



Prospect of Anterior Gradient 2 homodimer inhibition via repurposing FDA-approved drugs using structure-based virtual screening

Shafi Ullah¹ · Shafi Ullah Khan^{2,5} · Abbas Khan³ · Muhammad Junaid⁴ · Humaira Rafiq⁴ · Thet Thet Htar² · Yaxue Zhao¹ · Syed Adnan Ali Shah⁶ · Abdul Wadood⁴

Received: 25 March 2021 / Accepted: 21 June 2021 / Published online: 28 June 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Anterior Gradient 2 (AGR2) has recently been reported as a tumor biomarker in various cancers, i.e., breast, prostate and lung cancer. Predominantly, AGR2 exists as a homodimer via a dimerization domain (E60-K64); after it is self-dimerized, it helps FGF2 and VEGF to homo-dimerize and promotes the angiogenesis and the invasion of vascular endothelial cells and fibroblasts. Up till now, no small molecule has been discovered to inhibit the AGR2–AGR2 homodimer. Therefore, the present study was performed to prepare a validated 3D structure of AGR2 by homology modeling and discover a small molecule by screening the FDA-approved drugs library on AGR2 homodimer as a target protein. Thirteen different homology models of AGR2 were generated based on different templates which were narrowed down to 5 quality models sorted by their overall Z-scores. The top homology model based on PDB ID = 3PH9 was selected having the best Z-score and was further assessed by Verify-3D, ERRAT and RAMPAGE analysis. Structure-based virtual screening narrowed down the large library of FDA-approved drugs to ten potential AGR2–AGR2 homodimer inhibitors having FRED score lower than -7.8 kcal/mol in which the top 5 drugs' binding stability was counter-validated by molecular dynamic simulation. To sum up, the present study prepared a validated 3D structure of AGR2 and, for the first time reported the discovery of 5 FDA-approved drugs to inhibit AGR2–AGR2 homodimer by using structure-based virtual screening. Moreover, the binding of the top 5 hits with AGR2 was also validated by molecular dynamic simulation.

Shafi Ullah and Shafi Ullah Khan have contributed equally to this work

✉ Shafi Ullah
shafiullah16@sjtu.edu.cn

✉ Abdul Wadood
awadood@awkum.edu.pk

¹ Department of Medicinal Chemistry, School of Pharmacy, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

² School of Pharmacy - Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Selangor, Malaysia

³ Department of Bioinformatics and Biological Statistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

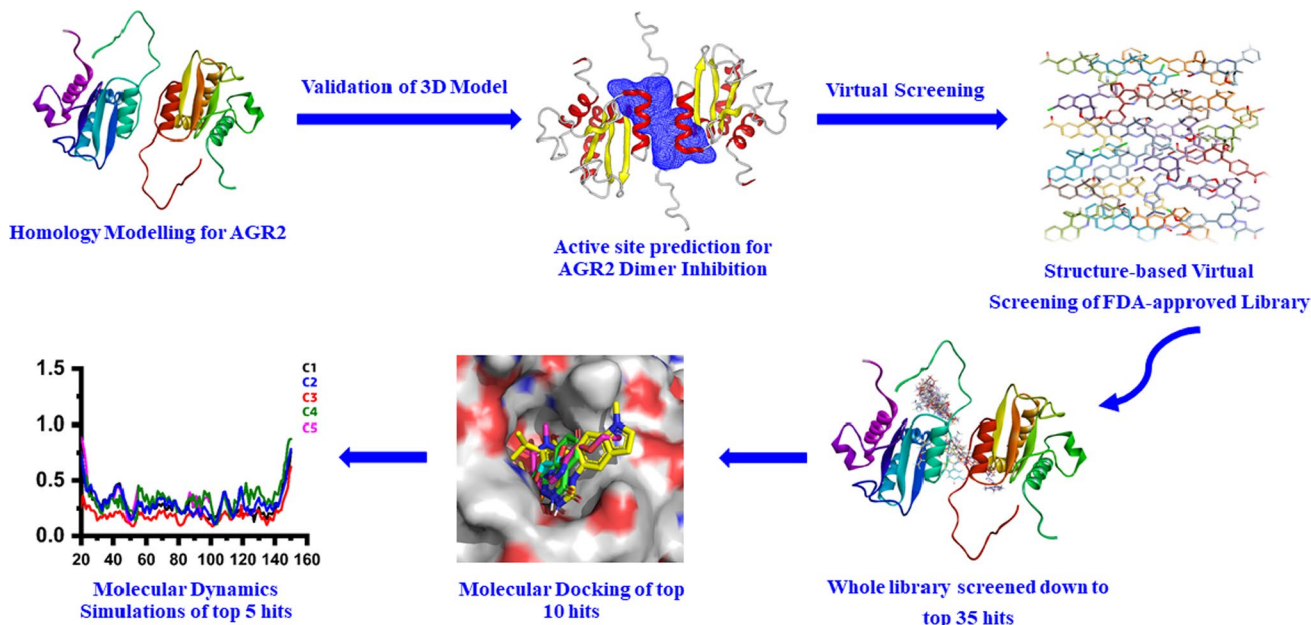
⁴ Department of Biochemistry, Abdul Wali Khan University, Mardan 23200, Pakistan

⁵ Department of Pharmacy, Abasyn University, Ring Road, Peshawar 25120, Pakistan

⁶ Faculty of Pharmacy, Atta-Ur-Rahman Institute for Natural Products Discovery (AuRIns), Universiti Teknologi MARA, Bandar Puncak Alam, Selangor Darul Ehsan, Malaysia

Graphic abstract

A validated 3D structure of Anterior Gradient 2 (AGR2) was prepared by homology modeling, which was used in virtual screening of FDA-approved drugs library for the discovery of prospective inhibitors of AGR2–AGR2 homodimer.



Keywords AGR2 homodimer · Cancer · Virtual screening · Drug repositioning · Dimer inhibitor

Introduction

Anterior Gradient 2 (AGR2), also known as secreted cement gland protein XAG-2, is a member of protein disulfide isomerase (PDI) superfamily [1], which is over-expressed in multiple human cancers, including breast, prostate, lung, gastric, ovarian and pancreatic cancers [2–6]. It has both intracellular and extracellular roles; it harbors a signal peptide (1 M–20 K) at the N terminus to import into ER. After maturation, it is released via the secretory pathway and acts as a signaling molecule in numerous pathways. By virtue of its C-terminal endoplasmic reticulum (ER) retention sequence (K172–L175), it resides in the lumen of ER and acts as a chaperone to assist other proteins to fold properly. Normally, it is expressed in secretory goblet cells of the intestine, where it is responsible for secreting MUC2, a cysteine-rich glycoprotein that protects the inner lining of the intestine [1]. AGR2 contains a cysteine-containing conserved thioredoxin-like domain, also called PDI domain or CPHS domain (C81–S84) which makes mixed disulfide bonds with its client proteins, and hence takes part in the maturation of mucins, e.g., MUC2, MUC1, MUC5AC [7] and other members of the secretory pathway and ultimately plays a pivotal role in the maintenance of ER homeostasis [8]. Predominantly,

AGR2 exists as a homodimer via a dimerization domain (E60–K64); after it is self-dimerized, it helps FGF2 and VEGF to homo-dimerize and promotes the angiogenesis (Fig. 1a and b) and the invasion of vascular endothelial cells and fibroblasts by augmenting the activities of vascular endothelial growth factor (VEGF) and fibroblast growth factor 2 (FGF2) [9].

Being an extracellular signaling molecule and intracellular ER chaperone, targeting AGR2 has been a bottleneck in discovering both anti-AGR2 small and large molecules for cancer therapy. Few approaches have been used to inhibit AGR2. Our collaborator group previously reported that a monoclonal antibody 18A4, targeting the dimerization domain (E60–K64) of AGR2, can inhibit different xenograft tumors in mice. The AGR2 inhibition activity of 18A4 is by virtue of its binding to amino acid sites spanning the regions of E60–H76 and A86–E153. However, the 18A4 was combined with bevacizumab (as an adjunct therapy) to produce a maximum effect in inhibiting the angiogenesis and tumor growth in SKOV3 cell lines [9, 10]. Another approach to inhibit AGR2 is by using a peptide; H10 peptide was discovered after screening millions of peptides by mRNA display library. The H10 peptide can inhibit AGR2 by binding with the interface of AGR2–AGR2 homodimer surrounding the amino acid residues P41, E60 and E96 [11].

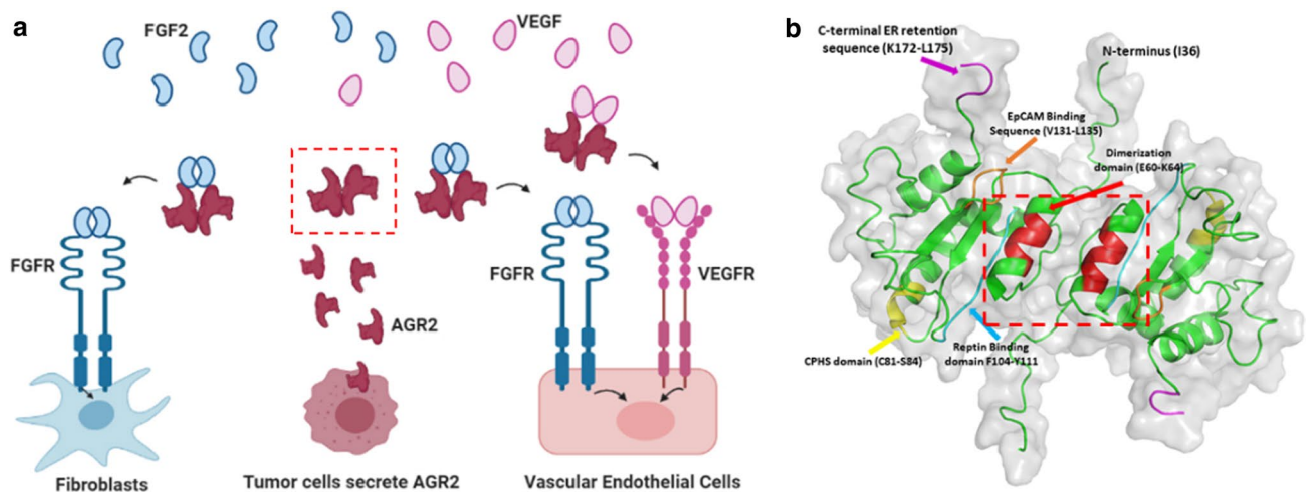


Fig. 1 a Pathway of AGR2 Dimerization [5] b Dimerization domain (E60-K64) of AGR2 [3]

Drug repurposing/repositioning offers a comparatively direct way to study the already approved or investigational drugs for a completely different disorder. It can provide a better hit without safety issues which can cut the time, technical and financial resources as compared to starting the drug discovery process from scratch [12].

Using structure-based drug design protocol, Rani et al. discovered several FDA-approved drugs against important enzymes of *Mycobacterium tuberculosis*; virtual screening of 1932 approved drugs from DrugBank and 1852 drugs from eLEA3D; using AutoDock Vina as the docking program and MurB and MurE enzymes as the protein targets, the study discovered that sulfadoxine and pyrimethamine showed stable interaction with MurB, while lifitegrast and sildenafil (−9.1 kcal/mol) showed the most reliable interaction with MurE [13].

Dakshanamurthy et al. reported that mebendazole (anti-parasitic) could structurally inhibit the vascular endothelial growth factor receptor 2 (VEGFR2), a mediator of angiogenesis; this finding was also supported by experimental data. They screened 3,671 FDA-approved drugs across 2,335 human protein crystal structures by using a high-throughput computational docking [14].

Recently, computer-aided drug design is widely used for rational drug discovery. Virtual screening of compounds against a valid target can lead to cost- and time-effective novel drug discovery [15]. For those proteins which crystal structure was not available, through homology modeling, the crystal structure was generated, and then, virtual screening was performed on the target protein. In order to discover novel antagonists of the endothelin-A receptor (ETAR), first, the 3D structure was developed by homology modeling because its X-ray crystal structure was not available. Then the target was virtually screened against

Traditional Chinese Medicine (TCM) database to identify novel natural ETAR antagonists resulting in the discovery of two potential antagonists; their binding with ETAR was validated by molecular dynamic simulation and molecular mechanics generalized born surface area [16].

To discover novel protein–protein interaction inhibitors of Nrf2-Keap1, structure-based virtual screening was performed on Specs database, which reported the compound 15 with an in vitro EC₅₀ of 9.80 μM in the fluorescence polarization (FP) assay [17].

Up till now, to the best of our knowledge, no small molecule has been discovered to inhibit the AGR2–AGR2 homodimer. Therefore, the present study was performed to prepare a validated 3D structure of AGR2 by homology modeling and for the first time discover a set of small molecules by screening the FDA-approved drugs library (<https://www.selleckchem.com/screening/fda-approved-drug-library.html>) on AGR2–AGR2 homodimer as a target protein. Since the X-ray crystal structure of AGR2 was not available, we prepared its structure by homology modeling and used the generated homology model for further virtual screening. Then we compared the list of screened drugs with that obtained from screening the NMR structure of AGR2 (PDB ID = 2lns). The obtained list of drugs was again screened by a molecular docking tool which narrowed down the list to 35 common drugs. The interaction of these drugs was determined and then ultimately validated by molecular dynamic simulation. To sum up, the present study reported the discovery of 5 FDA-approved drugs to inhibit AGR2–AGR2 homodimer for the first time by using structure-based virtual screening. Moreover, the binding of the top 5 FDA-approved drugs with AGR2 was also validated by molecular dynamic simulation.

Materials and methods

Homology modeling of AGR2

In the absence of suitable X-ray structure in RCSB protein data bank, “YASARA Structure’s homology modeling module” was used to generate the homology model of AGR2 [7, 18]. YASARA (Yet Another Scientific Artificial Reality Application) is a complete package of molecular modeling, molecular graphics and molecular dynamics. It is comparatively easy to use and reliable package for molecular modeling [19]. The most advanced version of YASARA (YASARA Structure) also includes a full homology modeling module that automate all steps from an amino acid sequence as an input (including alignment of amino acid sequence, building loops, rotamer selection, optimization of stereochemistry and subsequent validation of the homology model) to a refined high-resolution model as an output. It often generates a comprehensive scientific report for each modeling steps. The overall Z-scores obtained for AGR2 were within the range, “GOOD.” The Z-score is a matrix that can indicate how far a system deviates from the average of standard reference structures. The YASARA and

WHAT IF Twinset was used for the subsequent visualization and analysis [20].

Evaluation of the generated homology model

Following the generation of the AGR2 homology model, the RAMPAGE and ModRefiner servers were used to evaluate and optimize the AGR2 homology models [21].

Structure-based virtual screening

The workflow of docking-based virtual screening is depicted in Fig. 2. In brief, all FDA-approved drugs were subjected to molecular docking calculations using built-in utility of OpenEye Scientific Software’s FRED v3.2.0. Prior to docking, the AGR2 structure was optimized at pH 7.0 using the pdb2receptor tool in OEDocking. OMEGA 2.5.1 [22] was used for generating the multi-conformers of all compounds. During generation of the conformers, default OMEGA settings were used (maximum 200 conformers per molecule). For molecular docking calculations, binding site was portrayed around the entire AGR2 structure. FRED default

Method: Molecular Docking Studies

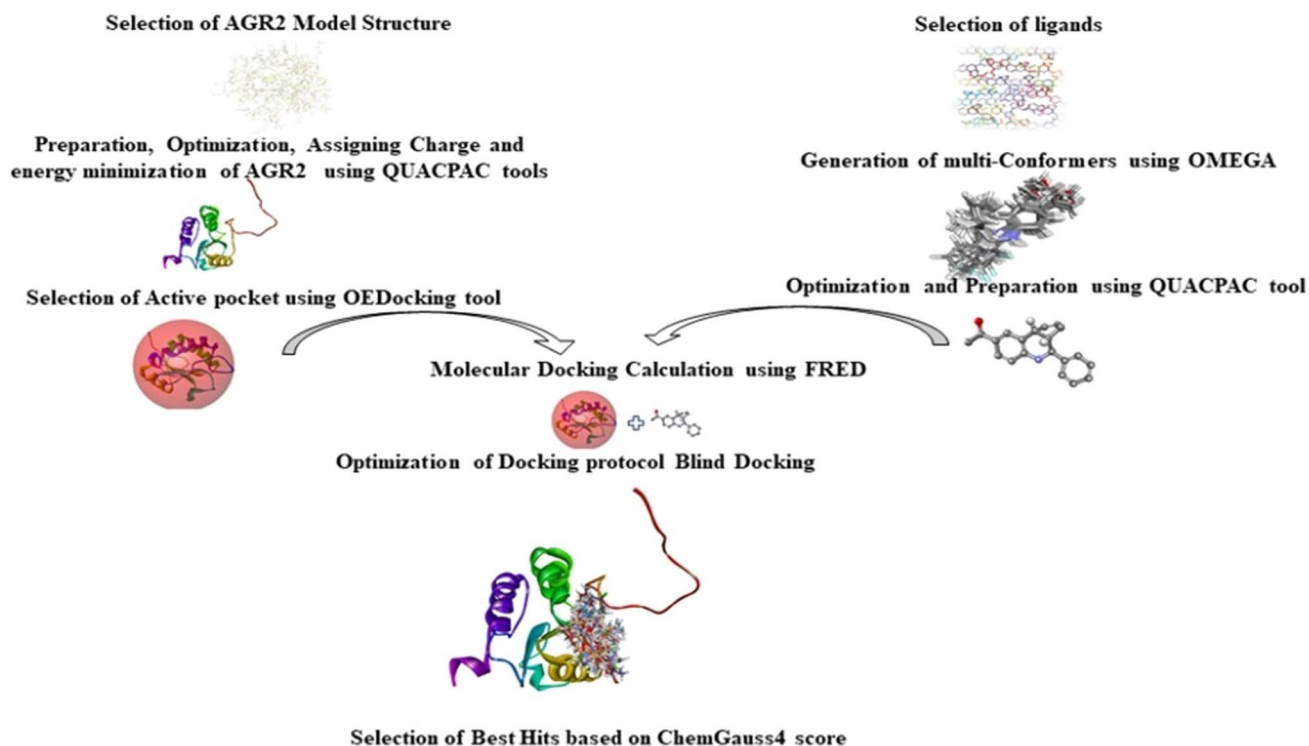


Fig. 2 Workflow of docking-based virtual screening

parameters were also used for predicting binding affinity of the compounds with AGR2 [23].

After protocol optimization, all FDA-approved compounds were docked using the previously described protocol [15]. A maximum of ten poses were created for each compound, and the best hits were chosen based on the lowest Chemgauss4 score. The Discovery Studio Visualizer [15] was used to illustrate the bonding orientation of docked poses within AGR2.

MD simulations

Amber14 package was used for MD simulation to understand the dynamic behavior of each ligand-bound system. Solvation with TIP3P water model with the addition of ions to neutralize the system was carried out. For drugs, topologies were generated with antechamber. AMBER14 FFSB force field was used for protein, whereas GAFF2 was used for drugs [24]. Two-step energy minimization followed by heating of the system was performed. We used default parameters such as temperature 300 K and 2 ps. PME (particle mesh Ewald) algorithm with cutoff distance 10 Å was used for long-range interactions. We used SHAKE algorithm for covalent interactions. A total of 100 ns simulation for each system was performed. CPPTRAJ and PYTRAJ were used for post-simulation analyses [22, 24, 25]. We prepared protein–ligand complexes of top 5 hits with AGR2, i.e., C1 (AGR2–AZD2281), C2 (AGR2–Emtricitabine), C3 (AGR2–flumazenil), C4 (AGR2–AGR2–ganetespib (STA-9090)) and C5 (AGR2–mercaptapurine) before running MD simulations.

Results and discussion

Homology models generation using YASARA

The homology modeling steps for generating AGR2 are depicted in Fig. S1. We used amino acid sequence of Anterior Gradient 2 (AGR2) protein, which was identified as “O95994” in UniProtKB/Swiss-Prot for homology modeling [26, 27].

The following procedure was used to generate the homology models; PSI-BLAST [28], which was built in YASARA, was used to locate the 13 closest templates in the PDB. Table 1 shows the PDB structures that have the greatest degree of similarity to our target sequence.

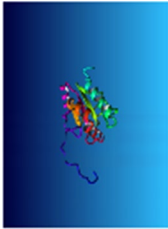
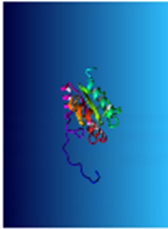
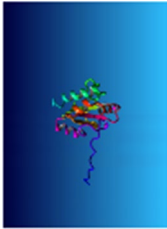
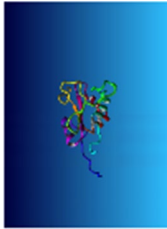
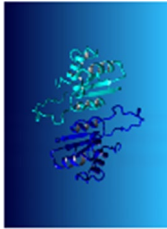
We built an AGR2 homology model utilizing both monomer and dimer. The Uniport ID “O95994” was used to retrieve AGR2’s full-length amino acid sequence. A secondary structure prediction for the target structure was needed to assist in the alignment correction and loop modeling. This was accomplished by using PSI-BLAST to generate a target sequence profile and then feeding it into the PSI-Pred secondary structure prediction algorithm [29].

The resulting prediction as shown in Fig. S2 indicated the estimated probability for each of the secondary structure classes, helix, strand and coil. To help in aligning the amino acid sequence of target and templates, a target sequence profile was constructed from multiple sequence alignment, which was built from similar UniRef90 sequences. Table 2 lists the 5 generated homology models sorted based on their overall quality Z-scores.

Table 1 List of top 13 closest templates identified for generating the homology model of AGR2

Template	Total score	BLAST E-value	Align score	Cover (%)	ID
1	357.11	1.00E-43	533	77	3PH9-B
2	266.71	2.00E-46	696	80	2LNT-A
3	188.7	4.00E-47	702	80	2LNS-B
4	127.94	4.00E-39	196	71	1SEN-A
3	38.03	4.00E-16	57	71	3FK8-A
4	27.78	7.00E-18	50	63	1UC7-A
5	26.95	8.00E-15	49	73	2JU5-A
6	24.79	3.00E-05	53	69	5UM7-B
7	24.59	7.00E-15	53	63	2LST-A
8	24.45	4.00E-15	54	59	1VRS-F
9	22.87	3.00E-17	46	58	4IP1-A
10	22.79	0.0005	42	68	4TW5-A
11	22.08	1.00E-16	44	58	4IP6-A
12	20.5	1.00E-16	39	59	2FWG-A
13	19.06	2.00E-36	28	81	3IRA-A

Table 2 Five representative homology models of AGR2

Entry	Z-score	Structure	State	Model ID	Residues	Comment
1	0.366		Monomer	3PH9-B01	41–174	Optimal
2	−0.463		Monomer	3PH9-B02	41–174	Good
3	−0.522		Monomer	1SEN-A	44–172	Good
4	−2.123		Monomer	2LNT-A	51–171	Poor
5	−2.265		Homodimer	2LNS-~	37–175	Poor

Evaluation of the generated homology model

The SAVES structure assessment software [<http://nihserver.mbi.ucla.edu/SAVES/>] was used to validate the reliability of the modeled AGR2 structure. SAVES is the integration of various tools for evaluating the protein structure such as Procheck, WhatCheck, ERRAT, Verify 3D and PROVE. The SAVES structure verification software was used to assess the quality of the modeled AGR2 structure, as shown in Fig. S3, S4 and S5. The Ramachandran plot in Fig. S3 reported that 90.3% of the amino acid residues were in the core region, 9.7% in the allowed region and 0% within the generously allowed regions. However, none of the residues was found in the disallowed region of the plot, confirming the stereochemical reliability of the AGR2 homology model.

Furthermore, a Verify-3D score of 88.88% (Fig. S4) and an ERRAT score of 97.04 percent (Fig. S5) verified the modeled structure's "structure sequence compatibility" and "non-bonded interactions," respectively. Similarly, the atomic volumes of the modeled AGR2 residues were compared to the corresponding residues in the PDB database, yielding an ideal Z-score mean value of 0.366 for the best AGR2 model. RAMPAGE generated Ramachandran plots of models for evaluation. The ModRefiner server was then used to refine the generated models. Figure S3 shows the Ramachandran plot after the refinement. The absence of any residues in the outlier area shows that each model is improved in stereochemistry after refinement.

Structure-based virtual screening

Molecular docking calculation of all FDA compounds was performed using blind docking in which whole structure was covered as an active site using FRED docking software. Detailed methodology of docking is illustrated in Fig. 2.

The FRED software has a strong track record of implementation in structure-based drug discovery. Huabin Hu et al. recently described FRED (Chemgauss4 score) as one of the best docking score functions among three separate docking score functions [30]. A maximum of 200 conformers were produced for each ligand and used as input. The protocol mentioned in methods was used to execute molecular docking calculation. For each compound, a maximum of ten poses were obtained, and the pose with the lowest Chemgauss4 score was chosen as a best hit compound. The chemgauss4 score was used as a criterion for detecting actives in a large pool of ligands. If a ligand's score was less than -7.8 kcal/mol, it was considered as active as shown in Table 3.

Table 3 List of top 10 drugs with FRED Chemgauss4 score lower than -7.8 kcal/mol

S no	Compounds	FRED CHEMGUASS4 Score
1	AZD2281.cdx	-11.0358
2	Flumazenil.cdx	-10.3697
3	Olsalazine sodium.cdx	-10.2100
4	Ganetespib(STA-9090).cdx	-9.94754
5	Ozagrel hydrochloride.cdx	-9.93677
6	Procodazole.cdx	-9.24921
7	Epinephrine HCl.cdx	-9.05115
8	Emtricitabine.cdx	-8.69452
9	Pramipexole.cdx	-7.84159
10	Mercaptopurine.cdx	-7.81037

Comparative analysis of binding interactions

Molecular docking was used to obtain insights into the binding interactions of identified best AGR2 ligands inside the target's active pocket. Figures 3, 4 and 5 and S6-S10 displays the detailed binding interactions of the top ten AGR2 ligands. Each of the top 10 hits was found interacting with the residues of dimerization domain of AGR2 (E60-K64).

Molecular dynamic simulation

The stability and flexible dynamic behavior of the AGR2, C1–C5 complexes

The current study aimed to identify the dynamic behavior and to evaluate the internal movement of the AGR2 with different ligand complexes (C1, C2, C3, C4 and C5). The structure of all the complexes was simulated in an explicit water environment. We further determined the effect of these ligands on the stability of the complex by predicting the thermodynamic state function (RMSD). The root mean square deviation (RMSD) method is commonly used to calculate the variation in a protein backbone from its initial structural conformation to its final position. The deviations observed during simulation period can be used to estimate the dynamics stability of a biological molecule relative to its conformation. The deviation that a protein faces during simulation is linked to its stability; a smaller variation indicates that the structure is more stable and is less likely to exceed the stability limit. Herein, the stability (RMSD) of C α backbone was estimated for 100 ns trajectory for each of the AGR2–ligand complex (C1–C5) which revealed that hits of the complex C1 and C2 strongly bound to AGR2 as compared to the rest of the hit compounds as shown in Fig. 4. The smaller deviation curve indicates higher stability and the opposite implies that of a lesser extent. The RMSD

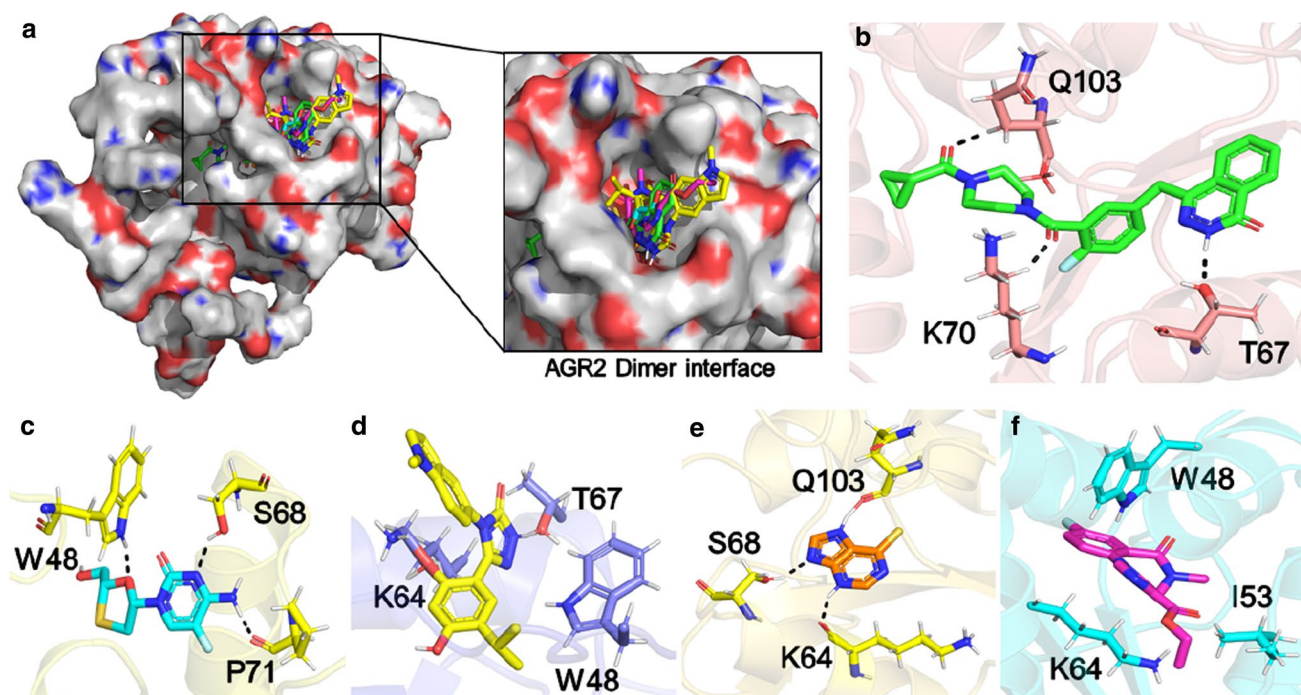


Fig. 3 a-f Binding interaction of top 5 hits (b AGR2–AZD228 complex, c AGR2–emtricitabine complex, d AGR2–ganetespi (STA-9090) complex, e AGR2–mercaptopurine complex, f AGR2–flumaze-

nil complex) docked in complex with AGR2. Hydrogen bonding are shown in black dotted lines

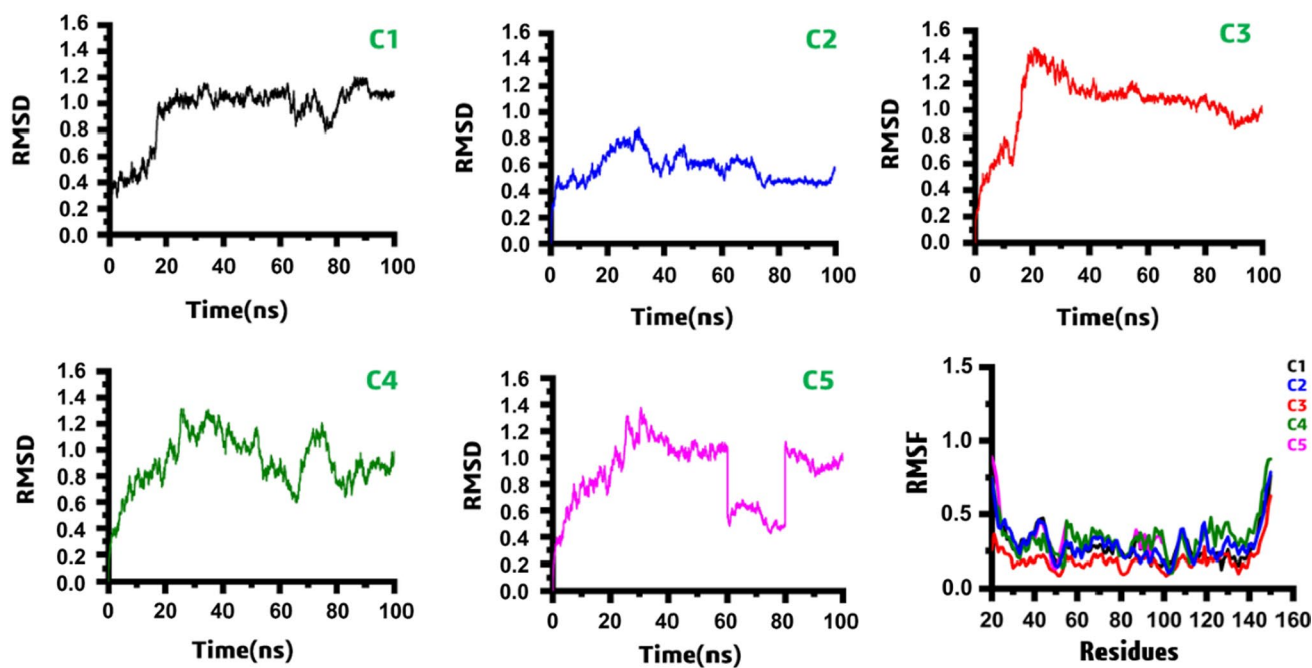


Fig. 4 RMSDs and RMSF of all five top hits in complex with AGR2

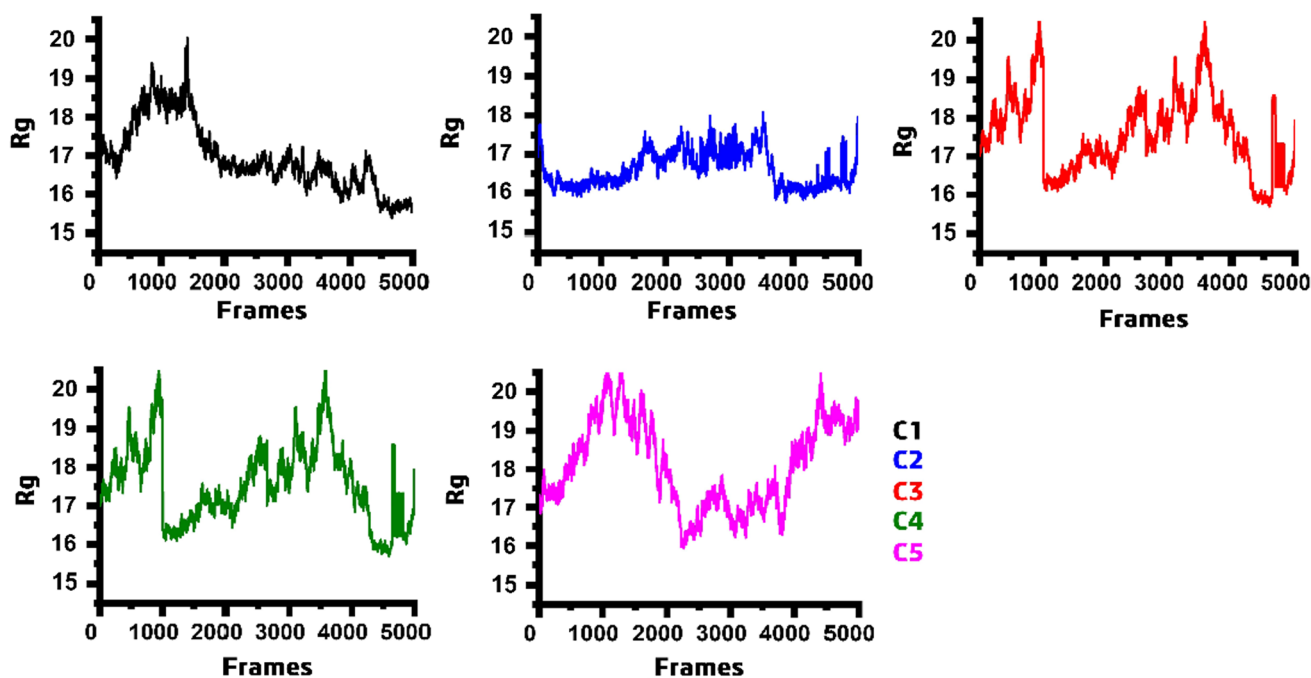


Fig. 5 RoGs of all five top hits in complex with AGR2

of all complexes in comparison with the existing structures demonstrated that a total of 100 ns of MD simulation time was sufficient to achieve equilibration at 310 K.

The RMSD dynamic analysis revealed that the AGR2 protein adopted diverse conformations in each of the 5 AGR2–ligand complexes (Fig. 4). The plot illustrates that for C1, the RMSD initially remained lower until 20 ns but then suddenly converged up to 1.0; then, it remained uniform until 60 ns. Later on, it converged again. Although in C2, the results of RMSD delineated, the AGR2 protein achieved equilibrium at ~80 ns of the simulation time, with lower variation at a specific site, while it continuously oscillated and showed stability in their behavior throughout the entire MD simulation time. Furthermore, the AGR2, C3 complex indicated that the RMSD pattern for backbone variation was steadily decreased initially, but later dynamically increased at 20 ns and remained constant throughout the simulation time (Fig. 4). Whereas in C4 and C5, the RMSD plot showed a similar pattern of backbone deviation, at the beginning, it was gradually decreased and then suddenly fluctuated higher up to 1.2 nm, with continued fluctuations throughout their MD simulation time owing to the instability of the systems as compared to the rest of complexes. Thus, hits of the C1, C4 and C5 complex had a higher affinity for binding to the targeted area, but did not completely diminish the activity of AGR2 protein, while hits of C2 and up to some extent C3 complex strongly bound to the dimerization cavity of AGR2, rendered the protein dynamic behavior steady and, to a larger level, inhibited the activity of the AGR2 protein.

Overall, the RMSD dynamic findings showed that all the (top 5) AGR2 inhibitor drugs had a higher affinity for binding to the intended location, but the hit compound of C2 complex bound to the protein more strongly than the others and stabilized the protein effectively.

To gain a better understanding of the impact of specific residues, that might provide information on residues flexibility in case of all ligand complexes of AGR2, i.e., C1–C5, RMSFs (root mean square fluctuations) of backbone C α were calculated and compared to gain insights into the dynamics association caused by drug binding of protein motions. The greater the RMSF value, the more flexible the region is, while the smaller the RMSF value, the less it moves from its average position during the simulation. We analyzed the C α RMSF for all of the proteins' side-chain atoms that frequently creates the structure of protein with high or less flexibility and provides insights into per residue flexibility (Fig. 4). To increase the accuracy of the analysis, we have retrieved the lowest/minimum energy structure coordinates from the equilibrium phase, and the structure was then aligned and was used as a reference to calculate the RMSF. The fluctuation is negatively correlated with the stability of the residues, i.e., the larger residue fluctuation shows the instability and vice versa. The results of RMSF indicated that fluctuation of the AGR2 protein significantly decreased in the presence of hit compounds of the complex C2 and C3, whereas the fluctuation patterns of C1, C4 and C5 implied different behaviors and led to excessive flexibility and, hence,

instability (Fig. 4). This may correspond to the different binding modes of C1, C4 and C5, which resulted in greater movement of residues and high flexibility to better match in the binding site to achieve the optimal binding mode. Thus, the overall results indicated that hit compounds of the complex C2 and C3 acted as better inhibitors than the rest of the compounds and increased the binding affinity for the desired protein.

Moreover, the compactness of the AGR2 protein in all complexes was assessed by using the radius of gyration (Rg). As a consequence of residue fluctuation and backbone variation, a more thorough examination of overall compactness in all complexes was needed. The Rg analysis demonstrated that all complexes of AGR2, i.e., C1–C5, showed a distinct pattern of compactness, as shown in Fig. 5. Significantly, the Rg of the C1–C2 complexes remained stable during the MD simulation, indicating a strong compact conformation, whereas C3–C5 complexes were found to be less compact over time. The dynamic behavior of the protein–ligand complexes showed that the binding altered the stability and residual flexibility, thus induced the therapeutic cloud.

Conclusion

In this report, we illustrated that the *in silico* virtual screening framework is an important tool for identifying hit compounds against AGR2–AGR2 homodimer. We reported a structure-based virtual screening protocol for identifying new AGR2–AGR2 homodimer modulators. Starting from the homology modeling of AGR2 protein, followed by the virtual screening of FDA-approved compounds, only 35 compounds were identified as a potential AGR2 ligands. Further validation by binding interactions and stability assessment resulted in only five FDA-approved drugs as new AGR2 modulators. Outcome of this study needs to be further ascertained by *in vitro* experimental validation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11030-021-10263-x>.

Acknowledgements All the authors acknowledge the "OpenEye Scientific Software" for providing a free academic license to perform *in silico* studies. The principal author is grateful to Professor Huchen Zhou and Professor Dawei Li for deep insights into the project.

Funding The lead author Shafi Ullah acknowledges China Scholarship Council (CSC) for funding this project.

Declarations

Conflict of interest The authors declare no conflict of interest with anyone.

References

- Park SW, Zhen G, Verhaeghe C, Nakagami Y, Nguyenvu LT, Barczak AJ, Killeen N, Erle DJ (2009) The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. *Proc Natl Acad Sci* 106:6950–6955. <https://doi.org/10.1073/pnas.0808722106>
- Tiemann K, Garri C, Lee SB, Malihi PD, Park M, Alvarez RM, Yap LP, Mallick P, Katz JE, Gross ME (2019) Loss of ER retention motif of AGR2 can impact mTORC signaling and promote cancer metastasis. *Oncogene* 38:3003–3018. <https://doi.org/10.1038/s41388-018-0638-9>
- Jia M, Guo Y, Zhu D, Zhang N, Li L, Jiang J, Dong Y, Xu Q, Zhang X, Wang M (2018) Pro-metastatic activity of AGR2 interrupts angiogenesis target bevacizumab efficiency via direct interaction with VEGFA and activation of NF- κ B pathway. *Biochim. Biophys. Acta. Mol Basis Dis* 1864:1622–1633. <https://doi.org/10.1016/j.bbadis.2018.01.021>
- Chevet E, Fessart D, Delom F, Mulot A, Vojtesek B, Hrstka R, Murray E, Gray T, Hupp T (2013) Emerging roles for the pro-oncogenic anterior gradient-2 in cancer development. *Oncogene* 32:2499–2509. <https://doi.org/10.1038/onc.2012.346>
- Brychtova V, Vojtesek B, Hrstka R (2011) Anterior gradient 2: a novel player in tumor cell biology. *Cancer Lett* 304:1–7. <https://doi.org/10.1016/j.canlet.2010.12.023>
- Negi H, Merugu SB, Mangukiya HB, Li Z, Zhou B, Sehar Q, Kamle S, Mashausi DS, Wu Z, Li D (2019) Anterior Gradient-2 monoclonal antibody inhibits lung cancer growth and metastasis by upregulating p53 pathway and without exerting any toxicological effects: a preclinical study. *Cancer Lett* 449:125–134. <https://doi.org/10.1016/j.canlet.2019.01.025>
- Patel P, Clarke C, Barraclough DL, Jowitt TA, Rudland PS, Barraclough R, Lian LY (2013) Metastasis-promoting anterior gradient 2 protein has a dimeric thioredoxin fold structure and a role in cell adhesion. *J Mol Biol* 425:929–943. <https://doi.org/10.1016/j.jmb.2012.12.009>
- Higa A, Mulot A, Delom F, Bouchecareilh M, Nguyễn DT, Bois-menu D, Wise MJ, Chevet E (2011) Role of pro-oncogenic protein disulfide isomerase (PDI) family member anterior gradient 2 (AGR2) in the control of endoplasmic reticulum homeostasis. *J Biol Chem* 286:44855–44868. <https://doi.org/10.1074/jbc.M111.275529>
- Guo H, Zhu Q, Yu X, Merugu SB, Mangukiya HB, Smith N, Li Z, Zhang B, Negi H, Rong R, Cheng K, Li D (2017) Tumor-secreted anterior gradient-2 binds to VEGF and FGF2 and enhances their activities by promoting their homodimerization. *Oncogene* 36:5098
- Guo H, Chen H, Zhu Q, Yu X, Rong R, Merugu SB, Mangukiya HB, Li D (2016) A humanized monoclonal antibody targeting secreted anterior gradient 2 effectively inhibits the xenograft tumor growth. *Biochem Biophys Res Commun* 475:57–63. <https://doi.org/10.1016/j.bbrc.2016.05.033>
- Garri C, Howell S, Tiemann K, Tiffany A, Jalali-Yazdi F, Alba MM, Katz JE, Takahashi TT, Landgraf R, Gross ME, Roberts RW (2018) Identification, characterization and application of a new peptide against anterior gradient homolog 2 (AGR2). *Oncotarget* 9:27363. <https://doi.org/10.18632/oncotarget.25221>

12. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, Doig A, Guilliams T, Latimer J, McNamee C, Norris A (2019) Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* 18:41–58
13. Rani J, Silla Y, Borah K, Ramachandran S, Bajpai U (2019) Repurposing of FDA-approved drugs to target MurB and MurE enzymes in *Mycobacterium tuberculosis*. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2019.1637280>
14. Dakshanamurthy S, Issa NT, Assefnia S, Seshasayee A, Peters OJ, Madhavan S, Uren A, Brown ML, Byers SW (2012) Predicting new indications for approved drugs using a proteochemometric method. *J Med Chem* 55:6832–6848. <https://doi.org/10.1021/jm300576q>
15. Khan SU, Ahemad N, Chuah L-H, Naidu R, Htar TT (2019) Sequential ligand-and structure-based virtual screening approach for the identification of potential G protein-coupled estrogen receptor-1 (GPER-1) modulators. *RSC Adv* 9:2525–2538. <https://doi.org/10.1039/C8RA09318K>
16. Gu X, Wang Y, Wang H, Wu H, Li W, Wang J, Li N (2020) Homology modeling, molecular dynamics and virtual screening of endothelin-A receptor for the treatment of pulmonary arterial hypertension. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1772106>
17. Sun HP, Jiang ZY, Zhang MY, Lu MC, Yang TT, Pan Y, Huang HZ, Zhang XJ, You QD (2014) Novel protein–protein interaction inhibitor of Nrf2–Keap1 discovered by structure-based virtual screening. *Med Chem Comm* 5(1):93–98. <https://doi.org/10.1039/C3MD00240C>
18. Venselaar H, Joosten RP, Vroiling B, Baakman C, Hekkelman ML, Krieger E, Vriend G (2010) Homology modelling and spectroscopy, a never-ending love story. *Eur Biophys J* 39:551–563. <https://doi.org/10.1007/s00249-009-0531-0>
19. Krieger E, Joo K, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins Struct Funct Bioinf* 77:114–122. <https://doi.org/10.1002/prot.22570>
20. Krieger E, Koraimann G, Vriend G (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins Struct Funct Bioinf* 47:393–402. <https://doi.org/10.1002/prot.10104>
21. Xu D, Zhang Y (2011) Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J* 101(10):2525–2534. <https://doi.org/10.1016/j.bpj.2011.10.024>
22. Hawkins PC, Skillman AG, Warren GL, Ellingson BA, Stahl MT (2010) Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model* 50(4):572–584. <https://doi.org/10.1021/ci100031x>
23. McGann M (2011) FRED pose prediction and virtual screening accuracy. *J Chem Inf Model* 51(3):578–596. <https://doi.org/10.1021/ci100436p>
24. Khan A, Sayed SA, Khan MT, Saleem S, Ali A, Suleman M, Babar Z, Shafiq A, Khan M, Wei DQ (2020) Combined drug repurposing and virtual screening strategies with molecular dynamics simulation identified potent inhibitors for SARS-CoV-2 main protease (3CLpro). *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1779128>
25. Roe DR, Cheatham TE (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* 9(7):3084–3095. <https://doi.org/10.1021/ct400341p>
26. Consortium, Uniprot (2010) The universal protein resource (UniProt) in 2010. *Nucleic Acids Res* 38:142–148. <https://doi.org/10.1093/nar/gkp846>
27. Apweiler R, Bairoch A, Wu CH (2004) Protein sequence databases. *Curr Opin Chem Biol* 8:76–80. <https://doi.org/10.1016/j.cbpa.2003.12.004>
28. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zheng Z, Webb M, David JL (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
29. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292(2):195–202. <https://doi.org/10.1006/jmbi.1999.3091>
30. Xia J, Hu H, Xue W, Wang XS, Wu S (2018) The discovery of novel HDAC3 inhibitors via virtual screening and in vitro bioassay. *J Enzyme Inhib Med Chem* 33:525–535. <https://doi.org/10.1080/14756366.2018.1437156>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.