The Free Internet Journal for Organic Chemistry

Review

Archive for Organic Chemistry

Arkivoc **2019**, part iv, 227-244

Challenges of antibacterial drug discovery

Folkert Reck, Johanna M. Jansen, and Heinz E. Moser

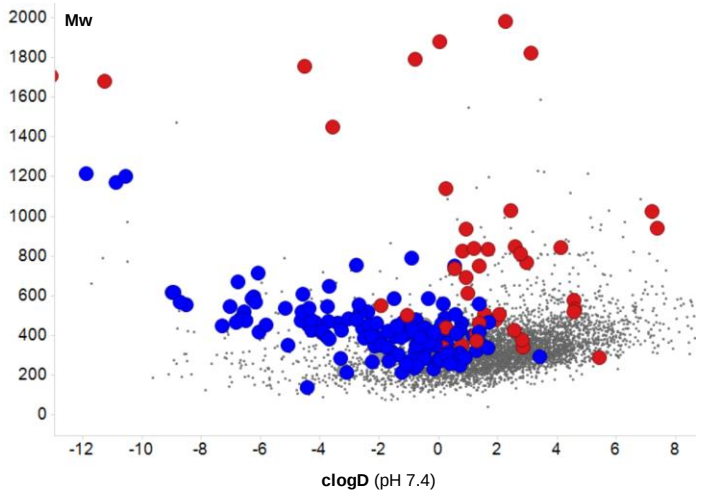
Novartis Institute for BioMedical Research, 5300 Chiron Way, Emeryville, CA 94608, U.S.A.
Email: heinz.moser@novartis.com

This article is dedicated to Stephen Hanessian; a long-term mentor, advisor, colleague and friend

Received 04-11-2019 **Accepted** 06-17-2019 **Published on line** 07-12-2019

Abstract

Efforts to discover novel antibacterial drugs over the past decades have been characterized by high attrition rates, driven mainly by technical challenges, high dose and safety requirements, regulatory difficulties and low-cost generic competition, resulting in a broken financial model. Consequently, most pharmaceutical and biotech companies have exited their research efforts in this field, leaving the required rejuvenation of antibacterial drugs under-resourced. This article sheds light on some of the technical difficulties and offers thoughts and suggestions on how to address and overcome these.



Keywords: Drug discovery, antibiotics, property space, safety, medicinal chemistry

Table of Contents

1. Introduction

Antibiotics are arguably one of the greatest achievements in modern medicine and have become an integral part of our lives. Serious, life-threatening infections have generally been successfully treated when using the appropriate drug. Antibiotics are a rather recent invention,[1] mainly driven by screening of natural product extracts for their inhibitory effect on bacterial growth. During the so-called golden era of antibacterial drug discovery, most antibacterial scaffolds still in use today were identified, including β-lactams (penicillins, cephalosporins, cephems, monobactams, carbapenems, etc.), lipopeptides, macrolides, tetracyclines, and aminoglycosides.[2]

Antibacterial drugs are unique and quite different from other drugs, mainly reflected by the difficulty to discover and optimize novel scaffolds and their limited lifespan to treat infections. Firstly, bacterial pathogens rapidly reproduce – roughly one hundred thousand times faster than humans – and therefore their populations readily adapt to a changing environment, including exposure to antibiotics. Therefore resistance to antibiotics will generally emerge following their clinical introduction which will ultimately limit their use throughout hospitals, hospices, and the community. This necessitates the continued development of new antibiotics to treat the same infections caused by organisms with increasing resistance levels to standard of care agents. Secondly, since antibiotics target hard-to-kill bacteria rather than human targets, there is often a high free drug exposure required to treat patients successfully. For β-lactams this can easily exceed ten grams per day, which also increases pressure on cost of goods, formulation, and especially safety requirements. Luckily, in the case of antibacterials, animal models are clinically validated and especially the murine thigh model usually allows an accurate prediction of drug exposure required in patients.[3] Finally, the financial model for antibiotics is broken not only due to the competition from a multitude of now low priced generics that were launched in the 80's and 90's (Ceftazidime, Piperacillin, Tazobactam, Meropenem, Cefepime and others) that society has become accustomed to, but also the fact that novel drugs with expanded spectrum will often be kept in emergency reserve even though resistance against generic antibiotics is rising, hampering sales after launch.

One successful approach to the discovery of new antibiotics, probably best exemplified with β-lactams, builds on originally identified scaffolds by using chemical modifications to improve not only resistance patterns and antibacterial spectrum, but also pharmacokinetics and safety profile. This exploits the well understood desirable properties of the scaffold itself which reduces risk, and this approach still yields good antibiotics.[4] Another approach is the identification of inhibitors for entirely novel target such as LpxC.[5-8]

Increased appreciation of technical issues specific to target based antibacterial discovery came with the advent of the genomic era. The pharmaceutical industry placed large bets on target-driven discovery, using corporate archives and compound collections to biochemically screen targets of interest. This approach turned out to be successful in many therapeutic areas, but attrition was much higher for antibacterial targets. This phenomenon was noted by experts in the field for some time, but was ultimately brought to general attention by Payne and co-authors in 2007. They described their genomic-driven, large-scale experience at Glaxo Smith-Kline on roughly 70 essential bacterial targets by screening the GSK corporate archive.[9] This immense effort only resulted in hits for 16 targets and further optimization yielded lead compounds for five programs, all with activity only against Gram-positive bacteria. To the best of our knowledge, none of these leads has progressed to a successful clinical candidate. This was likely to be due in part to the issues described above, but also driven by other factors specific to the early identification of target inhibitors.

In early 2005 Achaogen began to build a project portfolio dedicated to finding new therapeutics to fight multi-drug resistant Gram-negative bacteria. We believed, based on the experience at other companies as exemplified by the publication of Payne and co-authors,[9] that high-throughput screening (HTS) of corporate archives would not identify the required hits or chemical starting points and that we would fail in our attempt to convert them to clinical candidates. Inspired by the work of Christopher Lipinski at Pfizer,[10] who brought the attention to physico-chemical properties of drug candidates and their importance for progression of preclinical and clinical candidates, we initiated a thorough analysis of existing antibacterial drugs and drug candidates.[11] Our goal was to better understand which chemical matter would possess a higher chance of success as starting point for hit/lead optimization. Even Lipinski mentioned in his work that anti-infective drugs would often violate his rules of five and he considered them as special and outside the norm.

We reviewed the literature in 2005 and found that no satisfactory analysis of antibiotic chemical property space was available. Some data sets were analysed and published but lacked the insights we were seeking since the compound sets were too small and to the best of our understanding at the time, did not encompass an appropriate range of compounds leading to biases [see references in ref. 11]. Therefore, we assembled a collection of data and used it to conduct our first analysis,[11] relying on susceptible strains to avoid resistance biases.

It was of critical importance to take into account the differences in the architecture of bacterial cell envelopes, particularly between Gram-positive and Gram-negative bacteria. Gram-positive bacteria contain a phospholipid bilayer membrane (cytoplasmic membrane) that is covered with a peptido-glycan layer (cell wall) that can vary in thickness. The cell wall is not considered a permeability barrier to small molecules. Gram-negative bacteria also possess a cytoplasmic (inner) membrane and cell wall, although the latter is thinner and varies less in thickness. Importantly, Gram-negative bacteria are defined by the presence of a second, asymmetric bilayer on the outer side of the cell wall (the outer membrane, OM). The outer leaflet of the OM is comprised of a carbohydrate conjugated lipid A or lipopolysaccharide, which is unique to Gram-negative bacteria, while the inner leaflet is phospholipid. The asymmetric lipid bilayer of the outer membrane has low fluidity and provides for low passive permeability of lipophilic probes.[12] Nikaido and others have demonstrated that many compounds traverse the OM predominantly by three mechanisms: active uptake, membrane disruption, or entry through membrane-embedded porins. The latter are β-barrel proteins that form trimeric channels with a polar inner surface, restricting the entry of lipophilic and/or large compounds with a molecular cut-off at around 600 Dalton.[12] Therefore the OM can be thought of as a molecular sieve that reduces the rate of influx of many compounds. In conjunction with slow influx, bacteria also employ efflux pump machinery that actively extrudes compounds as they enter cells. This is particularly problematic in Gram-negative bacteria such as *P. aeruginosa* that contain a large complement of efflux pumps of the RND

(Resistance-Nodulation-Division) and other families. RND family pumps in particular can exhibit exceptionally broad and overlapping substrate profiles, and experience has taught us that avoidance of such pumps is one of the biggest challenges in antibacterial drug discovery.[13]

Therefore we derived our original analysis using *S. aureus* and *E. coli* as representative Gram-positive and Gram-negative bacteria, respectively. To avoid biases resulting from various resistance mechanisms, we selected data for minimal inhibitory concentrations (MICs) against susceptible ATCC strains or, if data was not available, using existing MIC_{50} (minimal inhibitory concentration preventing growth of 50% strains evaluated), with an MIC value of 8 μg/mL as a cut-off for defining susceptibility.[11] Compilation of data for both values, MIC_{50} and ATCC strains, revealed a difference of less than two-fold for most compounds. In addition, we assembled information on compound class, target location, mode of action, active transport mechanism, and a number of calculated properties and additional data to guide our analysis. In this report, we have refined the dataset by adding more compounds that have entered clinical trials, and used an internally trained model for pK_a prediction to calculate logD values (parameter for polarity),[14] giving a more accurate reflection of properties than discussed in our previous study. Finally, we present the results of three large discovery efforts at Novartis, two with target location in the cytoplasm (CP) and one in the periplasm (PP). The internally generated data allowed us to compare the property space for three different antibacterial scaffold series with the results from known antibiotics.

2. Discussion

2.1 Property space of known antibiotics

The analysis of the antibacterial drug property space is limited by data availability and, especially for Gram-negative organisms, largely driven by the five major classes of antibiotics. As fosfomycin and aminoglycosides possess active or self-promoted uptake mechanisms,[15,16] and the target for β-lactams resides outside the inner membrane, mainly sulfa drugs, tetracyclines and fluoroquinolones contribute to the Gram-negative property space for compounds interacting with targets in the cytoplasm. Consequently, this analysis has to be taken with a grain of salt and is not meant to provide rules but rather suggest a preferred property space for antibacterial compounds. We therefore present this data analysis of selected unpublished internal efforts in that context, to complement the publicly available data of known antibiotics.

Larger molecular weight (MW) and higher polarity (lower calculated logD at pH 7.4) characterize the property space of antibiotics as compared to the general drug space (adjusted Comprehensive Medicinal Chemistry, CMC, database), as described in our first analysis.[11] The difference between compounds active only against Gram-positive or active against Gram-negative bacteria (most of the latter are active against Gram-positive organisms as well) is striking: the former contain roughly 25% compounds with MW above 850 Dalton whereas the latter group has less large molecules and is more polar, roughly four orders of magnitude on average as compared to the CMC dataset. The visualization of these differences is captured in Figure 1 with the CMC set omitted for clarity reasons. The Gram-negative compounds populate mainly the mid- and left lower part of the graph with only a few compounds above the 650 Dalton line. Exceptions are indicated on the graph, including membrane-interacting polymyxins, siderophore-carrying β-lactams (active transport), and finally azithromycin and LTX-109, two compounds with borderline *E. coli* activity at 8 μg/mL.

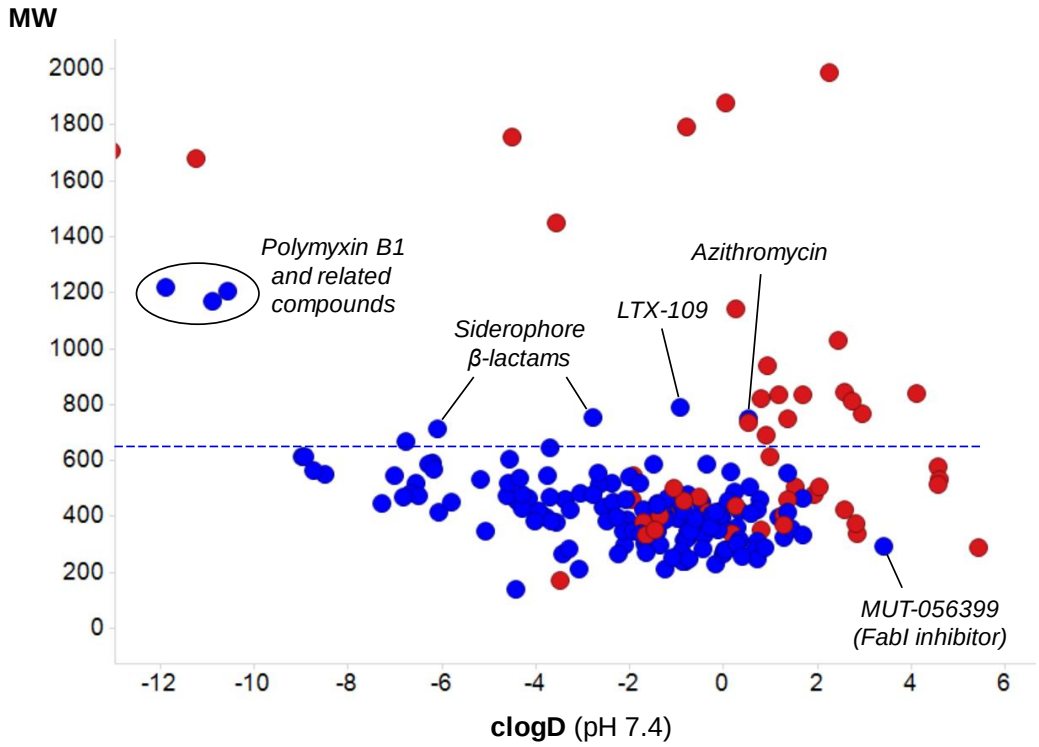


Figure 1. Property space of known antibiotics as visualized by molecular weight (MW) and calculated logD at pH 7.4 (see Supplementary Materials for calculation details). Antibiotics are colored by Gram-positive only (red) and Gram-negative (blue) activity, respectively.

In order to limit our analysis to property space correlating with passive uptake across the cell envelope, we had to further bin Gram-negative antibiotics: First, we eliminated compounds with self-promoted uptake (aminoglycosides and membrane disruptors) and the actively transported fosfomycin. Then we split compounds according to their target location; either in the CP with the requirement for molecules to permeate both membranes, or in the PP with the requirement to only permeate the outer membrane.

Table 1 captures some of the essential value ranges for some of the key properties and comparing them to the 'drug space' of the CMC database. The separation based on target location reveals the difference of is captured in the CP column. The MW is below 460 Dalton with a relatively narrow range of polarity. Experience demonstrates that these values are connected: for compounds on the more polar range, the MW has to be on the lower side as well to allow permeation through the inner membrane. In contrast, the size limitation for compounds targeting the PP space is less restricted with an average MW of roughly 100 Dalton higher as compared to targeting the CP. In addition, polarity of compounds is higher with little to no polarity restriction as seen by clogD and clogP values as well as by the polar surface area. The ranges listed in Table 1 should serve as guidelines only and especially for screening purposes, we recommend to populate the corresponding property space of compounds within libraries or archives of compounds.

Table 1. Property space ranges for individual parameters comparing Gram-negative antibiotics with a filtered CMC database. The ranges were determined from the datasets (supplementary materials) by cutting 10% on either end. Numbers of compounds are indicated in brackets. CP: Target within cytoplasm, PP: Target within periplasm

	CP (73) [a]	PP (60) [b]	CMC (4675)
MW	254-465	347-558	210-470
clogD (pH 7.4)	-1.4 to 1.1	-5.1 to –1.0	-1.3 to 4.7
clogp	-0.9 to 1.9	-3.4 to 0.9	-0.2 to 5.4
PSA [Å^2]	75 - 181	112 - 215	25 - 124
FRB	2 - 6	4 - 10	1 - 10
H-Donor	1 - 7	2 - 6	2 - 4
AROM	1 - 3	0 - 2	0 - 3

[a] Excluding aminoglycosides and Fosfomycin. [b] β-Lactams and β-lactamase inhibitors including DBOs (diazabicyclooctanes; avibactam scaffold)

2.2. Analysis of lead optimization programs

As an illustration of property space in antibacterial drug discovery, we decided to share data on multiple programs using scatter plot representation for size (MW), polarity (clogD) and adding an additional parameter for potency through coloring schemes for the dots (see Figure 2).

Figure 2a represents the property space for Gram-negative antibiotics (omitting aminoglycosides and Fosfomycin) as purple (CP) and blue (PP) circles with the CMC compounds as small grey dots. The purple square highlights the preferred space for Gram-negative antibiotics with targets in CP; this space is shifted to higher polarity and slightly larger MW as compared to the CMC compounds. The green ellipse depicts the most populated space within a typical corporate archive of compounds, which is shifted to higher MW and lipophilicity, largely driven by past synthetic efforts and chemical convenience in the past (vide infra). Based on this dataset of existing antibiotics it is clear how important the target location is and how this has to be taken into account both for hit finding and hit/lead optimization. Data published by other organizations active in this field are consistent with these findings as experimentally determined on a multitude of programs.[17,18]

Figure 2b contains compounds from an internal program with the target located intracellularly (CP). Only compounds are shown that were active on *P. aeruginosa* (MIC ≤ 8 µg/mL) and were not cytotoxic in two different cell lines (K562, HepG2) up to 100 µg/mL. The graph reveals the narrow range of property space for these lead compounds with the most promising compounds falling within a clogD space between minus one and one. Some compounds fall outside this illustrated preferred property space and remind us that this data analysis is based on a relative small data set of 151 Gram-negative antibiotics and as such can only provide us with some guidelines but not firm rules. In addition, calculation of polarity relies on accurate pKa prediction and this is – while improving over time – still inaccurate and will affect logD calculations especially with predicted pKa values close to the physiological pH. The supplementary material includes comparisons of the correlation between experimental and calculated logD values when using a pKa predictor trained with either a commercial training set or the commercial training set expanded with data from our in-house

knowledgebase.[14] The analysis shows that expansion of the training set for pKa prediction improves the performance of the subsequent logD prediction for the projects in this paper.

Figure **2a**

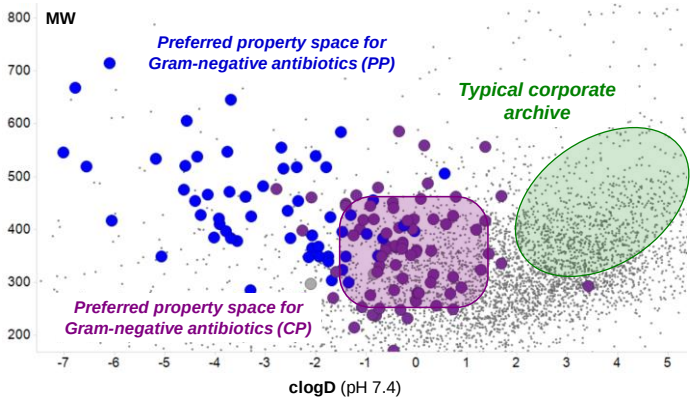


Figure **2b**

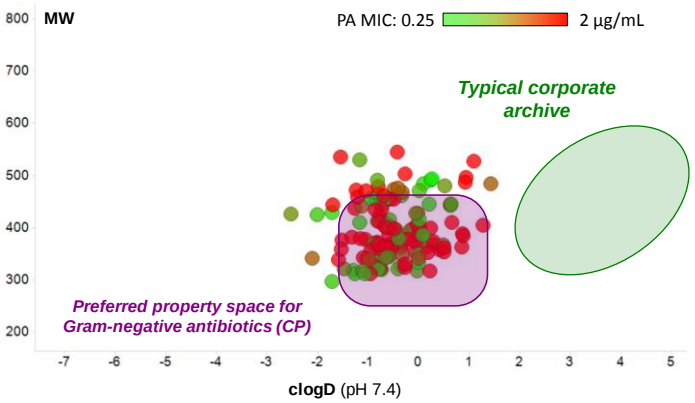


Figure **2c**

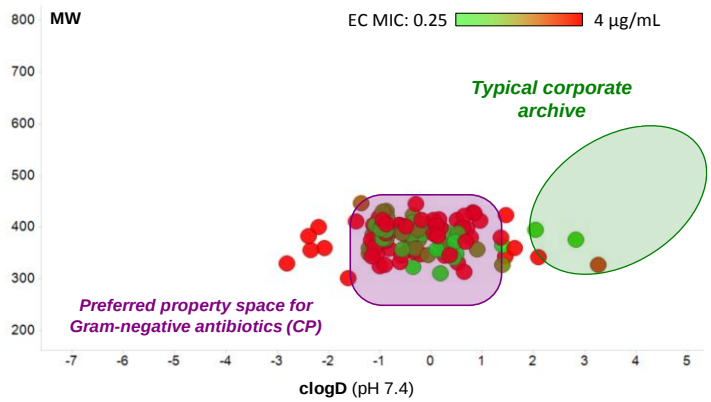


Figure **2d**

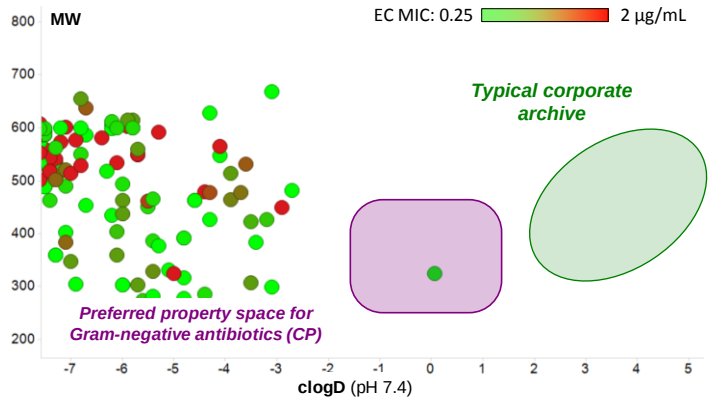


Figure 2. Representation of Gram-negative antibiotics in the MW/clogD property space. **2a**: Known antibiotics are colored according to target location for CP (purple) and PP (blue). Spaces occupied by typical archive (green) and Gram-negative antibiotics with CP target location (purple) are indicated. **2b**: First internal program with CP target location filtered by activity (*P. aeruginosa*) and cytotoxicity from over 2376 compounds made (816 compounds with MIC ≤8 µg/mL, 137 compounds without detected cytotoxicity). **2c**: Second internal program with CP target location filtered by activity (*E. coli*) and cytotoxicity from 1406 compounds made (803 compounds with MIC ≤ 8 µg/mL, 124 compounds without detected cytotoxicity). **2d**: Internal program with PP target location filtered by *E. coli* activity (MIC ≤ 2 µg/mL); 258 compounds from 541 made, including 26 reference compounds.

Compounds for a second intracellular target (CP) are visualized in Figure 2c. In this particular case, we used *E. coli* as activity cut-off (MIC ≤ 8 µg/mL) and colored the dots according to activity. Compounds with cytotoxicity at or below 100 µg/mL in three different cell lines (K562, MT4, and HepG2) were omitted from this analysis. Again, this data reveals a narrow distribution of active compounds in a similar property space range as compared to the first target. For both examples, lead compounds are positioned outside the most populated archive space, indicating these historic collections of compounds underrepresent the most desirable space.

Further support for the preferred property space defined here for cytoplasmic targets can be gleaned by contrasting it with the preferred property space for monobactam compounds that target PBPs in the

periplasm. Since these compounds do not need to permeate the inner membrane they are predicted to be much less constrained in terms of increased polarity than would be the case for inhibitors of cytoplasmic targets (see blue dots in Figure 2a). Figure 2d contains all non-cytotoxic compounds of our past monobactam program with *E. coli* activity of ≤8 µg/mL, and indeed these are concentrated to the left portion of Figure 2d, well outside of the box indicating preferred property space for cytoplasmic targets. This clear contrast based on one parameter provides support for both property spaces (that indicated in purple and one for periplasmic targets). However, it remains to be seen if relaxed polarity criteria would hold true for inhibitors of other periplasmic targets besides PBPs. If addition of more polarity is possible and supported by the target, we believe this to be beneficial as it typically adds to higher solubility of compounds (key importance for IV administration) and prevents them from entering mammalian cells; a clear benefit to the safety profile of such compounds.

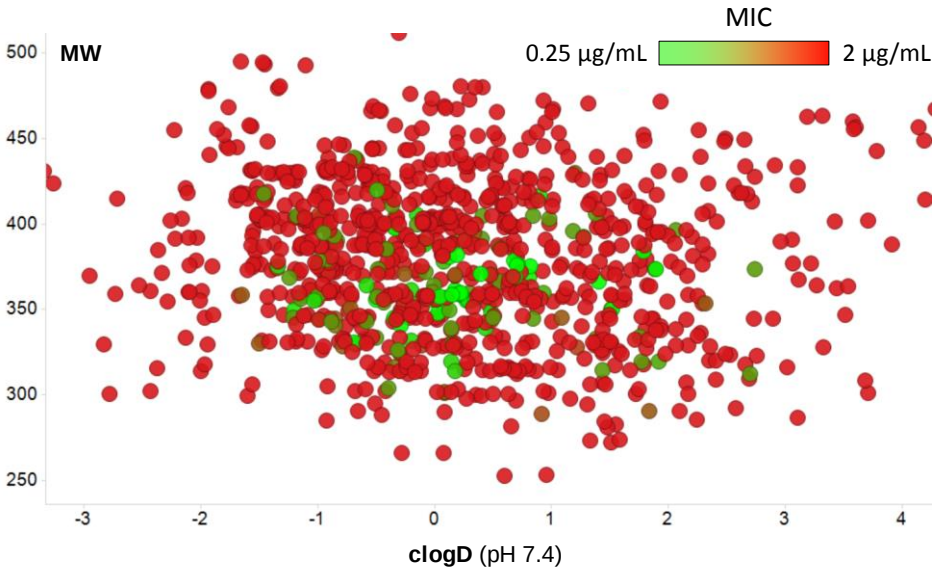


Figure 3. Second internal program with CP target localization. All compounds are displayed, irrespective of activity or cytotoxicity (1406 compounds without the reference standards).

The representations of compounds in Figure 2 are biased due to choosing non-toxic compounds with a minimal level of antibacterial activity. We therefore provide a visualization of all compounds made for the program shown in Figure 2c (CP target), irrespective of antimicrobial activity or cytotoxicity (Figure 3). This Figure nicely visualizes the activity hotspot around a MW of 360 Dalton and a calculated logD value between zero and one. Active compounds were prepared that deviated from these values, with potency rapidly decreasing above 400 Dalton whereas polarity is ranging from minus one to two (calculated logD). It is worth noting that compounds with lower polarity (log D ≥1) and good antibacterial activity do not show up in Figure 2c due to detected cytotoxicity (vide infra for discussion).

2.3 General safety considerations

Attrition of drug candidates at the candidate selection stage and in early clinical development is primarily due to safety concerns.[19] Waring and coauthors also found a link between physicochemical properties and failure due to safety issues; an observation which builds on Lipinski's in-depth analysis at Pfizer and the creation of rules of five.[10]

The importance of safety considerations for antibacterial drug candidates is further reinforced by the fact that safe and effective treatment options are generally available, except where eroded by resistance. The relative high doses needed for efficacy, especially for treatment of infections caused by Gram-negative organisms, results in high safety requirements for antibiotics.

This is in contrast to the success rate in late stage clinical trials, which is higher for infectious diseases[20] than for other disease areas, likely due to the availability of highly predictive animal efficacy models for key antibacterial indications[3].

An illustration of the safety challenges faced in development of novel antibacterial agents may be given by the struggle to avoid QTc prolongations due to inhibition of the human Ether-a-go-go-Related Gene product hERG. QTc prolongation is a class effect of the fluoroquinolones and has recently stalled the development of several non-fluoroquinolone inhibitors of bacterial topoisomerases as well.[21-23] Applying the conservative margin of 100-fold between the hERG IC_{50} and free C_{max} in the clinic, as recommended by Redfern,[24] the target hERG IC_{50} value for an optimized analog with a predicted free C_{max} in the clinic of 10 µM (a typcial value for a Gram-negative drug) would be 1 mM. Optimizing compounds for millimolar hERG IC_{50} values is a huge challenge, requiring sufficient solubility for *in vitro* hERG assays, a non-trivial challenge for most scaffold classes.

Off-target activity against ion channels like hERG, as well as against biogenic amine receptors like GPCRs, which have historically been important to avoid for safety reasons, is being driven to a large extent by physicochemical properties, namely lipophilicity, charge status and aromaticity, which will be discussed in more detail below.

2.3.1. Lipophilicity. Lipophilic interactions are structurally more forgiving and less ligand-protein complementarity is needed.[25] As a result, increased lipophilicity leads to a higher degree of pharmacological promiscuity.[26] Increased lipophilicity and pharmacological promiscuity correlates with worse toxicological outcomes.[27-29] Hansch has observed for approved drugs in the CNS space, that increased lipophilicity correlates with lethality, and has formulated the principle of minimal hydrophobicity: "Without convincing evidence to the contrary, drugs should be made as hydrophilic as possible without loss of efficacy".[30]

Gram-negative drugs are more polar than drugs of the CMC comparator set (Figure 2a), and the higher bar for safety is likely one of the key drivers in addition to better permeability of polar compounds through the porins of the outer membrane as discussed above.

Lipophilicity does not only affect the quality of drug candidates and drugs, but it also impacts the quality of hits from phenotypic screening. Figure 4 shows the hit-rates for *Pseudomonas aeruginosa*, as well as cytotoxicity against mammalian and yeast cells in high-throughput screening campaigns at Novartis that read out cell-death.

Hit rates for *P. aeruginosa* are relatively constant over a range for clogD of -1 to >6, while hit rates for mammalian and yeast cytotoxicity increase till reaching a plateau at around a clogD of >4. This difference is likely due to the relative impermeability of the outer membrane of *P. aeruginosa* to lipophilic probes.[31] Further analysis of the relationship between clogD and cytotoxicity includes a dataset of validated inhibitors of the growth of Gram-negative bacteria from the Novartis knowledge base. The analysis shows that at clogD of 3 and higher, 83% of these validated inhibitors of the growth of Gram-negative bacteria are also cytotoxic against mammalian cells (Figure 5), likely reflecting unspecific membrane action and higher degree of promiscuity for such lipophilic hits. The consequence of this is that compounds with a clogD > 3 are unlikely to specifically inhibit the growth of Gram-negative bacteria, and we tend to bias away from these lipophilic compounds when designing screening sets.[47]

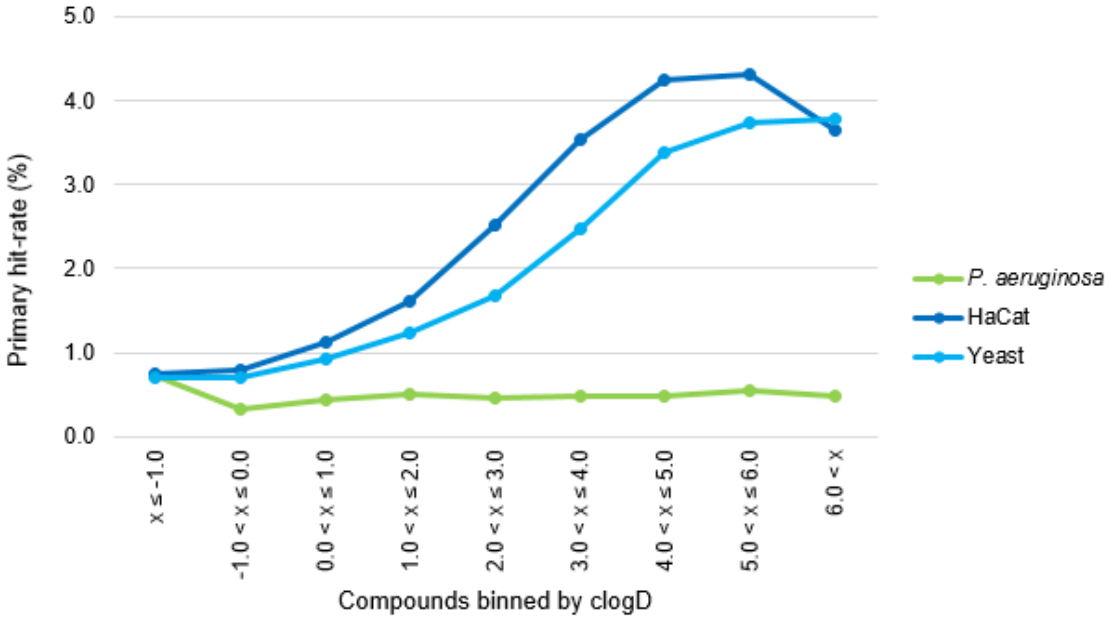


Figure 4. Primary hit-rate data for sets of compounds binned by clogD from three high-throughput screens (roughly 800k compounds each). Compounds were screened at 40 µM for activity against *P.aeruginosa* (efflux mutant; green), and at 20 µM and 24 µM for cytotoxicity against yeast and a human HaCat cell line respectively (blue); hits were called at >50% inhibition in all cases (see supplementary materials for further details).

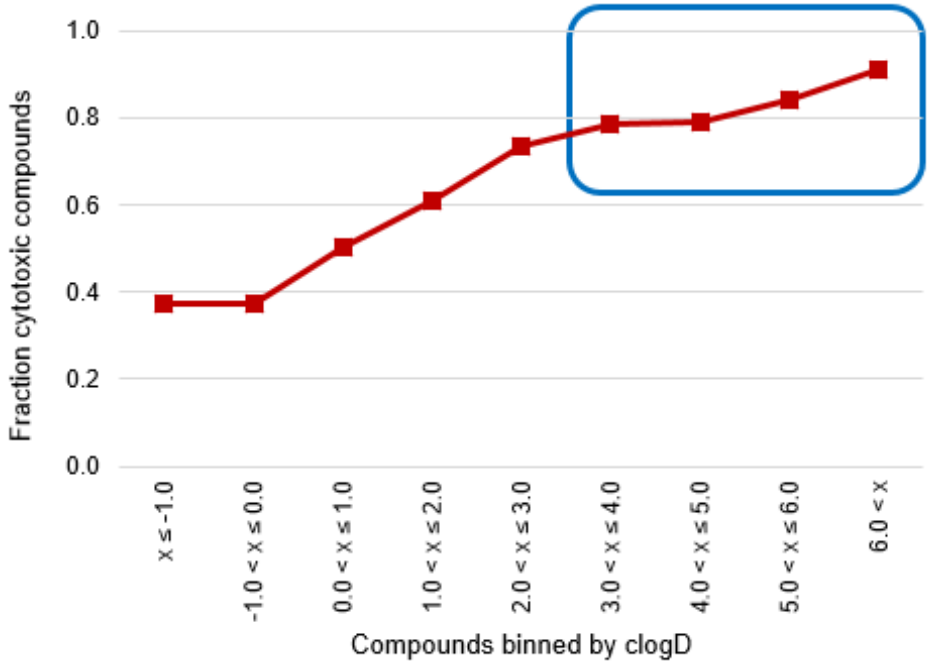
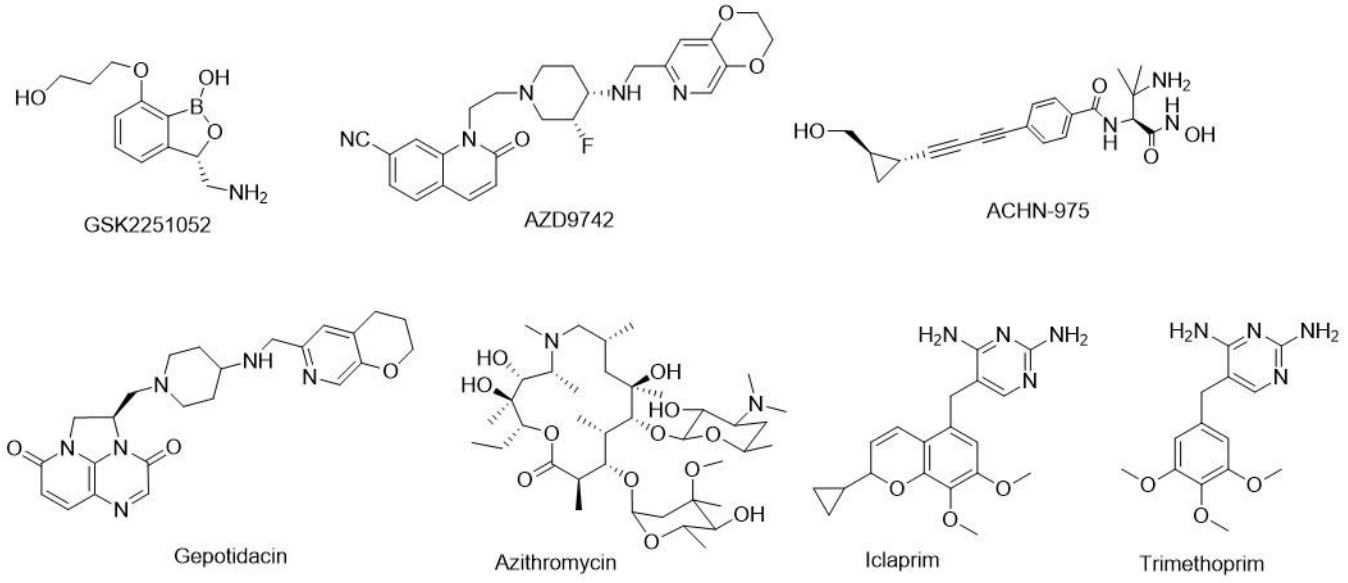


Figure 5. Fraction of cytotoxic compounds for sets of validated inhibitors of the growth of wild-type Gram-negative bacteria binned by calculated logD. There are a total of 292 compounds in the bins with clogD > 3 (blue box) and 241 of these (83%) are cytotoxic. Details of the analysis are shared in the supplementary materials.

2.3.2. Charge status. Pharmacological promiscuity with regard to target classes that have historically been implicated with safety issues, such as GPCRs and ion channels, is also driven by charge status.[26,32] Positively charged compounds tend to be most promiscuous, acids least and zwitterions and neutral molecules fall in-between. Basic compounds (positively charged at physiological pH) are frequent hits in phenotypic screens. The positive charge allows for interactions with negatively charged phospholipid head groups of the cytoplasmic membranes of bacteria,[33] as well as with negatively charged groups within lipopolysaccharides of the outer membrane of Gram-negative organisms to affect disruptions of bacterial membranes leading to compound entry.[16] The permeability advantage with primary amines has recently again been emphasized,[34] but the dilemma remains that such compounds are more difficult to optimize towards safe drugs.

We analyzed our database with regard to charge status for Gram-negative drugs (actively transported drugs excluded) (Table 1). For compounds targeting the cytoplasm, the majority are zwitterionic or weakly acidic, with only 10% (7 compounds) being bases. These 7 compounds (Scheme 1) are GSK2251052, AZD9742, ACHN-975, Gepotidacin, Azithromycin, Iclaprim and Trimethoprim. GSK2251052 showed QTc prolongation in man,[22] AZD9742 and ACHN-975 were discontinued in phase 1. Azithromycin, Iclaprim and trimethoprim are administered at a relative low dose (daily doses of 500, 160 and 240 mg respectively), indicating that if a low efficacious dose can be utilized, a basic compound may produce an acceptable safety profile. GSK2251052 is a small and polar (clogD -3.7) base dosed in phase 2 at 1.5g q12h. The compound failed due to rapid development of resistance in phase 2, but appears to have been well tolerated. It may thus be possible to develop a basic compound that is not actively transported at a relative high dose as long as it is small and sufficiently polar.



Scheme 1. Antibiotics carrying a positive charge.

Compounds targeting the periplasm are zwitterionic or strongly acidic. This also reflects the requirement of the targets penicillin binding proteins and β-lactamases for a negative charge. The charge status and low logD (see also Figure 2d) of these compounds explain the excellent safety profile of many members of the β-lactam class.

Basic compounds that have to permeate passively into the cytoplasm of Gram-negative organisms and require a high dose (>500 mg) are thus to our knowledge not precedented within the currently approved drug

space, even though basic compounds are frequently obtained as phenotypic sceening hits. However, if one could identify novel targets in the PP that can be inhibited with very polar amines, such molecules might show an acceptable safety profile. There are also no Gram-negative drugs targeting the CP that are strong acids, likely due to challenges for polar acids to effectively permeate through the inner membrane. Zwitterions and neutral compounds may provide a good balance between permeability and safety requirements, especially with pKa values close to physiological pH, allowing the coexistence of multiple species.

2.3.3. Aromaticity. Ritchie and Macdonald have shown that aromatic ring count (AROM) correlated with compound developability at GSK.[35] The mean AROM from candidate selection to phase 1 to Proof of Concept in phase 2 dropped from 3.3 to 2.5 to 2.3 respectively, reflecting challenges in areas of safety and ADME properties for compounds with high AROM. Increased AROM leads to higher plasma protein binding, hERG activity and cytochrome P_{450} (CYP_{450}) inhibition, and lower aqueous solubility. These observed effects are not only a function of higher lipophilicity for compounds with high AROM, but also reflect the higher content of aromatic rings. For example, receptor promiscuity increases with AROM independent of logD.

Higher aromatic ring count will make development of high-dose iv antibiotics especially difficult, due to formulation challenges for compounds with low solubility. The analysis of the iv subset of all antibacterial compounds in our database shows that the average AROM is 1.6 (113 compounds; AROM = 1.26 for the Gram-negative subset of 85 compounds), which is identical to the value for the CMC set, whereas the Novartis solution screening Archive shows an average AROM of 2.4. The typical screening decks in large pharma are too aromatic relative to existing drugs in general.[36,37] For Gram-negative antibiotics, especially iv agents, the low solubility of compounds with high aromaticity presents a challenge.

Conclusions

Antibacterial drug discovery is undoubtedly challenging, especially if targeting Gram-negative pathogens. Decades of intensive effort at biotech and large pharmaceutical companies as well as in academia have yielded only a few novel discoveries and the majority of new antibiotics with expanded spectrum originated from fine-tuning of existing scaffolds. A lesson learned from these past efforts is that screening of large company compound archives *in vitro* against essential targets will yield hits but their conversion into real drug candidates active against Gram-negative pathogens has a low probability of success.

One reason for this may relate to observations by Walters, who pointed out that compounds made by medicinal chemists have started to diverge from the chemical property space defined by drugs (clogP and fraction sp3) over time.[36] This trend began in the early 80's, coinciding with the second wave of palladium-catalyzed cross couplings which were discovered in the late 70's.[38] Other "workhorse" reactions used in medicinal chemistry include reductive aminations,[39] which appear to have enriched archives with basic compounds. The tendency to favor easy-to-synthesize compounds (presumably to increase output) is likely counterproductive, as many of these will reside outside the optimal property space for drug development in general. Table 1 and Figure 2A illustrate the notion that corporate compound archives also lack sufficient diversity of compounds within the preferred target property space associated with Gram-negative cellular activity. These issues manifest acutely in the case of antibacterial discovery, where compounds must be optimized for cellular potency, and therapeutics are usually dosed at comparatively high levels necessitating ever safer compounds with better solubility. Against this backdrop, less optimal chemical starting points can prohibitively widen the distance to the design of solid leads or therapeutic agents.

Given the hard-won knowledge outlined in this article and elsewhere, how can we improve the important process of antibacterial discovery and increase chances for success? We do not claim to have a patent solution, but provide suggestions we expect may go some way towards improving the odds. Target inhibitors are fairly easy to identify, either by *in vitro* biochemical or biophysical screens or by phenotypic screens and subsequent target identification.

The first challenge is therefore the conversion of the hit into an on-target, whole-cell active hit or lead compound, which usually boils down to improving permeability through two opposite membranes while avoiding a battery of efflux pumps. As we have illustrated, chemical starting points tend to reflect the suboptimal nature of compound archives. An obvious next step is therefore to improve the archives. One way to do this would be by purchasing diverse sets of compounds covering a favorable property space (Table 1) from commercial suppliers who realized the need for more polar, and therefore more functionalized, compounds and who are increasingly covering this need.

Another way to address the nature of chemical property space available to a particular target program is via fragment-based approaches that enable medicinal chemists to more specifically build and optimize program molecules. This also requires careful consideration when constructing fragment libraries for use in screens to provide useful starting points. Murray and Rees have proposed that for growing fragments in fragment-based drug design, new methodologies are still needed for more efficient transformation towards diverse and polar compounds.[40] There have been some notable recent advances in new methodologies suitable for generating compounds with higher sp^3 character (less flat). These include photoredox catalysis developed by David MacMillan's group,[41,42] bicycloalkyl bioisosteres of phenyl groups,[43] and a novel route to substituted piperidines with tunable regiochemistry.[44]

Along with augmenting traditional compound libraries with better starting points, an alternative is vastly increasing the size and correspondingly, the diversity of libraries. DNA-encoded libraries (DEL) offer an attractive way to interrogate very large libraries of compounds that directly interact with a given target, then enrich for binding library members and decode them through amplification and DNA sequencing.[45] Currently, DEL cannot be used for phenotypic screening due to the covalently attached DNA tag but technology advancements might eventually allow us to encode / decode large libraries in the context of phenotypic screening.

Natural products have been a main source for the antibiotics in use currently, and are still an attractive potential source of antibiotics, as it has been proposed that the vast chemical diversity encompassed by natural products remains unexplored. Future natural products efforts may benefit from broadening the range of novel producing organisms, including acquisition of producing strains from environments where evolution may have driven the generation of metabolites active against resistant bacteria (e.g. hospital sewage from geographic locations with high prevalence of antibiotic resistant organisms). Genomic and genetic approaches may provide insights into silent gene clusters, particularly if these lie in proximity to resistance genes that could hint at potential targets.[46] Lastly, exploring different product screening and isolation methods may enrich for previously unexploited compounds with attractive chemical properties or ease of creating synthetic analogs.

Historically, phenotypic or whole-cell screening has been used with success in antibacterial discovery, and indeed is responsible for identifying most antibacterial classes we know today. Since libraries have been extensively and repeatedly screened this way, the output from this approach has however dwindled. Indeed, many phenotypic screens employ highly compromised hypersusceptible bacteria which increases sensitivity but yields inhibitors that, like those identified by *in vitro* biochemical screens, possess little or no activity against wild type bacterial pathogens. With or without improved compound archives, screening conditions

that mirror a more physiologically relevant environment as compared to Mueller-Hinton broth might yield previously unidentified inhibitors from the portion of the archive with desirable properties and/or with activity against wild type cells.

Although there are several potentially promising avenues of antibacterial discovery it also needs to be appreciated that continued discovery of new antibiotics will require commitment and resourcing. Given the overall perception of a lack of financial incentive to continue discovery and especially development of novel antibiotics, this commitment, especially among large or medium sized pharmaceutical companies, has largely evaporated. The few products introduced to the market during the past two years (delafloxacin, meropenem/vaborbactam, ceftazidime/avibactam, plazomicin) had disappointing sales with a grim outlook for the companies promoting these products. This has placed a huge downward pressure on their stock price (Melinta traded at $0.83 and Achaogen at $1.30 as of December 21, 2018, 4 pm) which does little to improve the outlook. Unless substantial funding becomes available for novel antibacterial drug discovery and development, there is a worrying risk of fewer or even no treatment options against multi-drug resistant bacteria coming in the future. Furthermore, the resulting exodus from antibacterial research and development constitutes a substantial and possibly irretrievable loss of expertise in this area as many scientists have moved on to support other therapeutic areas or technology platforms.

Acknowledgements

We are grateful to many former and current colleagues who have contributed initially at different companies and over the past seven years at Novartis. Without their dedication to drug discovery and endless pursuit to generate high quality drugs, we could not have assembled this manuscript. Special thanks are due to Charles Dean who critically reviewed this manuscript and provided valuable scientific input.

Supplementary Material

The supplementary data provide specifics regarding logD calculations, additional information to Figures 4 and 5, the protocol for experimental logD determination, and scatter plots comparing experimental and calculated logD values for the two CP targets. In addition, information on the 210 antibacterial compounds used in this analysis, including structures, annotations and property space parameters, are provided in the supplied excel sheet.

References

1. Armstrong G.L.; Conn L.A.; Pinner, R. W. *J. Am. Med. Assoc.* **1999,** *281,* 61.
2. Aminov, R. *Biochem. Pharmacol.* **2017,** *133,* 4. https://doi.org/10.1016/j.bcp.2016.10.001
3. Craig, W. A. *Clinical Infectious Diseases* **1998,** *26,* 1. https://doi.org/10.1086/516284
4. Reck, F.; Bermingham, A.; Blais, J.; Capka, V.; Cariaga, T.; Casarez, A.; Colvin, R.; Dean, C. R.; Fekete, A.; Gong, W.; Growcott, E.; Guo, H.; Jones, A. K.; Li, C.; Li, F.; Lin, X.; Lindvall, M.; Lopez, S.; McKenney, D.;

Metzger, L.; Moser, H. E.; Prathapam, R.; Rasper, D.; Rudewicz, P.; Sethuraman, V.; Shen, X.; Shaul, J.; Simmons, R. L.; Tashiro, K.; Tang, D.; Tjandra, M.; Turner, N.; Uehara, T.; Vitt, C.; Whitebread, S.; Yifru, A.; Zang, X.; Zhu, Q. *Bioorg. Med. Chem. Lett.* **2018,** *28,* 748.
https://doi.org/10.1097/01.ccm.0000528084.08373.f2

5. Barb, A. W.; McClerren, A. L.; Snehelatha, K.; Reynolds, C. M.; Zhou, P.; Raetz, C. R. *Biochemistry* **2007,** *46,* 3793.
https://doi.org/10.1021/bi6025165
6. Piizzi, G.; Parker, D. T.; Peng, Y.; Dobler, M.; Patnaik, A.; Wattanasin, S.; Liu, E.; Lenoir, F.; Nunez, J.; Kerrigan, J. *J. Med. Chem.* **2017,** *60,* 5002.
https://doi.org/10.1021/acs.jmedchem.7b00377
7. Tomaras, A. P.; McPherson, C. J.; Kuhn, M.; Carifa, A.; Mullins, L.; George, D.; Desbonnet, C.; Eidem, T. M.; Montgomery, J. I.; Brown, M. F.; Reilly, U.; Miller, A. A.; O'Donnella, J. P. *mBio* **2014,** *5,* e01551.
https://doi.org/10.1128/mBio.01551-14
8. Onishi, H. R.; Pelak, B. A.; Gerckens, L. S.; Silver, L. L.; Kahan, F. M.; Chen, M. H.; Patchett, A. A.; Galloway, S. M.; Hyland, S. A.; Anderson, M. S.; Raetz, C. R. *Science* **1996,** *274,* 980.
https://doi.org/10.1126/science.274.5289.980
9. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nature Revs Drug Discov.* **2007,** *6,* 29.
https://doi.org/10.1038/nrd2201
10. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997,** *23,* 3.
https://doi.org/10.1016/S0169-409X(96)00423-1
11. O'Shea, R.; Moser, H. E. *J. Med. Chem.* **2008,** *51,* 2871.
https://doi.org/10.1021/jm700967e
12. Nikaido, H. *Microbiol. Mol. Biol. Rev.* **2003,** *67,* 593.
https://doi.org/10.1128/MMBR.67.4.593-656.2003
13. Lomovskaya, O.; Zgurskaya, H. I.; Totrov, M.; Watkins, W. J. *Nat. Rev. Drug Discov.* **2007**, *6*, 56.
https://doi.org/10.1038/nrd2200
14. Gedeck, P.; Lu, Y.; Skolnik, S.; Rodde, S.; Dollinger, G.; Jia, W.; Berellini, G.; Vianello, R.; Faller, B.; Lombardo, F. *J. Chem. Inform. Modeling* **2015,** *55,* 1449.
https://doi.org/10.1021/acs.jcim.5b00172
15. Kahan, F. M.; Kahan, J. S.; Cassidy, P. J.; Kropp, H. *Ann. N. Y. Acad. Sci.* **1974,** *235,* 364.
https://doi.org/10.1111/j.1749-6632.1974.tb43277.x
16. Hancock, R. E. W.; Bell, A. *Eur. J. Clin. Microbiol. Infect. Dis.* **1988,** *7,* 713.
https://doi.org/10.1007/BF01975036
17. Brown, D. G.; May-Dracka, T. L.; Gagnon, M. M.; Tommasi, R. *J. Med. Chem.* **2014,** *57,* 10144.
https://doi.org/10.1021/jm501552x
18. Tommasi, R.; Brown, D. G.; Walkup, G. K.; Manchester, J. I.; Miller, A. A. *Nature Revs Drug Discov.* **2015,** *14,* 529.
https://doi.org/10.1038/nrd4572
19. Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; Wallace, O.; Weir, A. *Nat. Rev. Drug Discov.* **2015,** *14,* 475.
https://doi.org/10.1038/nrd4609
20. Hwang, T. J.; Lauffenburger, J. C.; Wang, B.; Franklin, J. M.; Kesselheim, A. S.; Carpenter, D. *JAMA Intern. Med.* **2016,** *176,* 1826.
https://doi.org/10.1001/jamainternmed.2016.6008

21. Lesuisse, D.; Tabart, M. In *Comprehensive Chirality;* Yamamoto, H., Carreira, E. M., Eds.; Elsevier, 2012; Vol. 1, pp 8-29. https://doi.org/10.1016/B978-0-08-095167-6.00112-9
22. Hossain, M.; Zhou, M.; Darpo, B.; Tiffany, C.; Dumont, E.; Darpo, B. *Antimicrob; Agents Chemother.* **2017,** *61,* e01221. https://doi.org/10.1021/jo00884a049
23. Reck, F.; Alm, R. A.; Brassil, P.; Newman, J. V.; Ciaccio, P.; McNulty, J.; Barthlow, H.; Goteti, K.; Breen, J.; Comita-Prevoir, J.; Cronin, M.; Ehmann, D. E.; Geng, B.; Godfrey, A. A.; Fisher, S. L. *J. Med. Chem.* **2012,** *55,* 6916. https://doi.org/10.1021/jm300690s
24. Redfern, W. S.; Carlsson, L.; Davis, A. S.; Lynch, W. G.; MacKenzie, I.; Palethorpe, S.; Siegl, P. K. S.; Strang, I.; Sullivan, A. T.; Wallis, R.; Camm, A. J.; Hammond, T. G. *Cardiovasc. Res.* **2003,** *58,* 32. https://doi.org/10.1016/S0008-6363(02)00846-5
25. Peters, J. In *Polypharmacology in Drug Discovery;* J.-U. Peters Ed.; John Wiley & Sons Inc., 2012; pp 47-62. https://doi.org/10.1002/9781118098141.ch3
26. Leeson, P. D.; Springthorpe, B. *Nat. Rev. Drug Discov.* **2007,** *6,* 881. https://doi.org/10.1038/nrd2445
27. Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; DeCrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y. *Bioorg. Med. Chem. Lett.* **2008,** *18,* 4872. https://doi.org/10.1016/j.bmcl.2008.07.071
28. Greene, N.; Aleo, M. D.; Louise-May, S.; Price, D. A.; Will, Y. *Bioorg. Med. Chem. Lett.* **2010,** *20,* 5308. https://doi.org/10.1016/j.bmcl.2010.06.129
29. Rana, P.; Aleo, M. D.; Gosink, M.; Will, Y. *Chem. Res. Toxicol.* **2018,** .
30. Hansch, C.; Bjorkroth, J. P.; Leo, A. *J. Pharm. Sci.* **1987,** *76,* 663. https://doi.org/10.1002/jps.2600760902
31. Nikaido, H. *Microbiol. Mol. Biol. Revs* **2003,** *67,* 593. https://doi.org/10.1128/MMBR.67.4.593-656.2003
32. Peters, J.; Hert, J.; Bissantz, C.; Hillebrecht, A.; Gerebtzoff, G.; Bendels, S.; Tillier, F.; Migeon, J.; Fischer, H.; Guba, W.; Kansy, M. *Drug Discov. Today* **2012,** *17,* 325. https://doi.org/10.1016/j.drudis.2012.01.001
33. Epand, R. M.; Epand, R. F. *Journal of Peptide Science* **2011,** *17,* 298. https://doi.org/10.1002/psc.1319
34. Richter, M. F.; Drown, B. S.; Riley, A. P.; Garcia, A.; Shirai, T.; Svec, R. L.; Hergenrother, P. J. *Nature (London, U. K.)* **2017,** *545,* 299. https://doi.org/10.1038/nature22308
35. Ritchie, T. J.; MacDonald, S. J. F. *Drug Discov. Today* **2009,** *14,* 1011. https://doi.org/10.1016/j.drudis.2009.07.014
36. Walters, W. P.; Green, J.; Weiss, J. R.; Murcko, M. A. *J. Med. Chem.* **2011,** *54,* 6405. https://doi.org/10.1021/jm200504p
37. Lovering, F.; Bikker, J.; Humblet, C. *J. Med. Chem.* **2009,** *52,* 6752. https://doi.org/10.1021/jm901241e
38. Johansson Seechurn, C. C. C.; Kitching, M. O.; Colacot, T. J.; Snieckus, V. *Angew. Chem. Int. Ed.* **2012,** *51,* 5062.

https://doi.org/10.1002/anie.201107017
39. Roughley, S. D.; Jordan, A. M. *J. Med. Chem.* **2011,** *54,* 3451. https://doi.org/10.1021/jm200187y
40. Murray, C. W.; Rees, D. C. *Angew. Chem. Int. Ed.* **2016,** *55,* 488. https://doi.org/10.1002/anie.201506783
41. Zhang, P.; Le, C. C.; MacMillan, D. W. C. *J. Am. Chem. Soc.* **2016,** *138,* 8084. https://doi.org/10.1021/jacs.6b04818
42. Chu, L.; Ohta, C.; Zuo, Z.; MacMillan, D. W. C. *J. Am. Chem. Soc.* **2014,** *136,* 10886. https://doi.org/10.1021/ja505964r
43. Auberson, Y. P.; Brocklehurst, C.; Furegati, M.; Fessard, T. C.; Koch, G.; Decker, A.; La Vecchia, L.; Briard, E. *ChemMedChem* **2017,** *12,* 590. https://doi.org/10.1002/cmdc.201700082
44. Duttwyler, S.; Chen, S.; Takase, M. K.; Wiberg, K. B.; Bergman, R. G.; Ellman, J. A. *Science* **2013,** *339,* 678. https://doi.org/10.1126/science.1230704
45. Goodnow Jr, R. A.; Dumelin, C. E.; Keefe, A. D. *Nature Revs Drug Discovery* **2017,** *16,* 131. https://doi.org/10.1038/nrd.2016.213
46. Smanski, M. J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M. A.; Voigt, C. A. *Nature Revs Microbiology* **2016,** *14,* 135. https://doi.org/10.1038/nrmicro.2015.24
47. Jansen, J. M.; De Pascale, G.; Fong, S.; Lindvall; M.; Moser, H. E.; Pfister, K.; Warne, B.; Wartchow, C. *J. Chem. Inform. Modeling* **2019**, **DOI:** 10.1021/acs.jcim.9b00048 https://doi.org/10.1021/acs.jcim.9b00048

Authors' Biographies

Folkert Reck received his M.S. and Ph.D. in Organic Chemistry from the University of Hamburg (H. Paulsen) and did postdoctoral research at the Hospital for Sick Children (H. Schachter) and the Scripps Research Institute (A. Eschenmoser). He is a medicinal chemist with over 20 years of experience in antibacterial drug discovery, and is currently a director of chemistry at the Novartis Institutes for BioMedical Research.

Johanna M. ("Hanneke") Jansen received her Ph.D. in Computational Medicinal Chemistry from the School of Pharmacy, University of Groningen (H. Wikström, C. Grol)) and did postdoctoral research at the BioMedical Center of Uppsala University (U. Hacksell, A. Johansson). She is a director in Computer-Aided Drug Discovery at the Novartis Institutes for BioMedical Research. Hanneke's research interests include the design of relevant compound sets for hit-generation, predictive modeling that leverages large complex datasets, and developing computational chemistry and data-mining methods that can be integrated into state-of-the-art workflows to inform decision-making from target-ID and validation through hit generation and optimization. Prior to Novartis, Hanneke worked at Chiron Corporation and at Astra.

Heinz E. Moser received his M.S. and Ph.D. in Chemistry from ETH Zurich (A. Eschenmoser) and did postdoctoral research at Caltech (P. Dervan). He heads medicinal, analytical and structural chemistry in Emeryville, supporting the global NIBR portfolio in multiple areas. Prior to this position he worked as VP of chemistry at Achaogen, specializing in antibacterial drug discovery. Previously, he was CTO and VP of chemistry at Genesoft Pharmaceuticals, a company exploring sequence-specific DNA binding agents as novel therapeutic modality. He started his professional career in 1987 at CIBA in Basel where he initiated and later headed a project in the field of nucleic acid based therapeutics (antisense). During the merger to form Novartis in 1996 he relocated to Horsham UK, heading the medicinal chemistry department for respiratory diseases. His main professional interests are related to drug discovery and medicinal chemistry and specifically include antibacterials, cheminformatics, natural products and interdisciplinary approaches to improve the quality of novel drugs.

