Not Every Hit-Identification Technique Works on 1-Deoxy-D-Xylulose 5-Phosphate Synthase (DXPS)

Johannsen, Sandra; Gierse, Robin M.; Olshanova, Aleksandra; Smerznak, Ellie; Laggner, Christian; Eschweiler, Lea; Adeli, Zahra; Hamid, Rawia; Alhayek, Alaa; Reiling, Norbert

Published in:
ChemMedChem

DOI:
10.1002/cmdc.202200590

Publication date:
2023

ChemMedChem

Research Article
doi.org/10.1002/cmdc.202200590

Chemistry Europe
European Chemical Societies Publishing

www.chemmedchem.org

Not Every Hit-Identification Technique Works on 1-Deoxy-D-Xylulose 5-Phosphate Synthase (DXPS): Making the Most of a Virtual Screening Campaign

Sandra Johannsen,[a, b] Robin M. Gierse,[a, b, c] Aleksandra Olshanova,[a] Ellie Smerznak,[a] Christian Laggner,[d] Lea Eschweiler,[a] Zahra Adeli,[a] Rawia Hamid,[a, b] Alaa Alhayek,[a, b] Norbert Reiling,[e, f] Jörg Haupenthal,[a] and Anna K. H. Hirsch*[a, b, c]

In this work, we demonstrate how important it is to investigate not only on-target activity but to keep antibiotic activity against critical pathogens in mind. Since antimicrobial resistance is spreading in bacteria such as *Mycobacterium tuberculosis*, investigations into new targets are urgently needed. One promising new target is 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) of the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway. We have recently solved the crystal structure of truncated *M. tuberculosis* DXPS and used it to perform a virtual screening in collaboration with Atomwise Inc. using their deep convolutional neural network-based AtomNet® platform. Of 94 virtual hit compounds only one showed interesting results in binding and activity studies. We synthesized 30 close derivatives using a straightforward synthetic route that allowed for easy derivatization. However, no improvement in activity was observed for any of the derivatives. Therefore, we tested them against a variety of pathogens and found them to be good inhibitors against *Escherichia coli*.

Introduction

Antimicrobial resistance (AMR) is an increasing threat worldwide. The Covid-19 pandemic has raised awareness for infectious diseases, but while a lot of energy and money has been invested into vaccine and drug development against SARS-CoV-2, the same is not true for antibiotic research. Here, a lot more effort is needed to prevent the estimated 10 million deaths from AMR in 2050.[1] Bacteria have a variety of mechanisms to develop resistance against an antibiotic. Especially problematic are the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.), which have the highest priority for finding new antibiotics due to resistance development according to the World Health Organisation.[2] Another critical organism with rising AMR is *Mycobacterium tuberculosis*.[3] Its standard treatment takes six months and adherence to the protocol is crucial to prevent resistance development. However, infections with multidrug-resistant *M. tuberculosis* (MDR-TB) (resistant to the most effective tuberculosis drugs rifampicin and isoniazid) have continued to increase.[4] To be able to treat patients in the future, it is important to find new drugs with a new mode of action and addressing novel targets. A promising pool of targets is the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway.[5] It is utilized by many of the ESKAPE pathogens and by *M. tuberculosis* to synthesize the essential isoprenoid precursors dimethylallyl diphosphate (DMADP) and isopentenyl diphosphate (IDP). As this pathway is absent in humans, fewer side effects for the patient are expected. The first enzyme in the pathway is 1-deoxy-D-xylulose 5-phosphate synthase (DXPS). It catalyzes the condensation of pyruvate and D-glyceraldehyde-3-phosphate (D-GAP) to 1-deoxy-D-xylulose 5-phosphate (DOXP).[6] The product is not only used in the MEP pathway but is also involved in the biosynthesis of vitamins B_1 and B_6. DXPS is also the rate-limiting step and it has the highest flux-control coefficient, which means its activity influences the production of IDP and DMADP the most of all MEP enzymes. Hence, its inhibition will lower isoprenoid production most effectively.[7] We have recently solved the first crystal structure of truncated *M. tuberculosis* DXPS (Δ*Mt*DXPS, protein data bank (PDB):

[a] *Dr. S. Johannsen, Dr. R. M. Gierse, A. Olshanova, E. Smerznak, L. Eschweiler, Z. Adeli, R. Hamid, Dr. A. Alhayek, Dr. J. Haupenthal, Prof. Dr. A. K. H. Hirsch*
Helmholtz Institute for Pharmaceutical Research Saarland (HIPS)
Helmholtz Centre for Infection Research (HZI)
Campus Building E8.1, 66123 Saarbrücken (Germany)
E-mail: anna.hirsch@helmholtz-hips.de
[b] *Dr. S. Johannsen, Dr. R. M. Gierse, R. Hamid, Dr. A. Alhayek, Prof. Dr. A. K. H. Hirsch*
Department of Pharmacy, Saarland University
Campus Building E8.1, 66123 Saarbrücken (Germany)
[c] *Dr. R. M. Gierse, Prof. Dr. A. K. H. Hirsch*
Stratingh Institute for Chemistry
University of Groningen
Nijenborgh 7, 9747 AG Groningen (The Netherlands)
[d] *Dr. C. Laggner*
Atomwise Inc.
717 Market Street, Suite 800, San Francisco, CA 94103 (USA)
[e] *Dr. N. Reiling*
RG Microbial Interface Biology
Research Center Borstel
Leibniz Lung Center, Borstel (Germany)
[f] *Dr. N. Reiling*
German Center for Infection Research (DZIF)
Partner site Hamburg-Lübeck-Borstel-Riems, Borstel (Germany)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cmdc.202200590

7A9H).[8] *Mt*DXPS is usually difficult to purify and crystallize. By removing a flexible loop from the native structure, we were able to crystallize the truncated enzyme at 1.85 Å resolution. Before, there was only a homology model available to perform structure-based virtual screening.[9]

Deep convolutional neural networks in drug discovery

Atomwise developed the first deep convolutional neural network for molecular binding-affinity predication. Their Artificial Intelligence Molecular Screen (AIMS) Awards program ran from 2017 to 2022 with two major goals: to support researchers with resources to help advance their work and to demonstrate that the AI technology developed at Atomwise can indeed identify hits and add value to drug-discovery programs across all protein families. Through the AIMS program, Atomwise contributed AI technology and physical molecules to labs across the world while the academic partners contributed their biological knowledge and expertise in physically assessing compound activity. A manuscript summarizing the results of the program is currently in preparation.

The AtomNet platform is the first system that uses structural information of the target protein to make predictions.[10] The model takes millions of experimentally determined bioactivity values and thousands of protein structures from different families into account. Based on this information, it is even possible to perform AtomNet screens on targets for which no binding partners are known.

Virtual screening against truncated *M. tuberculosis* DXPS

Various approaches are available for structure-based virtual high-throughput screening (vHTS) of small-molecule databases. An overview and comparison of the different scoring methods – generally classified as physics-based, empirical, knowledge-based, and machine learning-based, the latter of which includes Atomwise's Deep Neural Network-based approach – is given by Li et al.[11] vHTS approaches can help identify bioactive molecules from an ever-increasing supply of available (either "off the shelf" or "make-on-demand") small drug-like molecules that can no longer be explored by traditional high-throughput screening methods.[12] Screening of such ultra-large spaces usually relies on accelerated GPU-based code, which has also been implemented into the AtomNet platform.[13]

After the crystal structure of Δ*Mt*DXPS (PDB: 7A9H) had been solved, Atomwise performed an AtomNet model virtual screen (VS) using the Enamine small-molecule library of several million compounds, as described previously.[8,10,14] The screen focused on the binding site of the second substrate D-GAP (gray mesh Figure 1).[15] Below the substrate pocket, thiamine diphosphate (ThDP) is bound to the protein (yellow and red sticks) and coordinates magnesium (green sphere). Pyruvate binds first to ThDP and they form a stable enamine intermediate that can be crystallized in the absence of D-GAP. Finding inhibitors for the D-GAP binding site has the benefit of not interfering with the stable and tightly bound enamine but preventing D-GAP from binding. In addition, this is the pocket that differs most from other ThDP-dependent enzymes, so the development of a DXPS-selective inhibitor should be possible. In total, 94 virtually binding compounds were identified and purchased from Enamine for testing.

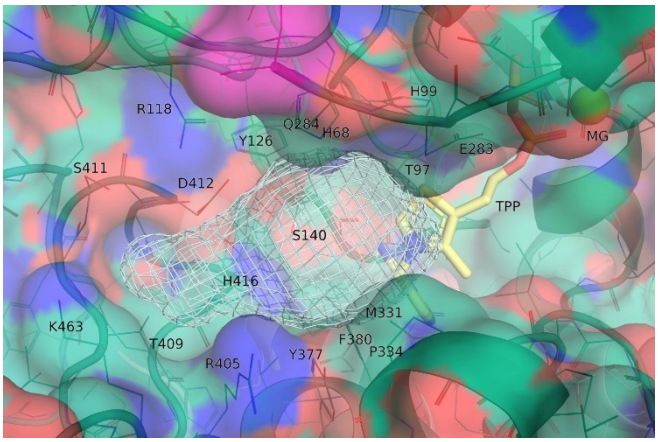


Figure 1. Binding site of D-glyceraldehyde 3-phosphate (gray mesh) with bound thiamine diphosphate (TPP, yellow and red structure) and Mg^{2+} (green structure) next to it. The amino acid residues lining the pocket in Δ*Mt*DXPS are labeled with one-letter code (protein data bank: 7A9H).

Results and Discussion

Hit-identification and validation attempts

Both *Mt*DXPS enzymes (native and truncated) were not suitable for activity assays as their substrate turnover in comparison to the model enzyme *Deinococcus radiodurans* (*Dr*)DXPS was very slow and a change in reaction velocity in the presence of inhibitors could not be observed reliably.[8,16] Hence, we used the two *Mt*-homologues in SPR binding studies, and *Dr*DXPS for the enzyme-activity assays.

To verify that the truncation of *Mt*DXPS not only provides a more stable and easier to crystallize protein, but that both homologues can be used interchangeably in binding studies, we immobilized both proteins on the surface of a surface plasmon resonance (SPR) chip. In both cases, the proteins gave a steady response signal (Δ*Mt*DXPS: Figure S**1**, *Mt*DXPS: Figure S2) over a period of four weeks. All 94 compounds and two dimethyl sulfoxide (DMSO) blanks were tested at 100 μM (Figure S**3**, Table S**1**), and for each protein the four compounds with the highest response unit (RU) were defined as hits (**1**, **2**, **3**, **4** and **1**, **3**, **5**, **6**). There is an overlap of two compounds (**1** and **3**), and a third compound (**4**) is also binding well to both although it is not among the best four against *Mt*DXPS. Dissociation constant (K_D) determinations were not possible as none of the compounds is soluble enough to reach saturation of the binding site.

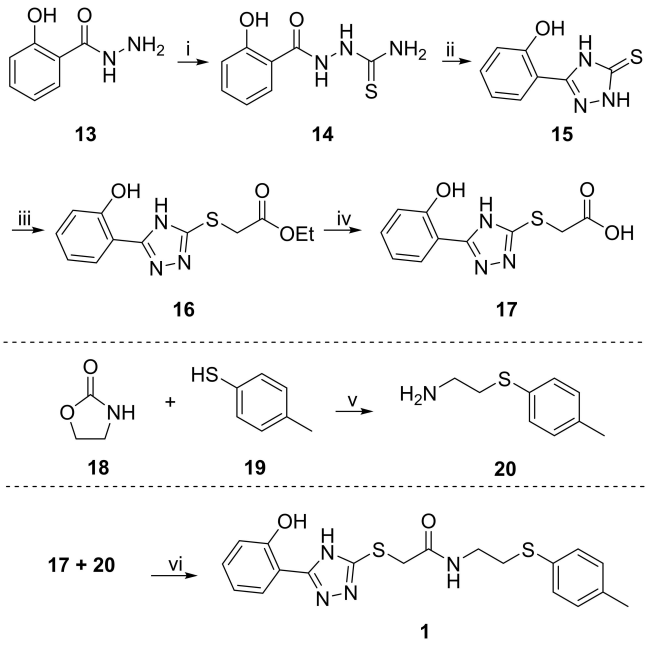
Table 1. Activity of selected compounds against *Dr*DXPS. Activity was measured at 100 μM in the DXPS-IspC-coupled assay and is reported as %inhibition.

SPR hits	Inhib. [%][a]	Assay hits	Inhib. [%][a]
1*	39 ± 16	**7**	28 ± 25
2	8 ± 10	**8**	29 ± 18
3	13 ± 2	**9**	36 ± 9
4	13 ± 8	**10**	31 ± 12
5	8 ± 10	**11**	35 ± 16
6	7 ± 2	**12**	23 ± 19

[a] Percent inhibition; values are the mean of three independent measurements.

After validating the binding of all 94 compounds, their inhibitory activity against *Dr*DXPS was measured in a DXPS-1-deoxy-D-xylulose 5-phosphate reductoisomerase (IspC)-coupled assay (Table S**2**). None of the six SPR hits show inhibition of over 50% at 100 μM (Table 1). The highest inhibition of *Dr*DXPS is seen using compound **1** at 39% inhibition that is also binding in SPR while the other SPR hits **2**–**6** inhibit below 20%. The most interesting compound **1** (Scheme 1) and five others inhibit *Dr*DXPS more than 23% (**7**–**12**).

Despite the low activity against *Dr*DXPS, the SPR results against *Mt*DXPS motivated us to synthesize a compound library with closely related derivatives of **1** to investigate whether this can be developed into a hit class for DXPS.



Scheme 1. Re-synthesis of hit **1**. i) NH_4SCN (1.0 equiv.), HCl (37%), ethanol (EtOH), 90 °C, 48 h; ii) NaOH 10%, 105 °C, 2 h, 68% over two steps (o2 s); iii) ethyl-2-bromo acetate (1.2 equiv.), $NaHCO_3$ (1.0 equiv.), EtOH, 25 °C, 24 h, 74%; iv) NaOH (4.0 equiv.), tetrahydrofuran, 25 °C, 16 h, 100%; v), NaOEt (2.0 equiv.), EtOH, 85 °C, 24 h, 89%; vi) 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC·HCl) (1.1 equiv.), 1-hydroxybenzotriazole (HOBt) (1.0 equiv.), *N*-methylmorpholine (NMM) (2.0 equiv.), dimethylformamide, 25 °C, 24 h, 83%.

Compound re-synthesis

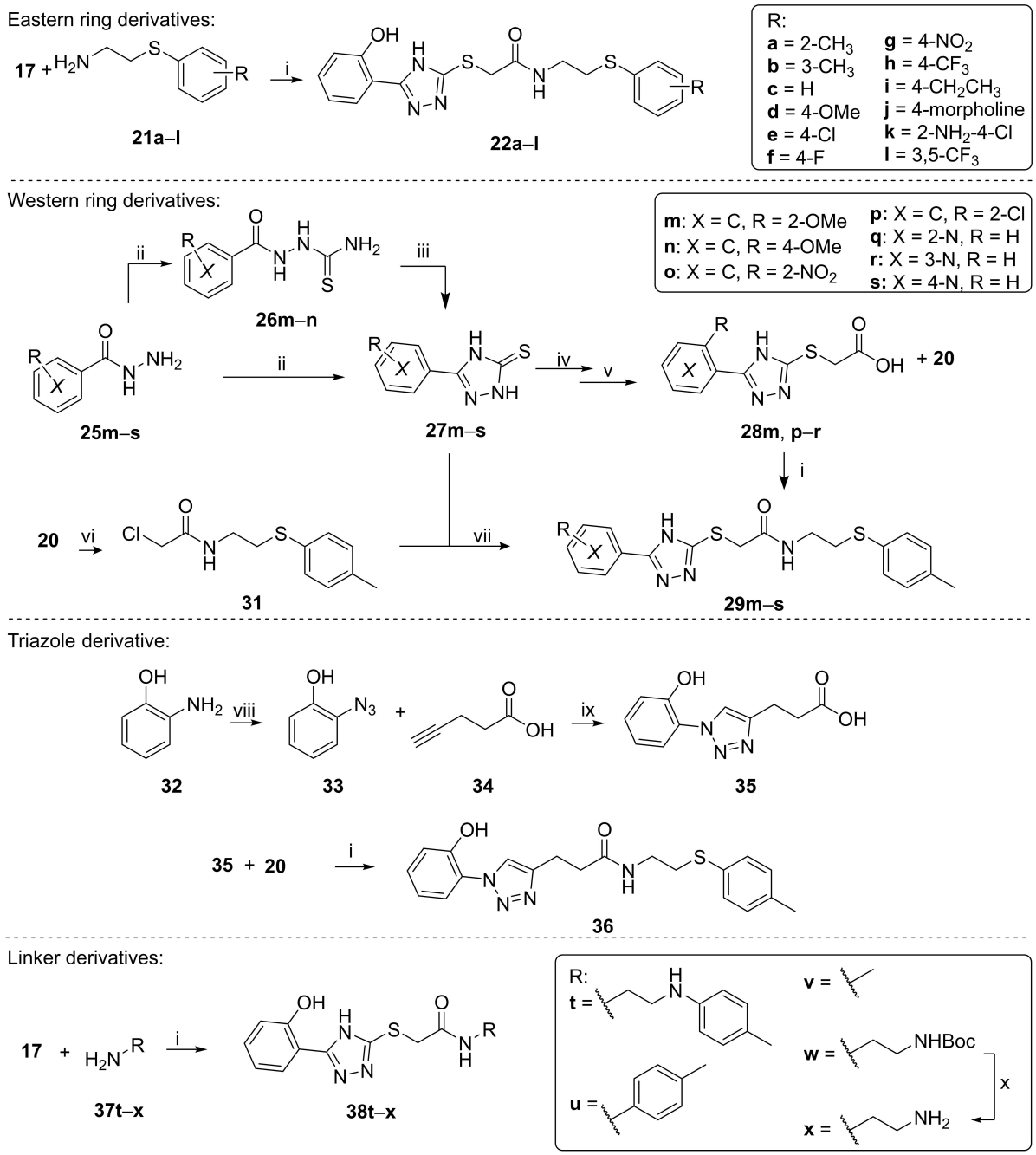
First, we developed a synthetic pathway that allowed easy derivatization on both sides of the molecule (Scheme 1).

The synthesis started from hydrazide **13**, which upon reaction with NH_4SCN and HCl (37%) in ethanol afforded the thioamide **14**.[17] Ring closure of thioamide **14** was achieved by refluxing in NaOH (10%). Ethyl-2-bromo acetate was introduced to form ester **16** followed by saponification using NaOH in tetrahydrofuran (THF).[18] Amine **20** was synthesized from 2-oxazolidone (**18**) and 1-phenylethanethiol (**19**) under decarboxylation.[19] The amide coupling between **17** and **20** proceeded smoothly with 83% yield.[20] The successfully resynthesized compound **1** was tested again in the DXPS-IspC coupled assay and an estimated IC_{50} value of 100 μM (Figure S**4**) was determined as the solubility limit of this compound is around 100 μM (Table S**3**).

Synthesis of derivatives

The synthesis of twelve derivatives with different substituents on the Eastern ring was performed using a variety of different thiols (Scheme 2). We based the selection of derivatives on the premise of making only small changes to get a clear picture of the structure–activity relationship. Therefore, the methyl group was moved to positions two and three of the ring (**22a**, **22b**) or removed completely (**22c**). A variety of substituents in 4-position (**22d**–**i**) was chosen to explore this position further before looking into substituents in *ortho* or *meta* position. To increase solubility, morpholine **22j** was synthesized starting by Boc-protection of **22e** (compound **23**), subsequent Buchwald Hartwig coupling (compound **24**) and final deprotection to amine **21j** that was then used for amide coupling. We also tested the effect of having two substituents on the Eastern ring (**22k** and **22l**).

The synthesis of the Western part was done following the same procedure as for **1** with slight modifications (Scheme 2). Hydrazides (**25m**–**s**) decomposed in concentrated hydrochloric acid and a 1 M aqueous solution had to be used instead. In addition, the thioamide precursors **26m** and **26n** were only found for triazoles **27m** and **27n** while all other reactions led to the cyclized compounds **27o**–**s** directly. This was unexpected in acidic solution and has not been reported before. If the thiadiazole is the intended isomer, the thioamide is normally treated with concentrated sulfuric acid.[17] To confirm that the



Scheme 2. Synthesis of derivatives of compound **1**. i) 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl) (1.1 equiv.), 1-hydroxybenzotriazole (HOBt) (1.0 equiv.), *N*-methylmorpholine (NMM) (2.0 equiv.), dimethylformamide (DMF), 25 °C, 24–72 h, 10%–88%; ii) HCl (1 M), reflux, 48 h, 74%; iii) NaOH (10%), reflux, 48 h, 34%–100%; iv) ethyl-2-bromoacetate (1.2 equiv.), $NaHCO_3$ (1.0 equiv.), ethanol (EtOH), 25 °C, 24 h, 63%–90%; v) NaOH (2.0 equiv.), H_2O, 25 °C, 24 h, 80%–100%; vi) 2-chloroacetic acid (1.0 equiv.), EDC·HCl (1.1 equiv.), HOBt (1.0 equiv.), NMM (2.0 equiv.), DMF, 25 °C, 24 h, 29%; vii) $NaHCO_3$ (1.0 equiv.) or $NaHCO_3$ (aq)/ethanol 1:1, ethanol, 25 °C, 24 h, 24%–53%; viii) $NaNO_2$ (3.0 equiv.), H_2O, HCl (6 N), 0 °C, 30 min, then NaN_3 (3.0 equiv.), 0 °C, 2 h; ix) sodium ascorbate (4.0 equiv.), $CuSO_4 \cdot 5\,H_2O$ (1.0 equiv.), MeCN, 25 °C, 72 h; x) HCl in dioxane (4 M), 0 °C to 25 °C, 16 h, 69%.

correct cycle was formed and not its thiadiazole isomer **30**, we compared the ^{1}H NMR and ^{13}C NMR chemical shifts of compounds **27 s** and **30** (Table S4). The differences especially of the proton chemical shifts confirmed that the correct isomer was formed.

Next, ethyl-2-bromoacetate was used to install the desired ester and subsequent saponification resulted in the free acids (**28 m**, **p**–**r**). The amide coupling to the final product was performed as described above (**29 m**, **p**–**r**).

As an alternative strategy, we tried to perform the amide coupling between **20** and 2-chloroacetic acid first to form chloride **31** and then add different triazoles directly (Scheme 2). As a model reaction, **27 s** was chosen and the reaction worked well with a yield of 53% for **29 s**. The same was true for **27 o**,

which reacts with ethyl-2-bromoacetate but in a low yield and the ester intermediate is difficult to purify. For the 4-methoxy derivative **27n**, the reaction also resulted in the desired final compound **29n**. In summary, the alternative route made it possible to circumvent saponification and more challenging purification steps and resulted in higher overall yield in fewer steps. Then, compound **29n** was mixed with BBr_3 to deprotect the hydroxyl group, but the compound decomposed.

One modification that would also allow for easy derivatization, is the replacement of the 1,2,4-triazole with a 1,2,3-triazole ring. The 1,2,3-triazole **35** can be synthesized by click-chemistry between an alkyne and an azide functional group, which reduces the number of steps from five to three. In the synthesis of **1**, most reactions need stirring over night while the synthesis of **36** comprises two quick steps. The azide **33** is formed from **32** within two hours and the click reaction with **34** is also quick if 1.0 equivalent of copper sulfate is added from the beginning. Only the amide coupling needs up to 72 h.

Two compounds with different linkers between the amide and the Eastern ring system were synthesized and three derivatives without the Eastern ring following the same amide coupling reaction as before (Scheme 2). First, the sulfur atom was replaced with a nitrogen atom (compound **38t**). For this, amine **37t** was synthesized following a literature procedure.[21] Removing the aliphatic linker led to derivative **38u** with the aromatic ring directly connected to the amide. The whole Eastern part of the molecule was replaced by just a methyl group (compound **38v**). In an attempt to improve the Gram-negative activity, we introduced a primary amine by amide coupling between **17** and *N*-Boc-ethylenediamine (compound **38w**) and subsequent deprotection (compound **38x**).[22]

Activity of derivatives against *Dr*DXPS

After the synthesis was completed, all compounds were tested for their inhibition of *Dr*DXPS. In order to see if the acid precursors are also active, they were included in all biological assays. First, the solubility limit for all compounds was determined (Table S3) and based on the results, the assay was performed at 200 μM, 100 μM or 50 μM (Table 2). However, none of the derivatives are active (more than 50% inhibition) at the respective highest soluble concentrations including compound **1**. Amides **22d**, **22j**, **38u** and acid **17** are the only compounds that show an inhibition higher than 30% at a concentration of 200 μM. The standard deviations for these measurements are high and the data difficult to rely on and to interpret.

Overall, these results suggest that the compounds are not DXPS inhibitors. When looking at the three different activity tests of compound **1**, we see low reproducibility. The first test of compound **1** suggested interesting inhibition (%inhi. = 39 ± 16). When the IC_{50} value of the resynthesized hit **1** was determined, 50% inhibition was detected at 100 μM, which was still promising, but when testing all synthesized compounds including **1** at 100 μM again %inhibition of **1** was only 27 ± 12. This could be explained by the fact that the solubility limit of **1** is around 100 μM, which can affect the assay results.

Discussion of VS results

Despite the recent progress in the areas of purchasable chemical space, scoring methods, and computational throughput, vHTS does not always manage to find bioactive molecules. In fact, bigger libraries pose the unique challenge that while increasing library size can provide better-fitting and higher-scoring, rare molecules that rank artifactually well likewise get enriched amongst the top-scoring molecules.[23] A recent paper

Table 2. Activity data of derivatives of compound **1** in enzyme activity assay against *Dr*DXPS and *E. coli* (*Ec*)IspC (activity is reported in %inhibition) and against *E. coli* K12 and ΔTolC (activity is reported as MIC or %inhi. at 100 μM).

Cpd.	DXPS-IspC inhibition [%]	Percent inhib. at 100 μM or MIC		Cpd.	DXPS-IspC inhibition [%]	Percent inhib. at 100 μM or MIC	
		E. coli K12	*E. coli* ΔTolC			*E. coli* K12	*E. coli* ΔTolC
1	27 ± 12%[b]	11 ± 2%[c]	28 ± 0 μM	**28q**	n.i.[a]	13 ± 4%	17 ± 3%
17	n.i.[a]	n.i.	n.i.	**28r**	n.i.[a]	n.i.	n.i.
22a	13 ± 4%[b]	16 ± 13%	49 ± 1 μM	**29 m**	10 ± 0%[c]	n.i.	n.i.
22b	16 ± 3%[a]	21 ± 3%	49 ± 0 μM	**29n**	n.i.[b]	n.i.[c]	48 ± 1%[c]
22c	22 ± 3%[a]	14 ± 1%	93 ± 2 μM	**29o**	n.i.[b]	20 ± 0%	99 ± 1 μM
22d	23 ± 2%[a]	6 ± 1%	96 ± 4 μM	**29p**	n.i.[c]	n.i.	50 ± 11%
22e	17 ± 4%[c]	30 ± 7%	25 ± 3 μM	**29q**	n.i.[b]	n.i.	82 ± 3%
22f	21 ± 4%[a]	16 ± 2%	96 ± 1 μM	**29r**	n.i.[a]	n.i.	44 ± 6%
22 g	12 ± 1%[b]	19 ± 0%	82 ± 3%	**29 s**	n.i.[b]	n.i.[c]	38 ± 15%[c]
22 h	n.i.[c]	22 ± 6%[c]	18 ± 3 μM	**36**	n.i.[b]	n.i.	n.i.
22i	13 ± 4%[c]	n.i.	19 ± 3 μM	**38 t**	n.i.[a]	n.i.	47 ± 2%[c]
22j	n.d.*	n.i.	63 ± 1%	**38u**	n.i.[b]	51 ± 10%	58 ± 6 μM
22k	n.i.[b]	23 ± 14%	48 ± 1 μM	**38v**	n.i.[a]	n.i.	n.i.
22 l	12 ± 15%[c]	11 ± 3%[d]	75 ± 5%[d]	**38w**	n.i.[b]	n.i.	28 ± 10%
28 m	n.i.[a]	n.i.	n.i.	**38x**	n.d.*	n.i.	n.i.
28p	n.i.[a]	n.i.	n.i.				

Activity measured at compound concentrations of [a] 200 μM, [b] 100 μM, [c] 50 μM, [d] 25 μM. * not enough compound available after biological testing. MIC = minimum inhibitory concentration, %inhi. = percent inhibition, n.d. = not determined, n.i. = no inhibition if inhibition < 10%.

by Céron-Carrasco discusses the results of two separate virtual screening campaigns for inhibitors of the main protease (Mpro) of SARS-Cov-2, both of which failed to provide hits that were active enough to be further optimized.[24] Since the 1000 highest-ranking compounds from the previous campaigns covered relatively narrow score ranges, the author attempted to recover active molecules from these groups by rescoring the molecules with a Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) approach. This method too provided limited success, finding only one compound with an IC_{50} = 0.8 mM against Mpro. However, combinations of high-throughput/lower precision initial screens with more refined (but lower throughput) methods seems like a reasonable approach to improve hit recovery, and improvement of the above-mentioned MM/GBSA and related MM/PBSA method is still ongoing.[25] It should also be pointed out that even for smaller libraries where physical high-throughput screening is an option, both physical and virtual screening approaches can lead to false positives (as well as false-negatives). However, combination of the two methods can ultimately lead to hits that overcome the shortcomings of both approaches, creating higher confidence in the results.[26]

The Atomwise AIMS projects intended to prospectively use and evaluate their proprietary AtomNet vHTS platform, and thus, we chose not to employ any complementary alternative prediction methods as additional filters. The selection of the final compounds purchased for testing uses a diversity-based clustering approach, to address the above-mentioned challenges that a number of top-scoring molecules are basically identical in their predicted score and may be enriched for problematic scaffolds, while also allowing to keep human interference to a minimum (i.e. no manual "hit-picking"). For DXPS this method proved to be insufficient and future virtual screening campaigns should take this into consideration.

DXPS is a difficult target to address virtually due to its large and flexible active site. However, during previous screenings, we have successfully used known binders to identify inhibitors against *Mt*DXPS that bind to the ThDP-binding site.[15] Freel-Meyers et al. found several compounds binding covalently to ThDP and extending into the D-GAP binding site.[27] The knowledge of their binding mode could be taken into consideration for future screenings when looking for inhibitors addressing the D-GAP binding site to improve hit quality.

Biological data

With 30 derivatives at hand, we decided to not give up on this compound class but to continue our investigations. AMR is especially problematic in Gram-negative bacteria. Therefore, we wanted to check if our VS hits are active in *E. coli* as a Gram-negative model organism that utilizes the MEP pathway. The twelve compounds that performed best in the SPR and *Dr*DXPS activity assay were tested against *E. coli* K12 and *E. coli* ΔTolC. The latter is a mutant of the former in which the efflux channel TolC is removed.[28] This mutant lacks the ability to pump the molecules out of the bacterial cells efficiently. Therefore, a compound that inhibits the growth of *E. coli* ΔTolC but not the wild-type strain K12 is probably eliminated by the efflux TolC system. With this knowledge good inhibitors are not excluded by accident but can be optimized to also inhibit *E. coli* K12.

From the twelve compounds, compound **1** has a MIC against *E. coli* ΔTolC of 28 μM and shows a 11% growth inhibition at 50 μM of *E. coli* K12 (Table 2). It was not possible to determine a MIC for any of the other compounds, although seven show some activity against *E. coli* ΔTolC (Table S**5**). All derivatives of **1** were tested against both *E. coli* strains and only three have a MIC lower than 28 μM in *E. coli* ΔTolC. They are all Eastern-ring derivatives and are substituted in the 4-position (**22e**: 4-Cl, **22h**: 4-CF_3, **22i**: 4-CH_2CH_3). In general, the Eastern ring derivatives are active against *E. coli*. Most have a MIC of 20–100 μM and only the morpholine derivative **22j** and the 4-NO_2-derivative **22g** show inhibition below 90% at 100 μM. Compounds **22i** and **22j** are the only two derivatives that do not inhibit *E. coli* K12, which indicates that aliphatic substitution in the 4-position leads to efflux. In contrast, all Western-ring derivatives are inactive against *E. coli* K12 and do not inhibit the growth of *E. coli* ΔTolC more than 82% at 100 μM. The only exception is compound **29o** with a NO_2-group in 2-position, which shows that the 2-hydroxy moiety can be replaced with a reduction, but not a complete loss of activity.

The linker derivatives are more diverse, but it can be seen that the sulfur atom is essential as compound **38t** (nitrogen instead of sulfur) loses activity against both strains in comparison to parent compound **1**. Removing the Eastern aromatic ring leads to complete loss of activity (**38v**–**x**). This was unexpected because we specifically installed a primary amine to improve Gram-negative activity. In 2017, Richter et al. coined the term eNTRy rules after analyzing the ability of over 180 diverse compounds to accumulate in *E. coli*.[22] They found that an ionizable, sterically non-hindered primary amine, a globularity (three-dimensionality) below 0.25 and fewer than five rotatable bonds help with accumulation in Gram-negative bacteria. Compound **38x** was designed to fulfill these criteria (primary amine, globularity = 0.063) except for the number of rotatable bonds which is six. However, from the analysis of the activities of the other compounds, the primary amine should have been installed at a different position as the removal of the Eastern rings system led to a loss of activity. Even if compound **38x** is now accumulating in the cell, it is not active anymore. This has to be taken into consideration when making more derivatives.

Surprisingly, compound **38u** with the aromatic ring directly attached to the amide has the highest inhibition of all derivatives against *E. coli* K12 and a moderate MIC of 58 μM against *E. coli* ΔTolC. Similar compounds have been reported to be New Delhi metallo-β-lactamase-1 (NDM-1) inhibitors (Figure 2).[29] NDM-1 is present in several bacteria such as *E. coli* and *K. pneumoniae*, and makes them resistant to β-lactam antibiotics. The only structural difference between the reported compound **39** and compound **38u** is the position of the methyl-group on the Eastern ring. The compounds in that study have only been tested on NDM-1 directly and their intrinsic inhibitory activity against bacteria has not been investigated

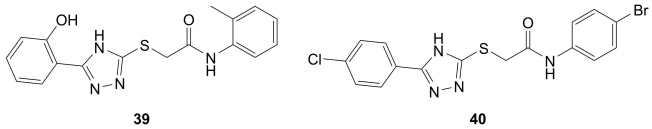


Figure 2. Two examples of derivatives of compound **38u** and their biological activities that have been reported before.

yet. However, another very similar compound class (representative **40** as the closest structural analogue to **38u**) has been tested against a variety of clinically relevant strains and its activity was promising in *E. coli*, although the target has not been determined.[30]

This could lead to the development of a dual inhibitor that can be used in combination with a β-lactam antibiotic. On the one hand, it can prevent the resistance against the antibiotic and on the other hand it kills the bacteria through a different mechanism of action. It might be interesting to investigate this class of compounds further by installing a primary amine at the methyl-substituent of the aromatic ring to see if accumulation can be improved.

Finally, the acid precursors have been tested but they have no interesting inhibitory activities.

We were also curious to see how the representative compounds **1**, **22e** and **22i** performed in other clinically relevant pathogens. However, the inhibition against *A. baumannii* and two different strains of *P. aeruginosa* (Table S6) was below 15% for all compounds making them selective *E. coli* inhibitors.

Compound **1** was also tested against human Hep G2 cells to determine its cytotoxicity. It does not inhibit cell growth at 25 μM, but at 100 μM, 49 ± 1% inhibition was measured. This value is acceptable as its MIC in *E. coli* ΔTolC is 28 μM. It can be assumed that the toxicity is comparable in all derivatives of the class, but this has to be confirmed if more diverse compounds are synthesized.

Lastly, **1** was also tested against *M. tuberculosis* H37Rv, but no inhibition was seen at 64 μM (data not shown).

In summary, the Western part of the molecule is essential for its activity in *E. coli* if the Eastern part contains a sulfur atom and a phenyl ring, preferably with a 4-CF_3 substituent. The 1,3,4-triazole cannot be replaced by a 1,2,3-triazole and the 2-hydroxy-substituent on the Western ring can only be replaced with a NO_2-group under loss of some activity. Derivatives of compound **38u** should be investigated further as potential dual inhibitors.

Conclusions

We have performed the first virtual screening on *Mt*DXPS based on its recently solved truncated crystal structure. For the first time, we used a program based on a convolutional neural network to identify a set of structurally diverse, potential inhibitors of a MEP pathway enzyme. The initial results from SPR and on-target inhibition data were ambiguous, but we wanted to give the most promising compound a chance because its estimated IC_{50} value was close to its solubility limit. We established a robust synthetic route that allowed for easy derivatization, and we synthesized 30 derivatives for further testing. Although, we improved the solubility in some of the derivatives, we were not able to improve the on-target activity against DXPS. Investigating antimicrobial activity of all derivatives revived our interest in the compound class as compound **1** showed a promising MIC of 28 μM against *E. coli* ΔTolC. Most of the Western-ring derivatives showed comparable MIC values, but most other structural changes were detrimental to the activity. Testing compound **1** in other pathogens only confirmed its selective inhibition of *E. coli*. Compound **38u** stood out by showing the best activity against *E. coli* K12 and a look into the literature revealed that it is a very close homologue of previously published β-lactamase NDM-1 inhibitors. We are now investigating compound **1** and its derivatives as potential dual inhibitors of NDM-1 and (an)other unknown target(s) in *E. coli*.

This long and difficult journey in the realm of drug discovery has not ended yet. We have learned a lot about the behavior of DXPS in different assay set-ups, how important it is to have a broad variety of assays available and that you should not give up on a compound class especially if you have established a robust synthetic route.

Experimental Section

Gene expression and protein purification of DXPS and IspC: Gene expression and protein purification of *D. radiodurans*, native and truncated *M. tuberculosis* DXPS and *E. coli* IspC followed previously reported protocols.[31,32]

Binding affinity determination by SPR: The SPR binding studies were performed using a Reichert SR7500DC surface plasmon resonance spectrometer (Reichert Technologies, Depew, NY, USA), and medium density carboxymethyl dextran hydrogel CMD500 M sensor chips (XanTec Bioanalytics, Düsseldorf, Germany). Milli-Q water was used as the running buffer for immobilization. Truncated *M. tuberculosis* DXPS (Δ*Mt*DXPS, 66.61 kDa) and native *Mt*DXPS (67.88 kDa) were immobilized in one of the two flow cells according to reported amine-coupling protocols.[33] The other flow cell was left blank to serve as a reference. The system was initially primed with borate buffer 100 mM (pH 9.0), then the carboxymethyldextran matrix was activated by a 1:1 mixture of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl) 100 mM and *N*-hydroxysuccinimide (NHS) 100 mM at a flow rate of 10 μL min^{-1} for 7 min. The Δ*Mt*DXPS or *Mt*DXPS was diluted to a final concentration of 1.1 μM and 1.4 μM, respectively, in 10 mM sodium acetate buffer (pH 4.5), and was injected at a flow rate of 5 μL min^{-1} for 10 min. Non-reacted surface was quenched by 1 M ethanolamine hydrochloride (pH 8.5) at a flow rate of 25 μL min^{-1} for 3 min. A series of ten buffer injections was run initially on both reference and active surfaces to equilibrate the system resulting in a stable immobilization level of approximately 5,000 (9,000 for *Mt*DXPS) μRIU (Figure S**1**, Figure S**2**), respectively. 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (50 mM HEPES, 150 mM NaCl, 0.05% v/v Tween 20, 1 mM $MgCl_2$, pH ^8.0) containing 5% v/v DMSO was used as the running buffer for binding studies. The running buffer was filtered and degassed prior to use. Binding

experiments were performed at 20 °C. Compounds dissolved in DMSO were diluted with the running buffer (final DMSO concentration of 5% v/v) and were injected at a flow rate of 30 μL min^{-1}. The association time was set to 60 s, and the dissociation phase was recorded for 120 s. Ethylene glycol 80% in the running buffer was used for regeneration of the gold-chip surface. Differences in the bulk refractive index due to DMSO were corrected by a calibration curve (nine concentrations: 3–7% v/v DMSO in the running buffer). Data processing and analysis were performed by Scrubber software (Version 2.0c, 2008, BioLogic Software). Sensorgrams were calculated by sequential subtractions of the corresponding curves obtained from the reference flow cell and the running buffer (blank). SPR responses are expressed in response units (RU).

DXPS-IspC coupled assay: The experiments to determine the inhibitory activity of compounds in the DXPS-IspC-coupled assay have been carried out as reported previously with minor adjustments.[32] The assay was conducted in transparent, round-bottom 96-well plates (Greiner BioOne), and the absorbance was measured using a microplate reader (PHERAstar, BMG Labtech). The assay mixture contained 100 mM HEPES with a pH of 7.0, 2 mM β-mercaptoethanol (BME), 100 mM NaCl, 0.5 mM ThDP, 1.0 mM $MgCl_2$, 0.05 mM nicotinamide adenine dinucleotide phosphate (NADPH), 1.0 mM pyruvate, 1.0 mM D/L GAP and 1.0 μM *Ec*IspC. The amount of DXPS used in the assays was determined experimentally by a dilution series of the enzyme. The concentration, which showed a reaction velocity in the range of −0.1 to −0.2 OD over a time range of 10 min was chosen for further experiments. The assay was prepared using two buffers, buffer A containing HEPES, NaCl, BME and all components of the reaction (except substrates) in 2x the final concentration and buffer B, consisting of only HEPES, NaCl, BME and the substrates, 2.0 mM pyruvate and 2.0 mM D/L-GAP. After the addition of 60 μL of buffer A to the plate with a pre-made inhibitor dilution series in 6 μL DMSO the reaction was started by the addition of 60 μL buffer B. The plate was centrifuged for 1 min at 2000 revolutions per minute (rpm) and 25 °C to remove possible air bubbles. The plate was then immediately supplied to the microplate reader, and the absorbance measured at 340 nm using the mode slow kinetics with a cycle time of 30 seconds and 60 cycles at 25 °C. Blank correction and linear fitting of the raw data were performed using the program Origin2021. The obtained initial velocities were converted to percent inhibition and plotted against the inhibitor concentration. The IC_{50} values were determined by nonlinear curve fitting using Origin2021.

Kinetic-solubility determination: The compounds were sequentially diluted in DMSO in a round-bottom 96-well plate. 3 μL of each well were transferred into another 96-well plate and mixed with 147 μL of phosphate-buffered saline (PBS) buffer. Plates were shaken at room temperature (r.t.) for 5 min at 600 rpm, and the absorbance at 620 nm was measured after 72 h. Absorbance values were normalized by blank subtraction and plotted using GraphPad Prism 8.4.2 (GraphPad Software, San Diego, CA, USA).

Antibacterial activity: Assays regarding the determination of the minimum inhibitory concentration (MIC) were performed as described previously.[34] The experiments were based on two *E. coli* strains/mutants (K12, ΔTolC) as well as *P. aeruginosa* (PA14, ΔmexA) and *A. baumannii* (DSM30007). In the case that no MIC value could be determined due to activity reasons, the percentage (%) of inhibition at 100 μM (or lower, depending on the solubility of the compounds) was determined. All samples were run in duplicate for each condition, and experiments were performed independently at least twice.

Determination of cytotoxicity of compound 1: HepG2 cells (2 x 105 cells per well) were seeded in 24-well, flat-bottomed plates. Culturing of cells, incubations and OD measurements were performed as described previously with small modifications.[35] 24 h after seeding the cells, the incubation was started by the addition of compounds in a final DMSO concentration of 1%. The metabolic activity of the living cell mass was determined after 48 h. Three independent measurements were performed. The IC_{50} values were determined during logarithmic growth using GraphPad Prism software and data reported represent the mean ± SD.

Antitubercular activity: Compound **1** was tested for its Mtb growth inhibitory capacity in liquid culture as previously described.[36] Tests were performed in 2-fold serial dilutions starting at 64 μM in triplicates (2×10^6 mCherry-Mtb H37Rv bacteria, volume 100 μL). Bacterial growth was measured after 7 days of culture as described. Obtained values were normalized to solvent control (DMSO-treated bacteria set to 100%).

Chemical synthesis: All procedures, characterization data and spectra are given in the Supporting Information.

Acknowledgements

The authors want to thank Dr. Andreas M. Kany, Selina Wolter and Jeannine Jung from HIPS for their technical support. A. K. H. H. gratefully acknowledges an Atomwise AIMS Award (No. A19-234), and funding from the European Research Council (ERC starting grant 757913), the Netherlands Organisation for Scientific Research (LIFT grant: 731.015.414) and from the Helmholtz Association's Initiative and Networking Fund. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

Christian Laggner is employed at Atomwise as stated in the manuscript.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Drug discovery · 2-*C*-methyl-D-erythritol 4-phosphate pathway · 1-deoxy-D-xylulose 5-phosphate synthase · inhibitors · neural networks

[1] J. O'Neill, "Tackling Drug-Resistent Infections Globally: Final Report and Recommendations", can be found under https://amr-review.org/sites/default/files/160518 Final%20paper with%20cover.pdf, **2016**.
[2] World Health Organisation, "WHO publishes list of bacteria for which new antibiotics are urgently needed", can be found under http://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed, **2017**.
[3] World Health Organisation, *Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis*, Geneva, **2017**.
[4] C. Lange, K. Dheda, D. Chesov, A. M. Mandalakas, Z. Udwadia, C. R. Horsburgh, *Lancet* **2019**, *394*, 953.
[5] a) S. Heuston, M. Begley, C. G. M. Gahan, C. Hill, *Microbiology* **2012**, *158*, 1389; b) X. Wang, C. S. Dowd, *ACS Infect. Dis.* **2018**, *4*, 278.

[6] F. J. Sangari, J. Pérez-Gil, L. Carretero-Paulet, J. M. García-Lobo, M. Rodríguez-Concepción, *PNAS* **2010**, *107*, 14081.
[7] a) R. Moreno-Sánchez, E. Saavedra, S. Rodríguez-Enríquez, V. Olín-Sandoval, *J. Biomed. Biotechnol.* **2008**, *2008*, 1; b) D. C. Volke, J. Rohwer, R. Fischer, S. Jennewein, *Microb. Cell Fact.* **2019**, *18*, 192.
[8] R. M. Gierse, R. Oerlemans, E. R. Reddem, V. O. Gawriljuk, A. Alhayek, D. Baitinger, H. Jakobi, B. Laber, G. Lange, A. K. H. Hirsch et al., *Sci. Rep.* **2022**, *12*, 7221.
[9] T. Masini, B. Lacy, L. Monjas, D. Hawksley, A. R. de Voogd, B. Illarionov, A. Iqbal, F. J. Leeper, M. Fischer, M. Kontoyianni et al., *Org. Biomol. Chem.* **2015**, *13*, 11263.
[10] I. Wallach, M. Dzamba, A. Heifets, "AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery", can be found under https://arxiv.org/pdf/1510.02855, **2015**.
[11] J. Li, A. Fu, Le Zhang, *Interdiscip. Sci.* **2019**, *11*, 320.
[12] W. A. Warr, M. C. Nicklaus, C. A. Nicolaou, M. Rarey, *J. Chem. Inf. Model.* **2022**, *62*, 2021.
[13] a) L. Solis-Vasquez, A. F. Tillack, D. Santos-Martins, A. Koch, S. LeGrand, S. Forli, *Parallel Comput.* **2022**, *109*; b) A. Morrison, G. Friedland, I. Wallach, *Efficient GPU Implementation of AutoDock Vina* **2020**; c) A. Morrison, *CUina: An Efficient Implementation of Autodock Vina Specially Crafted for CUDA GPUs*.
[14] a) H. K. Lichtenthaler, *Biochem. Soc. Trans.* **2000**, *28*, 785; b) C.-H. Hsieh, L. Li, R. Vanhauwaert, K. T. Nguyen, M. D. Davis, G. Bu, Z. K. Wszolek, X. Wang, *Cell Metab.* **2019**, *30*, 1131–1140.e7; c) S. Su, J. Chen, Y. Jiang, Y. Wang, T. Vital, J. Zhang, C. Laggner, K. T. Nguyen, Z. Zhu, A. W. Prevatte et al., *Adv. Sci. (Weinh)* **2021**, *8*, e2004846.
[15] Z. Di, S. Johannsen, T. Masini, C. Simonin, J. Haupenthal, A. Andreas, M. Awale, R. M. Gierse, T. van der Laan, R. van der Vlag et al., *Chem. Sci.* **2022**, *13*, 10686.
[16] R. M. Gierse, E. Redeem, E. Diamanti, C. Wrenger, M. R. Groves, A. K. Hirsch, *Future Med. Chem.* **2017**, *9*, 1277.
[17] E. M. Sarshira, N. M. Hamada, Y. M. Moghazi, M. M. Abdelrahman, *J. Heterocycl. Chem.* **2016**, *53*, 1970.
[18] a) A. Cristina, D. Leonte, V. Laurian, L. C. Bencze, S. Imre, V. Zaharia, *Farmacia* **2018**, *66*, 88; b) K. Sung, A.-R. Lee, *J. Heterocycl. Chem.* **1992**, *29*, 1101.
[19] H. Ishibashi, M. Uegaki, M. Sakai, *Synlett* **1997**, *1997*, 915.
[20] D. H. Steinman, M. L. Curtin, R. B. Garland, S. K. Davidsen, H. Heyman, J. H. Holms, D. H. Albert, T. J. Magoc, I. B. Nagy, P. A. Marcotte et al., *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2087.
[21] G. S. Poindexter, D. A. Owens, P. L. Dolan, E. Woo, *J. Org. Chem.* **1992**, *57*, 6257.
[22] M. F. Richter, B. S. Drown, A. P. Riley, A. Garcia, T. Shirai, R. L. Svec, P. J. Hergenrother, *Nature* **2017**, *545*, 299.
[23] J. Lyu, J. J. Irwin, B. K. Shoichet, *Nat. Chem. Biol.* **2023**.
[24] J. P. Cerón-Carrasco, *ChemMedChem* **2022**, *17*, e202200278.
[25] S. Zhong, K. Huang, S. Luo, S. Dong, L. Duan, *Phys. Chem. Chem. Phys.* **2020**, *22*, 4240.
[26] R. S. Ferreira, A. Simeonov, A. Jadhav, O. Eidam, B. T. Mott, M. J. Keiser, J. H. McKerrow, D. J. Maloney, J. J. Irwin, B. K. Shoichet, *J. Med. Chem.* **2010**, *53*, 4891.
[27] a) J. M. Smith, R. J. Vierling, C. F. Meyers, *MedChemComm* **2012**, *3*, 65; b) D. Bartee, S. Sanders, P. D. Phillips, M. J. Harrison, A. T. Koppisch, C. L. Freel Meyers, *ACS Infect. Dis.* **2019**, *5*, 406.
[28] V. Koronakis, J. Eswaran, C. Hughes, *Annu. Rev. Biochem.* **2004**, *73*, 467.
[29] Le Zhai, Y.-L. Zhang, J. S. Kang, P. Oelschlaeger, L. Xiao, S.-S. Nie, K.-W. Yang, *ACS Med. Chem. Lett.* **2016**, *7*, 413.
[30] S. Sahi, R. K. Sodhi, B. Jamwal, S. Paul, *J. Heterocycl. Chem.* **2018**, *55*, 1596.
[31] S. Hecht, J. Wungsintaweekul, F. Rohdich, K. Kis, T. Radykewicz, C. A. Schuhr, W. Eisenreich, G. Richter, A. Bacher, *J. Org. Chem.* **2001**, *66*, 7770.
[32] T. Masini, J. Pilger, B. S. Kroezen, B. Illarionov, P. Lottmann, M. Fischer, C. Griesinger, A. K. H. Hirsch, *Chem. Sci.* **2014**, *5*, 3543.
[33] R. B. M. Schasfoort, *Handbook of surface plasmon resonance*, Royal Society of Chemistry, Cambridge, **2017**.
[34] W. A. M. Elgaher, M. Fruth, M. Groh, J. Haupenthal, R. W. Hartmann, *RSC Adv.* **2013**, *4*, 2177.
[35] J. Haupenthal, C. Baehr, S. Zeuzem, A. Piiper, *Int. J. Cancer* **2007**, *121*, 206.
[36] R. P. Jumde, M. Guardigni, R. M. Gierse, A. Alhayek, Di Zhu, Z. Hamid, S. Johannsen, W. A. M. Elgaher, P. J. Neusens, C. Nehls et al., *Chem. Sci.* **2021**, *12*, 7775.

Manuscript received: November 2, 2022
Revised manuscript received: February 20, 2023
Accepted manuscript online: March 10, 2023
Version of record online: April 14, 2023