ACS Medicinal Chemistry Letters

pubs.acs.org/acsmedchemlett

Microperspectives

The Hitchhiker's Guide to Deep Learning Driven Generative Chemistry

Published as part of the ACS Medicinal Chemistry Letters virtual special issue "New Enabling Drug Discovery Technologies - Recent Progress".

Yan Ivanenkov, Bogdan Zagribelnyy, Alex Malyshev, Sergei Evteev, Victor Terentiev, Petrina Kamya, Dmitry Bezrukov, Alex Aliper, Feng Ren, and Alex Zhavoronkov*

Cite This: https://doi.org/10.1021/acsmedchemlett.3c00041

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: This microperspective covers the most recent research outcomes of artificial intelligence (AI) generated molecular structures from the point of view of the medicinal chemist. The main focus is on studies that include synthesis and experimental *in vitro* validation in biochemical assays of the generated molecular structures, where we analyze the reported structures' relevance in modern medicinal chemistry and their novelty. The authors believe that this review would be appreciated by medicinal chemistry and AI-driven drug design (AIDD) communities and can be adopted as a comprehensive approach for qualifying different research outcomes in AIDD.

KEYWORDS: *Artificial intelligence, Drug design, Deep learning*

Drug discovery is a highly complex process requiring a delicate balance of dozens of essential criteria. There are two different levels of usefulness for generative design. One goal of generative chemistry is *de novo* design of molecular structures with desirable properties such as high potency, good absorption, distribution, metabolism, excretion, and pharmacokinetics (ADME/PK) profiles and high synthetic feasibility through automatic or semiautomatic ways.[1] This capability is of great help for computational chemists and can advance a research program. Early examples of generative chemistry engines include synthetic combinatorial library enumeration tools, R-group enumeration software, and genetic algorithms.[2,3] The second goal is the generation of molecular structures to solve complex and challenging MPO tasks by proposing compounds that are patentable and sufficiently unobvious. However, the intellectual capacity of such engines is restricted and does not match the full diversity of requests made by medicinal chemists. Recent advances in artificial intelligence (AI) technologies, mostly related to the successes of deep learning (DL), prompted their adaptation to address some of the challenges associated with drug discovery.[4−6] The introduction of DL into drug discovery has caused an exponential growth of research outputs and revived generative chemistry, which is now often mostly associated with DL.[7]

There are several research papers describing various applications of ML and DL to generative chemistry, including recurrent neural networks (RNNs),[8] variational autoencoders (VAEs),[9−12] generative adversarial networks (GANs),[13] transformers, and hybrid approaches exploiting reinforcement learning (RL) methods.[14−16] Some algorithms have progressed into integrated early drug discovery pipelines available as SaaS and on-premise solutions.[17−21]

Cheminformaticians and AI/ML scientists in the field are typically interested in the metrics associated with the performance of generative algorithms including the coverage of chemical space, portion of the valid generated chemical representations (SMILES[22] (simplified molecular-input line-entry system) strings and two-dimensional (2D) structures), diversity of the generated structures, and "novelty" standing for nonsimilarity to the training set used by the generative model. Therefore, benchmarking solutions, such as MOSES[23] and GuacaMol,[24] aim to provide ML scientists with such analytical data.

Received: February 3, 2023
Accepted: June 9, 2023

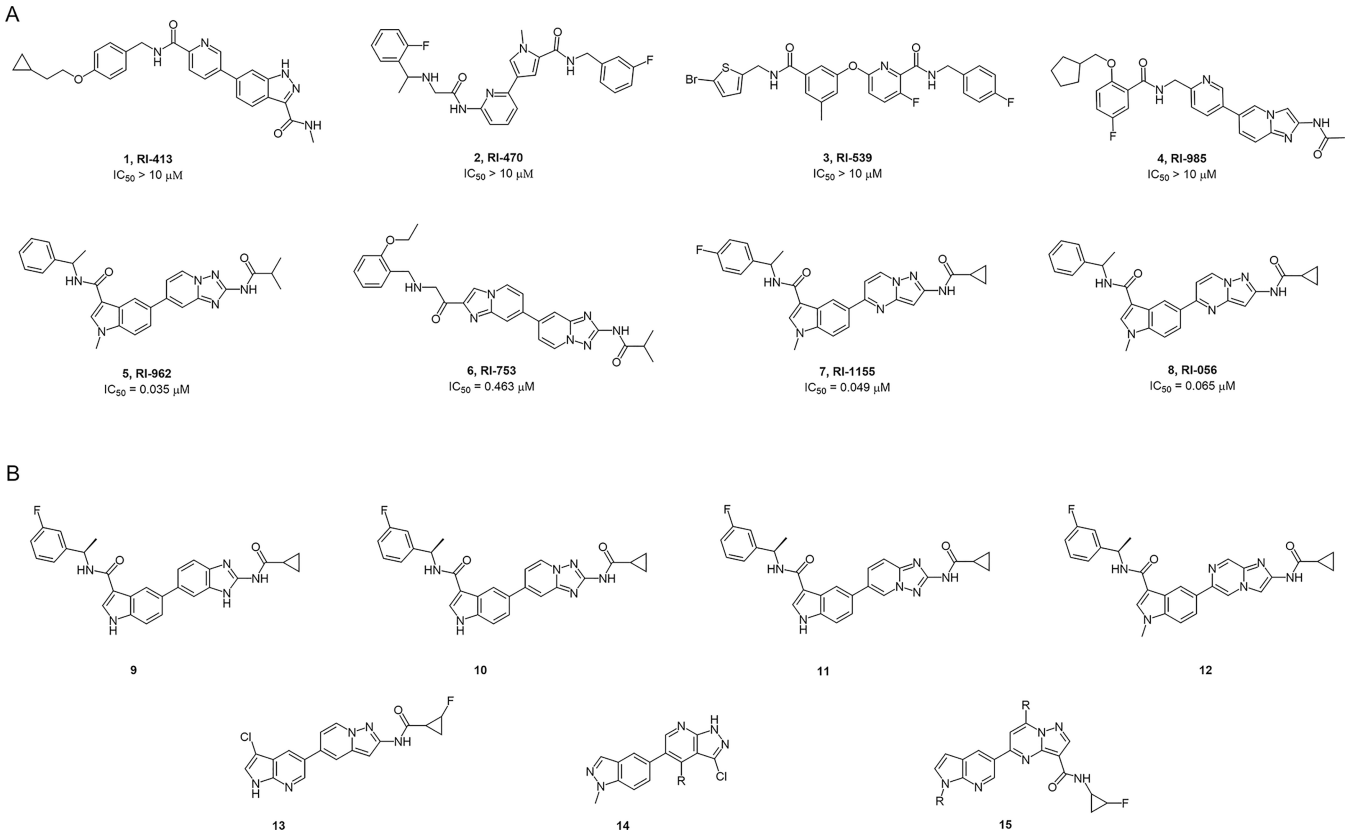

Figure 1. (A) Generated structures submitted for synthesis and *in vitro* evaluation. (B) Examples of RIPK inhibitors similar to compound **5** (RI-985).

However, medicinal chemists tend to look at the results of generative algorithms in a different way. Rather than looking at the numbers describing the validity of the generated SMILES strings, they are more enthusiastic about AI-generated structures, which are already synthesized "trendy" sp^3-rich and druglike compounds without structural alerts. Further, such compounds possess low-digit values of the half-maximal inhibitory concentration (IC_{50}) and good ADME/PK profiles. That is why experimental validation in the context of generative chemistry is very important to the medicinal chemist and why we pay more attention to research papers describing generation techniques reinforced by the synthesis and biological assessment of generated structures.

Although the total number of recent publications (from the last two years) and yet uncited research outcomes covering generative chemistry efforts was 55, in this review, we will focus on the eight papers supported by experimental validation. The discussion on the reported structures, synthesized compounds, and their measured properties and activities aims to reveal the impact of AI/ML methods on modern-day drug discovery.

Most generative DL models are goal-dependent and can provide numerous virtual structures. However, most of them do not satisfy the essential criteria for modern drug design. The majority of publications on generative chemistry do not address real validation issues and mainly focus on metrics. Moreover, the generated structures frequently resemble examples from the training data set, questioning the intellectual property (IP) position. Therefore, to overcome these limitations, Li et al. proposed a generative DL model based on a distribution-learning convolutional recurrent neural network with the long short-term memory (LSTM) algorithm that can generate previously unreported and diverse chemical scaffolds using SMILES strings as inputs/outputs.[25] Receptor-interacting protein kinase 1 (RIPK1) was used as a target for the generation workflow, followed by *in vitro* and *in vivo* validations. Transfer learning allowed the authors to shift the generation process close to the area of the target space in terms of cumulative properties, keeping the scaffold diversity high. Duplicated structures and items containing structural alerts or reactive moieties were excluded. The authors highlighted a good diversity among the generated scaffolds that was much higher than that observed in the source and target data. Druglike filters were applied for the structures with unique scaffolds, followed by pharmacophore-based virtual screening (PBVS) (11 features). Structures that matched at least four features were classified as fitting the constructed hypothesis. Further, molecular docking was applied to the selected virtual structures. Consequently, the authors obtained 50 high-scored structures, of which eight were selected for performing synthesis and subsequent biological evaluation (Figure 1A).

On the basis of the obtained results, four compounds (RI-413 (**1**), RI-470 (**2**), RI-539 (**3**), and RI-985 (**4**)) showed no inhibitory activity (IC_{50} > 10 μM). However, the most active compound, RI-962 (**5**), demonstrated an IC_{50} value of 35 nM against RIPK1 and a good selectivity profile among 408 human kinases (KINOME*scan* panel). With respect to scaffold novelty, Li et al. published a related patent application where a wide series of similar compounds were described as potent RIPK inhibitors two years before publishing the reviewed study.[26] Many isosteric scaffolds were synthesized, including those obtained via the atom-replacing approach, particularly by

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

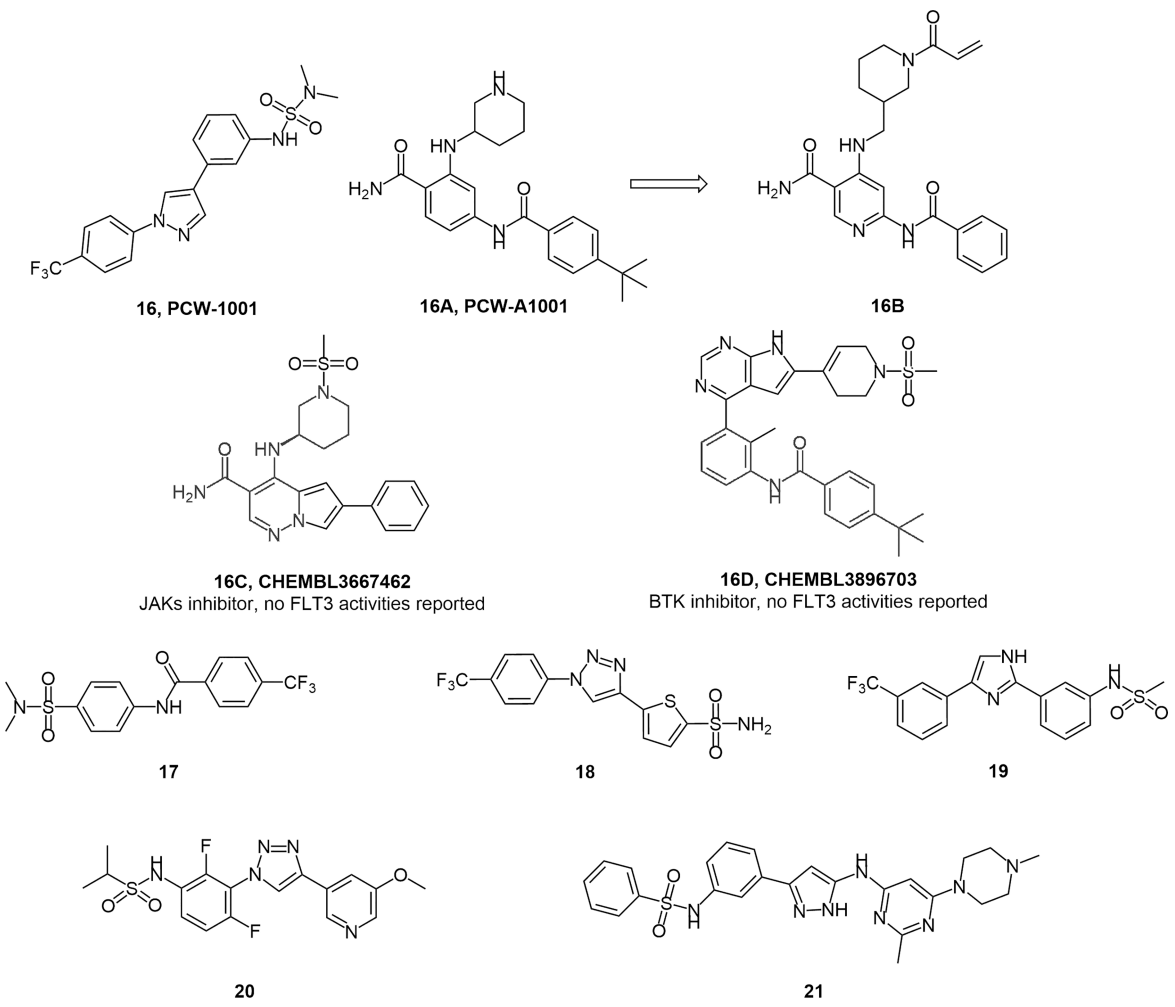

Figure 2. Generated compound **16A** and its comparison to the reported compounds **16**, **16B**−**16D**, and **17**−**21**.

changing the position of a heterocyclic nitrogen atom (compounds **9**−**12**, Figure 1B). However, Lee et al. were the first to claim closely related scaffolds (compound **13**) as RIPKs, c-Abl, and LRRK2 inhibitors.[27] In 2021, Hatcher et al. published a series of 5-(1*H*-indazol-5-yl)-1*H*-pyrazolo[3,4-*b*]pyridines (compound **14**) exhibiting nanomolar activity against cyclin-dependent kinases.[28,29] In 2020, Masse et al. described 5-(1*H*-pyrrolo[2,3-*b*]pyridin-5-yl)pyrazolo[1,5-*a*]-pyrimidine-3-carboxamide derivatives (compound **15**) as potent TYK2 inhibitors.[30] As shown in Figure 1B, the scaffold reported by Li et al. cannot be described as novel either among RIPK inhibitors or within an extensive category of kinase inhibitors. Many reported kinase scaffolds, such as pyridin-2-amines, pyridin-2-amides, their fused analogues, and 1*H*-imidazol-2-amides, could be found among the generated structures presented in the supporting information.[25] The authors used 1030 RIPK1 inhibitors to fine-tune the model, and within this set we revealed many compounds with high similarities to the generated structures. On the basis of a thorough analysis, we can conclude that the novelty of the generated structures is rather low and, as a consequence, the IP position is presumably weak. Li and colleagues also mentioned a "cleaning" procedure to remove bad structures within the generation pool. However, metabolically labile carbamates (RI-273), *N*-acyl derivatives (e.g., RI-251, RI-109, RI-111, and RI-166), hemiaminal-containing structures (RI-024, RI-452), esters (e.g., RI-441, RI-016, and RI-011), pyridin-4-ylmethanol (RI-17), and hydroxylamine (RI-109) are the unstable moieties present in the generated structures. Furthermore, several structures contain reactive fragments such as 4-chloropyridine (RI-111) and 1,4-MA (RI-373 and RI-535). At least one structure (RI-017) should be classified as rare because it contains a 1,4-dihydrocinnoline moiety, presumably prone to aromatization upon biological conditions. Although the study is attractive with exhaustive biological evaluation, the generative power of the presented approach should be evaluated for other targets, not from the kinase family, especially toward novel proteins with no reported ligands. Additionally, the active molecules contain a π-conjugated fused aromatic system similar to benzimidazole dimers or planar biaryl amides which are well-known minor groove DNA binders; therefore, the results of the related biological examination are also needed to characterize the active compounds more comprehensively.

Jang et al. developed a small-molecule compound, PCW-1001 (see Figure 2, **16**).[31] FLT3 kinase was initially considered one of the possible targets using an in-house network-based reverse target prediction module, ChEMBL database similarity searching, followed by docking against 2000 unique proteins. On the basis of the docking results, FLT3, JAK2, NTRK, MKNK2, and TGFBR1 (ALK5) were among the top 10 scored proteins selected as the possible targets for PCW-1001. It

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

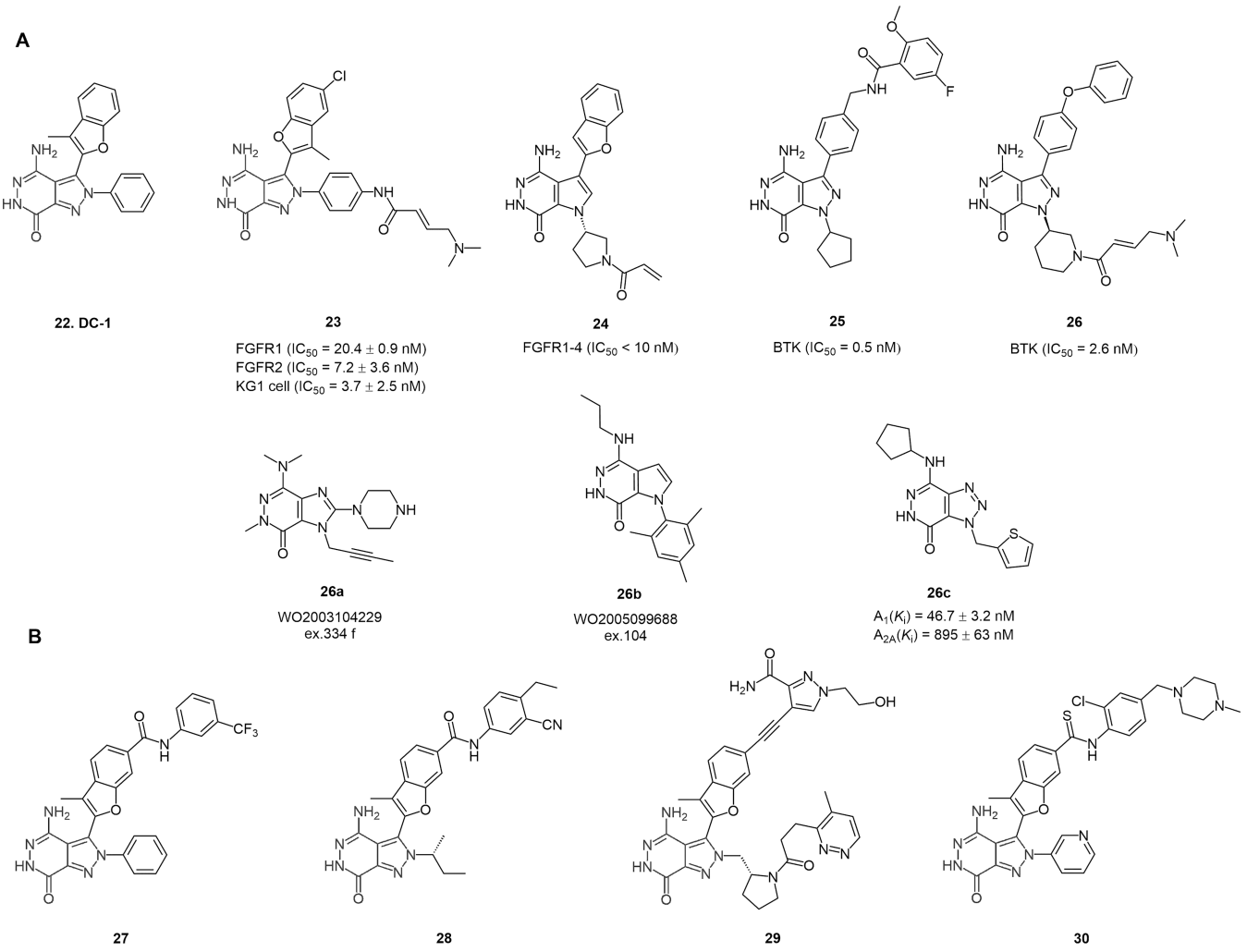

Figure 3. (A) Examples of reported compounds similar to the generated structures. (B) Examples of the generated structures targeted against DDR1.

showed IC_{50} values of 13.6 and 1.83 μM against FLT3 and D835Y, respectively. Notably, for high-throughput screening of novel kinase inhibitors, a compound with an activity of <500 nM is usually classified as the primary hit compound for a kinase with unknown inhibitors. The micromolar activity of a hit molecule is commonly considered a hard starting point. Then, a deep RNN with an LSTM approach was used to generate the optimized PCW-1001 analogues. Toxicities for the generated structures were predicted with the use of Pharmulator, and compounds containing structural alerts were removed from the pool. Only 190 structures, which fulfilled the criterion of synthetic feasibility (ML models), novelty, druglike properties, and PAINS filters, were subjected to a docking study with a visual inspection of the obtained binding poses. Then, high-scored structures were synthesized, and their activities were evaluated *in vitro*. PCW-A1001 (**16A**) showed good *in vitro* activity and was considered a potential and selective $FLT3^{WT}$ and $FLT3^{D835Y}$ inhibitor with IC_{50} values of 2.54 μM and 764 nM, respectively.

According to the related patent application, PCW-1001 (**16**) (KRCT-1) was tested against a panel of 104 kinases *in vitro* at a concentration of 10 μM.[32] Although it inhibited the activities of TRKA (~55%), FLT (~45%), MELK (~35%), FGFR (~35%), and PKC (~35%), the activities of other kinases, including MAPK2 and ALK, were unaffected. PCW-1001 showed an IC_{50} value of 5.8 μM against TRKA, which was considered the main target kinase. Furthermore, cell-based studies confirmed the inhibition of the TRKA-dependent downstream signaling route by PCW-1001 (**16**). Additional information about FLT3 inhibition was not presented. Five kinases (from a list of the top 10 kinases scored during the docking study) were mentioned as possible targets, including FLT3 and NTRK. However, it is not clear whether the kinase panel was manually designed based on the obtained docking results or whether the standard panel available in the Eurofins subdivision (Scotland, U.K., currently closed) was simply used as JAK2 kinase was not included in the panel. Considering unique preprocessing procedures and pose/scoring analysis, the time effectiveness of the docking study with 2000 proteins against the testing of kinase panel for 104 targets is not obvious. Moreover, the inclusion of kinases predicted by docking in the panel could be much better for estimating the predictive ability and to improve the docking model. We searched for similarities with the PCW-1001 structure and found at least four compounds (**17**–**20**) with higher topological and structural similarities with PCW-1001 than CHEMBL1807483 provided by the authors as the closest analogue (**21**). Compound **17** was described as the TRPV4

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

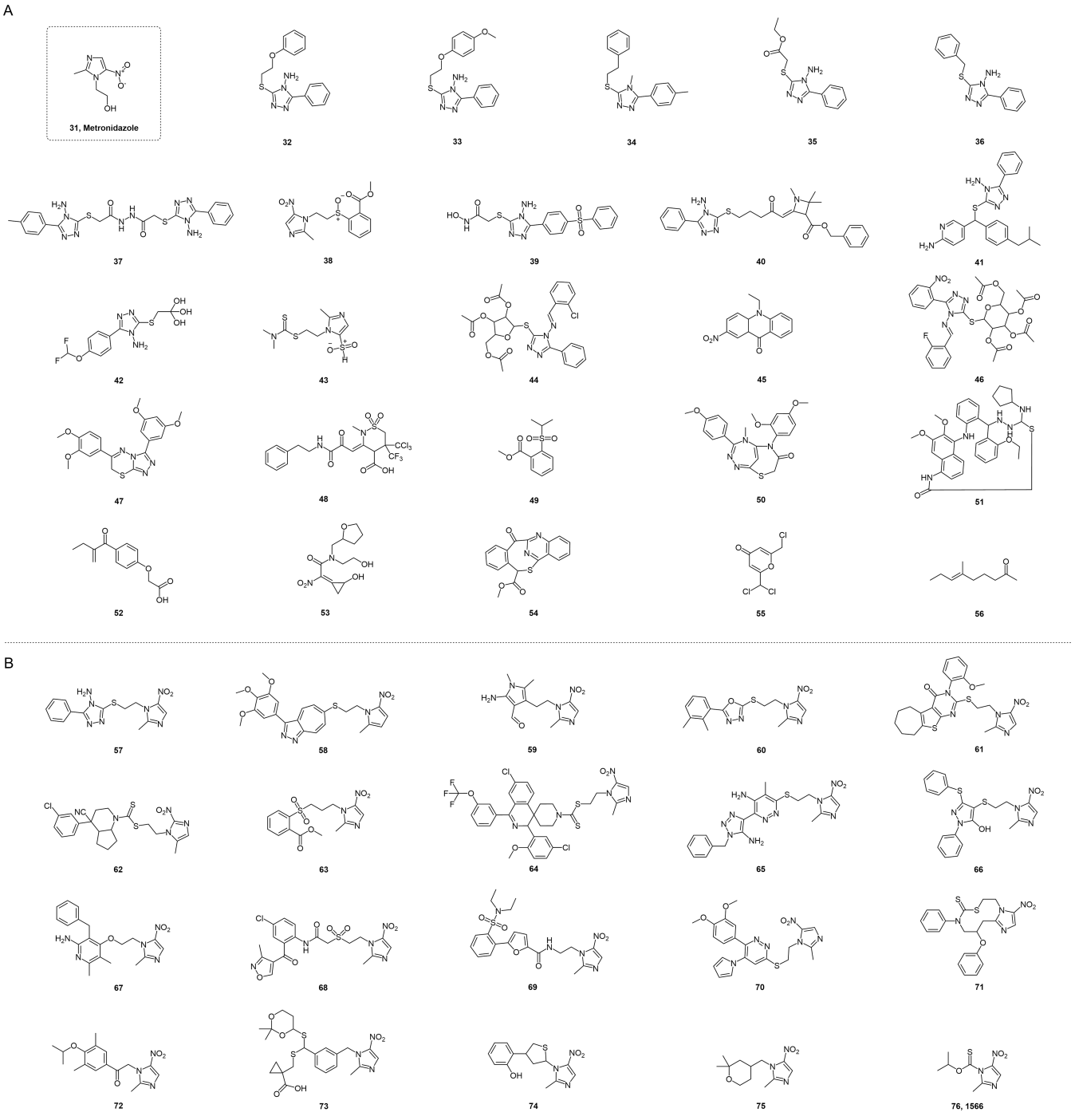

Figure 4. (A) Representative examples of structures generated by GRU-based algorithm and (B) 20 structures containing metronidazole substructure selected for further analysis.

receptor antagonist (IC_{50} = 4.22 μM, cell-based assay),[33] compound **18** showed activity against hCA-II (IC_{50} = 2.9 nM),[34] and compound **19** was reported as the neuropeptide-Y5 receptor binder with a K_i value of 1.4 nM.[35] Therefore, it would be valuable to test PCW-1001 against at least these targets to preevaluate the off-target space. Compound **20** was claimed as a kinase inhibitor, particularly acting against B-Raf kinase with a half-maximal effective concentration (EC_{50}) of 493 nM in a cell-based assay. However, PCW-1001 did not show considerable activity against B-Raf kinase according to the kinase panel results, presumably due to pyrazole being the only fragment which can be responsible for hinge binding, in contrast to compound **20** bearing a pyridine moiety in addition to a triazole fragment. In any case, these examples demonstrate that PCW-1001 could be obtained via scaffold hopping using bioisosteric amide or triazole moieties. The structure of PCW-A1001 resembles the topologies of several covalent and competitive JAK inhibitors containing 2-amino-4-benzamide fragments (compound **16B**)[36] and their isosteric analogues. Considering this, the generation process can be presumed to be partly based on this transformation or within the structural mixing between compounds **16C** and **16D** that are likely to be present in the training set of 438 552 structures for RNN, which unfortunately was not published. These inhibitors with

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

4-aminonicotinamide fragment bind JAK in two main ways: the terminal amide fragment placed close to the gatekeeper region ("amide in")[37] or located out of the pocket ("amide out").[37] The last binding mode is more realistic for PCW-A1001 because carbonyl oxygen, hypothetically, forms the H-bond with −NH− of Cys694, and the docking results provided by the authors correlate well with this hypothesis, although gilteritinib binds FLT3 via H-bonding between the terminal amide group and the hinge motif. The inhibition activity of PCW-A1001 against other kinases and examples of the generated structures were not provided by the authors. Additionally, the novelty, druglikeness, and potential toxicity of each obtained compound cannot be estimated as the authors did not provide this information. Although PCW-A1001 demonstrated a relatively low *in vitro* activity, its structures can be attributed to a novel category, at least for developing small-molecule FLT3 inhibitors. The authors presented good results; however, more detailed and versatile analysis of the obtained compounds and related biological activities could greatly enhance this study. It should be also noticed that the designed PCW-A1001 is not structurally related to the input PCW-1001 molecule. That is why the "optimization" term used by the authors of the paper is hardly applicable to the described design procedure, since the structural context of PCW-1001 is completely lost during the few indirect steps during this procedure.

Tan et al. developed a small-molecule selective DDR1 inhibitor using a deep generative scaffold decorator model.[38] They classified the most promising molecule as the lead compound and investigated its efficiency *in vivo* using a dextran sulfate sodium (DSS) induced colitis mouse model. The design was based on pyrazolo[3,4-*d*]pyridazinone-containing FGFR inhibitors, reported previously by the same authors as novel compounds. However, this scaffold and its isosteric derivatives were described as adenosine receptor ligands,[39,40] DPP-IV inhibitors,[41] CRF receptor antagonists,[42] and kinase inhibitors[43] (Figure 3A, **23**−**26**, **26a**−**26c**). An efficient route to obtain these compounds was also developed. However, the synthetic pathway, particularly for DC-1, has been described in many previous reports.[44−46] Consequently, DC-1 cannot be classified as novel (see Figure 3A, **22**), especially for kinase inhibitors. The authors revealed that DC-1 showed a weak cross activity against DDR1 with an inhibitory rate of 48% at 0.1 μM and used this template compound to optimize the scaffold toward DDR1 via the ML approach. Additionally, clarity is required regarding why the more active compound[47] that inhibited DDR1 up to 66% at the same concentration was not used. Moreover, it did not inhibit the activities of DDR2 and BTK.[47] Liang et al. published similar molecules (compound **24**, covalent inhibitor) as FGFR inhibitors, containing the 4-amino-dihydro-1*H*-pyrazolo[3,4-*d*]pyridazin-7-one scaffold,[48] whose close analogues were claimed as BTK inhibitors in 2014 (reversible compound **25** and covalent **26**).[43,49,50] The available X-ray data and the published results of the molecular docking studies showed that the binding modes of the scaffold with BTK and FGFR within the hinge region were identical. However, the 1,4-MA warhead was targeted on different cysteine residues.

The benzofuran moieties of compound **24** and DC-1 (**22**) and the phenyl fragments of the presented BTK inhibitors were located at the same positions. The predicted binding mode of the structures obtained via ML-based optimization of DC-1 in the DDR1 binding site was the same in the hinge region and proximal areas. Therefore, FGFR inhibitors containing the 4-amino-dihydro-1*H*-pyrazolo[3,4-*d*]pyridazin-7-one scaffold could be discovered based on similar BTK inhibitors via the trivial transposition of a covalent warhead.

Furthermore, regarding the design of DDR1 inhibitors (Figure 3B), the matched molecular pairs (MMP) algorithm was used to obtain a wide fragment library, which was applied to train the deep generative (decoration) model to investigate the links or decorations of scaffolds and fragments. The criteria applied during the training set accumulation were as follows: the scaffolds must contain at least one ring, and the decorations must meet the descriptor ranges of HBD $\leq$ 5, cLogP $\leq$ 5, and Rot bonds $\leq$ 5. The quality of the generated structures was assessed via the prediction/calculation of properties such as LogP, MW, and SA (synthetic accessibility). We performed visual analysis of the generated structures (Figure 3B) and concluded that the fragments, such as (trifluoromethyl)phenylamide and 1-benzyl-4-methylpiperazine, attached to the scaffold are present in the structures of the reported small-molecule DDR1 inhibitors. Therefore, whether this generation provided new fragments or the ML model was learned to recognize the most reliable scaffold within the MMP pool is unclear. Otherwise, the presented structures can be rapidly obtained using the available chemoinformatic software with the routine scaffold hopping approach. Although the subsequent pharmacophore searching and docking study can provide the same virtual hits, these arguments do not allow the assessment of the real role of the generative step in the design workflow.

Metronidazole is a widely used antibacterial agent (Figure 4A, **31**). However, it provokes severe side effects. Therefore, developing its analogues with a new skeleton and reduced toxicity is highly interesting. Chen et al. presented an approach to solving this problem.[51] First, a generative model based on gated recurrent unit (GRU) with transfer learning was used to produce novel structures (Figure 4A), of which structures containing a metronidazole substructure were selected. Then, 20 structures were sampled for further analysis after the application of clustering (Figure 4B), and structure **57** was chosen based on synthetic accessibility. Finally, a series of its analogues were synthesized and tested *in vitro*, revealing clear inhibitory activity against four bacteria species.

The structures generated by the GRU-based model seem questionable, as discussed further. For example, only 321 of the 3314 generated structures had the nitroimidazole moiety known for exhibiting antibacterial activity.[52] Consequently, most generated structures may show a loss of antibacterial activity. Moreover, multiple structures seemed potentially toxic (e.g., **45** looks similar to known DNA intercalators[53] and 48 contains a potentially mutagenic 1,2-dicarbonyl group[54]), unstable in water or under physiological conditions (e.g., **40** contains an ester group which is known to be prone to hydrolysis[55] and **42** contains an unstable trihydroxy-substituted carbon atom), or chemically unreasonable (e.g., **43** contains a terminal sulfone group and **50** contains an unusual bicyclic substructure).

Regarding the results of selecting nitroimidazole-containing structures for further analysis, such an approach does not find novel skeletons and all the selected structures contained the metronidazole moiety with different substituents in place of the hydroxyl group. Seemingly, selecting the generated structures with metronidazole cores played a crucial role in obtaining valuable examples. Compound **57** and its analogues

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

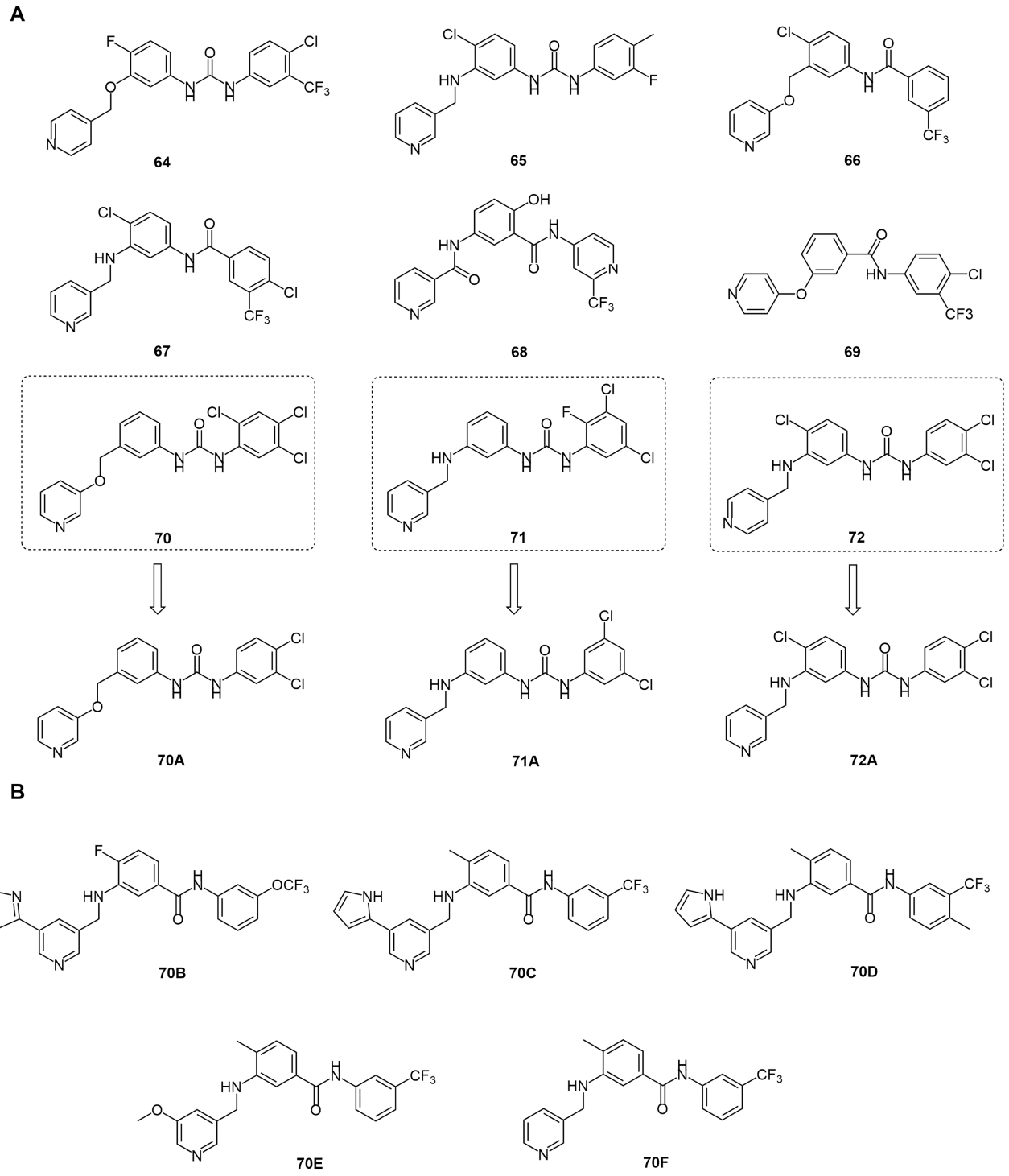

Figure 5. (A) Synthesized compounds evaluated as DDR1 inhibitors and (B) their previously described analogues.

were tested, showing moderate micromolar activity. Possibly, the effect of tested structures was mainly owing to the metronidazole substructure as metronidazole acts as a DNA binder after the metabolic activation of its nitro group. The metabolic reduction of the nitro group also leads to the generation of reactive oxygen species, providing a toxic effect.[56−58] Thus, it did not seem that the toxicity would be decreased by adding new fragments during the presence of the nitro group in the molecule structure. Moreover, the selected structures had some problems in terms of medicinal chemistry. For example, structures **57**, **58**, and **59** contained metabolically unstable aminotriazole fragments, unusual bicyclic fragments, and a highly reactive aldehyde group, respectively.

Additionally, some generated structures did not seem to be novel as their close structural analogues had been previously reported in the literature. For example, analogues of structure **60** were previously described as dual anticancer/antimicrobial agents with the highest minimum inhibitory concentration values of 1.56−3.13 μg/mL against the tested bacterial strains.[59] Structure **57**, selected for synthesis and biological

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

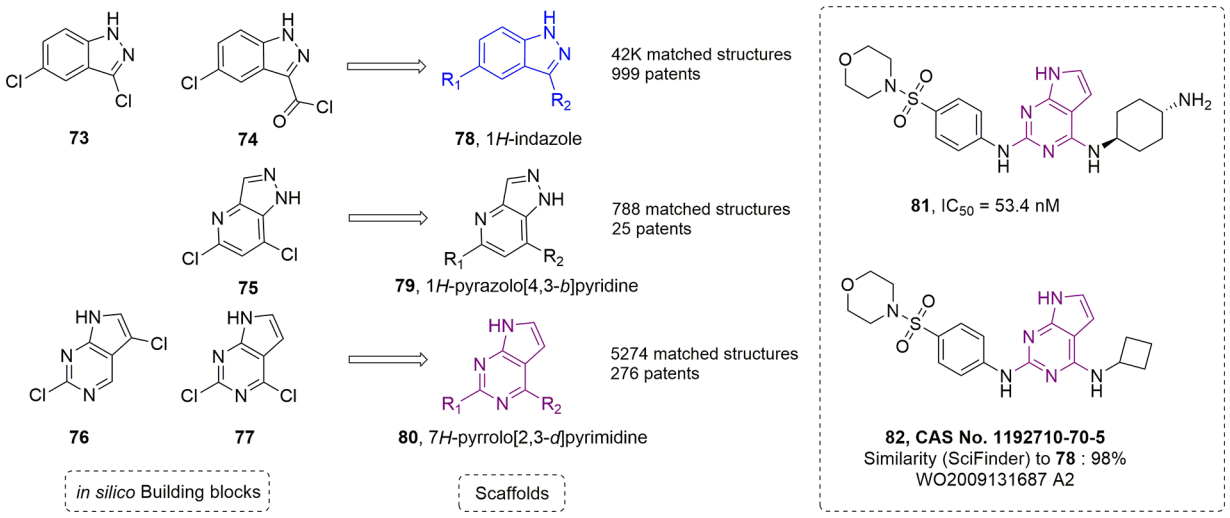

Figure 6. Generated and then selected building blocks (**73**−**77**), their scaffolds (**78**−**80**), and comparison of leading compound **81** to the reported compound **82**.

evaluation, was highly similar to the analogues of structure **60** and differed only by several atoms, obtaining predictable results. A close structural analogue of **61** was previously synthesized and evaluated as a potential antimalarial agent.[60] The structural derivative of **62** previously exhibited anti-trichomonas activity in a low micromolar range.[61] Structure **63** was described as a nanomolar antiamoebic agent.[62] The patent landscape of the substituted 5-nitro-1*H*-imidazoles as antibacterial agents is extensive, resulting in unattractiveness in developing novel molecules containing this core structure.

Yoshimori at al. applied a deep generative model and three-dimensional (3D) pharmacophore modeling and filtering based on the calculated binding affinity to obtain structures of novel DDR1 inhibitors.[63] Consequently, 570 542 structures were generated, 10 694 structures were selected based on the pharmacophore score, and 4731 structures were sampled based on the calculated binding affinity. Then, nine structures were selected, considering the pharmacophore score, binding score, and synthetic accessibility. Subsequently, nine compounds were synthesized based on the generated structures or their manually modified analogues (Figure 5A, **64**−**72**), and three showed submicromolar inhibitory activity. The developed approach was found suitable for hit identification and scaffold hopping.

However, there are several questionable points. A full list of the generated structures with high pharmacophore and binding affinity scores was not provided. Only nine finally selected compounds were provided. Therefore, the overall quality of generation cannot be assessed based on a few structures. It should be noted that less than 1% of generated structures met the requirements of binding affinity and pharmacophore filtering, which may indirectly indicate the low ability of the method to generate the desired structures. Furthermore, three of the analyzed compounds were manually modified, and the actual potential of the developed approach to generate active structures based on these three examples is hard to fully estimate. Structure **72A** was obtained from **72** via changing the position of the aromatic nitrogen in the pyridine ring to form a hydrogen bond with the hinge region. As this interaction is crucial for kinase targeting and compound **72A** itself showed only micromolar activity, it seems unlikely that compound **72** would be active against this target.

Among tested compounds, compounds **66**, **67**, and **70A** showed moderate activities (IC_{50}) of 92.5, 186.7, and 171.3 nM, respectively, in the submicromolar range. Seemingly, the hit rate of the developed approach is not as high as expected, especially considering that one of the discovered submicromolar inhibitors was obtained utilizing manual modification. It seems that structure **70** would not have the same activity as an inner geometry of phenylurea moiety would not be the same in these two compounds. The authors'[63] idea that the chlorine atom at the *ortho*-position in structure **70** should not affect the binding mode of the compound is somewhat confusing, because it would prevent the benzene ring from forming a favorable active conformation appropriate for binding due to forced rotation. Low nanomolar DDR1 inhibitors containing the same scaffold were previously described by Jeffries et al. (Figure 5B, **70B**−**70F**).[64] 3-Benzyloxypyridine, which includes hinge binding scaffold of compounds **66** and **70A**, is a well-known kinase inhibitor skeleton discovered more than 10 years ago.[65] Compound **66** and its close structural analogue[66] were described as p38α MAP kinase inhibitors in 2002.[67] Compound **66** was also claimed as a Raf kinase inhibitor in 2003.[68] Minor modifications in compound **66** led to a series of CSF-1R inhibitors.[69] Derivatives of compound **67** were also reported as nanomolar B-Raf kinase inhibitors more than 10 years ago.[70] Very close urea-containing analogues of compound **70** were also reported as p38α MAP[67] and Raf kinase inhibitors.[68] The developed approach allowed identifying a series of structures with low structural diversity and high similarity with the previously reported compounds. However, only a small fraction of the synthesized compounds showed inhibitory activity against DDR1.

Hua et al. applied RNN to generate functionalized building blocks, which can be utilized in Suzuki and Buchwald−Hartwig reactions *in silico* to get a virtual combinatorial library of the potential Mer tyrosine kinase (MerTK) inhibitors.[71] Besides this article, there have been reported few research papers utilizing direct accounting for synthetic context during the generation of molecular structures[72−74] since the SA of *de novo* outputs is still an "Achilles heel" of artificial intelligence driven drug design (AIDD). The outputs of generative chemistry are usually triaged before they are submitted for synthesis to focus on only feasible and promising structures. In this context,

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

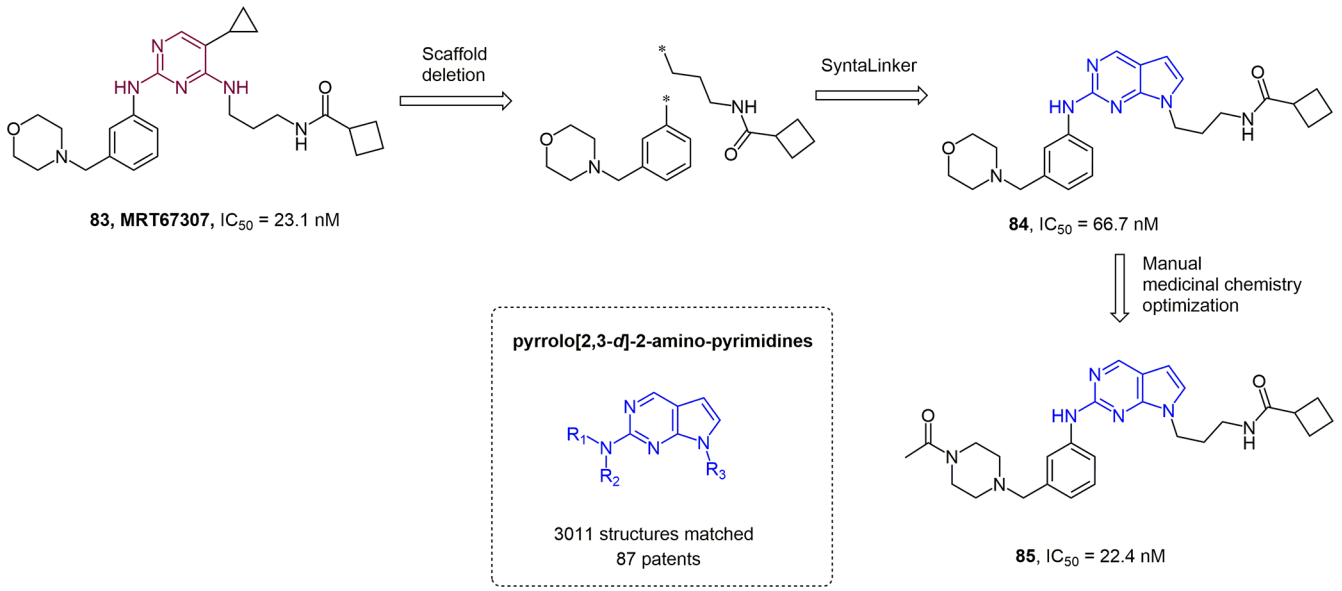

Figure 7. Process of compound **84** generation using SyntaLinker algorithm starting from the fragments of MRT67307 (**83**) and its manual modification to the potent TBK1 inhibitor **85**.

taking commercially available building blocks and feeding them to RNN to get a set of new "building blocks" is counterintuitive since the biggest value of a building block is its commercial availability, meaning fast-track delivery, which cannot be directly ensured by the RNN architecture or the synthetic accessibility score (SA score)[75] used by the authors for SA estimation. Since nothing more than the 17 selected structures, synthesized with a 100% success rate, were provided, it is hard to estimate the actual synthetic feasibility of the remaining top (by docking score) 19 983 structures. Thus, the synthetic accessibility of the generated structures that were not selected for synthesis is unclear. It is recommended not to make structural alterations of commercially available starting materials, even using AI/ML techniques, since even small alterations can render the compounds unavailable. The rational selection from the broader scope of *in silico* chemical reactions (not just two reactions) and rational iterative nomination of the existing building blocks based on the context of the target should be the focus of ML application rather than generating new building blocks.

The authors[71] claimed to perform scaffold analysis to find novel scaffolds applicable to the hinge binding region of MerTK after the docking studies for a designed reaction-based combinatorial library. Five building blocks (Figure 6, **73**−**77**) were picked as scaffolds that were considered to be related to MerTK. A brief analysis of the scaffolds that could be extracted from the selected five building blocks and 17 resulting synthesized compounds was carried out in SciFinder, showing that they could hardly be classified as novel. The indazole-based (**78**) disubstituted scaffolds of **73** and **74** were found in almost ∼1000 patents, including sample kinase inhibitor patents.[76,77] 7*H*-Pyrrolo[2,3-*d*]pyrimidines (**80**) in **76**, **77**, and the leading compound **81** are present as key scaffolds in 315 patents, most of them disclosing kinase inhibitors.[78,79] Finally, a routine similarity search in SciFinder also revealed that compound **82** (CAS No. 1192710-70-5) reported in the JAK tyrosine kinase-related patent[79] was very similar (98% by SciFinder similarity metrics) to compound **81** with only one trivial difference of a cyclobutane substituent changed by aminocyclohexane. Thus, the placed scaffold novelty is rather controversial.

Song and coauthors reported on the design of TANK-binding kinase 1 (TBK1) inhibitor utilizing the deep conditional transformer neural network SyntaLinker.[80] SyntaLinker[80,81] is an FBDD-based algorithm that attaches molecular fragments to linkers using deep transfer learning from the reported chemical space (ChEMBL database[82]). The focus was on replacing the 2-aminopyrimidine scaffold in a small-molecule TBK1 inhibitor, MRT67307 (Figure 7, **83**; IC_{50} = 23.1 nM), with a new one. The 2-aminopyrimidine ring from **83** was removed, and SyntaLinker was provided with two resulting terminal groups. The output of the SyntaLinker launch contained 1101 molecular structures, of which 276 structures contained both terminal groups of **83**. The hinge scaffold linkers proposed by SyntaLinker included the original cyclopropyl substituted 2-aminopyrimidine and 275 other linkers. The generated structures were then docked. In the top 10 high-scoring examples, seven structures were generated using differently substituted 2-aminopyrimidines originally present in **83**, which did not match with the scaffold hopping aim of the study. The remaining three structures from the pool contained two different bicyclic scaffolds: quinazolin-2-amine and pyrrolo[2,3-*d*]-2-aminopyrimidine. Compound **84** with the best docking score resulted from linkage within pyrrolo-[2,3-*d*]-2-aminopyrimidine scaffold. This compound was synthesized and tested against TBK1 (IC_{50} = 66.7 nM). The following optimization campaign led to a more potent and selective TBK1 inhibitor **85** (IC_{50} = 22.4 nM), and SyntaLinker was not exploited at this stage (see Figure 7). The only design contribution by SyntaLinker to this study was the scaffold hopping of 2,4-aminopyrimidine to 2-aminopyrrolo[2,3-*d*]pyrimidine. A closer look at the reported items revealed that this modification is well-described in the literature[83,84] and related examples could be found even in 2006.[85] Therefore, the authors of this review believe that trivial ring fusions adopted for scaffold hopping can be readily performed with one starting point, such as **83**, without any DL

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

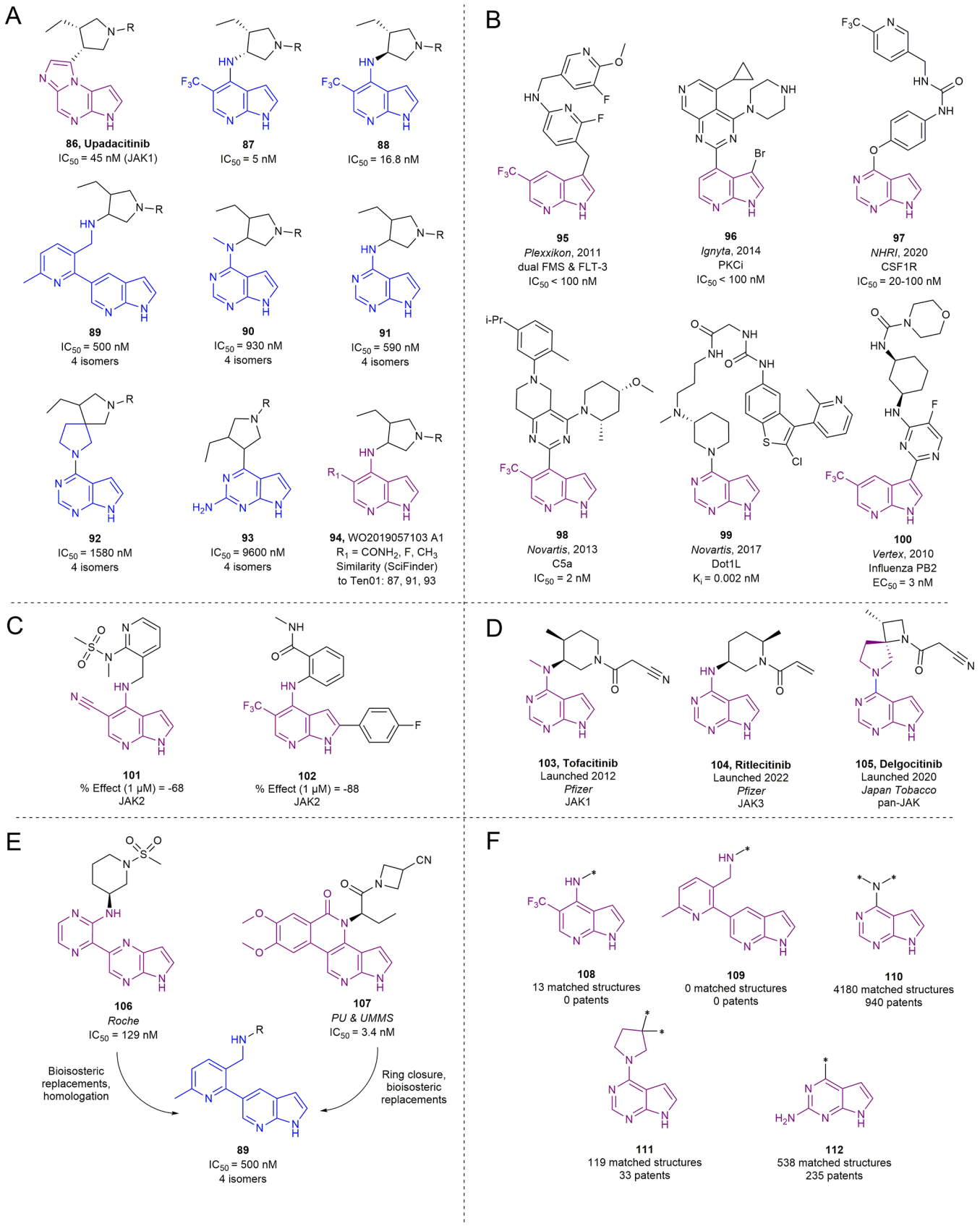

Figure 8. (A) Generated and synthesized compounds **87**–**93** and their comparison to upadacitinib (**86**) (R = $CONHCH_2CH_2CF_3$) and compound **94**. (B) Reported compounds **95**–**100** with profound structural proximity to compounds **87**–**93**. (C) Reported JAK2 inhibitors **101** and **102** structurally close to compound **87**. (D) Marketed JAK inhibitors sharing the same scaffolds with compounds **90**–**92**. (E) Potential bioisosteric ancestors (**106** and **107**) of compound **89**. (F) Patent analysis of the scaffolds (**108**–**112**) derived from the generated structures.

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

interventions. The rationale behind generating hundreds of chemical structures, docking them, and finally selecting the one that actually requires only one well-documented manual modification with the reported lead compound remains unclear. The authors[80] emphasized that SyntaLinker produced the novel pyrrolo[2,3-*d*]-2-aminopyrimidine scaffold for TBK1 inhibitors, which is an arguable statement with respect to the actual scaffold novelty, as it has been widely reported in 87 patents (according to SciFinder) and mostly in the context of the key part of a kinase inhibitor.

Yu et al. (Hitgen Inc. and Tencent AI Lab) reported a novel deep generative scaffold hopping algorithm (GraphGMVAE) for the *de novo* design of new JAK1 inhibitors.[86] It was pointed out that the main purposes of the scaffold hopping approach in medicinal chemistry are generating novel, diverse, and patentable molecules, selectivity, and PK optimization. Hence, the native activity and key periphery motifs of a template molecule should remain unchanged. It was also noted that the conventional way to perform scaffold hopping is to use commercial software instruments, which often have a limited scaffold database, long computational time, or expensive licenses. Thus, a quicker engine for scaffold hopping based on DL approaches was needed. Therefore, a generative pipeline comprising two main parts was developed. The first part was a generative model (GraphGMVAE with scaffold extraction algorithmics) that produced >30 000 molecular structures carrying novel scaffolds (as the authors claimed) distinct from the reported JAK inhibitors. Upadacitinib was used as a starting point (see Figure 8A, **86**). All the generated SMILES-coded structures are available in the authors' related supporting information. The second part was a prioritization pipeline with a set of filters (customized druglikeness parameters and MCFs, including PAINS and in-house rules) and scoring systems: consensus structure-based pharmacophore modeling and molecular docking, which exploited the cocrystals of the reported JAK1 inhibitors, and the in-house kinase 3D-CNN binding affinity predictor. The pipeline produced 25 high-score structures, which were selected for novelty assessment. Compounds **87**−**93** were synthesized and evaluated in an enzymatic JAK1 inhibition assay. Six molecules (**87**−**92**) exhibited good hit-level potencies.

However, the novelty of the scaffolds in terms of structure is arguable. For instance, 7-azaindole and pyrrolo[2,3-*d*]-pyrimidine are overused PKI scaffolds found in FMS (**95**, ChEMBL ID: CHEMBL3673933),[87] PKCi (**96**),[88] CSF1R (**97**),[89] and other kinase inhibitors (Figure 8B). These motives are also present in nonkinase modulators, e.g., C5a inhibitors (**98**, ChEMBL ID: CHEMBL364695),[90] Dot1L inhibitors (**99**, ChEMBL ID: CHEMBL4099771),[91] influenza PB2 inhibitors (**100**, ChEMBL ID: CHEMBL3317992),[92] etc. The side-chain novelty evaluation was out of the scope of further investigation owing to the study design. It was particularly emphasized that 97.9% of the 30 000 generated structures contained novel scaffolds that were distinct from the reported JAK inhibitors. It would be interesting for the medicinal chemistry experts to give opinions on the generated items as, however, the researchers did not provide the reference data set with JAKinibs used for the novelty calculation. The authors also argued that the structure of **87** was published,[93] two months after their algorithm generated the same structure (July 2020). Although it was checked that this scaffold was absent in the training data set, no information was provided regarding close analogues. The priority date of this application is March 14, 2019, which should be assigned to this structure. A similarity search in SciFinder unveiled a subnanomolar JAK1 inhibitor, a close analogue of **87** with a primary amide at the fifth position published a year earlier (Figure 8A, **94**).[94] In addition to the primary amide, methyl and fluorine were claimed as possible substituents at this position but not a bioisosteric or lipophilic trifluoromethyl group, making **87** the *de jure* novel JAK1 inhibitor with a relatively weak IP status. Yet another publication by Heinrich et al. reported substituted 5-trifluoromethyl-7-azaindoles as FAK inhibitors (ChEMBL ID: CHEMBL2311330).[95] Among them, compounds **101** (with no substituent and nitrile group, a bioisostere of $-CF_3$, in the second and fifth positions, respectively) and **102** (with a bulky 4-phenyl substituent and $-CF_3$ in the second and fifth positions, respectively) demonstrated high off-target activity (68 and 88% inhibition at 1 μM) against JAK2 kinase (Figure 8C). Scaffolds of **90**−**92** were presented in the structures of launched JAK inhibitors (Figure 8D). Although compound **89** might seem the topologically novel structure, it can be obtained via scaffold morphing and a couple of bioisosteric transformations starting from the described JAK inhibitors[96,97] (Figure 8E). The authors of this review also performed a preliminary check of **87**−**93** cores (**108**−**112**) in external databases, literature, and patents before July 2020 (Figure 8F). Based on the obtained result, almost all these scaffolds look tarnished.

Cliff's Notes. The conclusions and recommendations of this review are as follows:

1. The generated structures should be thoroughly inspected for their novelty with high attention to refute skepticism about their real IP position. In addition to free access to public databases like ChEMBL, PubChem, DrugBank, ChemSpider, and SureChEMBL, researchers are strongly recommended to seek resources for purchasing access to commercially available databases including SciFinder which many medicinal chemists routinely use.

2. Many of the generated structures contain well-documented structure alerts and confusing substructures with unsynthesizable moieties or even incorrect valencies. This problem can be overcome by using a set of rationally balanced and well-adapted for generative pipelines preprocessing rules and medicinal chemistry filters.

3. It is highly recommended for AI specialists to pay more attention to conventional terminology used in the field of medicinal chemistry and avoid misleading statements, especially "novel drug candidate" and "novel lead compounds", because these terms need to be supported with exhaustive biological data. In many cases, "primarily hit compound" is the only term that can be reasonably applied for active compounds of generative origin.

4. Frequently, AI-based generative pipelines produce active structures which can be readily obtained using routine medicinal chemistry approaches like trivial bioisosteric replacement and scaffold hopping, ring fusions, or cyclic analogues. To generate such a library of close structural homologues, a medicinal chemist uses available chemoinformatic software, and it is not a time-consuming process. Considering this, generative pipelines should be implemented at least with severe similarity metrics.

5. To correctly estimate the results of generative engines, a medicinal chemist needs to be provided with all the generated structures besides those presented by authors as the most promising ones. This considerably assists a medicinal chemist

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

to make more adequate and valuable feedback on a generation output.

6. In general, AI-based generative platforms are able to generate kinase inhibitors; however, this is a fairly narrow area of medicinal chemistry with historically predefined fragments at least within competitive hinge-targeted inhibitors. A plethora of patents and scientific publications describing kinase inhibitors significantly reduce the chance that trivial analogues will be attractive for IP profiles. Therefore, the results of generations beyond well-validated targets are actually needed to evaluate the ability of such AI-based pipelines to produce novel active structures, especially in the PPI area.

7. Active molecules of AI origin need to be evaluated at least with the use of the standard MTS or MTT tests to avoid nonspecific action and cytotoxicity. Furthermore, a preliminary evaluation of the selectivity and off-target activity of generated molecules allows AI scientists to get more valuable and comprehensive feedback.

8. The assessment of the synthetic accessibility of generated structures is a very important step in any generative platform for eliminating not valid structures and to reduce the cost of organic synthesis. The synthetic routes which are based on the commercially available building blocks can provide researchers with primarily hit compounds in a more rapid way. Furthermore, it should be pointed out that metrics such as the SA score are not good predictors of synthetic accessibility. While commonly used in publications as such those reviewed in this paper, these metrics actually measure molecular complexity[98] rather than synthetic accessibility.

9. Generative engines should be improved to provide computational chemists and medicinal chemists with novel ideas and truly novel molecular structures at least with more chances to be attractive and valuable for filling the IP profile. More attention should particularly be placed for a training set and similarity metrics. Undoubtedly, currently a medicinal chemist basically expects novel ideas from generative pipelines. Based on the selected one, he usually modifies a structure in order to optimize the required properties, especially the novelty.

Even though the current review covers only the recent and yet uncited research outcomes in other reviews, the distribution of molecular targets exploited for the reviewed studies is very biased toward protein kinases (seven out of eight), and this tendency can be noticed in general for much research in the field for the last five years. Although protein kinases are still the "hottest" drug targets, the scientific community is eager to see more challenging targets in AIDD, the attractive proteins from other target families, particularly those that have been considered undruggable so far. It is believed that AI/ML can contribute to the drug discovery process more sophisticatedly than now and produce novel and relevant molecular structures. AI is not required to conduct tasks, such as trivial scaffold replacement, that can be rapidly performed by an ordinary medicinal chemist as the scientific community expects valuable advances and breakthroughs from AIDD. From this perspective, the deeper involvement of skilled medicinal chemists in developing generative chemistry pipelines would highly benefit the overall evolution of drug discovery. In the context of the published studies, it would be highly appreciated if the training data sets and all the generated structures were provided rather than picking and publishing only a few selected, synthesized, and tested ones. Medicinal chemistry experts should iteratively provide feedback on the validity, relevance, and actual novelty of all the generated structures to AI specialists. The feedback should be intensively used for learning and enhancing generative algorithms. Otherwise, irrelevant AI-related research outputs with *vecchio* scaffolds, which can be easily produced by most medicinal chemists without any AI intervention, will be seen.

AUTHOR INFORMATION

Corresponding Author

Alex Zhavoronkov − *Insilico Medicine Hong Kong Ltd., Pak Shek Kok, Hong Kong;* orcid.org/0000-0001-7067-8966; Email: alex@insilico.com

Authors

Yan Ivanenkov − *Insilico Medicine Hong Kong Ltd., Pak Shek Kok, Hong Kong;* orcid.org/0000-0002-8968-0879

Bogdan Zagribelnyy − *Insilico Medicine AI Limited, Abu Dhabi, United Arab Emirates*

Alex Malyshev − *Insilico Medicine Hong Kong Ltd., Pak Shek Kok, Hong Kong*

Sergei Evteev − *Insilico Medicine Hong Kong Ltd., Pak Shek Kok, Hong Kong*

Victor Terentiev − *Insilico Medicine Hong Kong Ltd., Pak Shek Kok, Hong Kong*

Petrina Kamya − *Insilico Medicine Canada Inc., Montreal, Quebec, Canada H3B 4W8*

Dmitry Bezrukov − *Insilico Medicine Hong Kong Ltd., Pak Shek Kok, Hong Kong*

Alex Aliper − *Insilico Medicine AI Limited, Abu Dhabi, United Arab Emirates;* orcid.org/0000-0002-4363-0710

Feng Ren − *Insilico Medicine Shanghai Ltd., Shanghai 201203, China;* orcid.org/0000-0001-9157-9182

Complete contact information is available at: https://pubs.acs.org/10.1021/acsmedchemlett.3c00041

Notes

The authors declare the following competing financial interest(s): All authors work for Insilico Medicine, a commercial artificial intelligence company.

REFERENCES

(1) Goldman, B.; Kearnes, S.; Kramer, T.; Riley, P.; Walters, W. P. Defining Levels of Automated Chemical Design. *J. Med. Chem.* **2022**, *65* (10), 7073−7087.

(2) Sheridan, R. P.; Kearsley, S. K. Using a Genetic Algorithm To Suggest Combinatorial Libraries. *J. Chem. Inf. Comput. Sci.* **1995**, *35* (2), 310−320.

(3) Brown, N.; McKay, B.; Gilardoni, F.; Gasteiger, J. A Graph-Based Genetic Algorithm and Its Application to the Multiobjective Evolution of Median Molecules. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (3), 1079−1087.

(4) Zhavoronkov, A. Artificial Intelligence for Drug Discovery, Biomarker Development, and Generation of Novel Chemistry. *Mol. Pharmaceutics* **2018**, *15* (10), 4311−4313.

(5) Cao, C.; Liu, F.; Tan, H.; Song, D.; Shu, W.; Li, W.; Zhou, Y.; Bo, X.; Xie, Z. Deep Learning and Its Applications in Biomedicine. *Genomics Proteomics Bioinformatics* **2018**, *16* (1), 17−32.

(6) Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Mamoshina, P.; Zhavoronkov, A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Mol. Pharmaceutics* **2016**, *13* (7), 2524−2530.

(7) Vanhaelen, Q.; Lin, Y.-C.; Zhavoronkov, A. The Advent of Generative Chemistry. *ACS Med. Chem. Lett.* **2020**, *11* (8), 1496−1505.

(8) Grisoni, F.; Moret, M.; Lingwood, R.; Schneider, G. Bidirectional Molecule Generation with Recurrent Neural Networks. *J. Chem. Inf. Model.* **2020**, *60* (3), 1175−1183.
(9) Kadurin, A.; Nikolenko, S.; Khrabrov, K.; Aliper, A.; Zhavoronkov, A. druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico. *Mol. Pharmaceutics* **2017**, *14* (9), 3098−3104.
(10) Kadurin, A.; Aliper, A.; Kazennov, A.; Mamoshina, P.; Vanhaelen, Q.; Khrabrov, K.; Zhavoronkov, A. The Cornucopia of Meaningful Leads: Applying Deep Adversarial Autoencoders for New Molecule Development in Oncology. *Oncotarget* **2017**, *8*, 10883.
(11) Liu, Q.; Allamanis, M.; Brockschmidt, M.; Gaunt, A. L. Constrained Graph Variational Autoencoders for Molecule Design. *arXiv (Computer Science.Machine Learning)*, May 23, 2018, 1805.09076, ver. 1. http://arxiv.org/abs/1805.09076.
(12) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018**, *4* (2), 268−276.
(13) Bian, Y.; Wang, J.; Jun, J. J.; Xie, X.-Q. Deep Convolutional Generative Adversarial Network (dcGAN) Models for Screening and Design of Small Molecules Targeting Cannabinoid Receptors. *Mol. Pharmaceutics* **2019**, *16* (11), 4451−4460.
(14) Zhavoronkov, A.; Ivanenkov, Y. A.; Aliper, A.; Veselov, M. S.; Aladinskiy, V. A.; Aladinskaya, A. V.; Terentiev, V. A.; Polykovskiy, D. A.; Kuznetsov, M. D.; Asadulaev, A.; Volkov, Y.; Zholus, A.; Shayakhmetov, R. R.; Zhebrak, A.; Minaeva, L. I.; Zagribelnyy, B. A.; Lee, L. H.; Soll, R.; Madge, D.; Xing, L.; Guo, T.; Aspuru-Guzik, A. Deep Learning Enables Rapid Identification of Potent DDR1 Kinase Inhibitors. *Nat. Biotechnol.* **2019**, *37* (9), 1038−1040.
(15) Putin, E.; Asadulaev, A.; Ivanenkov, Y.; Aladinskiy, V.; Sanchez-Lengeling, B.; Aspuru-Guzik, A.; Zhavoronkov, A. Reinforced Adversarial Neural Computer for de Novo Molecular Design. *J. Chem. Inf. Model.* **2018**, *58* (6), 1194−1204.
(16) Putin, E.; Asadulaev, A.; Vanhaelen, Q.; Ivanenkov, Y.; Aladinskaya, A. V.; Aliper, A.; Zhavoronkov, A. Adversarial Threshold Neural Computer for Molecular de Novo Design. *Mol. Pharmaceutics* **2018**, *15* (10), 4386−4397.
(17) Ren, F.; Ding, X.; Zheng, M.; Korzinkin, M.; Cai, X.; Zhu, W.; Mantsyzov, A.; Aliper, A.; Aladinskiy, V.; Cao, Z.; Kong, S.; Long, X.; Man Liu, B. H.; Liu, Y.; Naumov, V.; Shneyderman, A.; Ozerov, I. V.; Wang, J.; Pun, F. W.; Polykovskiy, D. A.; Sun, C.; Levitt, M.; Aspuru-Guzik, A.; Zhavoronkov, A. AlphaFold Accelerates Artificial Intelligence Powered Drug Discovery: Efficient Discovery of a Novel CDK20 Small Molecule Inhibitor. *Chem. Sci.* **2023**, *14*, 1443.
(18) Zhavoronkov, A.; Kamya, P.; Pun, F. *From Paper to Industrial-scale Platform: a 3-Year Behind the Paper Journey from GENTRL to Chemistry42*. Springer Nature, August 17, 2022. http://bioengineeringcommunity.nature.com/posts/from-paper-to-industrial-scale-platform-a-3-year-behind-the-paper-journey-from-gentrl-to-chemistry42 (accessed 2023-01-31).
(19) Ivanenkov, Y. A.; Polykovskiy, D.; Bezrukov, D.; Zagribelnyy, B.; Aladinskiy, V.; Kamya, P.; Aliper, A.; Ren, F.; Zhavoronkov, A. Chemistry42: An AI-Driven Platform for Molecular Design and Optimization. *J. Chem. Inf. Model.* **2023**, *63* (3), 695−701.
(20) Makara, G. M.; Kovács, L.; Szabó, I.; Pőcze, G. Derivatization Design of Synthetically Accessible Space for Optimization: In Silico Synthesis vs Deep Generative Design. *ACS Med. Chem. Lett.* **2021**, *12* (2), 185−194.
(21) Bos, P. H.; Houang, E. M.; Ranalli, F.; Leffler, A. E.; Boyles, N. A.; Eyrich, V. A.; Luria, Y.; Katz, D.; Tang, H.; Abel, R.; Bhat, S. AutoDesigner, a De Novo Design Algorithm for Rapidly Exploring Large Chemical Space for Lead Optimization: Application to the Design and Synthesis of D-Amino Acid Oxidase Inhibitors. *J. Chem. Inf. Model.* **2022**, *62* (8), 1905−1915.
(22) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31−36.
(23) Polykovskiy, D.; Zhebrak, A.; Sanchez-Lengeling, B.; Golovanov, S.; Tatanov, O.; Belyaev, S.; Kurbanov, R.; Artamonov, A.; Aladinskiy, V.; Veselov, M.; Kadurin, A.; Johansson, S.; Chen, H.; Nikolenko, S.; Aspuru-Guzik, A.; Zhavoronkov, A. Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models. *Front. Pharmacol.* **2020**, *11*, 565644.
(24) Brown, N.; Fiscato, M.; Segler, M. H. S.; Vaucher, A. C. GuacaMol: Benchmarking Models for de Novo Molecular Design. *J. Chem. Inf. Model.* **2019**, *59* (3), 1096−1108.
(25) Li, Y.; Zhang, L.; Wang, Y.; Zou, J.; Yang, R.; Luo, X.; Wu, C.; Yang, W.; Tian, C.; Xu, H.; Wang, F.; Yang, X.; Li, L.; Yang, S. Generative Deep Learning Enables the Discovery of a Potent and Selective RIPK1 Inhibitor. *Nat. Commun.* **2022**, *13* (1), 6891.
(26) Yang, S.; Li, L. 5-Substituted Indole 3-Amide Derivative, Preparation Method and Use Thereof. WO 2022052861 A1, March 17, 2022.
(27) Lee, J.; Jo, S.; Kim, J. E.; Kim, M.; Yang, I. Compounds for Prevention or Treatment of Neurodegenerative Disorders. WO 2022038574 A1, February 24, 2022.
(28) Hatcher, J. M.; Vatsan, P. S.; Wang, E.; Jiang, J.; Gray, N. S. Development of Highly Potent and Selective Pyrazolopyridine Inhibitor of CDK8/19. *ACS Med. Chem. Lett.* **2021**, *12* (11), 1689−1693.
(29) Gray, N. S.; Hatcher, J. M.; Sreevatsan, P. Potent and Selective Azaindole Inhibitors of cdk8 and cdk19. WO 2021108581 A1, June 3, 2021.
(30) Masse, C. E.; Greenwood, J. R.; Mondal, S. Tyk2 Inhibitors and Uses Thereof. WO 2020081508 A1, April 23, 2020.
(31) Jang, S. H.; Sivakumar, D.; Mudedla, S. K.; Choi, J.; Lee, S.; Jeon, M.; Bvs, S. K.; Hwang, J.; Kang, M.; Shin, E. G.; Lee, K. M.; Jung, K.-Y.; Kim, J.-S.; Wu, S. PCW-A1001, AI-Assisted de Novo Design Approach to Design a Selective Inhibitor for FLT-3(D835Y) in Acute Myeloid Leukemia. *Front Mol. Biosci* **2022**, *9*, 1072028.
(32) Kim, J.-S.; Kim, A. Y.; Kang, M.; Jeong, K.-Y.; Jo, S. Composition for Preventing or Treating Cancer, Containing Novel Trifluoromethyl Phenyl Pyrazole Derivative as Active Ingredient. WO 2021060888 A2, April 1, 2021.
(33) Wei, Z.-L.; Nguyen, M. T.; O'Mahony, D. J. R.; Acevedo, A.; Zipfel, S.; Zhang, Q.; Liu, L.; Dourado, M.; Chi, C.; Yip, V.; DeFalco, J.; Gustafson, A.; Emerling, D. E.; Kelly, M. G.; Kincaid, J.; Vincent, F.; Duncton, M. A. J. Identification of Orally-Bioavailable Antagonists of the TRPV4 Ion-Channel. *Bioorg. Med. Chem. Lett.* **2015**, *25* (18), 4011−4015.
(34) Leitans, J.; Sprudza, A.; Tanc, M.; Vozny, I.; Zalubovskis, R.; Tars, K.; Supuran, C. T. 5-Substituted-(1,2,3-Triazol-4-Yl)thiophene-2-Sulfonamides Strongly Inhibit Human Carbonic Anhydrases I, II, IX and XII: Solution and X-Ray Crystallographic Studies. *Bioorg. Med. Chem.* **2013**, *21* (17), 5130−5138.
(35) Stamford, A. W.; Boyle, C. D.; Huang, Y. Substituted Imidazole Neuropeptide Y y5 Receptor Antagonists. WO 200144201 A1, June 21, 2001.
(36) Deng, Y.; Jia, Y.; Tian, Y. Acryloyl-Substituted Compound, Pharmaceutical Composition Containing Same, and Use Thereof. WO 2022199599 A1, September 29, 2022.
(37) Hamaguchi, H.; Amano, Y.; Moritomo, A.; Shirakami, S.; Nakajima, Y.; Nakai, K.; Nomura, N.; Ito, M.; Higashi, Y.; Inoue, T. Discovery and Structural Characterization of Peficitinib (ASP015K) as a Novel and Potent JAK Inhibitor. *Bioorg. Med. Chem.* **2018**, *26* (18), 4971−4983.
(38) Tan, X.; Li, C.; Yang, R.; Zhao, S.; Li, F.; Li, X.; Chen, L.; Wan, X.; Liu, X.; Yang, T.; Tong, X.; Xu, T.; Cui, R.; Jiang, H.; Zhang, S.; Liu, H.; Zheng, M. Discovery of Pyrazolo[3,4-D]pyridazinone Derivatives as Selective DDR1 Inhibitors via Deep Learning Based Design, Synthesis, and Biological Evaluation. *J. Med. Chem.* **2022**, *65* (1), 103−119.

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

(39) Biagi, G.; Giorgi, I.; Livi, O.; Scartoni, V.; Martini, C.; Tacchi, P.; Merlino, S.; Pasero, M. 1,2,3-Triazole[4,5-D]pyridazines-II. New Derivatives Tested on Adenosine Receptors. *Farmaco* **1994**, *49* (3), 175−181.
(40) Biagi, G.; Giorgi, I.; Livi, O.; Manera, C.; Scartoni, V.; Betti, L.; Giannaccini, G.; Lucacchini, A. 1,2,3-Triazolo[4,5-D]pyridazines: Part VI. New 1-Substituted-4-Amino Derivatives and Their Affinity towards A1 and A2A Adenosine Receptors. *Il Farmaco* **1999**, *54* (9), 615−623.
(41) Seiji, E.; Yoshikawa, F.; Emori, C.; Matsuura, R.; Hironori, K.; Ikuta, N.; Kira, T.; Yasuda, K.; Nagakura, Y. New Condensed Imidazole Derivatives. WO 2003104229 A1, October 6, 2005.
(42) Gyorkos, A. C.; Corrette, C. P.; Young Cho, S.; Turner, T. M.; Aso, K.; Kori, M.; Mochizuki, M.; Condroski, K. R.; Siedem, C. S.; Boyd, S. A. Cyclic Compounds. WO 2005099688 A2, October 27, 2005.
(43) Liu, D.; Zhang, M.; Hu, Q. Aminopyridazinone Compounds as Protein Kinase Inhibitors. WO 2016007185 A1, January 14, 2016.
(44) Dawood, K. M.; Farag, A. M.; Abdel-Aziz, H. A. Azoles and Azolo-Azines via 3-(3-Methylbenzofuran-2-Yl)-3-Oxopropanenitrile. *J. Chem. Res.* **2005**, *2005* (6), 378−381.
(45) Bastide, J. 1,3-Dipolar Cycloaddition of Diazomethane. II. Addition to Difunctionalized Acetylenic Compounds. Use of the Pyrazoles Obtained to Prepare Pyrazolo [3,4-D]pyridazines. *Bull. Soc. Chim. Fr.* **1971**, *4*, 1336−1343.
(46) Elmoghayar, M. R. H.; Elnagdi, M. H.; Ibrahim, M. K. A.; Sallam, M. M. M. Reactions with Carbo(heterylhydrazonoyl) Halides. I. Chemistry of carbo(3-Phenylpyrazol-5-Ylhydrazonoyl) Chlorides. *Helv. Chim. Acta* **1977**, *60* (7), 2171−2176.
(47) Wang, Y.; Dai, Y.; Wu, X.; Li, F.; Liu, B.; Li, C.; Liu, Q.; Zhou, Y.; Wang, B.; Zhu, M.; Cui, R.; Tan, X.; Xiong, Z.; Liu, J.; Tan, M.; Xu, Y.; Geng, M.; Jiang, H.; Liu, H.; Ai, J.; Zheng, M. Discovery and Development of a Series of Pyrazolo[3,4-D]pyridazinone Compounds as the Novel Covalent Fibroblast Growth Factor Receptor Inhibitors by the Rational Drug Design. *J. Med. Chem.* **2019**, *62* (16), 7473−7488.
(48) Liang, Y.; Xu, Z.; Zeng, Z.; Yan, W.; Xiong, F. Fgfr Kinase Inhibitor and Use Thereof. WO 2022166469 A1, August 11, 2022.
(49) He, C.; Wang, X.; Rong, Y.; Tan, R.; Zhang, H.; Tan, H.; Liu, Y.; Lin, S.; Li, T.; Zhao, X.; Wang, W. Compounds Useful as Kinase Inhibitors. WO 2021180107 A1, September 16, 2021.
(50) Zhang, X.; Sheng, X.; Shen, J.; Zhang, S.; Sun, W.; Shen, C.; Li, Y.; Wang, J.; Lv, H.; Cui, M.; Zhu, Y.; Huang, L.; Hao, D.; Qi, Z.; Sun, G.; Mao, W.; Pan, Y.; Shen, L.; Li, X.; Hu, G.; Gong, Z.; Han, S.; Li, J.; Chen, S.; Tu, R.; Wang, X.; Wu, C. Discovery and Evaluation of Pyrazolo[3,4-D]pyridazinone as a Potent and Orally Active Irreversible BTK Inhibitor. *ACS Med. Chem. Lett.* **2020**, *11* (10), 1863−1868.
(51) Chen, N.; Yang, L.; Ding, N.; Li, G.; Cai, J.; An, X.; Wang, Z.; Qin, J.; Niu, Y. Recurrent Neural Network (RNN) Model Accelerates the Development of Antibacterial Metronidazole Derivatives. *RSC Adv.* **2022**, *12*, 22893−22901.
(52) Dingsdag, S. A.; Hunter, N. Metronidazole: An Update on Metabolism, Structure-cytotoxicity and Resistance Mechanisms. *J. Antimicrob. Chemother.* **2018**, *73* (2), 265−279.
(53) Andrews, W. J.; Ray, S.; Panova, T.; Engel, C.; Panov, K. I. DNA Intercalators Inhibit Eukaryotic Ribosomal RNA Synthesis by Impairing the Initiation of Transcription. *Genes* **2021**, *12* (9), 1412.
(54) Huggins, D. J.; Venkitaraman, A. R.; Spring, D. R. Rational Methods for the Selection of Diverse Screening Compounds. *ACS Chem. Biol.* **2011**, *6* (3), 208−217.
(55) Lavis, L. D. Ester Bonds in Prodrugs. *ACS Chem. Biol.* **2008**, *3* (4), 203−206.
(56) Kumari, M.; Singh, P. Tribulus Terrestris Ameliorates Metronidazole-Induced Spermatogenic Inhibition and Testicular Oxidative Stress in the Laboratory Mouse. *Indian J. Pharmacol.* **2015**, *47* (3), 304−310.
(57) Menéndez, D.; Bendesky, A.; Rojas, E.; Salamanca, F.; Ostrosky-Wegman, P. Role of P53 Functionality in the Genotoxicity of Metronidazole and Its Hydroxy Metabolite. *Mutat. Res.* **2002**, *501* (1−2), 57−67.
(58) Menéndez, D.; Rojas, E.; Herrera, L. A.; López, M. C.; Sordo, M.; Elizondo, G.; Ostrosky-Wegman, P. DNA Breakage due to Metronidazole Treatment. *Mutat. Res.* **2001**, *478* (1−2), 153−158.
(59) Du, Q.-R.; Li, D.-D.; Pi, Y.-Z.; Li, J.-R.; Sun, J.; Fang, F.; Zhong, W.-Q.; Gong, H.-B.; Zhu, H.-L. Novel 1,3,4-Oxadiazole Thioether Derivatives Targeting Thymidylate Synthase as Dual Anticancer/antimicrobial Agents. *Bioorg. Med. Chem.* **2013**, *21* (8), 2286−2297.
(60) Antonova-Koch, Y.; Meister, S.; Abraham, M.; Luth, M. R.; Ottilie, S.; Lukens, A. K.; Sakata-Kato, T.; Vanaerschot, M.; Owen, E.; Jado, J. C.; Maher, S. P.; Calla, J.; Plouffe, D.; Zhong, Y.; Chen, K.; Chaumeau, V.; Conway, A. J.; McNamara, C. W.; Ibanez, M.; Gagaring, K.; Serrano, F. N.; Eribez, K.; Taggard, C. M.; Cheung, A. L.; Lincoln, C.; Ambachew, B.; Rouillier, M.; Siegel, D.; Nosten, F.; Kyle, D. E.; Gamo, F.-J.; Zhou, Y.; Llinás, M.; Fidock, D. A.; Wirth, D. F.; Burrows, J.; Campo, B.; Winzeler, E. A. Open-Source Discovery of Chemical Leads for next-Generation Chemoprotective Antimalarials. *Science* **2018**, *362* (6419), eaat9446.
(61) Kumar, L.; Jain, A.; Lal, N.; Sarswat, A.; Jangir, S.; Kumar, L.; Singh, V.; Shah, P.; Jain, S. K.; Maikhuri, J. P.; Siddiqi, M. I.; Gupta, G.; Sharma, V. L. Potentiating Metronidazole Scaffold against Resistant Trichomonas: Design, Synthesis, Biology and 3D-QSAR Analysis. *ACS Med. Chem. Lett.* **2012**, *3* (2), 83−87.
(62) Salahuddin, A.; Agarwal, S. M.; Avecilla, F.; Azam, A. Metronidazole Thiosalicylate Conjugates: Synthesis, Crystal Structure, Docking Studies and Antiamoebic Activity. *Bioorg. Med. Chem. Lett.* **2012**, *22* (17), 5694−5699.
(63) Yoshimori, A.; Asawa, Y.; Kawasaki, E.; Tasaka, T.; Matsuda, S.; Sekikawa, T.; Tanabe, S.; Neya, M.; Natsugari, H.; Kanai, C. Design and Synthesis of DDR1 Inhibitors with a Desired Pharmacophore Using Deep Generative Models. *ChemMedChem.* **2021**, *16* (6), 955−958.
(64) Jeffries, D. E.; Borza, C. M.; Blobaum, A. L.; Pozzi, A.; Lindsley, C. W. Discovery of VU6015929: A Selective Discoidin Domain Receptor 1/2 (DDR1/2) Inhibitor to Explore the Role of DDR1 in Antifibrotic Therapy. *ACS Med. Chem. Lett.* **2020**, *11* (1), 29−33.
(65) Kodama, Y.; Takeuchi, K.; Shimba, N.; Ishikawa, K.; Suzuki, E.-I.; Shimada, I.; Takahashi, H. Rapid Identification of Ligand-Binding Sites by Using an Assignment-Free NMR Approach. *J. Med. Chem.* **2013**, *56* (22), 9342−9350.
(66) Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; Woodhead, A. J.; Walker, M. T.; Congreve, M. S.; Devine, L. A.; Tisi, D.; O'Reilly, M.; Seavers, L. C. A.; Davis, D. J.; Curry, J.; Anthony, R.; Padova, A.; Murray, C. W.; Carr, R. A. E.; Jhoti, H. Identification of Novel p38alpha MAP Kinase Inhibitors Using Fragment-Based Lead Generation. *J. Med. Chem.* **2005**, *48* (2), 414−426.
(67) Murray, C. W.; Hartshorn, M. J.; Frederickson, M.; Congreve, M. S.; Padova, A.; Woodhead, S. J.; Gill, A. L.; Woodhead, A. J. 3-(hetero) Arylmethoxy ! Pyridines and Their Analogues as p38 Map Kinase Inhibitors. WO 2004004720 A1, January 15, 2004.
(68) Gill, A. L.; Woodhead, S. J.; Woodhead, A. J.; Frederickson, M.; Padova, A.; Apaya, R. P. Raf Kinase Inhibitors. WO 2005002673 A1, January 13, 2005.
(69) Feng, Y.; Wang, R.; Li, J.; Zheng, J.; Lian, X.; Gong, X.; Fu, Y.; Kang, X. Heteroaryl Compounds as Kinase Inhibitor. WO 2019174601 A1, September 19, 2019.
(70) Lyne, P. D.; Aquila, B.; Cook, D. J.; Dakin, L. A.; Ezhuthachan, J.; Ioannidis, S.; Pontz, T.; Su, M.; Ye, Q.; Zheng, X.; Block, M. H.; Cowen, S.; Deegan, T. L.; Lee, J. W.; Scott, D. A.; Custeau, D.; Drew, L.; Poondru, S.; Shen, M.; Wu, A. Identification of Amidoheteroaryls as Potent Inhibitors of Mutant (V600E) B-Raf Kinase with in Vivo Activity. *Bioorg. Med. Chem. Lett.* **2009**, *19* (3), 1026−1029.
(71) Hua, Y.; Fang, X.; Xing, G.; Xu, Y.; Liang, L.; Deng, C.; Dai, X.; Liu, H.; Lu, T.; Zhang, Y.; Chen, Y. Effective Reaction-Based De Novo Strategy for Kinase Targets: A Case Study on MERTK Inhibitors. *J. Chem. Inf. Model.* **2022**, *62* (7), 1654−1668.

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX

(72) Degen, J.; Wegscheid-Gerlach, C.; Zaliani, A.; Rarey, M. On the Art of Compiling and Using "Drug-like" Chemical Fragment Spaces. *ChemMedChem.* **2008**, *3* (10), 1503−1507.
(73) Horwood, J.; Noutahi, E. Molecular Design in Synthetically Accessible Chemical Space via Deep Reinforcement Learning. *ACS Omega* **2020**, *5* (51), 32984−32994.
(74) Zabolotna, Y.; Volochnyuk, D. M.; Ryabukhin, S. V.; Gavrylenko, K.; Horvath, D.; Klimchuk, O.; Oksiuta, O.; Marcou, G.; Varnek, A. SynthI: A New Open-Source Tool for Synthon-Based Library Design. *J. Chem. Inf. Model.* **2022**, *62*, 2151.
(75) Ertl, P.; Schuffenhauer, A. Estimation of Synthetic Accessibility Score of Drug-like Molecules Based on Molecular Complexity and Fragment Contributions. *J. Cheminform.* **2009**, *1* (1), 8.
(76) Bhagwat, S. S.; Satoh, Y.; Sakata, S. T. Indazole Derivatives as Jnk Inhibitors. WO 2002010137 A2, February 7, 2002.
(77) Jorand-Lebrun, C.; Crosignani, S.; Dorbais, J.; Grippi-Vallotton, T.; Pretre, A. Indazolyl Triazole Derivatives as IRAK Inhibitors. WO 2012084704 A1, June 28, 2012.
(78) Ancliff, R. A.; Atkinson, F. L.; Barker, M. D.; Box, P. C.; Daniel, C.; Gore, P. M.; Guntrip, S. B.; Hasegawa, M.; Inglis, G. G. A.; Kano, K.; Miyazaki, Y.; Patel, V. K.; Ritchie, T. J.; Swanson, S.; Walker, A. L.; Wellaway, C. R.; Woodrow, M. Pyrrolopyrimidine Derivatives as Syk Inhibitors. WO 2007042299 A1, April 19, 2007.
(79) Song, Y.; Xu, Q.; Bauer, S. M.; Jia, Z. J.; Mehrotra, M; Pandey, A. Inhibitors of Protein Kinases. WO 2009131687 A2, October 29, 2009.
(80) Song, S.; Tang, H.; Ran, T.; Fang, F.; Tong, L.; Chen, H.; Xie, H.; Lu, X. Application of Deep Generative Model for Design of Pyrrolo[2,3-D] Pyrimidine Derivatives as New Selective TANK Binding Kinase 1 (TBK1) Inhibitors. *Eur. J. Med. Chem.* **2023**, *247*, 115034.
(81) Yang, Y.; Zheng, S.; Su, S.; Zhao, C.; Xu, J.; Chen, H. SyntaLinker: Automatic Fragment Linking with Deep Conditional Transformer Neural Networks. *Chem. Sci.* **2020**, *11* (31), 8312−8322.
(82) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40* (D1), D1100−D1107.
(83) Harris, C. M.; Ericsson, A. M.; Argiriadi, M. A.; Barberis, C.; Borhani, D. W.; Burchat, A.; Calderwood, D. J.; Cunha, G. A.; Dixon, R. W.; Frank, K. E.; Johnson, E. F.; Kamens, J.; Kwak, S.; Li, B.; Mullen, K. D.; Perron, D. C.; Wang, L.; Wishart, N.; Wu, X.; Zhang, X.; Zmetra, T. R.; Talanian, R. V. 2,4-Diaminopyrimidine MK2 Inhibitors. Part II: Structure-Based Inhibitor Optimization. *Bioorg. Med. Chem. Lett.* **2010**, *20* (1), 334−337.
(84) Zeifman, A. A.; Titov, I. Y.; Svitanko, I. V.; Rakitina, T. V.; Lipkin, A. V.; Stroylov, V. S.; Stroganov, O. V.; Novikov, F. N.; Chilov, G. G. Rational Design and Synthesis of Novel Syk-Kinase Inhibitors. *Mendeleev Commun.* **2012**, *22* (2), 73−74.
(85) Choi, H.-S.; Wang, Z.; Richmond, W.; He, X.; Yang, K.; Jiang, T.; Sim, T.; Karanewsky, D.; Gu, X.-J.; Zhou, V.; Liu, Y.; Ohmori, O.; Caldwell, J.; Gray, N.; He, Y. Design and Synthesis of 7H-pyrrolo[2,3-D]pyrimidines as Focal Adhesion Kinase Inhibitors. Part 1. *Bioorg. Med. Chem. Lett.* **2006**, *16* (8), 2173−2176.
(86) Yu, Y.; Xu, T.; Li, J.; Qiu, Y.; Rong, Y.; Gong, Z.; Cheng, X.; Dong, L.; Liu, W.; Li, J.; Dou, D.; Huang, J. A Novel Scalarized Scaffold Hopping Algorithm with Graph-Based Variational Autoencoder for Discovery of JAK1 Inhibitors. *ACS Omega* **2021**, *6* (35), 22945−22954.
(87) Zhang, J.; Ibrahim, P. N.; Bremer, R.; Spevak, W.; Cho, H. Compounds and Methods for Kinase Modulation, and Indications Therefor. WO 2011057022 A1, May 12, 2011.
(88) Breslin, H. J.; Dorsey, B. D.; Dugan, B. J.; Fowler, K. M.; Hudkins, R. L.; Mesaros, E. F.; Monck, N. J. T.; Morris, E. L.; Olowoye, I.; Ott, G. R.; Pave, G. A.; Roffey, J. R. A.; Soudy, C. N.; Tao, M.; Zificsak, C. A.; Zulli, A. L. Azaquinazoline Inhibitors of Atypical Protein Kinase c. WO 2014052699 A9, July 16, 2015.
(89) Hsieh, H.-P.; Lee, K.-H.; Lin, W.-H.; Shih, C. Heterocyclic Compounds as Kinase Inhibitors for Therapeutic Uses. WO 2020210481 A1, November 19, 2020.
(90) Gelin, C.; Flyer, A.; Adams, C. M.; Darsigny, V.; Hurley, T. B.; Karki, R. G.; Ji, N.; Kawanami, T.; Meredith, E.; Serrano-Wu, M. H.; Rao, C.; Solovay, C.; Lee, G. T.-S.; Towler, C.; Har, D.; Shen, L.; Hu, B.; Jiang, X.; Cappaci-Daniel, C. Tetrahydropyrido-Pyridine and Tetrahydropyrido-Pyrimidine Compounds and Use Thereof as c5a Receptor Modulators. WO 2013016197 A1, January 31, 2013.
(91) Möbitz, H.; Machauer, R.; Holzer, P.; Vaupel, A.; Stauffer, F.; Ragot, C.; Caravatti, G.; Scheufler, C.; Fernandez, C.; Hommel, U.; Tiedt, R.; Beyer, K. S.; Chen, C.; Zhu, H.; Gaul, C. Discovery of Potent, Selective, and Structurally Novel Dot1L Inhibitors by a Fragment Linking Approach. *ACS Med. Chem. Lett.* **2017**, *8* (3), 338−343.
(92) Clark, M. P.; Ledeboer, M. W.; Davies, I.; Byrn, R. A.; Jones, S. M.; Perola, E.; Tsai, A.; Jacobs, M.; Nti-Addae, K.; Bandarage, U. K.; Boyd, M. J.; Bethiel, R. S.; Court, J. J.; Deng, H.; Duffy, J. P.; Dorsch, W. A.; Farmer, L. J.; Gao, H.; Gu, W.; Jackson, K.; Jacobs, D. H.; Kennedy, J. M.; Ledford, B.; Liang, J.; Maltais, F.; Murcko, M.; Wang, T.; Wannamaker, M. W.; Bennett, H. B.; Leeman, J. R.; McNeil, C.; Taylor, W. P.; Memmott, C.; Jiang, M.; Rijnbrand, R.; Bral, C.; Germann, U.; Nezami, A.; Zhang, Y.; Salituro, F. G.; Bennani, Y. L.; Charifson, P. S. Discovery of a Novel, First-in-Class, Orally Bioavailable Azaindole Inhibitor (VX-787) of Influenza PB2. *J. Med. Chem.* **2014**, *57* (15), 6668−6678.
(93) Xu, X.; Li, Y.; Zhang, Z.; Wang, G.; Zhang, L.; Yu, H.; Qin, H.; Zhang, N.; Liu, L.; Zhu, W.; Xi, C.; Li, Q.; Qu, M.; Wang, Y.; Ge, J. Jak Kinase Inhibitor, Preparation Method for Same, and Applications Thereof in Field of Medicine. WO 2020182159 A1, September 17, 2020.
(94) Xu, X.; Zhang, Z.; Li, Y.; Wang, G.; Zhu, W.; Li, Q.; Qu, M.; Zhang, L.; Song, J.; Liu, L.; Chen, H.; Liu, Q.; Wang, Y.; Ge, J. Jak Inhibitor, Preparation Method Therefor, and Application in the Field of Medicine. WO 2019057103 A1, March 28, 2019.
(95) Heinrich, T.; Seenisamy, J.; Emmanuvel, L.; Kulkarni, S. S.; Bomke, J.; Rohdich, F.; Greiner, H.; Esdar, C.; Krier, M.; Grädler, U.; Musil, D. Fragment-Based Discovery of New Highly Substituted 1H-pyrrolo[2,3-B]- and 3H-imidazolo[4,5-B]-Pyridines as Focal Adhesion Kinase Inhibitors. *J. Med. Chem.* **2013**, *56* (3), 1160−1170.
(96) De Vicente Fidalgo, J.; Hermann, J. C.; Lemoine, R.; Li, H.; Lovey, A. J.; Sjogren, E. B.; Soth, M. Inhibitors of Jak. WO 2011029804 A1, March 17, 2011.
(97) Elsayed, M. S. A.; Nielsen, J. J.; Park, S.; Park, J.; Liu, Q.; Kim, C. H.; Pommier, Y.; Agama, K.; Low, P. S.; Cushman, M. Application of Sequential Palladium Catalysis for the Discovery of Janus Kinase Inhibitors in the Benzo[c]pyrrolo[2,3- h][1,6]naphthyridin-5-One (BPN) Series. *J. Med. Chem.* **2018**, *61* (23), 10440−10462.
(98) Allu, T. K.; Oprea, T. I. Rapid Evaluation of Synthetic and Molecular Complexity for in Silico Chemistry. *J. Chem. Inf. Model.* **2005**, *45* (5), 1237−1243.

https://doi.org/10.1021/acsmedchemlett.3c00041
ACS Med. Chem. Lett. XXXX, XXX, XXX−XXX