Research Article

Fragment-based virtual screening to discover potential *Plasmodium* PI4KIIIβ ligands

Akachuwku Ibezim[1]*, Mbanefo S. Madukaife [2], Sochi C. Osigwe[1], Nadja Engel[3], Ramanathan Karuppasamy[4], and Fidele Ntie-Kang [5,6,*]

[1] Department of Pharmaceutical and Medicinal Chemistry, University of Nigeria, Nsukka; <u>akachukwu.ibezim@unn.edu.ng</u>; <u>sochimaosigwe@yahoo.com</u>; <u>ngozi.nwodo@unn.edu.ng</u>

[2] Department of Statistics, University of Nigeria, Nsukka;<u>mbanefo.madukaife@unn.edu.ng</u>

[3] Department of Oral, Maxillofacial and Plastic Surger, Rostock University Medical Center, Rostock, Germany; <u>nadja.engel@med.uni-rostock.de</u>

[4] Department of Biotechnology, Vellore Institute of Technology, Vellore - 632014, Tamilnadu, India; <u>kramanathan@vit.ac.in</u>

[5] Chemical Bioactivity Information Centre, Department of Chemistry, University of Buea, Cameroun; <u>fidele.ntie-kang@ubuea.cm</u>

[6] Institute of Pharmacy, Martin-Luther University Halle-Wittenberg, Halle (Saale), Germany; <u>fidele.ntie-kang@pharmazie.uni-halle.de</u>

* Correspondence: <u>akachukwu.ibezim@unn.edu.ng</u> (AI) and <u>fidele.ntie-kang@ubuea.cm</u> (FNK).

Abstract: *Plasmodium* species that cause malaria, a disease responsible for about half a million deaths per annum despite concerted efforts to combat it. The causative agent depends on type III beta phosphatidylinositol 4-kinase (*P*PI4K) during the development of merozoite. *P*PI4K is the only clinically validated *Plasmodium* kinase so far and its inhibitors are effective both *in vitro* and *in vivo*. In this work, a small library of ~22 000 fragments was virtually screened using *P*PI4K homology model to discover potential ligands of the enzyme. 16 virtual hits were selected based on ≤ -9.0 kcal/mol binding energy cut off and were subjected to similarity and substructure searching after they had passed PAINS screening. The derivatives obtained showed improved binding energies, which ranged from -10.00 to -13.80 kcal/mol. Moreover, the topmost ranking compound **31**, with interesting drug-like quality was stable within the protein's binding cavity during the 10 ns molecular dynamics simulation period. In addition, analysis of its binding pose revealed some unique binding interactions with *P*PI4K active site residues as the basis for the observed improved binding affinity. Overall, compound **31** appears to be a viable starting point for the development of *P*PI4K inhibitors with antimalarial activity.

Keywords: *Plasmodium*, malaria, type III beta phosphatidylinositol 4-kinase, virtual screening, homology modelling, molecular dynamics.

1. Introduction

Malaria is a fatal human parasitic disease caused by five species of *Plasmodium*, with *P. falciparum* as the deadliest species, and was responsible for more than 220 million clinical cases and about half a million deaths in 2018 alone [1-3]. Children under the age of five and pregnant women are most vulnerable and above 90% of the cases occurred in the sub-Saharan African region alone according to World Health Organization (WHO) 2019 report [4-5]. There has been progress in controlling the disease through drug therapies, the use of insecticides, and design of vaccines. However, it has persisted because of a plethora of reasons spanning from non-adherence to prescription by patients, emergence of *Plasmodium* drug-resistant strains to resistance to commonly used insecticides among the mosquito vector population [6-8]. Undoubtedly, the need for an effective antimalarial drug is urgent and many researchers tend to agree that targeting an enzyme

that is vital in all *Plasmodium* life cycle will do the job. An example of such a target is type III beta phosphatidylinositol 4-kinase (PI4K) [9-11]. PI4KIIIβ is a ubiquitous eukaryotic enzyme that plays an essential role in merozoite development through the regulation of intracellular signaling, cytokinesis, and membrane trafficking by catalyzing the conversion of phosphatidylinositol (PI) to phosphatidylinositol-4-phosphate (PI4P) [9-11]. Sternberg and Roepe [12] study identified *P*PI4K as an operational enzyme in all *Plasmodium* life-cycle stages and further provided genetic, chemical, and biochemical evidence to show that *P*PI4K is a suitable target for antimalarial drug and its inhibition can cure, prevent and block the transmission of the disease. In addition, the fact several scientists have virtually screened many small-molecule libraries [13-14] and chemical series such as aminopyridine/pyrazine series [15], imidazopyridazine series [16], and naphthyridine series [17] to discover *P*PI4K inhibitors underpins the relevance of this enzyme in antimalaria drug search. However, no study to the best of our knowledge has applied a fragment-based screening strategy to discover *P*PI4K ligands. Over the past two decades fragment-based drug discovery has demonstrated to be one of the most prominent alternatives to high throughput screening utilized in the identification of lead compounds in drug discovery. It begins with finding low fragments or low molecular weight compounds that bind with weak affinity to the target of interest. The fragments that form high quality interactions are maximized to lead compounds with high affinity and selectivity [18].

In the present study, a *P*PI4K homology model was constructed using PI4K of human crystal structure as a template, validated a molecular docking protocol using a series of PI4K inhibitors from PubMed, and applied it to screen a small library of fragment molecules. Pools of compounds from fragment hits were re-docked to investigate improved binding affinity and molecular dynamics simulation was applied to evaluate binding stability. The result of this study is capable of leading to new *P*PI4K inhibitor templates for designing novel antimalarial drugs.

2. Materials and Methods

2.1. Homology modelling

Since there was no X-ray crystal structure of *Plasmodium* PI4K (*P*PI4K), a model of it was constructed using its amino acid sequence which was taken from the NCBI website [19]. During BLAST, the target sequence was used to search RCSB Protein Data Bank to identify the template structure (PDB ID: 6GL3) [20]. The Modellerver 9.21 [21] was used to build the homology model and subsequently the selection of the homology model for this study was based on the assessment of the Modeller objective function and Discrete Optimization Protein Energy. PROCHECK [22] and Discovery Studio [23] programs were used to evaluate the quality of the best model.

2.2. Preparation of molecules and virtual screening

All the compounds used in this study (fragment and fragment-derivatives from the ZINC database [24]; inhibitors and non-inhibitors of PI4K from ChEMBL library) were energy minimized with Openbabel software.

The homology modeled *P*PI4K structure was prepared for molecular simulation purposes following the previous protocol described by Ibezim et al, 2017 [25]. Protonate 3D module and molecular mechanics MMFF94 force field [26] implemented in Molecular Operating Environment (MOE) [27] were respectively used to protonate the protein residues (with the pH set at 7.0 ± 2.0) and optimally orient the protein atoms to the lowest energy level (potential energy gradient of 10-15 kcal/mol).

To conduct molecular docking, first, the modelled protein was aligned to the template protein-ligand complex (PDB ID: 6GL3) to locate the binding site. Afterwards, a grid box measuring 42 x 22 x 20 points with 0.375 Å point spacing was generated around the protein active site and then both the inhibitors and non-inhibitors were docked to the grid using the AutoDock suite as described in our earlier work [28]. The AUC plot was generated from an R program. The validated virtual screening protocol was used to virtually screen the study compounds.

An online platform was used to filter the fragment hits against PAINS [29] while MACSS structural keys were used to assess their structural novelty. Candidates which passed both assessments were submitted for similarity and substructure analysis among compounds in the ZINC database.

2.3. Molecular dynamics simulation

The parameter file of ligand was generated by Automated Topology Builder (ATB) [30] whereas the molecular dynamics simulation was carried out by Gromacs ver4.5.5 using GROMOS96 (ffG53a6) force-field [31]. The protein-ligand complex was submerged into solvated cubic box simple point charge (SPC) comprising an explicit water model and sufficient sodium and chloride ions which serve to neutralize the system. Steepest-descent integrator and conjugate gradient algorithm were applied to perform energy minimization. The system was agitated for a period of 10 ns during which solvent molecules and ions were allowed to equilibrate around the solute molecule at 300 k and 1 bar while all the non-hydrogen atoms were subjected to position restraining force.

3. Results and Discussion

3.1. Construction of PPI4K homology model

The Plasmodium PI4K homology model was built using the protein sequence (sequence ID: KNG744841.1) with Modeler 9.21. Blast similarity search identified the human PI4K with an access ID: 6GL3 (Figure S1), as the suitable template for constructing the model since there, exists a sequence identity of 44.36% between them and the template structure (6GL3) has an inhibitor bound within the ATP-site which stretches beyond the ribose pocket to create ATP binding pocket with large conformation in the P-loops to accommodate molecules of varying structural sizes. After several models were built by the modeler program, the best model was selected based on the DOPE scoring function. PROCHECK identified that 77.0%, 19.70%, 2.60%, and 0.70% of the protein model residues are respectively located in the core, allowed, the general and disallowed region as shown by the Ramachandran plot in Figure 1Since the stereo-chemical quality of the homology model was satisfactory, it was then used in this study.
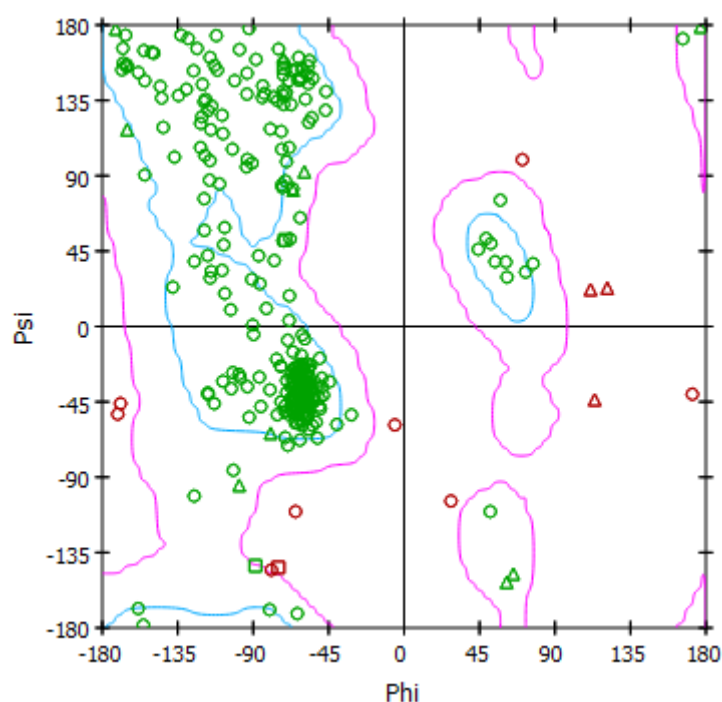
Figure 1. Ramachandran plot for the *Pf*PI4K receptor homology model.

3.2. Evaluation of virtual screening workflow

To ensure our virtual screening protocol is reliable, its ability to discriminate between PI4K actives and inactives were tested using a dataset of compounds extracted from the ChEMBL database in which compounds with IC50 < 10 μM are grouped as PI4K inhibitors and those with IC50 > 10 μM are grouped as non-inhibitors. The ratio of actives to inactives in the validation set was 1:5 (that is 30 out of 150 molecules). Figure 2 shows the enrichment curve and explains that the enrichment is better than random since according to the area under the ROC (AUR), the protocol demonstrated 82.53% overall prediction accuracy. Therefore, it can retrieve potential PI4K actives from a database of diverse molecules in a virtual screening scheme.
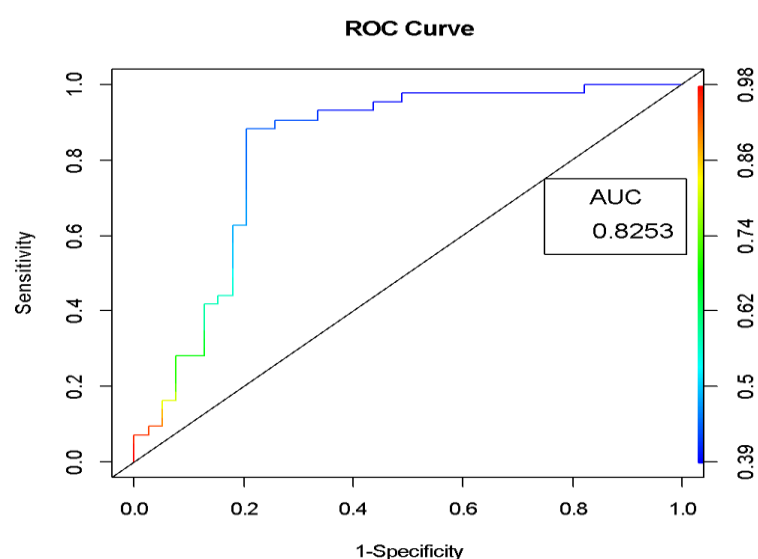
Figure 2. Assessment of the performance of virtual screening protocol: Area Under ROC curve.

3.3. Fragment-based virtual screening

A subset of 21 844 unique fragments was extracted from the ZINC database which is composed of 727 842 purchasable compounds and a total of over 34 million molecules. These fragments were virtually screened to identify new PI4K ligands. They showed binding energies which spanned from positive values to as low as -9.31 kcal/mol. Using binding energy ≤ -9.00 kcal/mol as cut-off, 16 best-ranked fragments were selected and submitted for pan-assay interference compounds (PAINS) filtration to identify frequent hits which often are promiscuous binders. Two molecules were caught by the filter while the remaining 14 fragments scaled through. To focus our study on molecules with unusual structural scaffolds different from already known PI4K inhibitors, Tanimoto coefficient calculation using MACCS structural fingerprints at the overlap of 60% to known PI4K ligands curated from the ChEMBL database was performed and resulted in seven fragments with 2 uncommon scaffolds (Figure 3). Through visual inspection, the two scaffolds were identified as isoquinoline and indazolone.
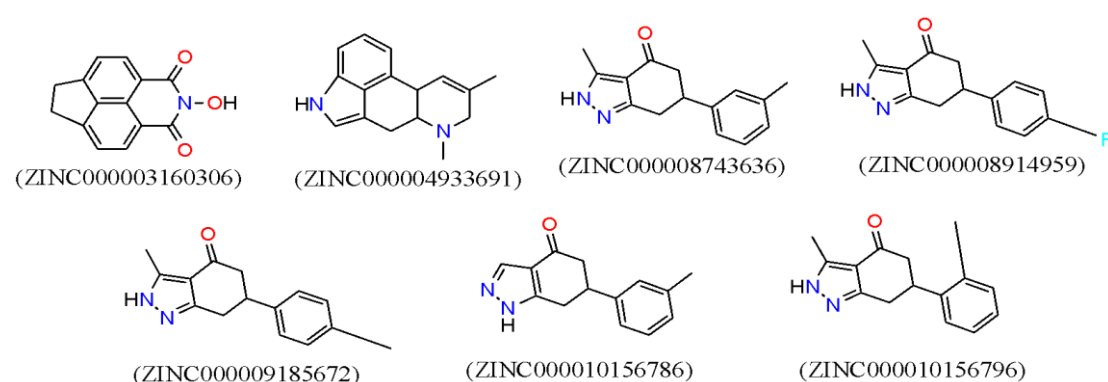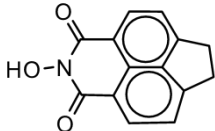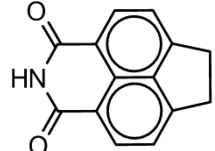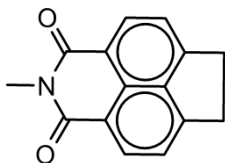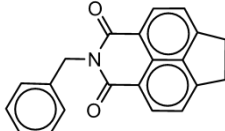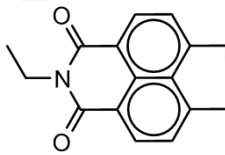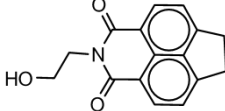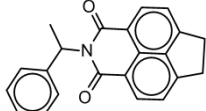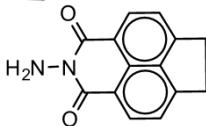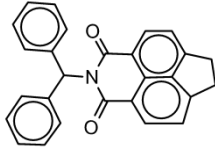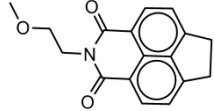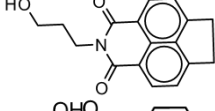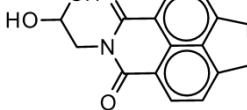

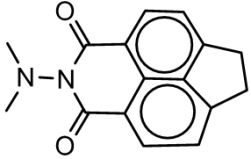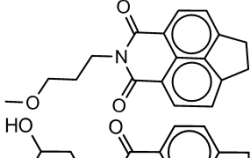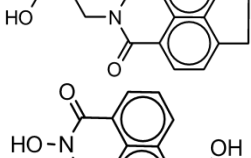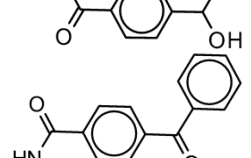**Figure 2.** The seven fragments with uncommon structural scaffolds from other known PI4K inhibitors.
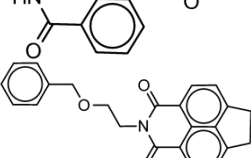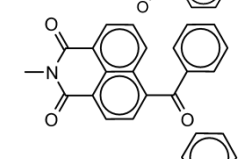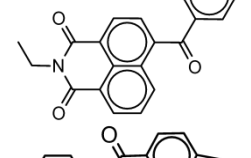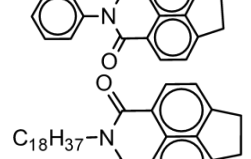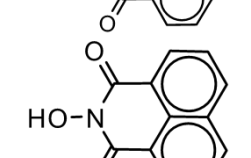
To explore ligand space of the uncommon scaffolds, they were used to conduct substructure and similarity search on the >34 000 000 molecules of the ZINC library. This led to 131 molecules that were docked into the *P*PI4K binding site. Although most of the derived molecules exhibited greater binding affinities than their starting materials, it appeared the protein preferred interaction with the indeno-isoquinolinedione and its derivatives since most of them ranked ≤ -11.00 kcal/mol (Table 1).

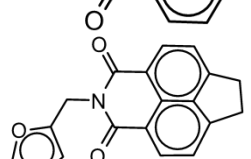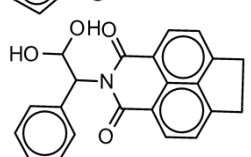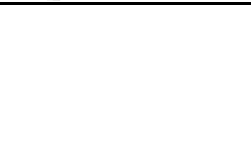There was no clear structure-activity relationship observed among the indeno-isoquinolinediones, however, it seems the N-2 position does not allow flexible moieties since all the derivatives (compounds **5**, **6**, **10**, **11**, **13**, **14**, **22**) with a flexible group at that position have relatively low scores, ≥ -10.00 kcal/mol. On the other hand, the presence of rigid group(s) and moieties capable of making electrostatic interaction at the N-2 and para positions of the two benzene rings tends to enhance binding affinity with PI4K.

That is probably the reason compound **29** has enhanced binding energy over those of **4** and **7** (Table 1). Of all the derivatives, compound **31** exhibited the highest binding affinity for *P*PI4K (-13.80 kcal/mol). In addition, results of the computed molecular descriptors (Table 2) suggest that this compound possesses attractive physicochemical properties. Hence, it is an appropriate candidate in this series worthy of further attention for PI4K inhibitor development.

Table 1. The derivatives of the best performing scaffold and their binding free energies

Symbol	Derivatives	Binding free energies (kcal/mol)
1		-11.50
2		-11.10
3		-11.10
4		-12.10
5		-10.90
6		-10.70
7		-12.30
8		-11.20
9		-11.40
10		-10.20
11		-10.40
12		-11.10

13		-10.90
14		-10.40
15		-11.10
16		-10.80
17		-12.60
18		-11.80
19		-12.50
20		-12.00
21		-12.20
22		-9.60
23		-10.30
24		-11.70
25		-11.50
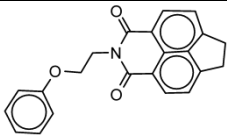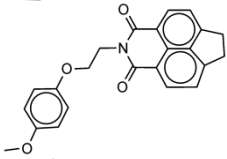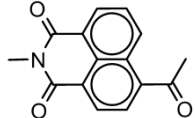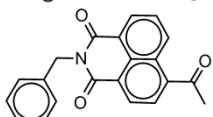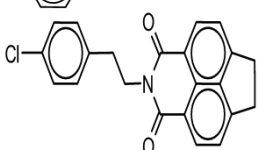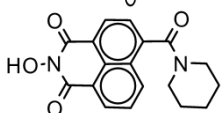
26		-11.80
27		-11.50
28		-12.20
29		-12.90
30		-12.20
31		-13.80

Table 2. Basic physical parameter of compound **31**.

Molecular descriptor	Value
MW	324.33
logP	2.45
HBA	6
HBD	1
NRB	2
TPSA	197.84
NHA	24

[1]TPSA = total polar surface area (Å^2), NRB = number of rotatable bonds, HBD = hydrogen bond acceptor, HBD = hydrogen bond donor, logP = lipophilicity, MW = molar weight (Da), NHA = number of heavy atoms.

Using MD simulations for a period of 10 ns, we investigated the stability of **31** in the PI4K binding pocket and the results showed that the compound maintained stable binding interaction with the protein's residues throughout the agitation time. According to Figure 4A and 4B respectively for the entire simulation period, compound **31** did not deviate more than ~2.0 Å away from the protein and the binding energy of the protein-ligand complex remained approximately the same. The pose of **31** (Figure 5A) retrieved after 9 ns of simulation revealed the compound fits tightly into PI4K binding pocket and Figure 5B showed the major intermolecular interactions which informed the observed stable binding of **31** to PI4K. Hydrogen bond was observed between compound **31** hydroxyl group and L15 backbone while carbonyl groups of the ligand made contact with the NH groups of W17 and N98 side chains. Overall, it appears polar bond is the main force behind the intercourse between the protein and ligand, even though the presence of tyrosine and tryptophan residues in the binding site suggests the pi-pi bond might have played a role.
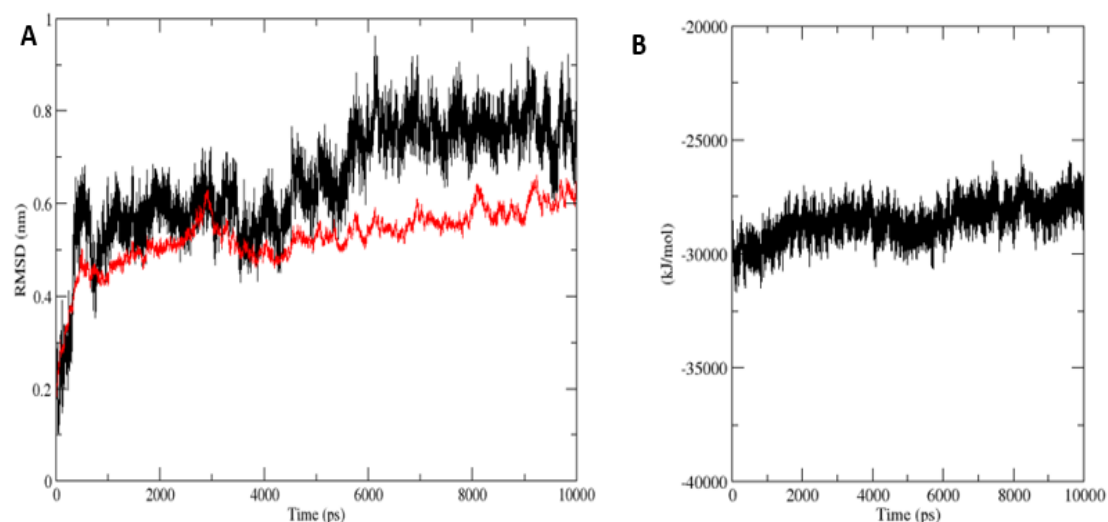
Figure 3. RMSD and energy plot for PfPI4K-**31** complex (A) The deviation of **31** (in red color) from the binding cavity of *P*PI4K Cα (in black color) (B) *P*PI4K-1HPD complex for 10 ns simulation in terms of binding free energy.
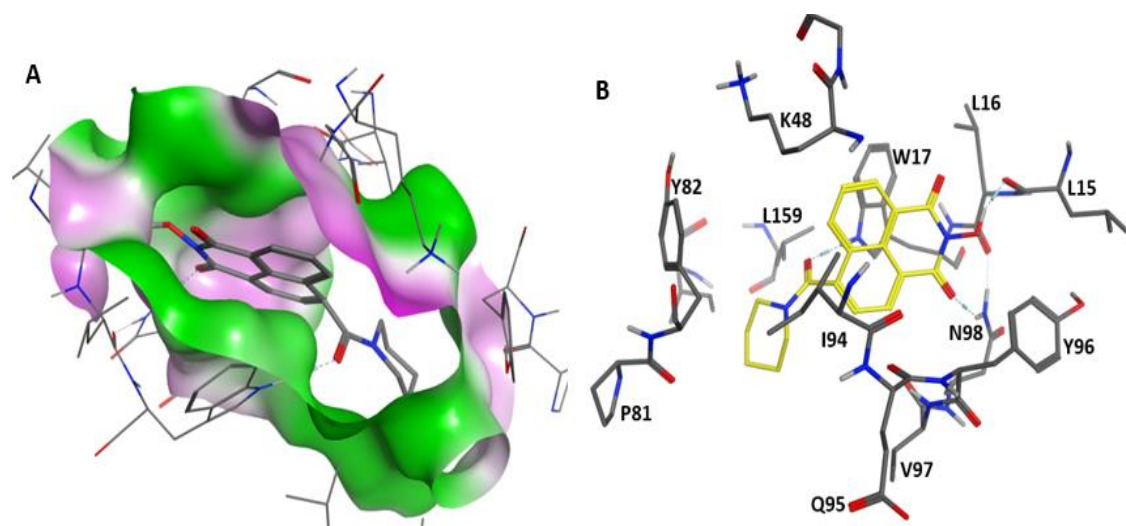
Figure 4. Pose of the best performing candidate in the *Pf*PI4K binding site (A) Surface map showing *Pf*PI4K binding cavity and the binding conformation of **31** in it. (B) Binding interaction of **31** with key residues of *P*PI4K.

5. Conclusions

A homology model of *P*PI4K based on the human PI4K template was built and a fragment-based virtual screen was performed to identify novel ligands of the protein. Two scaffolds uncommon among known PI4K ligands were identified amidst the fragment virtual hits. However, only the derivatives of one of them (indo) exhibited an interesting greater binding affinity. Importantly, the best-scored derivative has attractive physical features required for a drug candidate and was found to make a stable binding interaction with the *P*PI4K active site residues.

Author Contributions: A.I., S.O.C., M.S.M and N.E. did the formal analysis, investigation, data curation, provided administration, and wrote the original draft. R.K and FNK carried out the pro-

ject administration, validation of result and general supervision. A.I. and F.N.-K conceptualization and the final draft. The author(s) read and approved the final manuscript.

Data Availability Statement: Not applicable

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Onoabedje EA.; Ibezim A.; Okoro UC.; Sanjay Batra. Synthesis molecular docking, antiplasmodial and antioxidant activities of new sulfonamido-pepetide derivatives. *Heliyon*. **2020**, doi: 10.1016/j.heliyon. 2020.e04958

2. Ugwu DI.; Okoro UC.; Ukoha PO.; Okafor S.; Ibezim A.; Kumar NM.; Synthesis, characterization, molecular docking and *in vitro* antimalarial properties of new carboxamides bearing sulphonamide. *Eur. J. Med. Chem.* 2017; 135: 349-369, doi: 10.1016/j.ejmech.2017.04.029

3. Onoabedje EA.; Ibezim A.; Okoro UC.; Batra S.; New sulphonamidepyrrolidinecarboxamide derivatives: Synthesis, molecular docking, antiplasmodial and antioxidant activities. *PLoS ONE.***2021**; 16(2): doi: 10.1371/journal.pone.0243305

4. WHO Media Center, Fact Sheet on Malaria, **2020**. https://www.who.int/en/news-room/fact-sheets/detail/malaria. (Accessed 4 April 2020).

5. Ngo-Hanna J.; Ntie-Kang F.; Kaiser M.; Brun R.; Efange SMN, 1-Aryl-1,2,3,4-tetrahydroisoquinolines as potential antimalarials: synthesis, in vitro antiplasmodial activity and in silico pharmacokinetics evaluation, *RSC Adv.***2014**; 4:22856–22865**,** doi: 10.1039/C4QO00294F

6. De Rycker M.; Horn D.; Aldridge B.; AmewuRK.; Barry CE.; Buckner FS. Setting our sights on infectious diseases. *ACS Infect. Dis.* **2020**; 6:3–13, doi: 10.1021/acsinfecdis.9b00371

7. Uwimana A.; Legrand E.; Stokes BH.; Ndikumana JM.; Warsame M.; Umulisa N. Emergence and clonal expansion of in vitro artemisinin-resistant *Plasmodium falciparum* kelch13 R561H mutant parasites in Rwanda. *Nat. Med.***2020**; 26:1602, doi:10.1038/s41591-020-1005-2.

8. Bousema T.; Drakeley C. Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination. *Clin. Microbiol. Rev.* **2011**; 24:377–410, doi: 10.1128/CMR.00051-10.

9. McNamara CW.; Lee MCS.; Lim CS.; Lim SH.; Roland J.; Nagle A. Targeting *Plasmodium* PI(4)K to eliminate malaria. *Nature* **2013**;504:248-253, doi:10.1038/nature12782

10. Paquet T.; Le Manach C.; Cabrera DG.;Younis Y.; Henrich PP.; Abraham TS. Antimalarial efficacy of MMV390048, an inhibitor of *Plasmodium* phosphatidylinositol 4-kinase. *Sci. Transl. Med.* **2017**;9:9735, doi: 10.1126/scitranslmed .aad9735

11. Sternberg AR.; Roepe PD. Heterologous expression, purification, and functional analysis of the *Plasmodium falciparum* Phosphatidylinositol 4-Kinase IIIβ. *Biochemistry*. **2020**; 59:2494-2506, doi: 10.1021/acs.biochem.0c00259.

12. Fienberg S.; Eyermann CJ.; Arendse LB.; Basarab GS.; McPhail JA.; Burke JE.; Chibale K. Structural basis for inhibitor potency and selectivity of *Plasmodium falciparum* phosphatidylinositol 4-kinase inhibitors. *ACS Infect.* Dis.**2020**; doi:10.1021/acsinfecdis.0c00566.

13. Ibrahim MAA.; Abdelrahman AHM.; Hassan AMA. Identification of novel *Plasmodium falciparum* PI4KB inhibitors as potential anti-malarial drugs: Homology modeling, molecular docking and molecular dynamics simulations. *Comput. Biol. Chem*. **2019**; 80:79-89, doi: 10.1016/j.compbiolchem.2019.03.010

14. Younis Y.; Douelle F.; Feng TS.; Cabrera DG.; Manach CL.; Nchinda AT. 3,5-Diaryl-2-aminopyridines as a novel class of orally active antimalarials demonstrating single dose cure in mice and clinical candidate potential. *J. Med. Chem*. **2012**; 55:3479-3487, doi:10.1021/jm3001373

15. Le Manach C.; Gonzàlez-Cabrera D.; Douelle F.; Nchinda AT.; Younis Y.; Taylor D . Medicinal chemistry optimization of antiplasmodialimidazopyridazine hits from high throughput screening of a SoftFocus kinase library: Part 1. *J. Med. Chem*. **2014**; 57:2789-2798, doi: 10.1021/jm500098s.

16. Kandepedu N.; Gonzàlez-Cabrera D.; Eedubilli S.; Taylor D.; Brunschwig C.; Gibhard L. Identification, characterization, and optimization of 2,8-disubstituted-1,5-naphthyridines as novel *Plasmodium falciparum* phosphatidylinositol-4-kinase inhibitors with *in vivo* efficacy in a humanized mouse model of malaria. *J. Med. Chem*. **2018**; 61:5692-5703, doi: 10.1021/acs.jmedchem.8b00648.

17. National Center for Biotechnology Information (NCBI) (Internet). Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information: (1988) – (cited 2021 Apr 06). Available from: https://www.ncbi.nlm.nih.gov/

18. Kumar A.; Voet A.; Zhang KY. Fragment based drug design: from experimental to computational approaches. *Curr Med Chem*. **2012**;19(30):5128-47. doi: 10.2174/092986712803530467.

19. Berman HM.; Westbrook J.; Feng Z.; Gilliland G.; Bhat TN.; Weissig H.; Shindyalov IN.; Bourne PE. The protein data bank. Nucleic Acids Res. **2000**; 28:235–342.

20. Webb B.; Sali A. Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinform. **2016**; 54:561–567.

21. Laskowski RA.; MacArthur MW.; Moss DS.; Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*. **1993**; 26(2): 283-291, doi:10.1107/S0021889892009944

22. Discovery Studio Visualizer Software, Version 4.0, **2019**. http://www.accelrys.com

23. Irwin JJ.; Shoichet BK. ZINC – A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model*. **2005**; 45(1):177-182, doi:10.1021/ci049714+

24. Ibezim A.; Debnath B.; Ntie-Kang F.; MbahCJ.;Nwodo NJ. Binding of anti-*Trypanosoma* natural products from African flora against selected drug targets: a docking study. *Med. Chem. Res.***2017**; 26:562–579, doi:10.1007/s00044-016-1764-y

25. Halgren   TA.   Merck   molecular   forcefield*.   J   Comput   Chem*.   **1996**;   17(5-6):490–641, doi:10.1002/(SICI)1096-987X(199604)17:5/6%3C490::AID-JCC1%3E3.0.CO;2-P

26. Chemical Computing Group (2014) Molecular Operating Environment (MOE) software.

27. Ibezim A.; Onoabedje EA.; Adaka IC.; Omeje KO.; Onoabedje US.; Obi BC. Carboxamides bearing sulfonamide functionality as potential novel phospholipase A2 inhibitors. *ChemistrySelect*.**2020**; 5:14416– 14421, doi:10.1002/slct.202003784

28. Jonathan BB.; Georgian AH. New substructure filters for removal of Pan Assay Interference Compounds (PAINS) from screening libraries and for their exclusion in bioassay. *J. Med. Chem.***2010**; 53(7):2719-2740, doi:10.1021/jm901137j

29. Malde AK.; Zuo L.; Breeze M.; Stroet M.; Poger D.; Nair PC.;Oostenbrink C.; Mark AE. An Automated force field Topology Builder (ATB) and repository: version 1.0. *J. Chem. Theory Comput.***2011**; 7:4026-4037, doi:10.1021/ct200196m

30. Scott WRP.; Hunenberger PH.; Tironi IG.; Mark AE.; Bileter SR.; Fennen J.; Torda AE.; Huber T.; Kruger P.; vanGunsteren WF. The GROMOS biomolecular simulation program package. *J. Phys. Chem.* **1999**; 103:3596– 607, doi:10.1021/jp984217f