Combination of fragment growing and feed forward fragment searches for lead optimization campaigns

Helena Martín Hernández[1]

Scientific director: Martí Municoy[2]

[1,2] Nostrum Biodiscovery, Address: C/ Baldiri Reixac, 10, Barcelona, 08028 Spain.

Abstract

Motivation: During the past years, the number of compounds needed to identify a suitable lead have continuously rivesn. In order to innovate and accelerate the drug discovery process, we present Feed Forward Frag (F3), a new algorithm for filtering external molecule databases combining iterative searches and fragment growing to identify promising leads for the drug discovery process. The method is based on our in-house software FragPELE, a tool for in-silico hit-to-lead drug design, capable of growing a fragment from a bound core, adapted from PELE which uses implicit models to speed up conformational explorations. F3 successfully analyses millions of molecules in a few hours and identifies the bound core to enable the growth of the query ligand.

Results: Results show that the F3 algorithm could be useful to identify novel hits for a system and successfully explore the conformational space of the system at each iteration with PELE and efficiently sample the re-arrangement of the protein as the fragment is grown.

Supplementary information: The code of the algorithm is available at GitHub.

1 Introduction

The drug discovery process is initiated when, because of a clinical condition, there is a need to develop suitable medical products. The initial stage of research aims to define a hypothesis where a variation of a metabolic pathway results in a therapeutic effect in a disease state. Later on, in the basic research phase, a suitable target is selected. During this phase, an extensive search is performed in order to find a drug-like molecule, which after an exhaustive validation, will progress into preclinical and, if successful, into clinical development to finally be marketed. (Hughes et al., 2011)

Computational-aided drug design (CADD) methods play a major role in accelerating and economizing the drug discovery process. (Ou-Yang et al., 2012). The goal in drug discovery is to identify candidate molecules with improved biological potency and physicochemical properties.

A drug target is a biomolecule involved in signaling or metabolic pathways that are specific to a disease process. Modulation of the biological function performed by these molecules for therapeutic applications could be achieved by inhibiting their function with competitive molecules whose binding affinity is greater than the natural ligand, inhibiting the molecular interactions with small molecules or activating functionally deregulated molecules in some diseases. (Kumar S, 2013)

However, the number of compounds needed to identify a suitable lead has continuously risen over the past years. To reduce attrition and improve the productivity of the process it is crucial to have the best quality and quantity of hit classes in order to improve the efficiency and reduce the time required for successful hit identification. (Hao et al., 2016) The fundamental issue in this process is the high failure rate in clinical trials and the increase of development costs caused by molecules with inadequate properties. (Colombo and Peretto, 2008) Therefore, molecular properties such as toxicity, excretion, absorption and distribution must be considered at early stages of drug design. (Lagorce et al., 2008).

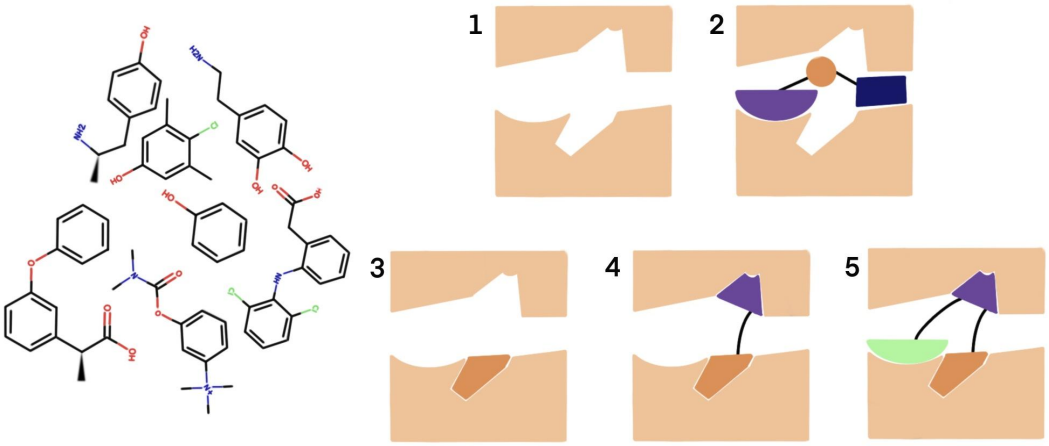


Figure 1: Schematic representation of HTS hits and FBDD hits as starting points for the drug discovery process. 1) *Cartoon representation of a protein active site where there are three pockets for inhibitors to bind.* ***2)*** *Cartoon representation of HTS hit binding to the active site. The hit is functionally complex and makes several but low-quality interactions with the protein.* ***3)*** *Cartoon representation of a fragment making few but highly efficient interactions with one of the pockets in the active site.* ***4, 5)*** *The lead has evolved by linking other fragments to it and makes further highly efficient interactions in the active site while retaining the key interactions from the first fragment.*

Furthermore, during the drug discovery process, there is an enormous chemical space to explore.

Therefore, screening approaches for assessing good ligand candidates require advanced techniques since small structural changes may result in huge differences in the interaction pattern between the ligand and the protein. The vast search space combined with the complexity of ranking each candidate easily leads to very challenging studies that require strategies to simplify and rationalize them. For example, in ligand-based methods, similarity and dissimilarity of previously known ligands is used to predict activity, develop molecular descriptors, quantitative structure-activity relationships, etc. to rationalize, reduce, and diversify the chemical space that is explored. (Sliwoski et al., 2013)

High-throughput screening (HTS) is the use of automated equipment to rapidly test millions of samples for biological activity at an organism, cellular, pathway or molecular level. HTS is commonly used in pharmaceutical and biotechnology companies to identify compounds with pharmacological or biological activity. (Wexler et al., 2014)

Due to their speed and low cost, computational models are an interesting approach to perform in-silico HTS on thousands or millions of compounds. During the drug discovery process, the industry has traditionally built large collections of highly functionalized compounds in an attempt to identify a sufficient number of hits, but even the largest conceivable compound collection falls far short of potential chemical diversity space. As molecular size decreases, the number of possible molecules decreases exponentially, so in theory it would be more efficient to screen collections of very small molecules and subsequently expand and merge them. (Erlanson, McDowell and O'Brien, 2004)

Fragment-based drug discovery (FBDD) has arisen as an alternative to HTS. (Murray and Rees, 2009) Fragments are small molecules that present low binding affinities against the target protein and therefore, would not be identified during a HTS strategy. However, once a hit has been identified, the attachment of small fragments to the molecular core can provide derivatives that might make further high-efficiency interactions in the active site while retaining the key interactions from the original ligand, thereby boosting the potency. **Figure 1** schematically represents the FBDD approach. In order to identify the best fragments and attaching points, a large library of small fragments with a diverse set of chemical scaffolds needs to be screened. (Fattori, 2004) To properly study and rank each fragment we also need to explore the conformational space of the small molecule and the protein residues that surround it in the binding site. (Congreve, Murray and Blundell, 2005) (Carr et al., 2005)

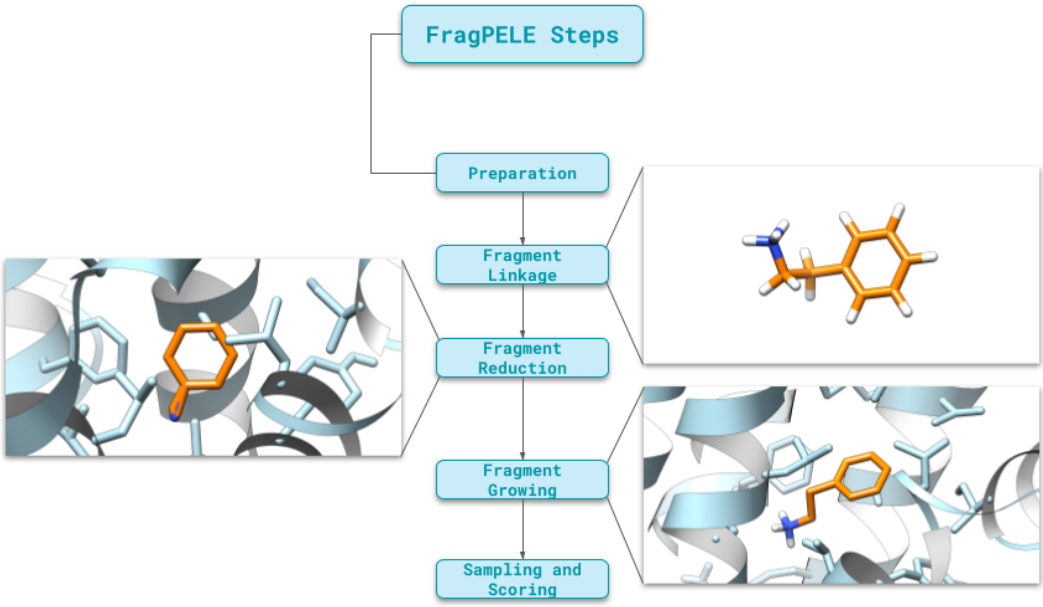


Figure 2: FragPELE simulation workflow.

In this scenario, we decided to utilise PELE, Protein Energy Landscape Exploration, as it is one of the leading algorithms to rapidly map protein–ligand interactions. (Borrelli et al., 2005) PELE relies on a Monte Carlo (MC) stochastic approach and a collection of algorithms that account for the flexibility of the protein, including the backbone and side chains, and small molecules. Its algorithm has been largely used in explorations of ligand binding and migration processes but also as a successful induced-fit docking approach.

In order to evolve the hits by merging other fragments, we used FragPELE, an algorithm that automatically grows one or more fragments onto different hydrogens of a ligand in multiple steps. (Perez et al., 2020) In order to allow the environment to adapt to the new fragment, it applies a slow-growing scheme based on running successive PELE simulations following an alchemical pathway; ie, the atoms of the fragment are shrinked in the first steps and then, progressively increased until reaching the regular atomic parameters.We have recently developed a new workflow for FragPELE in an attempt to ease and automatize the search, ranking and validation of new molecular fragments coming from external datasets. Specifically, our method takes as input hit compounds and looks for similar molecules from an external database, such as ZINC, in order to suggest fragments that can be incorporated to them in a FBDD routine. Then, employing FragPELE we can accordingly rank the resulting fragments to see whether the potency of our initial hit is improved with each of them or not. We can also concatenate multiple iterations of this algorithm to further increase the potency of our hit by adding multiple fragments on it.

Define objectives

The goal of the project is to develop an algorithm, working with PELE and FragPELE software, aiming to be a fast tool to locate and grow fragments to an initial seed compound from a protein-ligand complex to find suitable leads during the drug discovery process.

2 Methods

■ **Protein preparation:** The workflow requires a crystallographic structure corresponding to the hit already bound to the protein target. In order to use this complex structure in a FragPELE simulation, we need to represent it with a PDB that must have a proper format. Missing atoms and residues are handled with Schrodinger's Protein Preparation Wizard. (Madhavi Sastry et al., 2013) This tool is also used to fix wrong atom and residue names, protonate the system and remove water molecule

5A aways from the binding site. It is mandatory to ensure that the non-standard ligand has a unique chain ID and unique atom names. Non-standard residues are parameterized with OPLS2005.
Solvent is considered using the surface-generalized Born implicit model with a variable dielectric parameter.

■ **PELE:** PELE explores the conformational space of a biochemical system using a custom MC algorithm that relies on the following steps that can be executed in a multi-CPU infrastructure:

(1) **Perturbation:** PELE performs the local perturbation of the ligand. It can also perturb the protein backbone applying the normal modes found with the Anisotropic Network Model. If any contacts are found between the ligand and the backbone of the protein or with the ligand itself, the perturbation is directly rejected. Clashes between atoms that belong to rotatable groups or side chains are permitted since they can be relieved in the next steps.

(2) **Relaxation:** This step involves a side chain prediction algorithm and a global minimization. Firstly, bad contacts between the ligand and neighboring side chains are relieved by finding the best combination of rotamers. Finally, the global minimization applies a general relaxation to all the atoms of the system.

(3) **Acceptance or rejection of the step:** the new state of the system is only accepted if it fulfills the Metropolis criterion.

Recently, the exploration capabilities of PELE were improved with a protocol called Adaptive-PELE. (Lecina, Gilabert and Guallar, 2017) It relies on running short PELE simulations while combining them with a clustering/spawning algorithm to promote the exploration in the regions that have been less-explored in the previous runs.

■ **FragPELE:** FragPELE is based on the PELE software, it grows an R-group to a bound ligand in a series of steps, and is the basis of our approach.
As shown in **Figure 2**, the steps followed on a FragPELE simulation are:

(1) **Fragment Linkage:** At this stage, a given fragment is covalently linked to the chemical core at a position specified by the user.

(2) **Fragment Reduction:** The parameters of the atoms of the fragment are reduced to later be grown again within the binding site.

(3) **Fragment Growing:** during a series of steps, the fragment is grown iteratively increasing its parameters. This strategy along with the algorithms of PELE that account for the flexibility of both the ligand and the protein allows the binding site to adapt to the new chemical structure.

(4) **Sampling and scoring:** In this step, a PELE simulation is performed to score the grown molecule.

The combination of MC sampling with the growing algorithm used in FragPELE allows the complex to adapt while exploring the significant areas of the potential energy surface. FragPELE is based on the PELE software, which combines an MC stochastic approach with protein structure prediction techniques.
Several PELE simulations are run at each step to efficiently sample the re-arrangement of the system as the fragment is grown.

■ **Similarity search:** The similarity search within the algorithm uses the Tanimoto coefficient to calculate the degree of similarity between the query compound and the target structures.
The Tanimoto coefficient gives a value in the range of zero to one, one being the maximum similarity value, that is, the two molecules are equal. (Willett, 2006)
Thus, the Tanimoto coefficient takes the ratio of intersection over union, measuring the similarity between finite sample sets, being the molecular fingerprints in this project. It is defined as the size of the intersection divided by the size of the union of the fingerprints.

3 Results and Discussion

■ **The Feed Forward Frag algorithm:** Feed Forward Frag (F3) has been implemented as a new functionality of FragPELE. Our aim was to design a workflow able to locate, in an external dataset, new compounds similar to the small molecule (seed compound) that is already a known binder to a particular target. These new compounds must include the full molecular structure of the initial hit and only differ on the terminal molecular branches since it is important that we identify new molecules that can preserve the key interactions that the hit compound already has. New terminal branches will then be added to the ligand simulating its progressive growth with FragPELE. Finally, the new molecular candidates are ranked according to an energetic score, so we can identify

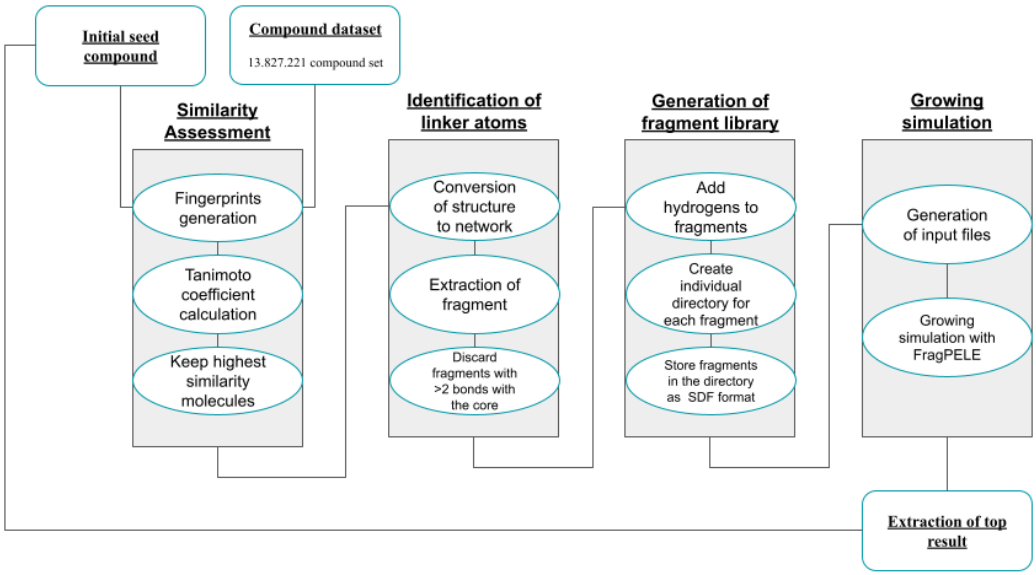


Figure 3: Diagram of the Feed Forward Frag algorithm steps.

those that could interact stronger with the protein cavity.

The illustration in **Figure 3** exhibits the main stages of the algorithm.

1. *Similarity Assessment:* The procedure starts with the computation of molecular fingerprints to encode the structure of the molecules and determine the similarity with the initial seed compound and the dataset compounds.
   Molecular fingerprints are representations of chemical structures that facilitate searching a substructure in a database and are used in multiple tasks such as similarity analysis, clustering and classification. (Hert et al., 2005)
   In this project we used extended connectivity fingerprints for similarity assessment during the filtering process.
   To collect the similarity value for each molecule without consuming the memory available, we used a heap queue algorithm. We
   used min heap algorithms to keep a list in a nearly sorted order as items are inserted and extracted. The nearly sorted order supported by heap queue algorithms is good for implementing priority queues but faster to establish and maintain than fully sorted data structures. (Martelli, 2017)
   Since FragPELE is only capable of growing one fragment at a time, we filtered out the molecules that had more than one R-group bound to the substructure by performing a substructure search of the query ligand in the database compound. The substructure search was also used to identify the core and the fragment of every potential hit.
2. *Identification of linker atoms:* FragPELE requires the information about the linker atoms from both the core and the fragment structures in order to attach them properly. To correctly identify the linker atoms, the structure of every molecule was represented as a network using the indexes of the core and the fragment obtained from the substructure search. Since linker atoms are identified through their PDB atom names., we aligned the core of the database compound with the query ligand to obtain the PDB atom name of the linker atom in the initial structure.
3. *Generation of the fragment library:* For each of the most similar compounds, the fragment extracted from the core in the previous step was stored as SDF format into an individual directory to create a library of fragments which FragPELE takes as input to run the growing simulation.
4. *Growing Simulation:* The last step consists of linking each one of the fragments of the fragment library to the query ligand with FragPELE.

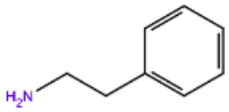
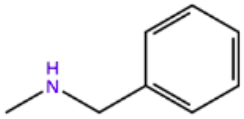
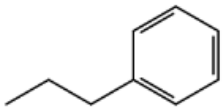
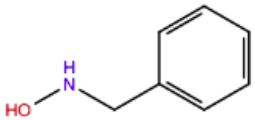
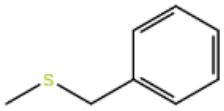
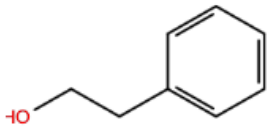
	Fragment	BE		Fragment	BE
1		-28.589	4		-28.089
2		-28.589	5		-28.074
3		-28.942	6		-27.048

Table 1: List of top results for T4 lysozyme (181L). *List of results with lowest binding energies from the first iteration of the F3 algorithm.*

5. *Sampling and Scoring:* Once the ligand is grown, a final PELE simulation is performed to score the molecule. The score is computed as the mean of the 25% lowest values of interaction energies.

■ Use Case: T4 Lysozyme

In order to study the capabilities of the F3 algorithm to suggest promising fragments that can be attached to the initial seed compound, along with the sampling potential of FragPELE, we applied our method to find new derivatives for the T4 Lysozyme. It is an endoacetyl-murmidase produced late in the infection of Escherichia coli by T4 bacteriophage (Madhavi et al., 2013). The structure consists of two domains with hydrolase activity, The interface between the two domains forms the active-site cleft.

A dataset of 13.827.221 compounds was analysed taking the ligand of the reference crystal structure (PDB id: 181L) as the initial seed compound, which is a benzene ring. A total of 38 compounds were selected according to their similarity with respect to the benzene. Then, extra chemical groups belonging to these compounds were grown and ranked with FragPELE. **Table 1** shows the best candidates according to the binding energy (BE) metric in the similarity assessment stage of the first iteration of the algorithm.

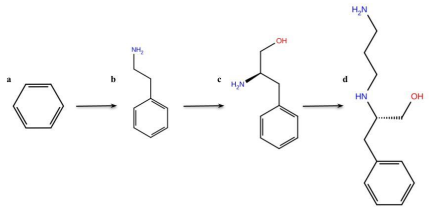


Figure 4: Optimization of query ligand into lead molecule d.

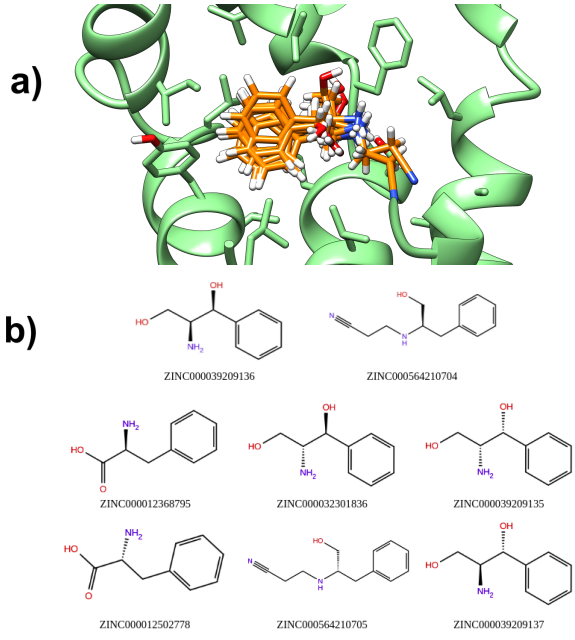


Figure 5: Structure of T4 lysozyme with overlap (a) of lowest BE results (b). *Showing a unique binding mode on the active site of the protein.*

A total of 3 iterations were performed. The best molecular candidates that were obtained after the full run are shown in **Figure 5b**. Thus, all the resulting structures contain a unique branch connected to the original benzene ring with some chemical groups able to establish hydrogen bonds. Moreover, results represented in **Figure 5a** shows that fragments occupied a similar binding mode after their growth with FragPELE. Particularly, all grown fragments are oriented towards the same cavity as seen in Figure 5. Besides, some of the polar groups in the grown branch could participate in new hydrogen bonds with the receptor. On the other hand, the same orientation predicted by our method is observed on crystallographic structures **2RBR, 2RB1, 2RAY** and **2OTZ** when similar fragments are attached to the benzene ring**,** as shown in **Figure 6**, thereby validating the binding modes founds by our FragPELE.

During the execution of the algorithm, we observed that as more iterations were performed, the BE values of the F3 results decreased, as shown in **Figure 8**. These results suggest that our method could improve the complementarity and specificity of the ligand towards the target after each iteration. In addition, since at each iteration the molecular weight of the initial seed compound increases, less results were obtained.

In the original protein-ligand system shown in **Figure 7a**, there are no interactions reported to be potent against the T4 lysozyme. The optimization of the initial seed compound (**Figure 4),** using the F3 algorithm, induces two new hydrogen bonds between **Val-111** and the hydroxyl group, and **Ser-117** and the amine group of the ligand as shown in **Figure 7b**. Moreover, in the native structure of the protein, Ser-117 and Asn-132 form a short hydrogen bond causing a steric strain. (Shoichet et al., 1995) However, the flexible induced-fit docking done by FragPELE is able to reorient Ser-117 to participate in a hydrogen bond with the ligand (**Figure 9**) in the majority of the top candidates that were obtained. In this way, this conformational change will probably improve the general stability of the protein.

In conclusion, the F3 method can be used not only to identify novel hits for a particular system, but also to open known pockets with new candidates of higher molecular weight suggested by the similarity-based searching algorithm. The conformational exploration power of the algorithms built in PELE provide the means to adapt the binding site to each of the molecules that are tried.

Regarding the computational time, the F3 algorithm takes 3 hours to scan the compound

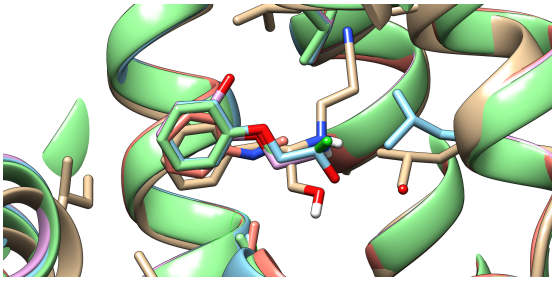


Figure 6: Overlap of structures 2RBR (blue), 2OTZ (red), 2RAY (green) and 2RB1 (purple) and F3 top result (brown).

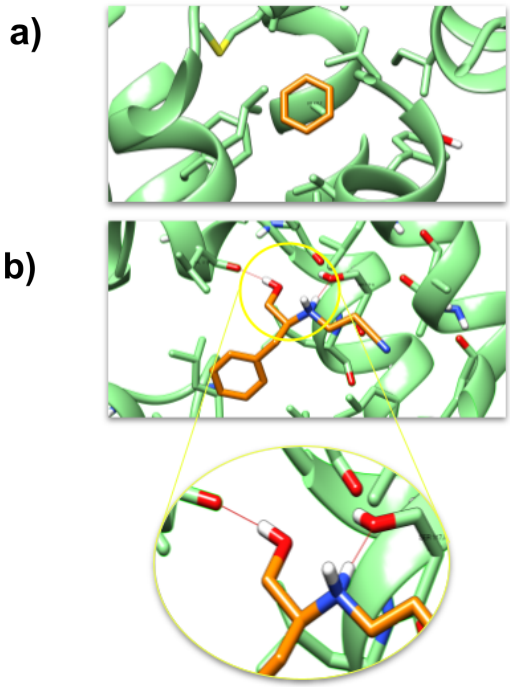


Figure 7: Interactions of the protein-ligand system before (a) and after (b) fragment optimization.

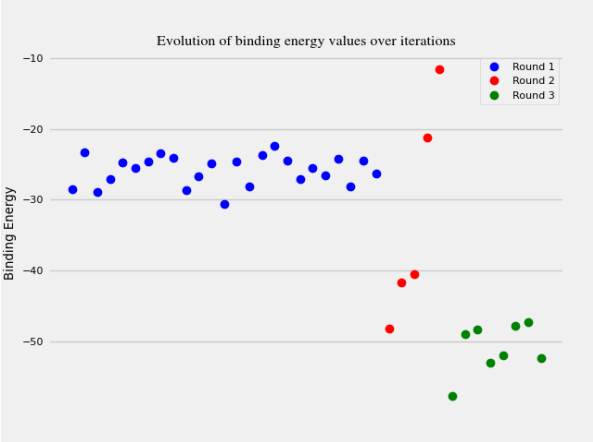


Figure 8: Evolution of BE values during F3 algorithm.

database and an average of 1 hour per fragment to run the fragPELE simulation using 20 CPUs on a Xeon Gold Processor.

4 Conclusions

FBDD has arisen as a powerful technique for identifying novel lead compounds. In this project, we have presented a novel algorithm to perform feed forward fragment searches to identify alternative lead compounds during the drug discovery process.

We tested the potential of the algorithm for finding suitable molecules and correctly identifying the linker atoms to efficiently perform the growing phase with FragPELE.

The early results of the F3 algorithm with the ZINC (Sterling and Irwin, 2015) database showed that it is a promising tool to identify new lead compounds being able to control the physical properties of the fragments and drug candidates.

Early structural results show that the F3 algorithm, along with PELE and FragPELE, show good identification of binding modes and improvement of the interactions in the protein-ligand complex.

However, further validation of the approach is needed. Some aspects to be improved are summarized in **Table 2**, such as the implementation of a cavity-wise growing that will allow the user to specify the cavity of the protein-ligand system that the new fragment must fill. This new feature is currently under development and will promote the adoption of a particular binding mode during the growing

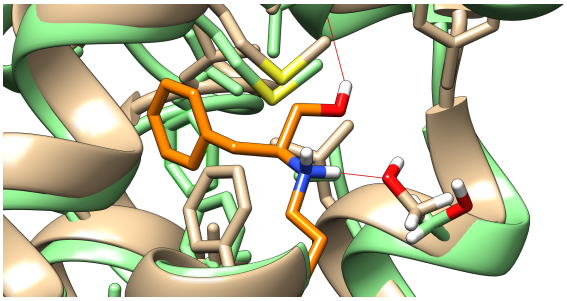


Figure 9: Serine 117 conformational shift.

simulation of new fragments while opening any cryptic pocket that the studied target might have.

Finally, the combination of AquaPELE (Municoy et al., 2020) and FragPELE to account for the perturbation of explicit water molecules during a growing simulation, could allow us to consider the mediation and displacement of buried water molecules. Therefore, we expect that the perturbation of water molecules at each iteration of the algorithm can boost the accuracy of the results on protein-ligand systems where interfacial water plays crucial roles.

Table 2. Aspects of improvement of the F3 algorithm.

	Improvement points for the F3 algorithm
1	For a given set of molecules, improve the assessment of the best result instead of only focusing on BE values
2	Parallelization of the similarity assessment step
3	Add pharmacophore analysis to filter out specific chemical groups
4	Include point analysis to choose fragments with a specific orientation inside the protein pocket

Acknowledgements

I would like to express my deepest appreciation to my scientific director, Martí Municoy, for providing guidance throughout the project and his support.

Furter, I would like to thank the rest of the members of the IT team, Anna Gruszka and Rosario Gago, without whom I would not have been able to complete this research.

Finally, I could not forget to thank Ezequiel Mas, for providing me the opportunity to work at Nostrum Biodiscovery.

References

(1) Hughes, J., S., Kalindjian, S. and Philpott, K. (2011). Principles of early drug discovery. British Journal of Pharmacology, 162(6), pp. 1239-1249.10.1111/j.1476-5381.2010.01127.x

(2) Ou-Yang, S., Lu, J., Kong, X., Liang, Z., Luo, C. and Jiang, H. (2012). Computational drug discovery. Acta Pharmacologica Sinica, [online] 33(9), pp. 1131–1140. 10.1038/aps.2012.109

(3) Kumar S, C.K.S. (2013). An Insight to Drug Designing by In Silico approach in Biomedical Research.

(4) Hao, G.-F., Jiang, W., Ye, Y.-N., Wu, F.-X., Zhu, X.-L., Guo, F.-B. and Yang, G.-F. (2016). ACFIS: a web server for fragment-based drug discovery. Nucleic Acids Research, pp.W550–W556. 10.1093/nar/gkw393

(5) Colombo, M. and Peretto, I. (2008). Chemistry strategies in early drug discovery: an overview of recent trends. Drug Discovery Today, [online] 13(15-16), pp.677–684. 10.1016/j.drudis.2008.03.007

(6) Lagorce, D., Sperandio, O., Galons, H., Miteva, M.A. and Villoutreix, B.O. (2008). FAF-Drugs2: Free ADME/tox filtering tool to assist drug discovery and chemical biology projects. BMC Bioinformatics. 10.1186/1471-2105-9-396

(7) Sliwoski, G., Kothiwale, S., Meiler, J. and Lowe, E.W. (2013). Computational Methods in Drug Discovery. Pharmacological Reviews. 10.1124/pr.112.007336

(8) Sommer, K., Flachsenberg, F. and Rarey, M. (2019a). NAOMInext – Synthetically feasible fragment growing in a structure-based design context. European Journal of Medicinal Chemistry. 10.1016/j.ejmech.2018.11.075

(9) Sterling, T. and Irwin, J.J. (2015). ZINC 15 – Ligand Discovery for Everyone. Journal of Chemical Information and Modeling, 55(11), pp.2324–2337. 10.1021/acs.jcim.5b00559

(10) Cereto-Massagué, A., Ojeda, M.J., Valls, C., Mulero, M., Garcia-Vallvé, S. and Pujadas, G. (2015). Molecular fingerprint similarity search in virtual screening. Methods. 10.1016/j.ymeth.2014.08.005

(11)Perez, C., Soler, D., Soliva, R. and Guallar, V. (2020). FragPELE: Dynamic Ligand Growing within a Binding Site. A Novel Tool for Hit-To-Lead Drug Design. Journal of Chemical Information and Modeling, 60(3), pp.1728–1736. 10.1021/acs.jcim.9b00938

(12) Montaruli, M., Alberga, D., Ciriaco, F., Trisciuzzi, D., Tondo, A.R., Mangiatordi, G.F. and Nicolotti, O. (2019). Accelerating Drug Discovery by Early Protein Drug Target Prediction Based on a Multi-Fingerprint Similarity Search Molecules. 10.3390/molecules24122233

(13) Hert, J., Willett, P., Wilton, D.J., Acklin, P., Azzaoui, K., Jacoby, E. and Schuffenhauer, A. (2005). Enhancing the Effectiveness of Similarity-Based Virtual Screening Using Nearest-Neighbor Information. Journal of Medicinal Chemistry, 48(22), pp.7049–7054. 10.1021/jm050316n

(14) Rogers, D. and Hahn, M. (2010). Extended-Connectivity Fingerprints. Journal of Chemical Information and Modeling, 50(5), pp.742–754. 10.1021/ci100050t

(15) Willett, P. (2006). Similarity-based virtual screening using 2D fingerprints. Drug Discovery Today, 11(23-24), pp.1046–1053. 10.1016/j.drudis.2006.10.005

(16) Murray, C.W. and Rees, D.C. (2009). The rise of fragment-based drug discovery. Nature Chemistry, 1(3), pp.187–192. 10.1038/nchem.217

(17) Congreve, M., Murray, C.W. and Blundell, T.L. (2005). Keynote review: Structural biology and drug discovery. Drug Discovery Today, 10(13), pp.895–907. 10.1016/s1359-6446(05)03511-7

(18) Fattori, D. (2004). Molecular recognition: the fragment approach in lead generation. Drug Discovery Today, 9(5), pp.229–238. 10.1016/s1359-6446(03)03007-1

(19) Carr, R.A.E., Congreve, M., Murray, C.W. and Rees, D.C. (2005). Fragment-based lead discovery: leads by design. Drug Discovery Today, 10(14), pp.987–992. 10.1016/s1359-6446(05)03511-7

(20) Wexler, P. (2014). Encyclopedia of toxicology. Amsterdam: Academic Press.

(21) Martelli,A ., Ravenscroft, A,. Holden, S. *Python in a Nutshell. A Desktop Quick Reference* ISBN: 9781449392925

(22) Erlanson, D.A., McDowell, R.S. and O'Brien, T. (2004). Fragment-Based Drug Discovery. Journal of Medicinal Chemistry, 47(14), pp.3463–3482. 10.1021/jm040031v

(23) Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. and Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. Journal of Computer-Aided Molecular Design, [online] 27(3), pp.221–234. 10.1007/s10822-013-9644-8

(24) Weaver, L.H. and Matthews, B.W. (1987). Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. Journal of Molecular Biology, [online] 193(1), pp.189–199. 10.1016/0022-2836(87)90636-x

(25) PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. 10.1021/ct0501811

(26) Shoichet, B.K., Baase, W.A., Kuroki, R. and Matthews, B.W. (1995). A relationship between protein stability and protein function. Proceedings of the National Academy of Sciences 10.1073/pnas.92.2.452

(27) Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. and Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. Journal of Computer-Aided Molecular Design, [online] 27(3), pp.221–234. 10.1007/s10822-013-9644-8

(28) Lecina, D., Gilabert, J.F. and Guallar, V. (2017). Adaptive simulations, towards interactive protein-ligand modeling. Scientific Reports, 7(1) 10.1038/s41598-017-08445-5

(29) Municoy, M., Roda, S., Soler, D., Soutullo, A. and Guallar, V. (2020). aquaPELE: A Monte Carlo-Based Algorithm to Sample the Effects of Buried Water Molecules in Proteins. Journal of Chemical Theory and Computation, 16(12), pp.7655–7670 10.1021/acs.jctc.0c00925