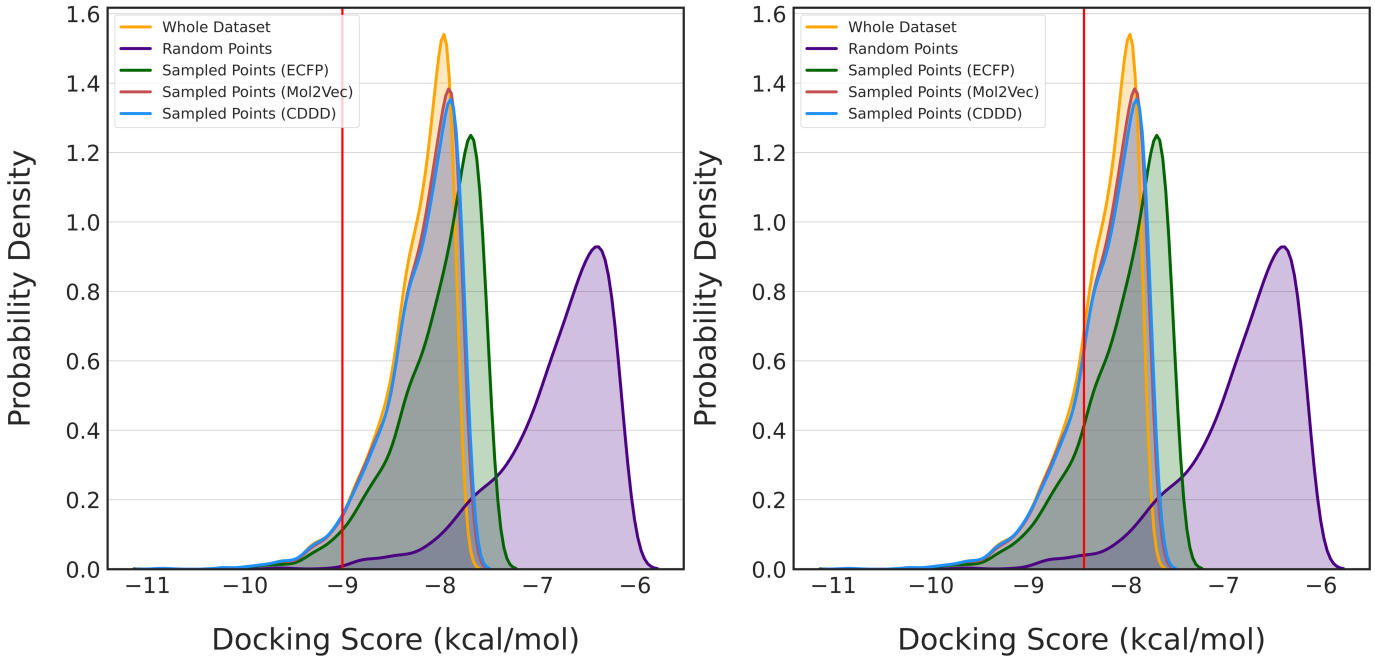


(c) Target Protein: TTBK1



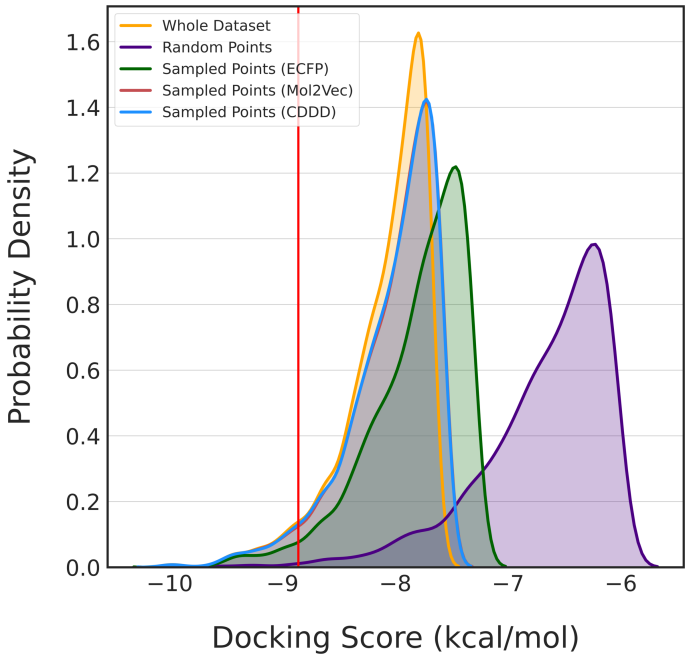
(d) Target Protein: SARS-CoV-2 M^{pro}

Figure S15: Performance of DeepMEMES on Zinc-250K against protein receptor TTBK1 and SARS-CoV-2 M^{pro}. Figure shows the fraction of top molecules sampled by DeepMEMES that matches with actual top hits of complete library. (a) and (b), and (c) and (d) shows the plots for top 100 molecules and top 500 molecules, respecitvely. The reported trial results are average of 3 runs and the shaded region represent standard deviation across these runs.

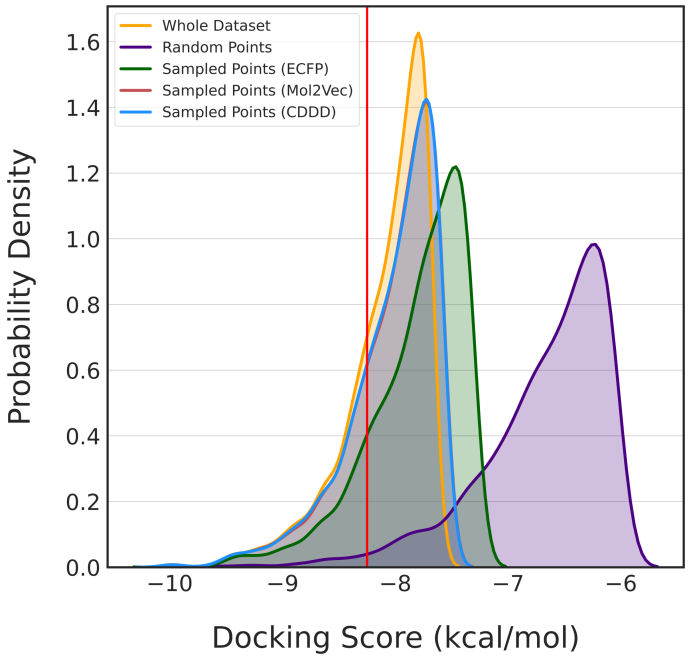


(a) Target Protein: TTBK1

(b) Target Protein: TTBK1

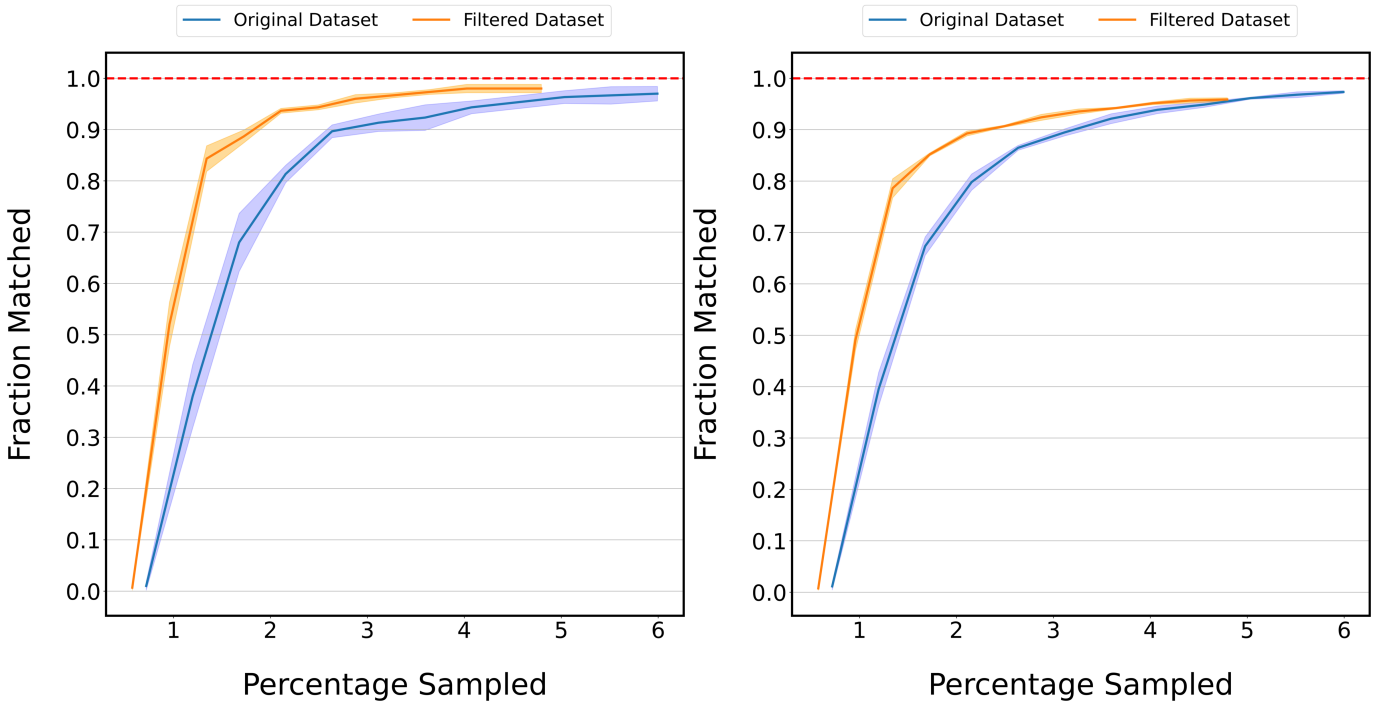


(c) Target Protein: SARS-CoV-2 M^{pro}



(d) Target Protein: SARS-CoV-2 M^{pro}

Figure S16: Figure shows the distribution of docking scores for top 2000 molecules sampled by DeepMEMES, random sampling, and actual top 2000 docking hits of the complete Zinc-250K molecular library. Vertical red line shows the cut-off docking score for top 100 molecules in (a) and (c), and for top 500 molecules in (b) and (d).



(a) Top 100 molecules

(b) Top 500 molecules

Figure S17: Figures shows the fraction of top molecules sampled by DeepMEMES (with Mol2Vec as molecular embedding technique) from Zinc-250K drug library that are actual top hits against target receptor TTBK1. The reported trial results are average of 3 runs and the shaded region represent standard deviation across these runs.

Supplementary Discussion

1. Performance of DeepMEMES on Zinc-250K

Statistics on the performance of DeepMEMES on Zinc250K against both target receptors is given in this section.

Fraction Matched vs Percentage of docking library sampled

Figure S15 shows the fraction of top 100 and top 500 sampled molecules that are actual top hits for both target proteins against the percentage of molecules sampled from the docking library using DeepMEMES and random sampling.

Statistics on top sampled molecules

Table S4 represents the overlap of top 100 and top 500 molecules identified by DeepMEMES framework (CDDD, Mol2Vec and ECFP) with actual top hits from the whole dataset.

Distribution of docking scores

Figure S16 compares distribution of docking scores for top molecules sampled using DeepMEMES framework (CDDD, Mol2Vec and ECFP) against random selection for both the protein receptors.

2. Rule Based Screening followed by application of DeepMEMES

It is a common practice in drug discovery to apply rule based filters such as a molecular volume or polar surface area filter to remove molecules and reduce the size of the dataset to be screened. Rule based screening followed by application of MEMES framework on the Enamine dataset against TTBK1 target receptor is shown in this section. Topological Polar Surface Area (TPSA) filter was applied on the dataset to filter out the ligands that have TPSA > 90 $\mathring{A}^2$. This reduces the dataset size from $\sim$2 million compounds to $\sim$1.6 million.

To asses the performance of DeepMEMES framework with rule-based filtering, it was applied on the filtered dataset.

Figure S17 compares the performance of the DeepMEMES framework on the Enamine dataset before and after applying the TPSA filter. Figure shows the fraction of top molecules sampled by DeepMEMES that are actual top hits from the corresponding dataset. Note that percentage sampled shown on the x-axis is calculated by taking the size of complete dataset. It is quite evident from the figure that the proposed framework was able to identify high percentage of top hits from the filtered dataset by performing less number of docking calculations. Conclusively we can say that applying such rule based filter increases the efficiency of the proposed framework.

3. Comparison with Deep Docking

Recently, Gentile et al. proposed a deep learning based method "Deep Docking" to augment the process of SBDD.[1] In this section, the MEMES framework performance was compared with Deep Docking. To effectively compare the performance of both methods the computational budget of docking was kept 6% of the docking library size. The intersection of the top molecules identified by the respective models with the top hits identified by docking the whole dataset is used as a metric for comparison.

Table S5 shows the intersection of top 500 molecules identified by DeepMEMES framework (CDDD) and Deep Docking method with actual top hits from the Zinc-250K dataset. Time taken by both methods is also compared in table S6. Conclusively it can be said that DeepMEMES has a superior performance over deep docking in the terms of identifying top hits and have almost same performance as that of deep docking in terms of computational time taken. Table S7 compares performance in the terms of identifying top hits from Enamine dataset that contains approximately 2 million compounds against target protein TTBK1.

4. Analysis of time taken by MEMES

In this section, the time taken by different steps of the proposed method is reported and is compared against the time taken to perform docking of all the compounds in the library. For this analysis, DeepMEMES was chosen as the method and Zinc-250K as the docking library. We used 1 Nvidia GeForce RTX 2080 TI for GPU computations and 24 Intel(R) Xeon(R) CPU E5-2640 v4 @ 2.40GHz for CPU computations.

The time reported in the table S6 is the computational hours taken by different steps of proposed method. The clustering and training step requires GPU computation, and CPU computation is required in embedding and docking calculations. It is quite evident from the table that the most of the time taken in by the docking calculation. The proposed method was able to reduce the computational cost by approximately 10 times.

Supplementary Methods

Docking Methodology

Ligand Preparation

Structure of the ligands was obtained from docking library and converted to the pdb format for AutoDock. The conversion was carried out by OpenBabel 2.3.1 and the energy minimization was done using a MMFF94 forcefield.[2]Preparation of the selected ligands for docking was carried out using AutoDock 4.2 (AD 4).[3]

Protein Preparation

Target protein, Tau-Tubulin Kinase 1 (PDB ID:4BTK) and SARS-CoV-2 Mpro (PDB ID:6LU7) with resolution 2.00 Å and 2.16 Å respectively were obtained from Research Collaboratory for Structural Bioinformatics – Protein Data Bank (RCSB – PDB).[4] All bound waters and cofactors were removed from the protein manually, Kolmann charges were computed, polar

hydrogen atoms were subsequently added and the AutoDock atom types were defined using AutoDockTools (ADT), Graphical User Interface (GUI) of AutoDock implemented in Molecular Graphics Laboratory (MGL) Tools.[3]

Grid Generation and Docking

Docking methodologies are utilized at initial stages of the drug discovery process to swiftly determine fitting molecules that could act as potential leads against a desired protein.[5] This methodology further gains significant prominence in circumstances when new targets emerge for which hits are ascertained. Docking protocols aid in elucidation of the most energetically favorable binding pose of a ligand to a receptor by ranking the ligands based on their estimated binding energy. The objective of our current docking study is to establish the crucial interactions responsible for inhibition of SARS-CoV-2 Mpro and Tau-Tubulin Kinase 1. Besides, to give way for development of potential leads capable of halting the coronavirus outbreak.

Docking was done using AD4, implementing Lamarckian Genetic Algorithm (LGA). LGA has enhanced performance compared to simulated annealing or the simple genetic algorithm, the other search algorithms available in AutoDock4 for ligand conformational searching.[6] Grid maps were generated for each atom type along with electron density maps and desolvation maps using the Autogrid4 utility in AutoDock. The grid spacing was changed from 0.375 Å to 0.600 Å and the size of the docking grid was fixed at 40 Å x40 Å x40 Å for 4BTK while for 6LU7 thegrid spacing was changed from 0.375 Å to 0.500 Å and the size of the docking grid was fixed at 40 Å x40 Å x40 Å. The X, Y, and Z coordinates of the grid box for 4BTK were fixed at 61.955,18.155,24.305 respectively while for 6LU7 they were fixed at -10.063, 16.667, 67.294 respectively, thus encompassing the active site. The active site of receptor was kept rigid and docking was carried out. The docking parameters for AD4 were kept at their default values. The 10 independent GA runs from AD4 were processed using the built-in clustering analysis with a 2.0 Å cutoff.

Choice of protein receptors

The massive growth and improvement in the field of healthcare over the years has halted the advent of fatal diseases. However, there are a class of diseases which have been going against the tide and have become more prevalent over the years; neurodegenerative diseases. In lieu of this, there is a desperate need of effective treatments/drugs that could combat them by blocking or delaying the progressive loss of the neurons.[7] In many neurodegenerative diseases, a notable feature is the presence of neurofibrillary tangles (NFT), the chief component of which is hyperphosphorylated tau protein (pTau). The presence of this pTau in NFT hints towards an imbalance between tau kinase and phosphatase activity. Among many kinases that are involved in phosphorylation of Tau, glycogen synthase kinase 3-β (GSK3β) and cycline-dependent kinase 5 (Cdk5) have been suspected to play an intricate and major role in phosphorylation of tau in the brain.[8] Substantial efforts have been made to develop their inhibitors, however due to their ubiquitous expression, there is a possibility of some serious side effects[9] . Thus, brain-specific tau kinases like tau-tubulin kinase 1 (TTBK1) are an attractive target to combat a myriad of neurodegenerative diseases.[10] Very few TTBK 1 inhibitors have been reported till date and consequently we selected TTBK1 as protein of choice.

At the time of writing this article, more than 4.7 million people across the globe have been infected by COVID-19 and there is still no available treatment/drug which has been able to curb the outbreak.[11] Among the drug targets available to combat SARS-CoV-2 like RNA-dependent RNA-polymerase and spike protein, the main protease (M pro) has been most extensively characterized.[12] M pro is a critical enzyme which plays a vital role in mediating viral transcription and replication of coronaviruses. Over and above, the absence of closely related homologues in humans, coupled with an urgent need to find a suitable small molecule inhibitor, made this target an attractive choice.[12] Present study was aimed to establish scientific insights while designing small molecule inhibitors and also plausible inhibitors targeting TTBK1 and SARS-COV-2 M pro proteins. Figure S13 shows the protein-

ligand complex and Figure S14 shows the prtoein-ligand interaction Ligplot, respectively, for target receptor (both TTBK1 and SARS-CoV-2 M^{pro}) and best ligand from Zinc-250K dataset.

References

(1) Gentile, F.; Agrawal, V.; Hsing, M.; Ton, A.-T.; Ban, F.; Norinder, U.; Gleave, M. E.; Cherkasov, A. Deep Docking: A Deep Learning Platform for Augmentation of Structure Based Drug Discovery. **2020**,

(2) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3*, 33.

(3) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(4) RCSB. `https://www.rcsb.org.`, rcsb.

(5) Sethi, A.; Joshi, K.; Sasikala, K.; Alvala, M. *Drug Discovery Dev. New Adv*; IntechOpen, 2019.

(6) Venkatesan, S. K.; Shukla, A. K.; Dubey, V. K. Molecular docking studies of selected tricyclic and quinone derivatives on trypanothione reductase of Leishmania infantum. *J. Comput. Chem.* **2010**, *31*, 2463–2475.

(7) Gitler, A. D.; Dhillon, P.; Shorter, J. Neurodegenerative disease: models, mechanisms, and a new hope. *Dis. Models Mech.* **2017**, *10*, 499 – 502.

(8) Ikezu, S.; Ikezu, T. Tau-tubulin kinase. *Front. Mol. Neurosci.* **2014**, *7*, 33.

(9) Brunden, K. R.; Ballatore, C.; Crowe, A.; Smith III, A. B.; Lee, V. M.-Y.; Trojanowski, J. Q. Tau-directed drug discovery for Alzheimer's disease and related tauopathies: a focus on tau assembly inhibitors. *Exp. Neurol.* **2010**, *223*, 304–310.

(10) Sato, S.; Cerny, R. L.; Buescher, J. L.; Ikezu, T. Tau-tubulin kinase 1 (TTBK1), a neuron-specific tau kinase candidate, is involved in tau phosphorylation and aggregation. *J. Neurochem.* **2006**, *98*, 1573–1584.

(11) WHO. `https://www.who.int/emergencies/diseases/novel-coronavirus-2019`.

(12) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **2020**, *368*, 409–412.