Journal of Medicinal Chemistry

pubs.acs.org/jmc

Article

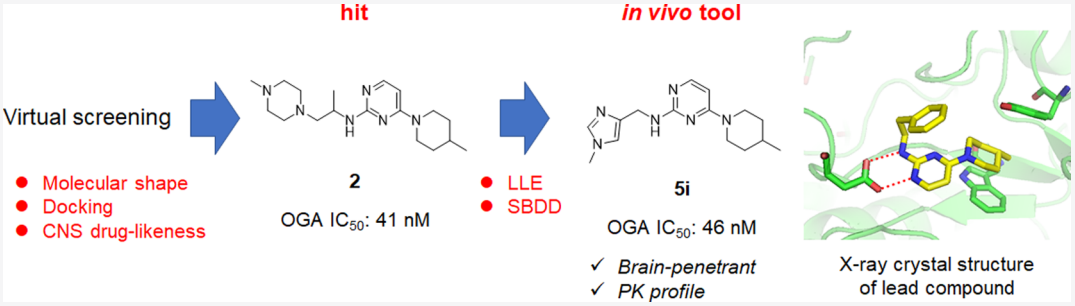
Discovery of a Novel and Brain-Penetrant *O*-GlcNAcase Inhibitor via Virtual Screening, Structure-Based Analysis, and Rational Lead Optimization

Michiko Tawada,* Makoto Fushimi, Kei Masuda, Huikai Sun, Noriko Uchiyama, Yohei Kosugi, Weston Lane, Richard Tjhen, Satoshi Endo, and Tatsuki Koike

Cite This: *J. Med. Chem.* 2021, 64, 1103−1115

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: *O*-GlcNAcase (OGA) has received increasing attention as an attractive therapeutic target for tau-mediated neurodegenerative disorders; however, its role in these pathologies remains unclear. Therefore, potent chemical tools with favorable pharmacokinetic profiles are desirable to characterize this enzyme. Herein, we report the discovery of a potent and novel OGA inhibitor, compound **5i**, comprising an aminopyrimidine scaffold, identified by virtual screening based on multiple methodologies combining structure-based and ligand-based approaches, followed by sequential optimization with a focus on ligand lipophilicity efficiency. This compound was observed to increase the level of *O*-GlcNAcylated protein in cells and display suitable pharmacokinetic properties and brain permeability. Crystallographic analysis revealed that the chemical series bind to OGA via characteristic hydrophobic interactions, which resulted in a high affinity for OGA with moderate lipophilicity. Compound **5i** could serve as a useful chemical probe to help establish a proof-of-concept of OGA inhibition as a therapeutic target for the treatment of tauopathies.

INTRODUCTION

Aggregation of the tau protein is one of the hallmarks of the progression of neurodegenerative disorders, such as Alzheimer's disease, Pick's disease, and Parkinson's disease.[1,2] Results from recently published studies indicate that *O*-GlcNAcylation, which consists of the glycosylation of serine or threonine residues with *O*-linked N-acetylglucosamine (*O*-GlcNAc), plays an important role in aggregation of the tau protein.[3,4] *O*-GlcNAcylated tau protein displays improved stability and solubility compared with its unmodified state, which in turn leads to a suppression of the aggregation of the tau protein. The results of some clinical studies indicated that the *O*-GlcNAcylation level of tau protein decreases in patients with Alzheimer's disease.[5,6] Notably, *O*-GlcNAcylation is regulated by two enzymes: *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA).[7,8] OGT catalyzes the addition of *O*-GlcNAc to protein substrates, whereas OGA catalyzes the cleavage of *O*-GlcNAc from the *O*-glycosylated proteins. Since inhibition of OGA has been reported to increase the level of *O*-GlcNAcylation of tau proteins, OGA has been the object of increasing attention as a potential therapeutic target for the treatment of tau-related neurodegenerative diseases.

OGA consists of two domains: a C-terminal domain displaying histone acetyltransferase activity and an N-terminal domain displaying *O*-GlcNAcase activity. The N-terminal domain, which is classified into the glycoside hydrolase family 84 (GH84),[9] breaks the glycosidic bond of the *O*-GlcNAc group attached to proteins. Crystal structures of OGA in complex with substrate analogues acting as enzyme inhibitors determined by X-ray diffraction have been reported over the past few decades. These structures enabled us to understand the interactions between OGA and its protein substrates as well as the catalytic mechanism of the hydrolysis of *O*-

Received: September 29, 2020
Published: January 6, 2021

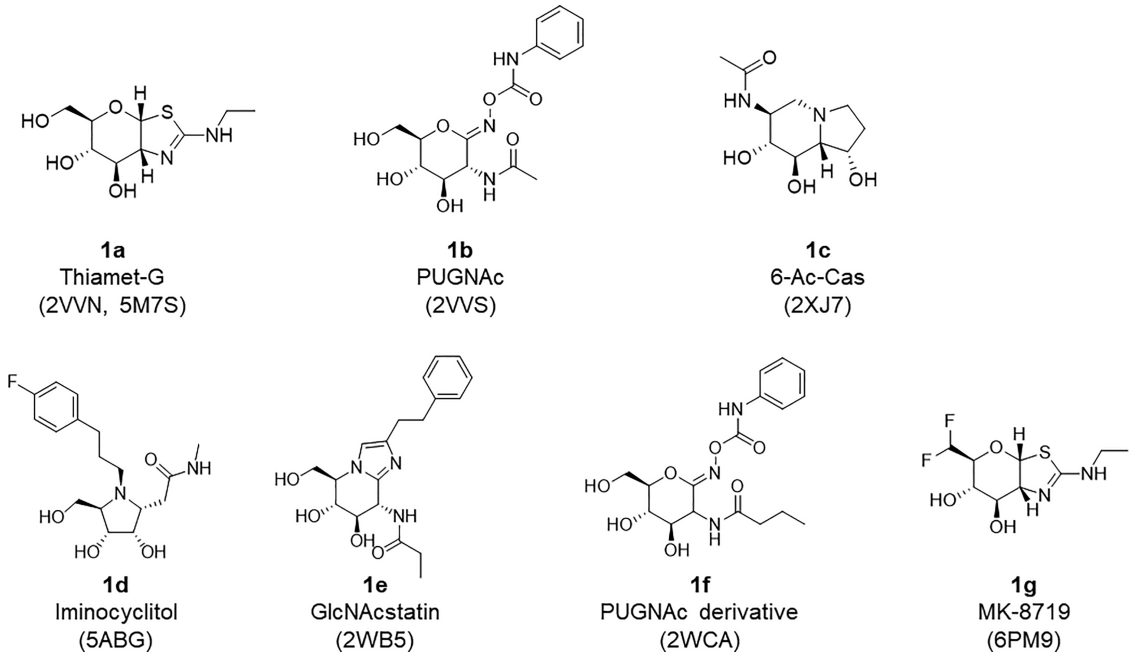


Figure 1. Reported inhibitors of *O*-GlcNAcase (OGA). Crystal structures of the various compounds in complex with bacterial OGAs (human and bacterial in the case of inhibitor **1a**) are indicated, with their corresponding letter designations specified in brackets.

glycosylated proteins.[10−12] *O*-GlcNAcylated proteins bind to OGA through hydrogen bonds with polar residues such as Asp285 and Asn313. The residue pair of Asp174 and Asp175 assists the hydrolysis reaction by donating and accepting electrons and protons with the substrate.

Based on the structural insights thus achieved, several OGA inhibitors have been discovered to date, with some representative compounds depicted in Figure 1.[13−18] Notably, most of these inhibitors possess sugar-like substructures that mimic the natural substrate of the enzyme, *O*-GlcNAc. As these compounds comprise a number of hydroxy groups in their sugar-like scaffold, these inhibitors might exhibit a high polarity, which might be a concern about low brain penetration. In 2019, a sugar-mimetic OGA inhibitor MK-8719, with improved pharmacokinetic profile and brain penetration, was reported as a compound to be tested in a clinical trial;[19] however, non-sugar-like compounds with high potency and better pharmacokinetic profiles are still awaited to help characterize OGA.

As X-ray crystal structures of OGA in complex with several OGA inhibitors have been reported, hit compounds can be explored by virtual screening efficiently.[20] Structure-based virtual screening using crystal structures has been widely performed;[21] in fact, a study of hit discovery for OGA by docking-based virtual screening was published in 2019.[22] In it, multiple docking-based methods with different accuracy levels were implemented sequentially, resulting in the identification of a non-sugar-type OGA inhibitor. However, this compound exhibited lower OGA inhibitory potency than traditional sugar-type inhibitors, which suggests that the development of a new strategy could be beneficial to the search for potent and novel OGA inhibitors. Indeed, known OGA inhibitors allow us to implement a ligand-based approach, such as similarity search for analogues or bioisosteres.[23] When three-dimensional (3D) conformations of query compounds are determined or predicted, pharmacophore search or shape similarity search can also be applied.[24] Ligand-based virtual screening allows active compounds to be identified with a high hit rate because this type of method explores compound libraries focusing on common chemical features or pharmacophores with known active compounds, based on the fact that compounds displaying structural similarities with respect to the mentioned features often have similar biological profiles. A structure-based approach sequentially added to the described ligand-based approach could be a powerful method for identifying potentially active compounds for unfavorable contact with amino acids or quantitated binding affinity to the target protein. Therefore, we applied a multiple hit exploration approach combining a structure-based method and a ligand-based one. In the first step of the approach, hit compounds were extracted by ligand-shape-based virtual screening, and thus, the obtained compounds were subsequently narrowed down by conducting a sequential docking study. Furthermore, filtering of the compounds by physicochemical parameters for druglike properties and central nervous system (CNS) penetration was incorporated in the approach to enrich the qualities of a compound set to be examined in biological assays. The identified hit compounds were optimized to improve their pharmacokinetic profile focusing on ligand lipophilicity efficiency (LLE).[25−27] By this approach, we were able to identify a brain-penetrable in vivo tool compound. We hereby report the discovery of a novel in vivo tool compound starting from hit identification achieved by computer-aided virtual screening and subsequent hit/lead optimization.

RESULTS AND DISCUSSION

Homology Modeling. Although about 20 X-ray crystal structures of human OGA have been reported so far,[12,19,28] only the crystal structures of OGAs from *Bacteroides thetaiotaomicron* (BT_4395), *Clostridium perfringens*, and *Oceanicola granulosus* had been published at the time of the present study.[13−18,29] Among these three bacterial proteins, the amino acid sequence of BT_4395 has the highest identity with the sequence of human OGA. Thus, BT_4395 was selected for

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

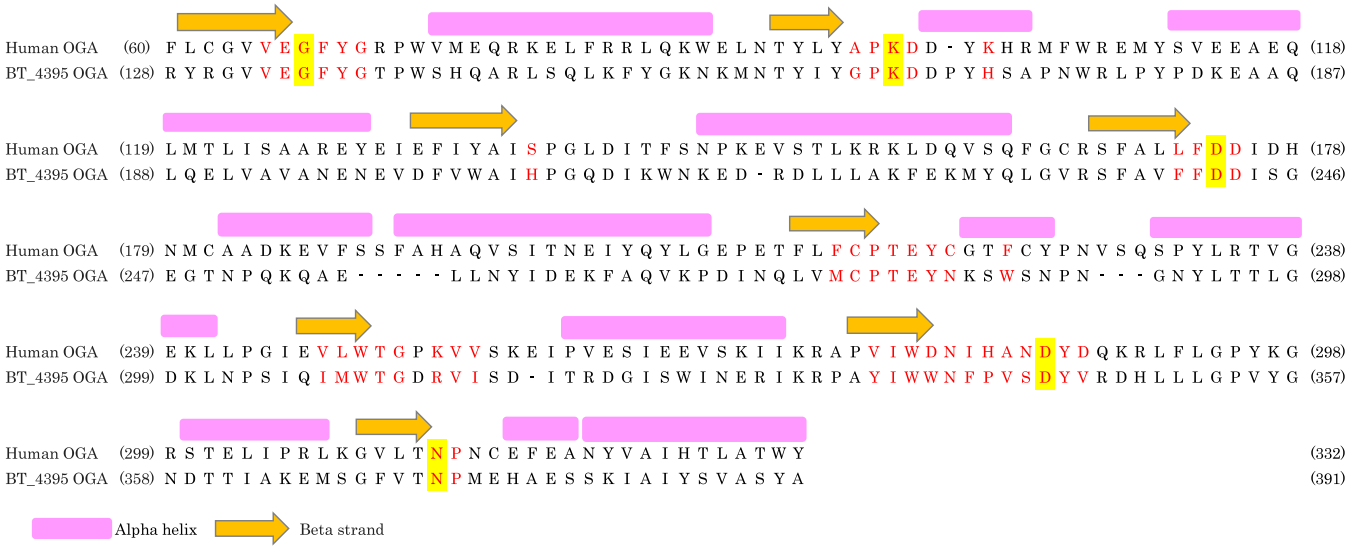


Figure 2. Alignment of the amino acid sequences between human *O*-GlcNAcase (OGA) and bacterial BT_4395 OGA. Amino acid residues known to be within 8 Å of the ligand Thiamet-G in the crystal structures of the complexes between human OGA and the said substrate analogue and BT_4395 OGA and the same ligand are colored in red; moreover, the residues involved in direct interactions with Thiamet-G are highlighted in yellow.

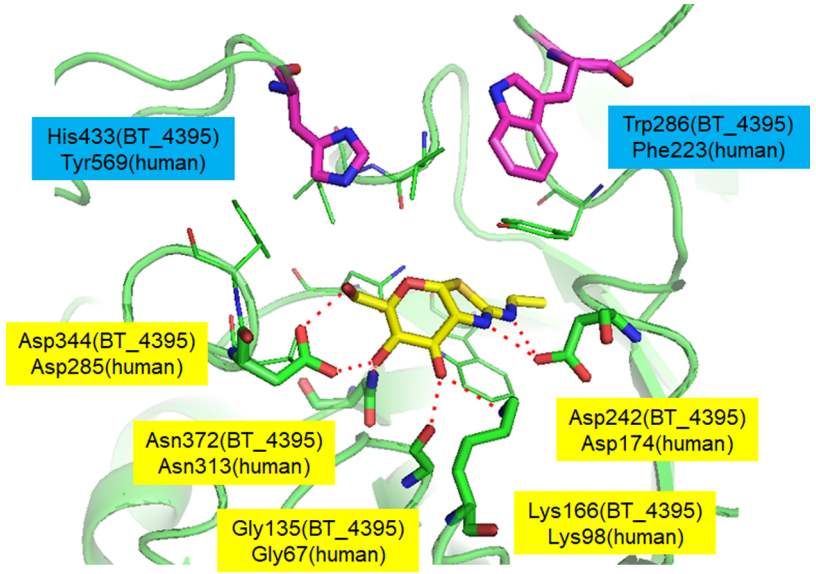


Figure 3. Interactions between BT_4395 *O*-GlcNAcase (OGA) and Thiamet-G (PDB ID: 2VVN). Hydrogen bonds are indicated by red dashed lines. Amino acid residues interacting directly with Thiamet-G are made evident by green-colored thick sticks and their letter designations are highlighted in yellow boxes. Amino acid residues that are not conserved between BT_4395 OGA and human OGA are indicated by magenta-colored stick, with their letter designations highlighted in cyan boxes.

sequence and structure analysis to compare with human OGA. Notably, the amino acid identity between human OGA and BT_4395 OGA in the catalytic domain in N-terminal domains of these proteins was 33%, whereas the X-ray crystal structure of BT_4395 OGA in complex with an OGA inhibitor, Thiamet-G (PDB code: 2VVN[13]), indicated an approximately 70% conservation of amino acid residues within 8 Å from Thiamet-G (Figure 2). Furthermore, all of the seven amino acid residues interacting with Thiamet-G are identical between human and BT_4395 OGA: Asp174, Asp285, Asn313, Gly67, Lys98, Tyr219, and Trp278 (human OGA numbering). The major differences around the ligand binding sites are Trp286 (BT_4395 OGA)/Phe223 (human OGA) and His433 (BT_4395 OGA)/Tyr569 (human OGA) (Figure 3). Since these residues are located around the entrance to the substrate-binding pocket and the values for their B-factors are higher than those of other residues, the residues specified above should only make a small contribution to protein−ligand binding. This speculation is supported by the crystal structures of OGA in complex with many potent inhibitors, which bind to the catalytic site in the absence of interactions with Trp286 and His443. Therefore, a homology model of human OGA based on the structure of bacterial BT_4395 OGA is applicable with enough accuracy for docking-based virtual screening. Thus, we constructed a homology model of human OGA using the crystal structure of BT_4395 OGA in complex with Thiamet-G (PDB code: 2VVN) as a template. In fact, by superposing and comparing the structure of BT_4395 OGA in

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

complex with Thiamet-G (PDB code: 2VVN[14]) to that of human OGA in complex with Thiamet-G (PDB code: 5M7S[28]), it can be concluded that not only the backbones of the two proteins but also the side chains around the ligand binding sites adopted essentially identical conformations.

Virtual Screening. As mentioned above, a number of sugar-derived OGA inhibitors are currently known. However, their high polarity might be a concern from the standpoint of pharmacokinetics. To estimate the possibility of identifying hit compounds with moderate lipophilicity by virtual screening, a protein druggability assessment was carried out to examine the profile of the ligand binding site of OGA using SiteMap[30,31] module in Maestro.[32] SiteMap computes SiteScore and DScore values in terms of the size, tightness, and hydrophobic and hydrophilic characters of a pocket. The values of SiteScore and DScore for the said pocket in the structure of the complex between BT_4395 OGA and Thiamet-G (PDB code: 2VVN) were 1.15 and 1.20, respectively, which suggested that OGA has a highly druggable pocket. Therefore, a hit exploration strategy starting from a ligand-based approach followed by the reranking of the deliverables by a sequential docking-based approach could identify high-quality hit compounds with lead-like properties.

The crystal structures of the complexes between OGA and several known OGA inhibitors have been reported over the past decade. The active conformations of these ligands, as determined through the said crystal structures, highlight the steric and electrostatic features of the ligand that induce OGA inhibition. Therefore, a 3D ligand-based approach utilizing these active conformations was applied as the first step of the hit exploration process. Multiple ligands used as queries of the ligand-based approach could expand the chemical space to be explored. We selected six inhibitors, compounds **1a**–**f**, that exhibited high potency and for which the crystal structure of their complexes with OGA were available (Figure 1). Although the crystal structures of human OGA had not yet been published at the time of the present study, it was speculated that there was only small difference between the protein conformations around the active site of human OGA and bacterial OGA according to structural comparison analysis mentioned in the previous section, which indicated that the bound conformations of the ligands are almost identical. Therefore, we used the coordinates from the crystal structures in complex with bacterial OGA. The 3D ligand-based approaches mainly fall into two categories: the molecular shape similarity-based approach and the pharmacophore-based approach. Given that the reported OGA inhibitors have characteristic sp^3-rich frameworks and bind to the pocket via several hydrogen bonds with amino acid residues, both the shape and the chemical characteristics of a molecule would be important factors that obtain binding affinity of the said molecule to OGA. Based on the fact that the utilization of ROCS, one of standard methods for shape similarity searches, yields overlay patterns whereby not only the shape of the molecule but also their chemical features are matched,[33,34] ROCS was used for ligand-based virtual screening.

In the first stage of the virtual screening, a 3D shape similarity search of a corporate library was carried out against six query compounds using ROCS. Subsequently, we conducted the docking of approximately 100 000 compounds selected by 3D molecular shape to prioritize compounds exhibiting favorable contacts or forming additional interactions with amino acid residues. According to the published crystal structures, OGA inhibitors form hydrogen bonds with polar amino acids, such as Asp285, Asp174, or Asn313. This evidence led us to pick out 70 000 compounds that formed hydrogen bonds with at least one residue among the mentioned three. Finally, to focus on compounds with CNS-drug likeness, the extracted compounds were filtered based on the range of physicochemical properties used to define the CNS multiparameter optimization (CNS-MPO) index[35] (molecular weight $\leq$ 350, $A\log P \leq 5$, number of hydrogen donors ≤ 4, and topological polar surface area ≤ 125) and the indicators of lead likeness that are not included in the CNS-MPO index (number of hydrogen acceptors ≤ 8 and number of rotatable bond ≤ 6); by this process, 55 000 compounds were retained. This set of compounds was subsequently filtered applying the pan assay interference compound (PAINS) filter. To submit a diverse compound set to our biological assay of choice, the compounds were divided into 2681 clusters based on molecular similarity and a total of 2681 compounds, the top-ranking compounds of each cluster, were subjected to a biological assay. The workflow of the virtual screening is depicted in Figure 4.

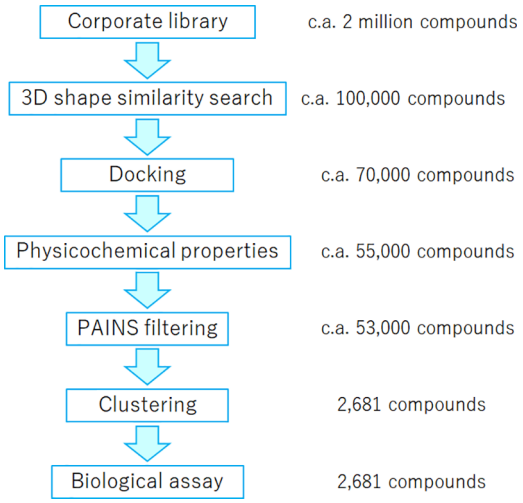
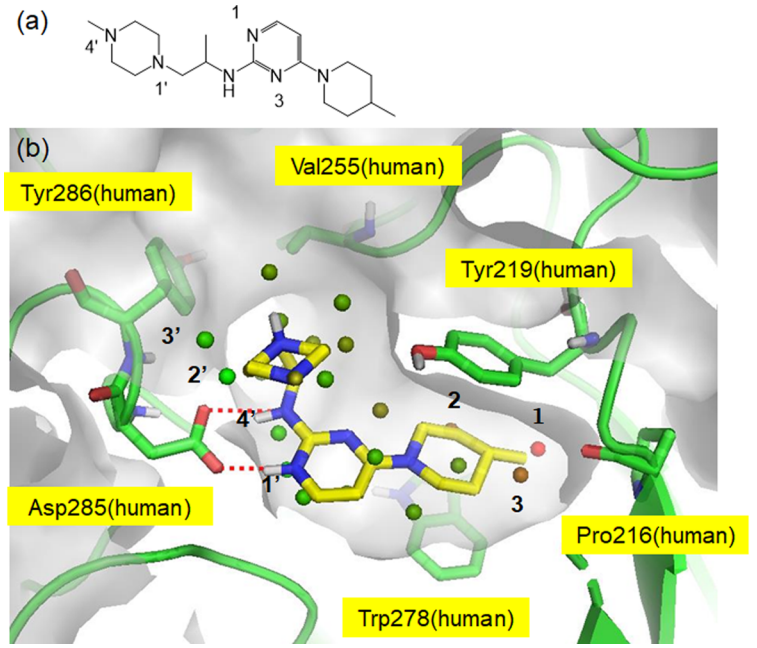


Figure 4. Workflow of virtual screening.

Structure-Based Analysis of Hit Compound. The OGA inhibitory activities of the 2681 compounds were measured using the full-length human OGA protein and 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide dihydrate as fluorogenic substrate. The initial screening at 10 μM concentration led to the identification of 13 compounds displaying more than 50% inhibition. The subsequent dose–response assay allowed us to identify compound **2** (Figures 5a and S1), which exhibited a potent inhibitory activity with a measured IC_{50} value of 41 nM. Compound **2** was confirmed to exhibit target-specific inhibitory activity toward OGA via performance of a counter-assay involving β-hexosaminidase, which, similar to OGA, catalyzes the hydrolysis of *O*-glycosylated proteins. Notably, **2** did not exhibit any inhibitory activity against β-hexosaminidase at 10 μM concentration. To begin lead optimization utilizing the structure-based drug design, the docking model of compound **2** was constructed (Figure 5b). The binding mode obtained in the virtual screening was carefully checked to compare each generated binding pose of all of the stereoisomers, *R*-form and *S*-form, and protonation states that exist in a possible pH range. Notably, the pyrimidine

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103–1115



(c)

Water ID	ΔH	-TΔS	ΔG	#HB (PW)
1	4.03	3.65	7.68	0.55
2	2.58	2.05	4.63	0.83
3	1.48	2.94	4.42	0.84
4'	-3.39	1.91	-1.48	1.12
3'	-2.62	1.02	-1.60	0.00
2'	-6.50	2.94	-3.56	0.98
1'	-5.69	1.36	-4.33	2.06

ΔH: Enthalpy
TΔS: Entropy
ΔG: Free energy
#HB(PW): Average hydrogen bonds between protein and waters

Figure 5. (a) Chemical structure of compound **2**. (b) Docking model of **2** in *S*-configuration into human *O*-GlcNAcase (OGA) homology model and WaterMap analysis for the docked protein. The most stable water molecules, which exhibit the lowest Δ*G* values, are colored in green, and the most unstable water molecules, which exhibit the highest Δ*G* values, are colored in red. The other water molecules are color-coded from green to red by gradation according to their Δ*G* values. Notably, the unstable water molecules are labeled as 1, 2, and 3, and the stable ones are labeled as 1′, 2′, 3′, and 4′. (c) Enthalpy, entropy, and free-energy values for water molecules as computed by WaterMap.

and the piperazine rings could exist in three and two distinct possible protonation states, respectively: the cationic states whereby the N-1 or the N-3 positions are protonated and the neutral state, for the pyrimidine; the cationic states whereby the N-1′ or the N-4′ positions are protonated, for the piperazine. The pK_a values of the nitrogen atoms are 4.9 for N(1) and 6.7 for N(3) in the case of the pyrimidine ring and 8.2 for N(1′) and 8.4 for N(4′) in the case of the piperazine ring.[32,36] As the acidity might be relatively high around the pyrimidine ring due to the presence of the acidic amino acids, Asp285 and Asp174, in the active site of OGA, even the protonated form of N(1) could exist in a stable form in this specific region. Based on the results of the docking study performed taking into consideration all possible combinations of the two stereoisomers and various protonation states by LigPrep,[32] the *S*-isomer of **2**, whereby N(1) and N(4′) were protonated, exhibited the most reasonable binding mode. In this docking mode, the N(1) of the pyrimidine and the neighboring amine form a bidentate hydrogen bond with the carboxyl acid group of Asp285. The methylpiperidine group binds to the deep pocket stacking with Tyr219 and Trp278. The piperazine group is positioned at the entrance of the ligand binding site, where Val255 and Tyr286 are located, without getting involved in direct interactions with any amino acid residues.

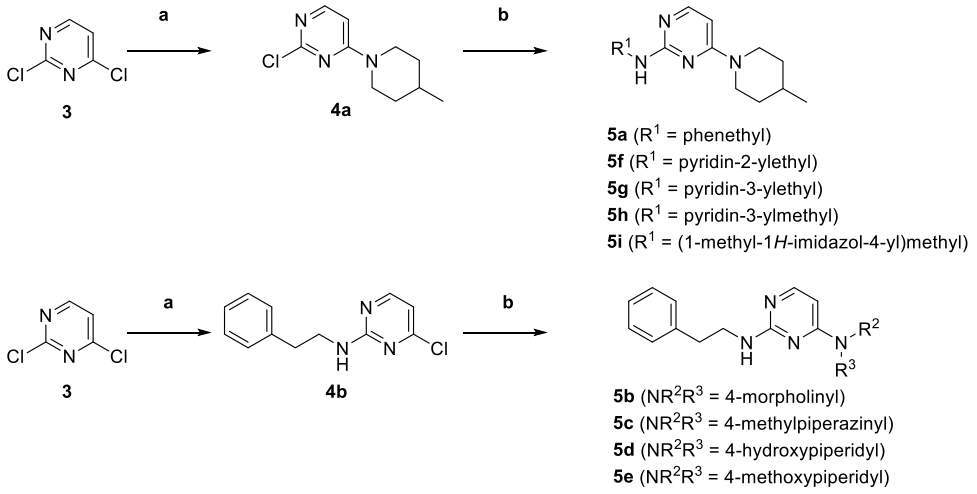
To elucidate the properties of the ligand binding site and the protein−ligand interaction in more detail, a WaterMap calculation was applied to the docking model of compound **2** with the hOGA homology model.[37,38] In this approach, water molecules present in a ligand binding site of the protein are displaced and their free-energy values are computed. WaterMap provides not only the values for enthalpy (Δ*H*) and entropy (−*T*Δ*S*), which are the components of free energy (Δ*G*), but also the number of hydrogen bonds between protein and water [#HB(PW)], so that it provides physicochemical insight of the binding site.

The predicted thermodynamic parameters of the water molecules are reported in Figure 5b,c. According to the result of WaterMap calculations, the three most unstable water molecules, 1, 2, and 3, with Δ*G* values in the range of 4−8 kcal/mol gather in the deep site consisting of Tyr219 and Trp278. Since the displacement of unstable water molecules to the bulk solution leads to a gain in free energy, this deep site makes a large contribution to the protein−ligand binding. As speculated by the surrounding hydrophobic side chains of Tyr219 and Trp278, SiteMap also suggested this site to be hydrophobic; this site is mainly covered with the hydrophobic surface in the surface map drawn by SiteMap (Figure S2). In spite of the presence of the said hydrophobic pocket, the #HB(PW) values of the three water molecules ranged from 0.5 to 0.9, which means that these water molecules were on average engaged in 0.5−0.9 hydrogen bond interactions during the simulation. The oxygen atom in the main chain of Pro216, which is slightly facing the pocket surface, could be a hydrogen bond counter at part of simulation time. Although the deep hydrophobic pocket displays a tendency to prefer hydrophobic substituents of the ligand, the properties of the water molecules as computed by WaterMap indicate that a heteroatom is also acceptable.

On the other hand, the four most stable water molecules with Δ*G* values lower than −1 kcal/mol are observed near Asp285. The water molecules in this site show a large enthalpy gain (Δ*H* ranged from −7 to −2 kcal/mol); moreover, here enthalpy is related mainly to nonbonding interaction. The water molecules 1′ and 4′ (Figure 5b,c) form hydrogen bonds with Asp285. Since the nitrogen atoms of the aminopyrimidine group overlap well with these water molecules, the amino-pyrimidine group is predicted to find itself in the proper position to form hydrogen bonds with Asp285.

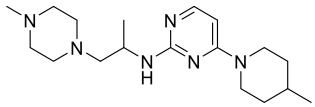
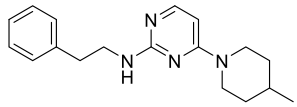
https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

Scheme 1. Synthetic Route of Pyrimidine Derivatives[ab]



[a]Amines, triethylamine, tetrahydrofuran (THF), room temperature. [b]Amines, *N*,*N*-diisopropylethylamine (DIPEA), solvents, heating.

Table 1. Inhibitory Activities of Compounds 2 and 5a against OGA

Compound	Structure	OGA IC_{50} (nM)[a]	CNS-MPO
2		41 (39-43)	5.69
5a		24 (22-26)	4.47

[a]IC_{50} values and 95% confidence intervals (given in parentheses) were analyzed by duplicate measurements with GraphPad Prism V.5 (GraphPad Software, San Diego, CA) statistical analysis software using a nonlinear regression for a sigmoidal dose−response curve with a variable slope.

Water molecules exhibiting moderate values for the free energy (ΔG = −1−1 kcal/mol) were observed around the residues, Val255 and Tyr286, at the entrance of the pocket. This observation indicates that the said entrance is a solvent-accessible region, which allows various substituents. The methyl group attached to the chiral carbon also seems to engage in no specific interactions with amino acids. Consistent with this observation, no unstable water molecules were located in the position of the methyl group, supporting the speculation that this methyl group made no significant contribution to binding affinity. Therefore, this methyl group could be removed from compound **2**, with the ligand maintaining its inhibitory potency toward OGA.

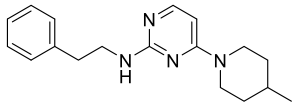
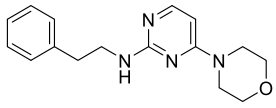
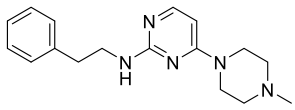
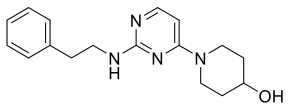
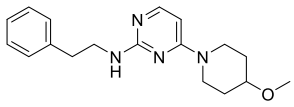
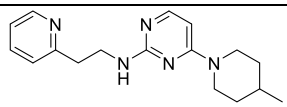
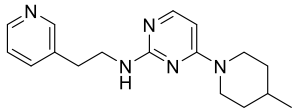
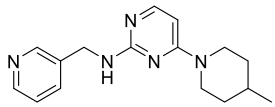
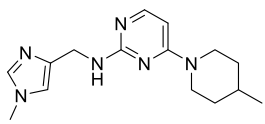
Chemistry. The designed compounds were synthesized according to Scheme 1. 2,4-Diaminopyrimidines **5a**−**i** were synthesized by the sequential nucleophilic substitution of 2,4-dichloropyrimidine **3** and corresponding amines (Scheme 1).

Biological Evaluation. To initiate structure−activity relationship (SAR) studies, a chemically accessible scaffold would be preferable. WaterMap analysis results, reported in Figure 5b,c, suggested that the piperazine group might be replaced by various substituents, with little influence on the affinity. Replacement of the piperazine group with a phenethyl group produced compound **5a**, which retained the inhibitory potency of compound **2** toward OGA (IC_{50} = 24 nM; Table 1). This simple compound was a good starting point for further optimization. To characterize the inhibitory mechanism of the chemical series, a kinetic assay of **5a** was carried out. The substrate-competitive profile of **5a** was investigated by estimating the change in K_m values as a result of changes in compound concentration. As can be observed from Figure S3, increase in the K_m value was associated with increase in **5a** concentration, indicating that **5a** inhibits the OGA activity in a substrate-competitive manner.

Ligand lipophilicity efficiency (LLE) was considered during further optimization. Notably, LLE has been reported to be a useful concept in lead discovery and optimization to make more specific interactions with a target molecule.[25−27] Compound **5a** exhibited a moderate LLE value of 3.5 due to its relatively high lipophilicity ($A\log P$[39,40] = 4.2). We explored the substituent targeting the deep pocket of OGA (Table 2). Although the results of the WaterMap analysis suggested the deep pocket to be mostly hydrophobic, the opportunity still seems to exist for the said pocket to accommodate a hydrophilic substituent (Figure 5b,c). To reduce the lipophilicity of the ligand, 6-membered cyclic amines were introduced, such as morpholine (**5b**; LLE = 2.7) and methylpiperazine (**5c**; IC_{50} > 10 000 nM), that had decreased

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

Table 2. OGA Inhibitory Activities and Ligand Lipophilic Efficiency of Compounds 5a–i[b]

Compound	Structure	OGA IC_{50} (nM)[a]	AlogP	LLE[b]	CNS-MPO
5a		24 (22-26)	4.2	3.5	4.47
5b		4,700 (4,400-5,000)	2.7	2.7	5.77
5c		>10,000	2.9	-	5.71
5d		>10,000	2.4	-	5.33
5e		970 (910-1,000)	2.8	3.2	5.51
5f		360 (350-380)	3.2	3.3	5.63
5g		31 (29-33)	3.0	4.5	5.69
5h		18 (17-19)	2.7	5.1	5.83
5i		46 (41-52)	2.1	5.2	5.83

[a]IC_{50} values and 95% confidence intervals (given in parentheses) were analyzed by duplicate measurements with GraphPad Prism V.5 (GraphPad Software, San Diego, CA) statistical analysis software using a nonlinear regression for a sigmoidal dose–response with a variable slope. [b]LLE = $pIC_{50} - A \log P$.

LLE. Further modification of the piperidine was explored by introducing polar substituents. The introduction of a hydroxy or a methoxy group produced no beneficial effects (4-hydroxypiperidine **5d**; IC_{50} > 10 000 nM, 4-methoxypiperidine **5e**; LLE = 3.2). Improving the LLE by introducing a polar substituent targeting the deep pocket proved difficult to achieve. On the other hand, various modifications could be applied to the phenethyl group, as suggested by the WaterMap calculation data and inhibitory activities of **2** and **5a**. To improve the LLE value, the phenyl ring of **5a** was replaced with a pyridine ring, which is characterized by lower lipophilicity. The 2-pyridine derivative of **5a** (**5f**; LLE = 3.3) had a similar LLE value to **5a**, whereas the 3-pyridine derivative (**5g**; LLE = 4.5) displayed an increased LLE value. The corresponding methylene derivative **5h** (LLE = 5.1) was characterized by an even higher value of LLE and a lower value of $A \log P$. Further modification with nitrogen-containing heterocycles yielded compound **5i** (IC_{50} = 46 nM; LLE = 5.2), which had the

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103–1115

characteristics of a potent OGA inhibitor with a desirable LLE value. The CNS-MPO values of the mentioned compounds were also monitored throughout the optimization. Although the lead compound **5a** was characterized by a modest CNS-MPO value, due to its high lipophilicity, the value of CNS-MPO was observed to increase mainly as a result of the decrease of the *A* log *P* value; ultimately, **5i** was observed to display a substantial increase in the CNS-MPO value.

A detailed characterization of compound **5i** was conducted in terms of its efficacy in cells, its selectivity, and its pharmacokinetic profile. The levels of *O*-GlcNAcylated protein were evaluated by in-cell Western assay. Compound **5i** induced glycosylation in cells with an EC_{50} value of 450 nM for causing an increase of *O*-GlcNAcylated protein to half-maximal level (95% confidence intervals = 190−1100 nM, E_{max} = 143%). In terms of the assessment of selectivity, the in vitro inhibitory activity of **5i** was evaluated against β-hexosaminidase and various kinases. In detail, compound **5i** did not display any significant inhibitory activity against β-hexosaminidase (IC_{50} > 10 μM) and 277 kinases (IC_{50} > 1μM) under the assay condition described previously (Figure S4).[41]

Finally, a pharmacokinetic study of compound **5i** was performed in mice. The high exposure levels of **5i** were observed in the plasma and the brain at the 0.5 and 1 h marks after oral administration to mice (Table 3).

Table 3. Pharmacokinetic Properties of 5i in the Plasma and Brain of Mouse

time (h)	C_{plasma} (μg/mL)	C_{brain} (μg/g)	$K_{p,uu}$[a]
0.5	2.8	2.7	0.25
1.0	1.9	2.6	0.36

[a]Brain-to-plasma unbound concentration ratio ($K_{p,uu}$) calculated from the values for the unbound fraction in plasma ($f_{u,plasma}$) and brain ($f_{u,brain}$) of 0.53 and 0.14, respectively.

Crystal Structure of Compound ***5a****.* The structure of compound **5a** bound to BT_4395 OGA was determined by X-ray crystallography (Figure 6a). The protein structure adopts a folded conformation that is essentially identical to those of the other reported structures of bacterial OGA. Compound **5a** mostly displays a common protein−ligand interaction pattern in the crystal structure compared with that of the human OGA homology model; the aminopyrimidine core was involved in a bidentate hydrogen bonds with Asp344 (corresponding to Asp285 in human OGA). However, one exception was found; a difference between the two structures lies in the conformation of Asp242 (corresponding to Asp174 in human OGA), which is located around the gate region of the deep pocket. Not only the methylpiperidine group bound to the deep pocket stacks with the side chains of Tyr282 and Trp337 (corresponding to Tyr219 and Trp278 in human OGA) as the docking model suggested, but it also kicks out Asp242. Although most reported OGA inhibitors like Thiamet-G were observed to pull in Asp242 to form hydrogen bonds with this residue in the gate region of the deep pocket, the crystal structure of the complex between BT_4395 OGA and compound **5a** suggested that **5a** engaged in a tight hydrophobic interaction with the protein by causing the methylpiperidine group to be placed at the bottom of the deep pocket kicking out Asp242. By utilizing the characteristic hydrophobic interaction, this compound achieved high affinity with OGA despite the reduced number of hydrogen bond donors and the moderate lipophilicity. Notably, an intramolecular $\pi-\pi$ stacking between the pyrimidine core and the terminal benzene ring was also observed in the mentioned crystal structure. This stacking interaction helped in keeping **5a** in a rigid conformation. Replacement of the benzene ring with the pyridine ring (see **5f** and **5g**), which is characterized by lower π-electron density, reduced the $\pi-\pi$ stacking effect.[42] In contrast, incorporation of 3-pyridine to produce **5g** led to the retention of the potency of the inhibitor, possibly because the nitrogen atom at the 3-position of the pyridine ring might interact with the adjacent Tyr219 in human OGA via a water-mediated hydrogen bond. Although the predicted docking model was in good agreement with the experimentally determined structure, with respect to protein−ligand interactions, the crystal structure suggested a more detailed

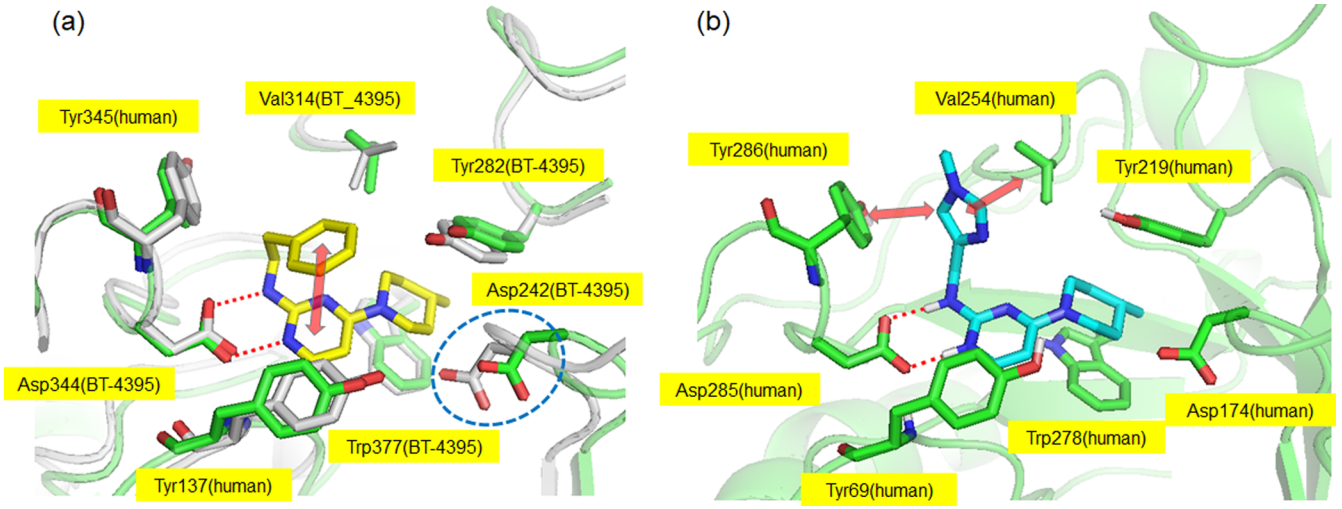


Figure 6. (a) Crystal structures of the complexes between BT_4395 *O*-GlcNAcase (OGA; green) and compound **5a** (yellow) overlaid with the crystal structure (white) of BT_4395 OGA in complex with Thiamet-G (Thiamet-G not shown; PDB code 2VVN), and (b) docking model of **5i** (cyan) into the human OGA homology model (green). Hydrogen bonds are represented by red dashed lines. CH−π or $\pi-\pi$ interactions are identified by red arrows.

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

interpretation for the SAR of the terminal aryl ring, which was attributed to the intramolecular $\pi-\pi$ stacking.

*Docking Model of Compound **5i**.* To interpret the SAR and high LLE of compound **5i**, the docking model of **5i** was constructed based on the crystal structure of **5a** in complex with BT_4395 OGA (Figure 6b). In addition to the common interactions with Asp285, Tyr219, and Trp278 (corresponding to Asp344, Tyr282, and Trp337, respectively, in BT_4395 OGA) observed for **5a**, the terminal methylimidazole moiety was predicted to engage in CH$-\pi$ interactions with Val254 and Tyr286 at the entrance of the pocket. Due to the decreased length of the alkyl linker compared with that of the original compound **5a**, intramolecular $\pi-\pi$ stacking involving the terminal aryl ring should be lost; however, favorable contacts with Val254 and Tyr286 should be gained, so that the inhibitory potency toward OGA is retained. To confirm the stability of the predicted binding mode and pivotal interactions, a molecular dynamics (MD) simulation of **5i** in complex with the human OGA homology model was carried out. The binding mode was found to be stable during the MD simulation since the root-mean-square deviation values of the protein backbone and the heavy atoms of the ligand were kept in the range of 2.8−3.2 Å and 1.2−2.4 Å, respectively (Figure S5a). As can be observed from the interaction diagram during the MD simulation, hydrogen bonds were observed to exist between the protonated pyrimidine and Asp285 throughout the entire duration of the simulation (Figure S5b). The hydrogen bonds were inferred to be pivotal for binding to OGA. Furthermore, the MD simulation revealed an interaction with Tyr69, which might arise as a byproduct of hydrogen-bonding interactions between Asp285 and the nitrogen atoms of the pyrimidine ring. Tyr69 was located 4.8 Å from the center of the pyrimidine ring. Although the two aromatic rings, the phenol ring of Tyr69 and the pyrimidine ring of **5i**, are too distant from each other to engage in ideal $\pi-\pi$ interactions, the hydrogen bonds to Asp285 fixed the position of the pyrimidine ring and helped to create the interaction with Tyr69 over a long time. Consistently with the insight afforded by the docking model, the MD simulation supported the existence of $\pi-\pi$ stacking interactions of **5i** with Tyr219 and Trp278 and CH$-\pi$ interactions with Val254 and Tyr286.

Since polar amino acids such as Asp174 and Asp285 are located in the active site of OGA, the reported hydrophilic sugar-derived OGA inhibitors bind to the active site mainly via polar interactions. On the other hand, compound **5i**, with its moderate lipophilicity, achieved potent OGA inhibition through tight binding with the hydrophobic deep pocket comprising Tyr219 and Trp278 with kicking out of Asp174 and CH$-\pi$ interaction with Val254 and Tyr286, which are located at the entrance of the pocket. Furthermore, the aminopyrimidine scaffold, which would be in its stable neutral form under physiological conditions (pH ∼ 7.4), would be protonated at the nitrogen atom in the specific environment, where it is surrounded by several Asp residues, and thus, it would act as a hydrogen bond donor. The superior brain penetration of compound **5i** would be achieved by the aminopyrimidine core structure engaging in a crucial interaction with Asp285 accompanying a reduced number of hydrogen bond donors.

CONCLUSIONS

We have discovered a novel OGA inhibitor, compound **5i**, utilizing virtual screening based on multiple methodologies combining ligand-based and structure-based approaches followed by optimization focusing on LLE. Compound **5i** exhibited efficacy in cells and a favorable pharmacokinetic profile with brain penetration. The OGA inhibitor **5i** could be a useful in vivo chemical tool to elucidate the function of OGA in tauopathies and neurodegenerative events, and it has the potential to be developed as a therapeutic agent in the treatment of tau-mediated neurodegenerative diseases.

EXPERIMENTAL SECTION

Chemistry: General. ^{1}H NMR spectra were obtained at Bruker AVANCE 300 (300 MHz). Chemical shifts are given in δ values (ppm) using tetramethylsilane as the internal standard. Peak multiplicities are expressed as follows: s, singlet; d, doublet; t, triplet; dd, doublet of doublets; tt, triplet of triplets; ddd, doublet of doublet of doublets; brs, broad singlet; m, multiplet. Reaction progress was determined by thin-layer chromatography (TLC) analysis on Merck Kieselgel 60 F254 plates or Fuji Silysia NH plates. Chromatographic purification was carried out on silica gel columns (Merck Kieselgel 60, 70−230 mesh, or 230−400 mesh, Merck; Chromatorex NH-DM 1020, 100−200 mesh, Fuji Silysia Chemical; Inject column and Universal column, YAMAZEN; or Purif-Pack Si or NH, Shoko Scientific). Low-resolution mass spectra (MS) were acquired using an Agilent LC/MS system (Agilent1200SL/Agilent6130MS) or Shimadzu UFLC/MS (Prominence UFLC high-pressure gradient system/LCMS-2020) operating in an electrospray ionization mode (ESI+). The column used was an L-column 2 ODS (3.0 mm × 50 mm ID, 3 μm, CERI) with a temperature of 40 °C and a flow rate of 1.2 or 1.5 mL/min. Condition 1: mobile phases A and B under acidic conditions were 0.05% trifluoroacetic acid (TFA) in water and 0.05% TFA in MeCN, respectively. The ratio of mobile phase B was increased linearly from 5 to 90% over 0.9 min, 90% over the next 1.1 min. Condition 2: mobile phases A and B under neutral conditions were a mixture of 5 mM $AcONH_4$ and MeCN (9/1, v/v) and a mixture of 5 mM $AcONH_4$ and MeCN (1/9, v/v), respectively. The ratio of mobile phase B was increased linearly from 5 to 90% over 0.9 min, 90% over the next 1.1 min. Chemical intermediates were characterized by ^{1}H NMR or mass spectral data or both. The purities of all tested compounds in biological systems were confirmed to be more than 95% purity as determined by analytical high-performance liquid chromatography (HPLC). Analytical HPLC was carried out using a Corona CAD (charged aerosol detector) or a photodiode array detector. The column was a Capcell Pak C18AQ (50 mm × 3.0 mm ID, Shiseido, Japan) or an L-column 2 ODS (30 mm × 2.0 mm ID, CERI, Japan) with a temperature of 50 °C and a flow rate of 0.5 mL/min. Mobile phases A and B under neutral conditions were a mixture of 50 mM $AcONH_4$, water, and MeCN (1:8:1, v/v/v) and a mixture of 50 mM $AcONH_4$ and MeCN (1:9, v/v), respectively. The ratio of mobile phase B was increased linearly from 5 to 95% over 3 min, 95% over the next 1 min. Mobile phases A and B under acidic conditions were a mixture of 0.2% formic acid in 10 mM ammonium formate and 0.2% formic acid in MeCN, respectively. The ratio of mobile phase B was increased linearly from 14 to 86% over 3 min, 86% over the next 1 min. Yields refer to isolated and purified products derived from nonoptimized procedures. Reagents and solvents were obtained from commercial sources and used without further purification.

*Structural Characterization of Hit Compound (**2**).* ^{1}H NMR (300 MHz, DMSO-d_6) δ 0.85−1.12 (8H, m), 1.54−1.71 (3H, m), 2.05−2.48 (13H, m), 2.68−2.86 (2H, m), 3.90−4.08 (1H, m), 4.18−4.36 (2H, m), 5.97 (1H, d, J = 6.0 Hz), 6.06 (1H, d, J = 7.2 Hz), 7.73 (1H, d, J = 6.0 Hz). MS: $[M + H]^+$ 333.3. HPLC purity 91%.

*General Procedure A: Combinatorial Nucleophilic Substitution Reaction of Compound (**4a**).* The mixture of compound **4a** (0.017 g, 0.08 mmol), corresponding amine (0.120 mmol), DIPEA (0.042 mL, 0.240 mmol), and dimethyl sulfoxide (DMSO, 1 mL) was heated at 150 °C for 1 h under microwave irradiation. The reaction mixture was diluted with MeCN (0.5 mL) and purified by preparative HPLC (Actus Triart C18, eluted with MeCN/0.1% TFA aqueous solution).

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

Pure fractions were combined and concentrated by blowing away with the air at 60 °C to yield a product.

*General Procedure B: Combinatorial Nucleophilic Substitution Reaction of Compound (**4b**).* The mixture of compound **4b** (0.019 g, 0.08 mmol), corresponding amine (0.160 mmol), DIPEA (0.070 mL, 0.400 mmol), and tert-AmOH (1 mL) was heated at 130 °C for 1 h under microwave irradiation. The reaction mixture was diluted with MeCN (0.5 mL). The mixture was purified by preparative HPLC (Actus Triart C18, eluted with MeCN/0.1% TFA aqueous solution). Pure fractions were combined and concentrated by blowing away with the air at 60 °C to yield a product.

*2-Chloro-4-(4-methylpiperidin-1-yl)pyrimidine (**4a**).* To a solution of 2,4-dichloropyrimidine (8.15 g, 54.71 mmol) and 4-methylpiperidine (7.77 mL, 65.65 mmol) in THF (300 mL) was added triethylamine (11.41 mL, 82.06 mmol) at room temperature. The mixture was stirred at room temperature over the weekend. The mixture was quenched with water at room temperature and extracted with EtOAc. The organic layer was separated, washed with brine, dried over Na_2SO_4, and concentrated in vacuo. The residue was purified by column chromatography (silica gel, eluted with 40−100% EtOAc in hexane) to yield compound **4a** (9.09 g, 78%) as colorless crystals. ^{1}H NMR (300 MHz, DMSO-d_6) δ 0.91 (3H, d, J = 6.0 Hz), 0.95−1.13 (2H, m), 1.58−1.76 (3H, m), 2.91 (2H, t, J = 12.2 Hz), 4.30 (2H, brs), 6.82 (1H, d, J = 6.0 Hz), 8.01 (1H, d, J = 6.4 Hz). MS: $[M + H]^+$ 212.1.

*4-Chloro-N-phenethylpyrimidin-2-amine (**4b**).* To a solution of 2,4-dichloropyrimidine (10 g, 67.12 mmol) and phenethylamine (12.64 mL, 100.69 mmol) in THF (250 mL) was added triethylamine (18.66 mL, 134.25 mmol) at room temperature. The mixture was stirred at room temperature for 5 h. The mixture was quenched with water at room temperature and extracted with EtOAc. The organic layer was separated, washed with brine, dried over Na_2SO_4, and concentrated in vacuo. The residue was purified by column chromatography (silica gel, eluted with 30−100% EtOAc in hexane) to yield compound **4b** (3.48 g, 22%) as colorless crystals. ^{1}H NMR (300 MHz, DMSO-d_6) δ 2.82 (2H, t, J = 7.5 Hz), 3.47 (2H, d, J = 6.4 Hz), 6.66 (1H, d, J = 5.3 Hz), 7.12−7.34 (5H, m), 7.73 (1H, t, J = 5.7 Hz), 8.23 (1H, brs). MS: $[M + H]^+$ 234.0.

*4-(4-Methylpiperidin-1-yl)-N-phenethylpyrimidin-2-amine (**5a**).* To a mixture of compound **4a** (200 mg, 0.94 mmol) and phenethylamine (0.178 mL, 1.42 mmol) in tert-AmOH (4 mL) was added DIPEA (0.977 mL, 5.67 mmol) at room temperature. The mixture was heated at 130 °C for 1 h under microwave irradiation. The mixture was concentrated in vacuo. The residue was purified by column chromatography (NH silica gel, eluted with 10−80% EtOAc in hexane) to yield compound **5a** (260 mg, 93%) as a yellow oil. ^{1}H NMR (300 MHz, DMSO-d_6) δ 0.91 (3H, d, J = 6.4 Hz), 0.93−1.11 (2H, m), 1.49−1.71 (3H, m), 2.67−2.88 (4H, m), 3.34−3.48 (2H, m), 4.30 (2H, d, J = 13.2 Hz), 6.00 (1H, d, J = 6.0 Hz), 6.48 (1H, brs), 7.14−7.34 (5H, m), 7.76 (1H, d, J = 6.0 Hz). MS: $[M + H]^+$ 297.3.

*4-Morpholino-N-phenethylpyrimidin-2-amine (**5b**).* Compound **5b** was synthesized by general procedure B combinatorially in a 54% yield. MS: $[M + H]^+$ 285.2.

*4-(4-Methylpiperazin-1-yl)-N-phenethylpyrimidin-2-amine (**5c**).* Compound **5c** was synthesized by general procedure B combinatorially in a 67% yield. MS: $[M + H]^+$ 298.3.

*1-(2-(Phenethylamino)pyrimidin-4-yl)piperidin-4-ol (**5d**).* Compound **5d** was synthesized by general procedure B combinatorially in a 65% yield. MS: $[M + H]^+$ 299.3.

*4-(4-Methoxypiperidin-1-yl)-N-phenethylpyrimidin-2-amine (**5e**).* Compound **5e** was synthesized by general procedure B combinatorially in a 58% yield. MS: $[M + H]^+$ 313.3.

*4-(4-Methylpiperidin-1-yl)-N-(2-(pyridin-2-yl)ethyl)pyrimidin-2-amine (**5f**).* Compound **5f** was synthesized by general procedure A combinatorially in a 62% yield. MS: $[M + H]^+$ 298.3.

*4-(4-Methylpiperidin-1-yl)-N-(2-(pyridin-3-yl)ethyl)pyrimidin-2-amine (**5g**).* Compound **5g** was synthesized by general procedure A combinatorially in a 27% yield. MS: $[M + H]^+$ 298.3.

*4-(4-Methylpiperidin-1-yl)-N-(pyridin-3-ylmethyl)pyrimidin-2-amine (**5h**).* To a mixture of compound **4a** (0.17 g, 0.80 mmol) in *tert*-AmOH (1.606 mL) were added 3-picolylamine (0.082 mL, 0.80 mmol) and DIPEA (0.421 mL, 2.41 mmol) at room temperature. The mixture was heated to 150 °C for 8 h under microwave irradiation in a sealed tube. The mixture was concentrated in vacuo and purified by column chromatography (NH silica gel, eluted with 40−100% EtOAc in hexane and 0−20% MeOH in EtOAc) to yield compound **5h** (0.090 g, 40%) as a white solid after recrystallization from EtOAc/hexane. ^{1}H NMR (300 MHz, DMSO-d_6) δ 0.81−1.06 (5H, m), 1.60 (3H, d, J = 10.2 Hz), 2.66−2.83 (2H, m), 4.24 (2H, d, J = 13.2 Hz), 4.41 (2H, d, J = 6.2 Hz), 6.01 (1H, d, J = 6.2 Hz), 7.09 (1H, brs), 7.30 (1H, dd, J = 7.7, 4.7 Hz), 7.68 (1H, dd, J = 10.0, 1.7 Hz), 7.75 (1H, d, J = 6.0 Hz), 8.39 (1H, dd, J = 4.7, 1.7 Hz), 8.51 (1H, d, J = 1.7 Hz). MS: $[M + H]^+$ 284.3.

*N-((1-Methyl-1H-imidazol-4-yl)methyl)-4-(4-methylpiperidin-1-yl)pyrimidin-2-amine (**5i**).* To a solution of (1-methyl-1*H*-imidazol-4-yl)methylamine (50 mg, 0.45 mmol) and compound **4a** (105 mg, 0.49 mmol) in *tert*-AmOH (3 mL) was added DIPEA (0.775 mL, 4.50 mmol) at room temperature. The mixture was sealed and stirred at 130 °C overnight. The mixture was concentrated in vacuo. The residue was purified by column chromatography (NH silica gel, eluted with 50−100% EtOAc in hexane and 0−20% MeOH in EtOAc) to yield compound **5i** (55.0 mg, 43%) as colorless crystals. ^{1}H NMR (300 MHz, $CDCl_3$) δ 0.95 (3H, d, J = 6.0 Hz), 1.04−1.19 (2H, m), 1.51−1.65 (2H, m), 1.70 (1H, brs), 2.71−2.85 (2H, m), 3.63 (3H, s), 4.33 (2H, d, J = 12.8 Hz), 4.52 (2H, d, J = 5.7 Hz), 5.12 (1H, brs), 5.88 (1H, d, J = 6.4 Hz), 6.79 (1H, s), 7.34 (1H, s), 7.86 (1H, d, J = 6.4 Hz). MS: $[M + H]^+$ 287.2.

Biology: Assays for Determination of IC_{50} Values for Inhibition of Human OGA and β-Hexosaminidase Activity. Human OGA was generated from Expi293 cells that had been transfected with the full-length gene. The enzyme was purified by an anti-FLAG agarose column and gel filtration. Human β-hexosaminidase enzyme from placenta was purchased from Sigma-Aldrich (Cat#A6152, St. Louis, MO). Enzymatic reactions were carried out in a reaction containing 50 mM NaH_2PO_4, 100 mM NaCl and 0.01% bovine serum albumin (BSA), 0,01% Tween 20, 1 mM dithiothreitol (DTT, pH 7.4) using 2 mM 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide dihydrate (Sigma-Aldrich, St. Louis, MO, Cat#M2133) as a substrate. The amount of human OGA or β-hexosaminidase enzyme used in the reaction was 1 nM or 10 μ units unit/mL. Test compound of varying concentrations was added to the enzyme prior to initiation of the reaction. The reaction was performed at room temperature (RT) in a 384-well plate (Greiner, Frickenhausen, Germany, Cat#784076) and was initiated by the addition of substrate. The production of the fluorescent product was measured every 5 min for 60 min by a SpectraMax M5/M5e Microplate Reader (Molecular Devices, San Jose, CA) with excitation at 355 nm and emission detected at 460 nm. Inhibition rate was calculated based on 0% control wells with enzyme and DMSO, and 100% control wells without enzyme. Data analysis was done with GraphPad Prism V.5 (GraphPad Software, San Diego, CA) statistical analysis software using a nonlinear regression for a sigmoidal dose−response with a variable slope. The IC_{50} value was defined as the concentration inhibiting the activity by 50%.

The kinetics experiments of compound **5a** or Thiamet-G were performed with a final concentration of 1 nM human OGA and various concentrations of 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide dihydrate obtained by serial 2-fold dilutions starting from 2 mM in an assay buffer (50 mM NaH_2PO_4, 100 mM NaCl and 0.01% BSA, 0,01% Tween 20, 1 mM DTT (pH 7.4)). Fluorescence for each reaction was measured for 60 min using a SpectraMax M5/M5e Microplate Reader (Molecular Devices, San Jose, CA) with excitation at 355 nm and emission detected at 460 nm. V_{max} and K_m values were calculated using the Michaelis−Menten nonlinear fit in GraphPad Prism V.5 (GraphPad Software, San Diego, CA).

In-Cell Western (ICW) Assay for Determination of EC_{50} Values for Cell-Based Inhibition of *O*-GlcNAcase Activity. Inhibition of *O*-GlcNAcase, which removes *O*-GlcNAc from cellular

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

proteins, results in an increase in the level of *O*-GlcNAcylated protein in cells. An increase in the amount of *O*-GlcNAcylated protein was measured by the ICW assay using an anti-*O*-GlcNAc antibody RL2 (Abcam, Cambridge, MA) and the Aerius Automated Infrared Imaging System (LI-COR Biosciences, Lincoln, NE) according to the manufacturer's instructions.

Briefly, human neuroblastoma SH-SY5Y cells were plated in 96-well plates (Corning, Tewksbury, MA, BioCoat Poly-D-Lysine, Cat# 354640) with approximately 1.8×10^4 cells/well. After 48 h incubation with the compounds at 37 °C, the cells were fixed with ice-cold methanol for 30 min at RT. Then, the cells were permeabilized with 0.1% Triton X-100 for 15 min at RT and blocked with an LI-COR Odyssey Blocking Solution (LI-COR Biosciences, Lincoln, NE, Cat#927-40000) for 60 min. The cells were stained with an *O*-GlcNAc Ab (RL2) Alexa Fluoro 700 (1:500 dilution, Cat#NB300-524AF700, NOVUS) at RT for 60 min. After three washes with PBS-T (0.1% Tween/phosphate-buffered saline (PBS(−))), the integrated fluorescence intensities representing the protein expression levels were acquired using the software provided with the imager station. Values were calculated to DMSO-treated cells as the positive control (100%) and culture medium as the negative control (0%). The EC_{50} value was defined as the inflection point of the curve, the concentration of the test compound increasing the amount of *O*-GlcNAcylated proteins by 50%. E_{max} was defined as the maximum % of the increased *O*-GlcNAcylated protein by compounds. The EC_{50}, 95% confidence intervals, and E_{max} were analyzed by triplicate measurements with GraphPad Prism V.5 (GraphPad Software, San Diego, CA) statistical analysis software using a nonlinear regression for a sigmoidal dose−response with a variable slope.

Protein Crystallography. The BT_4395 construct was cloned into a pFastBac vector containing amino acid residues 22−737 (UniProt: Q89ZI2) with an N-terminal 6XHis tag. The protein was expressed in baculovirus-infected B. thetaiotamicron cells. Purification was with Ni-affinity, gel filtration, and 2006 NSMB concentration steps followed by cleavage step (TEV) followed by NiRev and SEC. 200. Protein buffer was 20 mM 4-(2-hydroxyethyl)-1-piperazineetha-nesulfonic acid (HEPES) pH 7.5, 300 mM NaCl at a final concentration of 11 mg/mL. Hanging drop crystallization using a matrix around the PDB 2CHN condition: 0.5 M NaAcetate, 15% PEG 3350, 0.1 M MES pH 6.0, 20% glycerol. The final conditions of the crystal chosen for complexing were 15−19% PEG 3350, 10% glycerol, 0.34−0.4 M NaAcetate, 0.1 M MES 6.0. The crystal was soaked for 60 h with compound at a concentration of 5 mM. X-ray diffraction was collected at Advanced Light Source beamline 5.0.3 under cryo-temperature. Diffraction images were processed with HKL2000 and solved and refined with CCP4 tools Phaser and Refmac, respectively.[43] Model was built into the electron density using Coot.[44]

Brain Distribution in Mice. Compound **5i** was suspended in 0.5% (w/v) methylcellulose solution and orally administered to C57Bl/6J mice ($n = 3$). At 0.5 and 1 h after oral administration (30 mg/kg), blood and brain samples were collected. The blood samples were centrifuged to obtain the plasma fraction. The brain samples were homogenized in saline to obtain the brain homogenate. The plasma and brain homogenate samples were deproteinized with acetonitrile containing an internal standard. After centrifugation, the supernatant was diluted with 0.2% (v/v) formic acid in 10 mmol/L ammonium formate (pH 3). The compound concentrations were measured by LC/MS/MS. After determination of concentration in plasma (C_{plasma}) and brain (C_{brain}), $K_{p,uu}$ was calculated by the equation as follows

$$K_{p,uu} = (f_{u,brain} \times C_{brain})/(f_{u,plasma} \times C_{plasma})$$

where $f_{u,brain}$ and $f_{u,plasma}$ represent the unbound fraction in the plasma and brain, respectively. The $f_{u,brain}$ and $f_{u,plasma}$ were determined by an equilibrium dialysis.

Computational Chemistry: Homology Modeling. The amino acid sequence of human OGA was collected from UniProt Knowledgebase[45] (UniProt accession number: O60502). The crystal structure of BT_4395 OGA in complex with Thiamet-G was downloaded from the Protein Data Bank[46] (PDB ID: 2VVN). This structure includes the A-chain and B-chain of OGA. The A-chain was used as a template for the homology model of human OGA. The homology model of human OGA was built using [Homology modeling/Prime] in Maestro version 10.4.[32] The constructed homology model was prepared using Protein Preparation Wizard in Maestro.

Virtual Screening. An in-house compound library comprising a few million compounds was screened by virtual screening. The compounds were prepared using LigPrep to generate all possible isomers and ionization states and build one 3D structure per isomer/ionization state. For ligand-based virtual screening, multiple conformations were generated using OMEGA version 2.5.[47] with the following parameter sets: RMS = 0.3 and number of conformations = 20 000. As templates for the ligand-based virtual screening, the crystal structures of complexes between various bacterial OGA and their inhibitory ligands were downloaded from the Protein Data Bank (PDB IDs: 2VVN, 2VVS,[14] 2XJ7,[15] 5ABG,[16] 2WB5,[17] and 2WCA[18]) and the coordinates of the ligands were extracted. Shape-based similarity search of in-house compounds was carried out against reported compounds using ROCS version 3.2[34] and then the top 100 000 molecules, in terms of their TanimotoCombo scores, were extracted per one query compound. After removing duplicate molecules, the docking of the remaining compounds was performed into the constructed homology model of human OGA employing the Glide SP mode.[32] Compounds forming hydrogen bonds with at least one residue among Gly67, Asp174, and Asp285 were selected by the Pose Filter function in Maestro. Thereafter, the selected compounds were filtered on the basis of their physicochemical properties (molecular weight ≤ 350, $A \log P \leq 5$, number of hydrogen bond donors ≤ 4, number of hydrogen bond acceptors ≤ 8, number of rotatable bonds ≤ 8, and topological polar surface area ≤ 125) and via PAINS queries using the Pipeline Pilot version 17.2.[40] The remaining compounds were divided into 2681 clusters employing the fingerprint-based clustering method using ECFP_6 with Pipeline Pilot.

Docking Model. Compounds were prepared using LigPrep in Maestro and docked into the homology model using the Glide SP mode. The top-ranked binding pose was refined by energy minimization including the flexible amino acid residues within 6 Å of the compound.

WaterMap Analysis. WaterMap[32,37,38] was run for the docking model of compound **2** with the hOGA homology model, where the compound was removed in the MD simulation.

Molecular Dynamics (MD) Simulation. Desmond[32] was used for 50 ns MD simulation. The docking model of compound **5i** with the human OGA homology model was used as the initial structure of the MD simulation. During simulation, the temperature and the pressure were maintained at 300 K and 1.01 bar, respectively.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.0c01712.

Molecular formula strings (CSV)

Dose−response curve of **2** for hOGA inhibition (Figure S1); surface map characterized by SiteMap (Figure S2); kinetic assay to determine K_m for substrate-competition of **5a** (Figure S3); kinome map of tested kinase for **5i** (Figure S4); MD calculation for the docking model of **5i** in complex with hOGA (Figure S5); data collection and refinement statistics of crystallography (Table S1); and docking model of 2 in complex with human OGA (PDB) (PDF)

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115

Accession Codes

Atomic coordinates and structure factors for the crystal structure of BT_4395 OGA in complex with compound **5a** have been deposited with the RCSB Protein Data Bank (PDB code 7K41).

AUTHOR INFORMATION

Corresponding Author

Michiko Tawada − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan;* orcid.org/0000-0002-1349-0059; Email: michiko.tawada@takeda.com

Authors

Makoto Fushimi − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan*

Kei Masuda − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan*

Huikai Sun − *Takeda California, Inc., San Diego, California 92121, United States;* orcid.org/0000-0002-4593-1306

Noriko Uchiyama − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan*

Yohei Kosugi − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan;* orcid.org/0000-0001-9318-4723

Weston Lane − *Takeda California, Inc., San Diego, California 92121, United States*

Richard Tjhen − *Takeda California, Inc., San Diego, California 92121, United States*

Satoshi Endo − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan*

Tatsuki Koike − *Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa 251-8555, Japan;* orcid.org/0000-0002-3264-1406

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.0c01712

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge Koichi Iwanaga, Atsuko Ochida, and Izumi Nomura for the excellent technical assistance with conducting combinatorial chemistry; Takanobu Kuroita for valuable suggestions to drug design; Terufumi Takagi for helpful suggestions to computational chemistry; Yoshihiko Hirozane for conducting kinase panel screening; Shoichi Okubo and Takashi Ito for preparing the human OGA cDNA and full-length proteins; and Koji Murakami, Shunya Suzuki, and biology group members for the valuable scientific discussion. We thank the staff of the Berkeley Center for Structural Biology, which operates Advanced Light Source beamline 5.0.3. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health and National Institute of General Medical Sciences. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences of the U.S. Department of Energy under contract number DEAC02-05CH11231.

ABBREVIATIONS USED

CNS, central nervous system; CNS-MPO, central nervous system-multiparameter optimization; dba, DIPEA, *N,N*-diisopropylethylamine; DMSO, dimethyl sulfoxide; HPLC, high-performance liquid chromatography; LLE, ligand lipophilicity efficiency; MeCN, acetonitrile; OGA, *O*-GlcNAcase; O-GlcNAc, *O*-linked *N*-acetylglucosamine; OGT, *O*-GlcNAc transferase; PDB, Protein Data Bank; SAR, structure−activity relationship; *tert*-AmOH, 2-methyl-2-butanol; TFA, trifluoroacetic acid; THF, tetrahydrofuran

REFERENCES

(1) Ballatore, C.; Lee, V. M. Y.; Trojanowski, J. Q. Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nat. Rev. Neurosci.* **2007**, *8*, 663−672.

(2) Lee, V. M. Y.; Goedert, M.; Trojanowski, J. Q. Neurodegenerative tauopathies. *Annu. Rev. Neurosci.* **2001**, *24*, 1121−1159.

(3) Yuzwa, S. A.; Shan, X.; Macauley, M. S.; Clark, T.; Skorobogatko, Y.; Vosseller, K.; Vocadlo, D. J. Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation. *Nat. Chem. Biol.* **2012**, *8*, 393−399.

(4) Zhu, Y.; Shan, X.; Yuzwa, S. A.; Vocadlo, D. J. The emerging link between O-GlcNAc and Alzheimer disease. *J. Biol. Chem.* **2014**, *289*, 34472−34481.

(5) Liu, F.; Iqbal, K.; G.-Iqbal, I.; Hart, G. W.; Gong, C. O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10804−10809.

(6) Liu, F.; Shi, J.; Tanimukai, H.; Gu, J.; Gu, J.; G.-Iqbal, I.; Iqbal, K.; Gong, C. X. Reduced O-GlcNAcylation links lower brain glucose metabolism and tau pathology in Alzheimer's disease. *Brain* **2009**, *132*, 1820−1832.

(7) Gao, Y.; Wells, L.; Comer, F. I.; Parker, G. J.; Hart, G. W. Dynamic O-Glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neural, cytosolic β-N-acetylglucosaminidase from human brain. *J. Biol. Chem.* **2001**, *276*, 9838−9845.

(8) Wells, L.; Gao, Y.; Mahoney, J. A.; Vosseller, K.; Chen, C.; Rosen, A.; Hart, G. W. Dynamic O-Glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocytoplasmic β-N-acetylglucosaminidase, O-GlcNAcase. *J. Biol. Chem.* **2002**, *277*, 1755−1761.

(9) Liu, T.; Yan, J.; Yang, Q. Comparative Biochemistry of GH3, GH20 and GH84 β-N-acetyl-Dhexosaminidases and recent progress in selective inhibitor discovery. *Curr. Drug Targets* **2012**, *13*, 512−525.

(10) Schimpl, M.; Borodkin, V. S.; Gray, L. J.; van Aalten, D. M. F. Synergy of peptide and sugar in O-GlcNAcase substrate recognition. *Chem. Biol.* **2012**, *19*, 173−178.

(11) Marcelo, F.; He, Y.; Yuzwa, S. A.; Nieto, L.; Jimenez-Barbero, J.; Sollogoub, M.; Vocadlo, D. J.; Davies, G. D.; Blériot, Y. Molecular basis for inhibition of GH84 glycoside hydrolases by substituted azepanes: conformational flexibility enables probing of substrate distortion. *J. Am. Chem. Soc.* **2009**, *131*, 5390−5392.

(12) Roth, C.; Chan, S.; Offen, W. A.; Hemsworth, G. R.; Willems, L. I.; King, D. T.; Varghese, V.; Britton, R.; Vocadlo, D. J.; Davies, G. J. Structural and functional insight into human O-GlcNAcase. *Nat. Chem. Biol.* **2017**, *13*, 610−612.

(13) Yuzwa, S. A.; Macauley, M. S.; Heinonen, J. E.; Shan, X.; Dennis, R. J.; He, Y.; Whitworth, G. E.; Stubbs, K. A.; Mceachern, E. J.; Davies, G. J.; Vocadlo, D. J. A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nat. Chem. Biol.* **2008**, *4*, 483−490.

(14) Macauley, M. S.; Bubb, A.; Martinez-Fleites, C.; Davies, G. J.; Vocadlo, D. J. Elevation of global O-GlcNAc levels in 3T3-L1 adipocytes by selective inhibition of O-GlcNAcase does not induce insulin resistance. *J. Biol. Chem.* **2008**, *283*, 34687−34695.

(15) Macauley, M. S.; He, Y.; Gloster, T. M.; Stubbs, K. A.; Davies, G. J.; Vocadlo, D. J. Inhibition of O-GlcNAcase using a potent and cell-permeable inhibitor does not induce insulin resistance in 3T3-L1 adipocytes. *Chem. Biol.* **2010**, *17*, 937−948.

(16) Bergeron-Brlek, M.; Goodwin-Tindall, J.; Cekic, N.; Roth, C.; Zandberg, W. F.; Shan, X.; Varghese, V.; Chan, S.; Davies, G. J.;

Vocadlo, D. J.; Britton, R. A convenient approach to stereoisomeric iminocyclitols: generation of potent brain-permeable OGA inhibitors. *Angew. Chem., Int. Ed.* **2015**, *54*, 15429−15433.
(17) Dorfmueller, H. C.; Borodkin, V. S.; Schimpl, M.; Van Aalten, D. M. F. GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation. *Biochem. J.* **2009**, *420*, 221−227.
(18) Balcewich, M. D.; Stubbs, K. A.; He, Y.; James, T. W.; Davies, G. J.; Vocadlo, D. J.; Mark, B. L. Insight into a strategy for attenuating AmpC- mediated beta-lactam resistance: structural basis for selective inhibition of the glycoside hydrolase Nagz. *Protein Sci.* **2009**, *18*, 1541−1551.
(19) Selnick, H. G.; Hess, J. F.; Tang, C.; Liu, K.; Schachter, J. B.; Ballard, J. E.; Marcus, J.; Klein, D. J.; Wang, X.; Pearson, M.; Savage, M. J.; Kaul, R.; Li, T.-S.; Vocadlo, D. J.; Zhou, Y.; Zhu, Y.; Mu, C.; Wang, Y.; Wei, Z.; Bai, C.; Duffy, J. L.; McEachern, E. J. Discovery of MK-8719, a Potent O-GlcNAcase inhibitor as a potential treatment for tauopathies. *J. Med. Chem.* **2019**, *62*, 10062−10097.
(20) Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **2004**, *303*, 1813−1818.
(21) Lyne, P. D. Structure-based virtual screening: an overview. *Drug Discovery Today* **2002**, *7*, 1047−1055.
(22) Dong, L.; Shen, S.; Chen, W.; Xu, D.; Yang, Q.; Lu, H.; Zhang, J. Discovery of novel inhibitors targeting human O-GlcNAcase: docking-based virtual screening, biological evaluation, structural modification, and molecular dynamics simulation. *J. Chem. Inf. Model.* **2019**, *59*, 4374−4382.
(23) Geppert, H.; Vogt, M.; Bajorath, J. Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation. *J. Chem. Inf. Model.* **2010**, *50*, 205−216.
(24) Kirchmair, J.; Ristic, S.; Eder, K.; Markt, P.; Wolber, G.; Laggner, C.; Langer, T. Fast and efficient in silico 3D screening: toward maximum computational efficiency of pharmacophore-based and shape-based approaches. *J. Chem. Inf. Model.* **2007**, *47*, 2182−2196.
(25) Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discovery* **2007**, *6*, 881−890.
(26) Edwards, M. P.; Price, D. A. Role of physicochemical properties and ligand lipophilicity efficiency in addressing drug safety risk. In *Annual Reports in Medicinal Chemistry*; Academic Press, 2010; Vol. *45*, pp 380−391.
(27) Tarcsay, Á.; Nyíri, K.; Kaseru, G. M. Impact of lipophilic efficiency on compound quality. *J. Med. Chem.* **2012**, *55*, 1252−1260.
(28) Roth, C.; Chan, S.; Offen, W. A.; Hemsworth, G. R.; Willems, L. I.; King, D. T.; Varghese, V.; Britton, R.; Vocadlo, D. J.; Davies, G. J. Structural and functional insight into human O-GlcNAcase. *Nat. Chem. Biol.* **2017**, *13*, 610−612.
(29) Schimpl, M.; Schuttelkopf, A. W.; Borodkin, V. S.; Van Aalten, D. M. F. Human OGA binds substrates in a conserved peptide recognition groove. *Biochem. J.* **2010**, *432*, 1−12.
(30) Halgren, T. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377−389.
(31) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146−148.
(32) *Maestro*, version 10.4; Schrödinger, LLC: New York, NY, http://www.schrodinger.com.
(33) Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74−82.
(34) *ROCS version 3.2, OMEGA version 2.5*; OpenEye Scientific Software: Santa Fe, NM, http://www.eyesopen.com.
(35) Wager, T. T.; Hou, X.; Verhoset, P.; Villalobos, A.; et al. Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem. Neurosci.* **2010**, *1*, 420−434.
(36) Shelley, J. C.; Cholleti, A.; Frye, L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: a software program for pKa prediction and protonation state generation for drug-like molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681−691.
(37) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the active-site solvent in the thermodynamics of Factor Xa ligand binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.
(38) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein−ligand binding. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808−813.
(39) Ghose, A. K.; Crippen, G. M. Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships I. partition coefficients as a measure of hydrophobicity. *J. Comput. Chem.* **1986**, *7*, 565−577.
(40) *Pipeline Pilot*, version 17.2; BIOVIA: San Diego, CA, https://www.3dsbiovia.com.
(41) Hirozane, Y.; Toyofuku, M.; Yogo, T.; Tanaka, Y.; Sameshima, T.; Miyahisa, I.; Yoshikawa, M. Structure-based rational design of staurosporine-based fluorescent probe with broad-ranging kinase affinity for kinase panel application. *Bioorg. Med. Chem. Lett.* **2019**, *29*, 126641−126643.
(42) Hohenstein, E. G.; Sherrill, C. D. Effects of heteroatoms on aromatic π-π interactions: benzene−pyridine and pyridine dimer. *J. Phys. Chem. A* **2009**, *113*, 878−886.
(43) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235−242.
(44) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.
(45) The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506−D515.
(46) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(47) *OMEGA*, version 2.5; OpenEye Scientific Software: Santa Fe, NM, http://www.eyesopen.com.

https://dx.doi.org/10.1021/acs.jmedchem.0c01712
J. Med. Chem. 2021, 64, 1103−1115