UDC 577.218+29 *https://doi.org/10.7124/FEEO.v36.1728*

STESHENKO A. O.[1, 2], OZHEREDOV S. P.[1], SPIVAK S. I.[1], BLUME Ya. B.[1,2] KARPOV P. A.[1]✉
[1] *Institute of Food Biotechnology and Genomics of Natl. Acad. Sci. of Ukraine,*
Ukraine, 04123, Kyiv, Baida-Vyshnevetsky str., 2A, ORCID: 0000-0003-4710-0706, 0000-0002-1928-2744, 0000-0001-7078-7548, 0000-0002-6876-642X
[2] *Educational and Scientific Center "Institute of Biology and Medicine" of Taras Shevchenko National University of Kyiv,*
Ukraine, 01601, Kyiv, Volodymyrska str., 64/13, ORCID: 0009-0008-2503-5206
✉ *karpov@nas.gov.ua, 0505446903*

THE VIRTUAL SCREENING FOR POTENTIAL ALLOSTERIC INHIBITORS TARGETING PIF-POCKET OF MAST PROTEIN KINASES

Aim. The objective of current study was the virtual screening for potential allosteric inhibitors of MAST protein kinases, as well as revision of their isotype-specific potential. ***Methods.*** Literature and database search. Molecular modeling, pharmacophore screening, artificial intelligence, molecular docking, etc. ***Results***. Based on pharmacophore screening of Enamine Ltd chemical space, 23 potential effectors of the PIF site of MAST protein kinases were selected. Based on AI reconstruction of protein-ligand complexes, the ATP-, PIF- and K-sites of MAST protein kinases where predict for potential binding of selected compounds. Subsequent CCDC GOLD molecular docking, identify the PIF-pocket as the priory site for selected compounds binding, and the leaders of screening where selected. ***Conclusions.*** Based on the complex protocol of Enamine Ltd. (4,117,328 compounds/60,516,302 conformers) chemical space virtual screening, 23 perspective PIF-site specific inhibitors of MAST protein kinases were selected. The possibility of isotype-specific inhibition of representatives of the MAST family was revealed. The leading compounds were selected for further combinatorial design and laboratory testing.

Keywords: MAST family, protein kinases, inhibitors, ligands, binding site.

Microtubule-Associated Serine/Threonine Kinases (MAST) play a critical role in cellular functions and in pathogenesis of a wide range of diseases, including types of cancer, venous thrombosis, neurological disorders (encephalopathy, epilepsy, mental retardation), inflammatory bowel disease (including Crohn's disease), metabolic disorders, type 2 diabetes, autoimmune disorders, etc [1]. Also, the members of the MAST family have been shown to be involved in cytoskeletal regulation [2]. Since discovery of MAST2, in 1993, four new family members have been identified, and currently, MAST family is divided into two subfamilies: MAST (MAST1, MAST2, MAST3, MAST4) and MASTL (MAST-like), also known as GreatWalL (GWL) protein kinase [3]. However, despite of the fact that these protein kinases are sited as promising molecular targets, selective inhibitors of MAST family members are currently unknown [4].

It is worth noting that small molecule inhibitors are a convenient and affordable tool for experimental study of protein kinases function and role. The vast majority of inhibitors of these enzymes are ATP-competitive. Because of common substrate (evolutionarily conserved from *Escherichia coli* to human), the ATP-pocket demonstrates strong similarity not only at the AGC family, but also at the level of whole kinome. Inevitably, it entails to undesirable nonspecific and alternative interactions of ATP-pocket effectors. Our previous revision of protein-ligand pockets showed complete a.a. conservation in ATP-binding site in all MAST subfamily members and minor differences in MASTL [5]. Despite the fact that inhibition of MAST protein kinases by the ATP-competitive compounds is quite effective, these interactions are not specific even at AGC superfamily [6]. Despite differences in MASTL, these variations are not sufficient to ensure selectivity of protein-ligand interactions. Nevertheless, in addition to ATP-pocket, the inhibition of AGC protein kinases is possible to processes through alternative sites of ligand binding [5]. For example, site PIF (PDK1 Interacting Fragment), firstly identified in PDK1, is currently considered as one of the most promising sites of allosteric inhibition. Appearing in all AGC-kinases, it demonstrates significantly more obvious variances in amino acid composition, and its effectors reveal strong suppressing effect on different superfamily members [7].

There are several protocols of selective inhibition of various AGC family members based on specific ligand interaction with PIF-pocket [8]. Consequently, such selectivity, opens up prospects for clinical implementation of such allosteric effectors. However, it should be noted that the vast majority of known PIF-inhibitors were developed for other members of the AGC superfamily: PDK1 (PS210, 1F8, Aurothiomalate, RS1, RS2, SBF1), PKC (PS267, MP7, PS432, PS423, PS168, PS171), Pkh2 (PS77, PS48), and S6K (PS423).

The objective of the study was the virtual screening for MAST-specific inhibitors and analysis of their isotype-binding variability, grounded on PIF-pocket differences in MAST1, 2, 3, 4 and MASTL.

Materials and methods

Amino acid sequences of MAST kinases were obtained from the UniProtKB database (www.uniprot.org): MAST1_HUMAN (Q9Y2H9), MAST2_HUMAN (Q6P0Q8), MAST3_HUMAN (O60307), MAST4_HUMAN (O15021) and GWL_HUMAN (MASTL, Q96GX5).

Structural studies were grounded on experimental models from the RCSB Protein Data Bank (www.rcsb.org), as well as, on computational models reconstructed with artificial intelligence methods and tools. In particular, 3D-models of proteins and protein-ligand complexes, were computed with Protenix service (https://protenix-server.com/), assembled on AlfaFold3 (https://deepmind.google/; https://alphafoldserver.com/).

The ligands (in *.*SD/*.sdf and *.mol2 formats) were downloaded from Enamine Ltd (https://enamine.net/) and PubChem (https://pubchem.ncbi.nlm.nih.gov/) databases. Also, in this study we used structures of compounds, deposed in MedChemExpress (https://www.medchemexpress.com/).

Pharmacophore-based virtual screening was performed with Pharmit on-line service (https://pharmit.csb.pitt.edu/), integrating academic and commercial databases of compounds, including libraries of Enamine Ltd. and PubChem. Shaping pharmacophore descriptors, their radii (Radius – «r») and vectors of interaction (θ: /φ:) were taken into account. Steric constraints were applied to shape target sites pockets (Exclusive Shape: Receptor; Tolerance=1), the radius of the descriptors were amounted to 1 (r = 1), and the number of hits per ligand was reduced to single upper-score conformation (Max Hits per Mol: 1). Upon pharmacophore screening, standard filters and thresholds were applied: MaxScore = 0; Max mRMSD = 1; Single conformer = ON. The best hits of screening were selected based on the rate of *Score* («minimisedAffinity») and *mRMSD* («minimisedRMSD») indicator.

Ligand-protein docking was performed using CCDC GOLD 2023.2.0 (www.ccdc.cam.ac.uk). The ranking of docking complexes was performed based on the rates of ChemPLP and ASP fitness functions.

The final processing and analysis of pharmacophore screening and molecular docking data were performed in MS Excel. The resulting heat map of ligand-protein interactions was constructed based on resulting scores of ChemPLP (www.ccdc.cam.ac.uk), using original tools, inbuilt into MS Excel.

Visualization and analysis of molecular structures were performed using PyMOL v.3.1.3 software (Schrödinger LLC, www.pymol.org) and BIOVIA Discovery Studio 2024 Client (https://discover.3ds.com/).

Results and discussion

Currently, there are no any PIF-effectors, reported for protein kinases of MAST family. At the same time, the PIF pocket (Fig. 2) is considered to be one of the most promising experimentally proven sites for allosteric inhibition of AGC protein kinases [9]. Preliminary studies of the MAST protein kinases PIF pocket were performed using reference ligand 78W (PSE10 = 2-oxidanylidene propyl~{N}-(2-chloroanil-6-fluorophenyl)carbonyl-~{N}'-(4-chlorophenyl) carbamimidothioate) transferred from the PDB structure 5LVO of the human PDK1 (PDPK1). At the same time, under the distance of 6 Å from refrence structure (78W), for all members of the MAST family, it was selected 13 a.a. amino acids of predicted PIF pocket. At the same time, some a.a. variations, were discovered (K/LI/I/[AV]5F/R/SF[ED]10/LCM), allowing us to expect certain isotype differences in protein-ligand interactions. Since PDB structures of MAST protein kinases are not available, the reconstruction of the primary complexes of MAST1, 2, 3, 4 and GWL protein kinases, in complex with 9EJ (PDB: 78W) were performed with application of artificial intelligence tools (Protenix + AlphaFold3 service). After structural optimization of complexes, the generalized pharmacophore was constructed, and Pharmit virtual screening of Enamine Ltd. chemical space was performed (Fig. 1). At the time of the study, the library of Enamine Ltd. provided 4117328 individual com-

pounds, and considering conformational and stereoisomerism, the screening library consisted of 60516302 exclusive conformers.

Performing pharmacophore screening, some steric constraints were applied based on the shape of the site pocket, radius of descriptors, and the maximum number of hits per molecule (see Materials and Methods). Subject to the application of standard filters and threshold restrictions (MaxScore=0; Max mRMSD=1; Single conformer=ON), as well as restrictions on the number of screening hits by affinity (minimisedAffinity) and root mean square deviation (minimisedRMSD), the list of virtual screening hits comprised 23 compounds (Fig. 1).

The subsequent AI reconstruction of protein-ligand complexes with Protenix network service revealed, that selected compounds, can bind to three sites located in catalytic domain of MAST protein kinases (Fig. 2): allosteric PIF pocket, the pocket of ATP and pocket K (predicted by homology to the Aurora-A/TPX2 complex, PDB: 5ORL).

The interaction potential of the selected compounds was tested based on the results of molecular docking in the PIF, ATP and K pockets. The group of inhibitors, with proved crystallographic evidence of interaction with AGC protein kinases, was used as a positive control (see earlier). As mentioned before, protein-ligand docking of joint library of control ligands and the hits of pharmacophore screening, was performed with application of CCDC GOLD, supported with ChemPLP and ASP fitness functions. The *Genetic Algorithm* of the program was preconfigured for «*Very Flexible*» docking with double *Search Efficiency* = 200 %. Upon completion of ligands docking into the PIF, ATP and K sites, the reconstructed complex hypotheses were ranked with inner filters of CCDC GOLD (Fig. 3). For each compound, the program generated up to 10 hypotheses of complex, and the initial hierarchical ranking of the results was performed according to the best ChemPLP and ASP values. Based on the ranking, the single top hypothesis of the complex was selected for each «compound/ site» pair. The final results were transferred into MS Excel, finally ranked and heatmap of interactions was constructed. The key measure for affinity prediction was the main of the evaluation function of

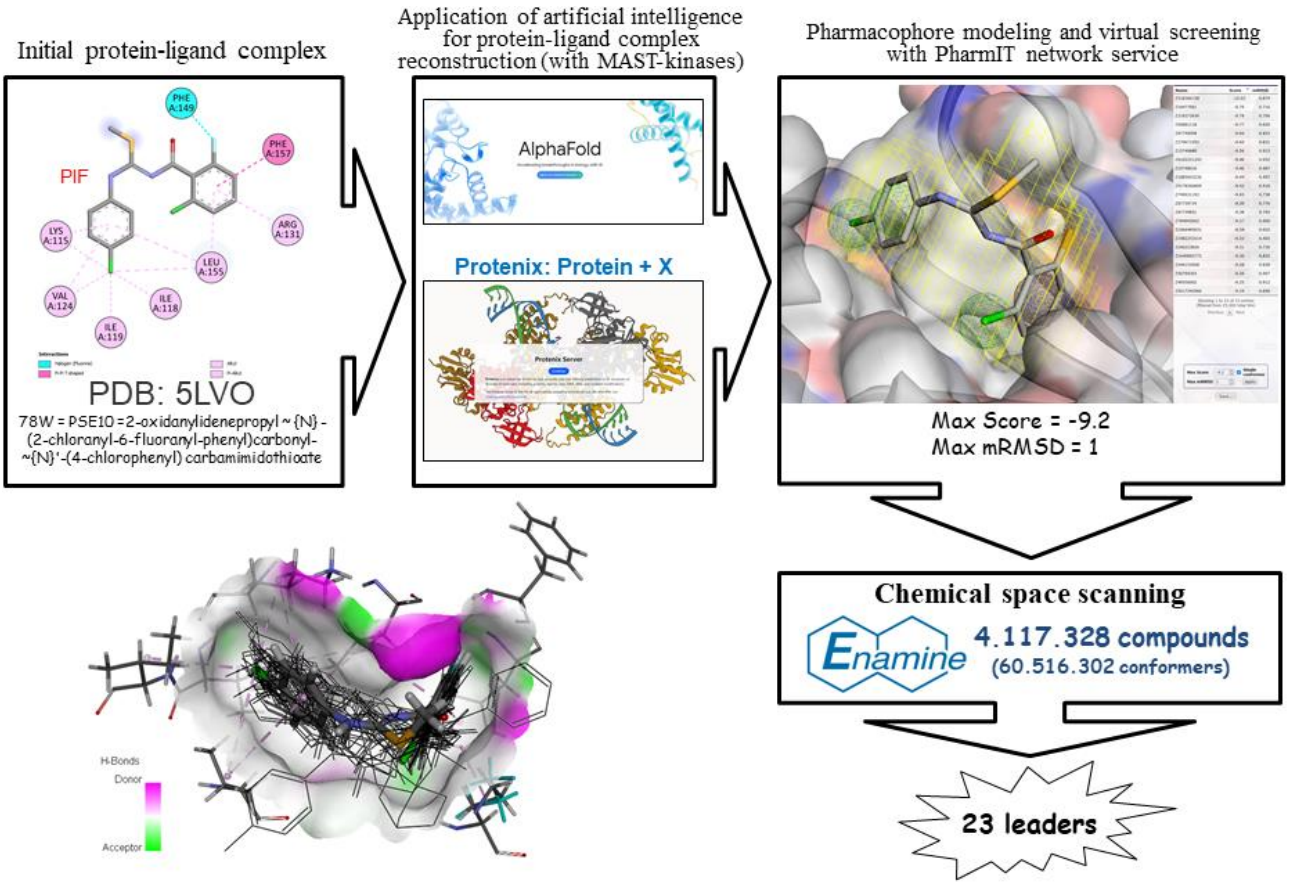

Fig. 1. The screening for inhibitors of the allosteric site (PIF-pocket = 'C').

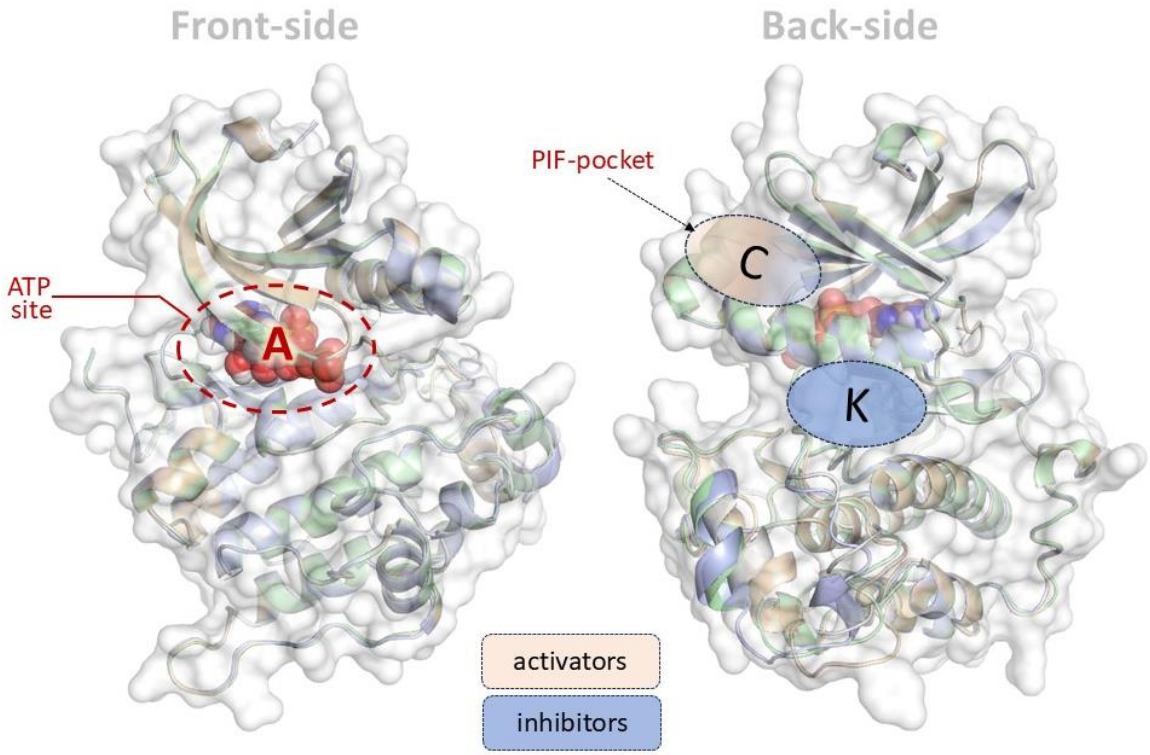

Fig. 2. The ligand-binding pockets, predicted for pharmacophore search leaders, based on artificial intelligence methods of complex reconstruction (AlphaFold3 + Protenix). A – ATP site, C – PIF-pocket and K – the pocket predicted by homology to the Aurora-A/TPX2 complex (reference PDB structure: 5ORL).

The analysis of the joint database revealed that all selected inhibitors demonstrate preferential binding to the PIF site of MAST protein kinases. This conclusion remains true even for those compounds that also revealed alternative affinity for the K pocket and ATP-binding site. It is noteworthy that the predicted affinity of all selected compounds for the studied MASTL protein kinase sites was inferior to other family members. At the same time, unique interaction profiles inherent in individual ligands were also identified. For example, SBF1 almost does not interact with the PIF-site of MASTL and the ATP-binding pocket of MAST2. However, SBF1 shows potential strong affinity for the PIF and K sites of other family members. The highest potential affinity of SBF1, was predicted for the PIF site of MAST4. Also, a slightly lower value was observed in the case of a similar site of MAST3. A potential inhibitor from the Enamine library, Z18477561, was characterized by the absence of interaction with the ATP-binding site of MAST3 protein kinase. The control compound PS423 showed the highest potential affinity for the PIF site of MAST1, MAST2 and MAST3 protein kinases. The commercial compound Z87739851 mainly interacts with the PIF site of MAST4, while Z19748016 displays increased potential affinity for MAST2 and MAST4. Among all the compounds from the Enamine chemical library, compound Z15749680 demonstrates the best potential binding parameters, making it a reference for further combinatorial optimization and *in vitro* testing. It is assumed that compound Z318373636 is able to effectively interact with the PIF-site of MAST1 and MAST3, and compound Z56794203 is most closely related to the mentioned site of MAST1, MAST2 and MAST3. At the same time, compound Z279671592 was selected as the most versatile candidate for isotype-neutral inhibitor of MAST family members.

Conclusions

Based on a comprehensive virtual screening of the Enamine Ltd. chemical space (4,117,328 compounds/60,516,302 conformers), 23 promising inhibitors of the allosteric PIF site of MAST family protein kinases were selected. The possibility of isotype-specific inhibition of MAST family members was shown. The leader-compounds were selected for further combinatorial design and laboratory testing.

The molecular docking study was compiled using CCDC GOLD software, kindly provided by the Cambridge Crystallographic Data Center (CCDC) as part of a grant from the Frank H. Allen International Research and Illumination Program. Allen International Research and Education Program, FAIRE) – ID: 21554.

	PIF					ATP					K				
	MAST1	MAST2	MAST3	MAST4	MASTL	MAST1	MAST2	MAST3	MAST4	MASTL	MAST1	MAST2	MAST3	MAST4	MASTL
9EJ	-68,81	-76,7	-76,7	-84,87	-65,79	-77,86	-78,37	-80,13	-79,36	-65,91	-64,04	-73,45	-51,69	-79,33	-44,4
SBF1	-54,67	-61,45	-80,02	-93,79	147,71	-75,79	-7,47	-74,67	-72,37	-75,1	-75,37	-81,3	-78,22	-70	-52,8
PS48	-64,18	-67,85	-67,83	-59,69	-48,5	-52,18	-53,91	-54,54	-53,92	-50,51	-55,03	-55,24	-53,98	-53,1	-41,9
PS171	-59,52	-63,82	-59,51	-64,65	-56,87	-57,34	-57,49	-59,09	-59,3	-58,07	-52,14	-56,3	-51,31	-57,07	-42,02
PS432	-69,81	-70,28	-70,28	-68,44	-34,13	-63,4	-62,01	-56,02	-68,24	-59,73	-54,91	-62,35	-48,24	-58,19	-49,63
PS77	-70,02	-71,53	-71,78	-69,71	-54,57	-61,99	-63,33	-59,91	-61,72	-62,04	-70,23	-73,47	-55,37	-59,58	-44,53
PS423	-83,01	-86,38	-86,38	-75,82	-57,88	-61,53	-66,44	-61,55	-61,6	-63,51	-66,59	-71,61	-60,48	-58,58	-45,25
Z18477561	-56,96	-56,69	-63,95	-68,53	-54,68	-60,92	-57,89	28,86	-56,6	-46,64	-65,91	-62,47	-48,49	-62,07	-55,37
Z19822533314	-64,29	-63,82	-64,12	-58,14	-55,24	-60,81	-62,53	-62,95	-61,75	-55,38	-55,74	-52,63	-47,16	-61,69	-40,89
Z749631192	-72,69	-73,15	-73,15	-56,13	-52,25	-60,55	-60,63	-62,43	-60,55	-55,57	-52,07	-52,34	-46,47	-60,5	-42,74
Z240223836	-64,56	-63,77	-64,49	-64,91	-52,74	-59,26	-65,18	-58,74	-58,43	-54,54	-62,06	-58,26	-46,89	-59,21	-49,87
Z4178366809	-60,81	-62,37	-68,49	-66,52	-52,96	-64,08	-66,65	-61,34	-63,48	-53,27	-66,15	-62,56	-50,57	-62,56	-48,84
Z87739851	-46,51	-42,99	-48,23	-57,61	-47,66	-57,34	-59,63	-57,05	-57,05	-48,25	-51,22	-56,48	-47,14	-57,58	-47,83
Z318366728	-65,81	-68,22	-68,22	-56,81	-51,63	-56,99	-57,56	-58,9	-57,85	-51,63	-52,54	-55,39	-55,18	-56,78	-39,2
Z56881118	-56,06	-58,76	-58,76	-63,2	-49	-56,84	-55,19	-52,61	-53,11	-50,19	-62,51	-59,37	-46,89	-51,82	-42,53
Z78494360	-59,74	-62,31	-65,77	-59,73	-54,41	-56,81	-60,05	-58,88	-55,24	-53,86	-57,32	-58,05	-50,46	-56,28	-40,54
Z344153008	-68,94	-70,17	-70,6	-57,93	-53,27	-56,19	-50,22	-53,57	-57,3	-53,13	-50,32	-54,53	-48,49	-57,64	-44,71
Z19748016	-50,18	-55,2	-63,37	-58,27	-46,76	-56,05	-60,5	-54,06	-56,9	-49,37	-59,19	-56,94	-47,18	-56,66	-43,22
Z87740098	-50,19	-48,4	-50,17	-57,45	-47,29	-55,93	-59,12	-48,83	-56,75	-48,39	-59,69	-57,5	-52,29	-56,21	-38,6
Z15749680	-76,66	-79,13	-79,13	-68,9	-62,63	-72,7	-73,4	-63,59	-72,15	-65,5	-65,77	-77,11	-53,77	-65,68	-54
Z318373636	-63,42	-59,33	-63,63	-53,98	-47,4	-51,86	-53,03	-53,44	-52,97	-47,57	-50,36	-52,11	-56,28	-53,42	-38,98
Z87739734	-54,53	-53,96	-54,49	-55,54	-46,27	-51,42	-57,42	-50,75	-51,21	-48,52	-52,09	-54,75	-51,19	-51,42	-41,58
Z1885603236	-52,58	-65,75	-65,75	-60,17	-55,79	-51,03	-50,2	-52,58	-52,81	-52,39	-55,75	-55,55	-36,61	-52,83	-38,2
Z8102221243	-50,16	-51,8	-51,77	-43,93	-38,26	-53,1	-40,98	-52,37	-48,32	-48,17	-42,58	-53	-35,15	-35,8	-34,3
Z49556002	-49,32	-47,84	-49,01	-52,45	-40,49	-52,18	-52,15	-52,5	-50,6	-49,93	-51,12	-53,31	-51,36	-37,69	-39,09
Z56794203	-74,39	-76,9	-74,33	-48,47	-48,93	-66,62	-67,31	-66,91	-65,28	-56,35	-60,07	-59,79	-55,23	NO	-43,51
Z279671592	-70,18	-68,8	-71,41	-72,96	-20,25	-58,26	-58,22	-57,04	-57,6	-54,22	-53,69	-58,31	-48,04	NO	-44,35
Z1966495631	-44,58	-46,49	-47,35	-45,33	-43,95	-46,56	-49,58	-46,55	-45,65	-41,61	-54,04	-53,68	-44,92	NO	-37,97
Z5017242066	-50,36	-48,99	-50,11	-49,26	-43,98	-47,5	-51,17	-47,72	-47	-40,84	-51,2	-45,17	-44,67	NO	-33,22

Fig. 3. The heatmap of predicted affinity, calculated for virtual screening hits against three priority sites (PIF, ATP and K) of MAST-kinases. The ranking was processed based on the ChemPLP fitness function of the CCDC GOLD. NO – no hypothesis of ligand binding for indicated site.

References

1. Rumpf M., Pautz S., Drebes B., Herberg F. W., Müller H. J. Microtubule-Associated Serine/Threonine (MAST) kinases in development and disease. *International Journal of Molecular Sciences*. 2023. Vol. 24 (15). P. 11913. https://doi.org/10.3390/ijms241511913.
2. Chudinova E. M., Karpov P. A., Fokin A. I., Yemets A. I., Lytvyn D. I., Nadezhdina E. S., Blume Y. B. MAST-like protein kinase IREH1 from *Arabidopsis thaliana* co-localizes with the centrosome when expressed in animal cells. *Planta*. 2017. Vol. 246 (5). P. 959–969. https://doi.org/10.1007/s00425-017-2742-4.
3. Pearce L. R., Komander D., Alessi D. R. The nuts and bolts of AGC protein kinases. *Nature Reviews Molecular Cell Biology*. 2009. Vol. 11 (1). P. 9–22. https://doi.org/10.1038/nrm2822.
4. Thaimattam R., Banerjee R., Miglani R., Iqbal J. Protein kinase Inhibitors: Structural Insights into Selectivity. *Current Pharmaceutical Design*. 2007. 13 (27). P. 2751–2765. https://doi.org/10.2174/138161207781757042.
5. Karpov P. A., Steshenko A. O., Ozheredov S. P., Blume Y. B. Prediction of protein-ligand binding sites modulating activity of MAST protein kinases. *Faktori Eksperimentalnoi Evolucii Organizmiv*. 2024. Vol. 35. P. 151–157. https://doi.org/10.7124/feeo.v35.1677.
6. Jester B. W., Gaj A., Shomin C. D., Cox K. J., Ghosh I. Testing the promiscuity of commercial kinase inhibitors against the AGC kinase group using a split-luciferase screen. *Journal of Medicinal Chemistry*. 2012. Vol. 55 (4). P. 1526–1537. https://doi.org/10.1021/jm201265f.
7. Engel M. The PIF pocket of AGC kinases: a target site for allosteric modulators and Protein–Protein interaction inhibitors. *Methods and Principles in Medicinal Chemistry*. 2013. P. 187–223. https://doi.org/10.1002/9783527648207.ch9.
8. Busschots K., Lopez-Garcia L.A., Lammi C., Stroba A., Zeuzem S., Piiper A., Alzari P. M., Neimanis S., Arencibia J.M., Engel M., Schulze J. O., Biondi R. M. Substrate-Selective Inhibition of Protein Kinase PDK1 by Small Compounds that Bind to the PIF-Pocket Allosteric Docking Site. *Chemistry & Biology*. 2012. Vol. 19 (9). P. 1152–1163. https://doi.org/10.1016/j.chembiol.2012.07.017.
9. Busschots K., Lopez-Garcia L. A., Lammi C., Stroba A., Zeuzem S., Piiper A., Alzari P. M., Neimanis S., Arencibia J. M., Engel M., Schulze J. O., Biondi R. M. Substrate-Selective Inhibition of Protein Kinase PDK1 by Small Compounds that Bind to the PIF-Pocket Allosteric Docking Site. *Chemistry & Biology*. 2012. Vol. 19 (9). P. 1152–1163. https://doi.org/10.1016/j.chembiol.2012.07.017.

СТЕШЕНКО А. О.[1, 2], ОЖЕРЄДОВ С. П.[1], СПІВАК С. І.[1], БЛЮМ Я. Б.[1, 2], КАРПОВ П. А.[1]

[1] Державна установа «Інститут харчової біотехнології та геноміки НАН України»,
Україна, 04123, м. Київ, вул. Байди-Вишневецького, 2А
[2] Навчально-науковий центр «Інститут біології та медицини» Київського національного університету імені Тараса Шевченка,
Україна, 01601, м. Київ, вул. Володимирська, 64/13

ВІРТУАЛЬНИЙ СКРИНІНГ ПОТЕНЦІЙНИХ АЛОСТЕРИЧНИХ ІНГІБІТОРІВ ПРОТЕЇНКІНАЗ РОДИНИ MAST, ЩО ВЗАЄМОДІЮТЬ З САЙТОМ PIF

Мета. Метою даного дослідження був віртуальний скринінг потенційних алостеричних інгібіторів протеїнкіназ MAST, а також перевірка їхньої ізотипової специфічності. ***Методи.*** Пошук в літературі та базах даних. Молекулярне моделювання, фармакофорний скринінг, штучний інтелект, молекулярний докінг тощо. ***Результати.*** На основі фармакофорного скринінгу хімічного простору Enamine Ltd було відібрано 23 потенційних ефектори PIF-сайту протеїнкіназ MAST. На основі AI-реконструкції ліганд-білкових комплексів було визначено ATP-, PIF- та K-сайти протеїнкіназ MAST, де було передбачено потенційне зв'язування відібраних сполук. Подальший молекулярний докінг за допомогою CCDC GOLD виявив PIF-кишеню як пріоритетний сайт для зв'язування обраних сполук та визначив лідерів скринінгу. ***Висновки.*** На основі комплексного протоколу компанії Enamine Ltd. (4 117 328 сполук/60 516 302 конформерів) віртуального скринінгу хімічного простору відібрано 23 перспективних PIF-сайт-специфічні інгібітори протеїнкіназ MAST. Виявлено можливість ізотип-специфічного інгібування представників родини MAST. Провідні сполуки відібрано для подальшого комбінаторного дизайну та лабораторного тестування.

Ключові слова: родина MAST, протеїнкінази, інгібітори, ліганди, сайти зв’язування.