ELSEVIER

feature

Accelerated drug discovery by rapid candidate drug identification

Fredrik Bergström, Fredrik.h.bergstrom@astrazeneca.com **and Bo Lindmark**

The eventual candidate drug (CD) is often already synthesized during early drug discovery but not nominated until much later. To facilitate the rapid identification of a potential CD, a thoroughly worked-out CD target profile (CDTP) with criteria acceptable for the disease target product profile (TPP) is required at the start of lead generation (LG). In addition to driving the compound property optimization, the preclinical project team has to understand the ultimate goal to be able to rapidly identify and progress a potential CD. A screening cascade with meaningful and well-balanced progression criteria based on the CDTP is required to rapidly filter out unwanted compounds and to progress a potential CD through the cascade to candidate selection.

Introduction

After a highly successful period from the beginning of the 1980s until the end of the millennium, the pharmaceutical industry then faced a decline in productivity, where the cost and time to develop a new drug both increased substantially [1]. During the early 2000s, the clinical development program for a new chemical entity (NCE) typically took 9 years [2]. To address the issue of long cycle times in drug development, the US Food and Drug Administration (FDA) introduced a variety of programs in which new drug submissions have benefitted from expedited pathways, such as fast track, break-through, priority review and accelerated approval. The number of projects benefitting from one or several of these programs is increasing year on year [3,4]. To be able to take full advantage of these programs, the drug discovery phase in the pharmaceutical industry must keep up and supply high-quality CDs at increasing speed. For a successful drug discovery project, the time from selection of the target until delivery of a CD is ~5 years [5]. This is ~30–40% of the total time for drug development and shortening this phase will significantly contribute to the accelerated delivery of new medicines for the benefit of patients. Drug discovery has changed dramatically over the past few decades, with technological revolutions in several areas, including combinatorial chemistry, DNA sequencing, X-ray crystallography, and high-throughput screening (HTS) [2]. In numerous publications, authors argue that the technical revolution has come at a price; a lowering of scientific quality in favor of too much focus on quantity with HTS in simple in vitro systems and the use of different metrics to measure and further enhance productivity [6–8]. Even if this is correct, the solution is not to turn back time but rather to combine high scientific quality with an efficient drug discovery process. At AstraZeneca, this mindset has been exemplified by an improved design–make–test–analyze (DMTA) workflow with multidisciplinary design teams, enhanced throughput, and strict turn-around times in compound synthesis and in vitro screening [9,10]. With an effective DMTA cycle in place, the next step is to reduce the time between first synthesis and selection of the CD. Here we present lessons learnt from a retrospective analysis of AstraZeneca drug discovery projects, providing tools to facilitate the identification of the actual CD.

Retrospective analysis

With the purpose of finding ways to accelerate projects during the drug discovery phase, a retrospective analysis of drug projects within AstraZeneca that had recently delivered a CD for clinical development was performed. In the analysis, we noted when the compound that was eventually selected

1359-6446/
https://doi.org/10.1016/j.drudis.2019.03.026

as a CD was first synthesized and the time between first synthesis and CD selection. The lead times from first synthesis until the compound was selected as a CD were variable between projects (ranging from 9 to 36 months). We believe that it would be beneficial from a portfolio planning and resourcing point of view to reduce lead times, and are confident that this can be accomplished. A nonobvious finding was that, for several projects, the CD had already been synthesized in the lead generation (LG) phase before starting lead optimization (LO). Contraintuitively, these projects did not deliver a CD faster than other projects, and an in-depth analysis was carried out to understand the underlying reasons. A common denominator for these projects was that the compound that eventually became the CD was either not fully evaluated against the CD criteria or considered not to be of adequate quality against one or several key properties. For the latter case, project teams judged it likely that these insufficiencies would be improved during the LO phase and a CD would be delivered with overall better quality and lower potential risk for attrition. However, at the end of the LO phase, when shortlisted compounds were evaluated, the assessment came out in favor of the earlier synthesized compound. Interestingly, the liabilities were still there, but the overall risk assessment had changed when compared with other shortlist compounds. An example of a recent AstraZeneca project in which the CD was synthesized before the start of the LO phase is shown in Fig. 1, with timelines and events from when the compound was first synthesized to its selection as a CD.

An additional finding was that certain projects nominated the CD a very short time after it was first synthesized. An in-depth look at what distinguished these projects from the others helped us to identify factors promoting rapid delivery in drug discovery. Among these factors were well-established in vivo models and solid human target validation in addition to chemical starting points with drug-like properties, which is typically the case with back-up or fast follower approaches. Furthermore, we also noted that projects benefit from being highly resourced and from having the freedom to run parallel activities at risk.

Identifying the CD

To rank order compounds within and between chemical series and to drive compound design, a variety of ligand efficiency metrics (LEM) are widely used in the pharmaceutical industry [11–17]. The idea behind these metrics is to normalize target potency and affinity to physicochemical properties and/or molecular size, or composites thereof. These metrics are expected to be predictive of the pharmacokinetic (PK) and toxicological properties of compounds. As DMPK scientists, we are the first to acknowledge the importance of optimizing compounds not only on potency, but also on PK and physicochemical properties. This leads to improvements in compound quality and, thus, reduces the risk of attrition during clinical development. The literature that illustrates the importance of multiparameter optimization is overwhelming and several strategies describing how to incorporate different metrics into the drug design have been presented [17,18]. During the early discovery phase, these metrics are useful because they can be used to identify clusters and series with the highest potential for further optimization into CDs. Once the optimization phase of a chemical series has started, LEMs can be applied as simple measures to follow the evolution of the compound series and to guide the design of new molecules. However, LEMs do not provide any measure of the distance to the target properties for the CD or information about whether the properties of a compound are of CD quality. To identify the potential CD, more refined parameters and risk assessments are needed. A reliable human dose prediction is the most important of these parameters, in which knowledge about the PK and pharmacodynamic (PD) properties of the compound are incorporated. The link between toxicological risk and high dose has been described extensively in the literature [19–21] and a low human dose is also desirable from the manufacturing and cost of goods perspectives. The predicted human dose is a composite parameter of target potency and exposure profile of the compound. Human PK parameters can be assessed from relatively limited experimental data [22], such as in vitro intrinsic metabolic clearance and plasma protein binding, which allows the prediction of human clearance (CL) and in vivo rat PK experiments to estimate human volume of distribution (Vss). In combination with in vitro potency data, a human dose can be predicted for all compounds

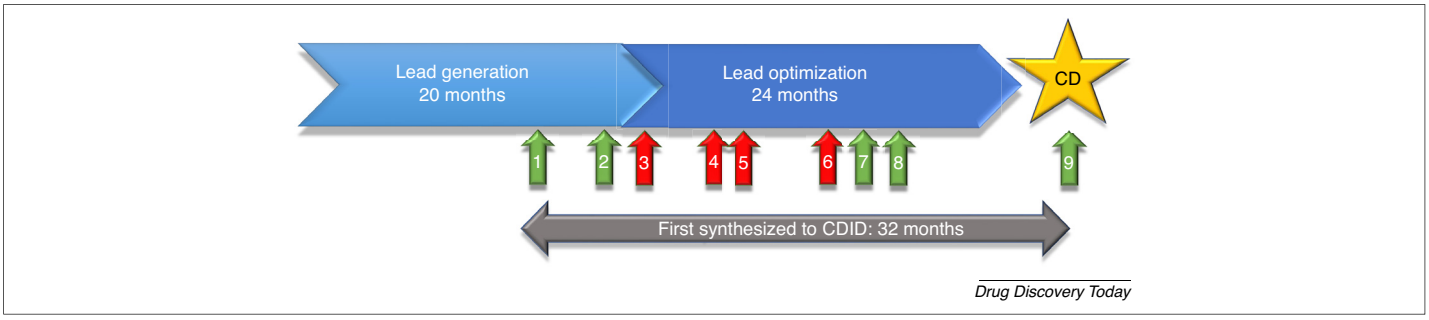

FIGURE 1

Timelines and events for a recent AstraZeneca project where the candidate drug (CD) had already been synthesized during the lead generation (LG) phase. Green arrows are associated with compound [A], eventually selected as the CD, and red arrows with an alternative compound [B], synthesized soon after the lead optimization investment decision (LOID). (1) Compound [A] first synthesized. (2) Demonstrated effect of [A] in a preclinical pharmacodynamic (PD) model, but considered to carry overall suboptimal properties with regards to target potency, drug–drug interaction potential as a substrate for CYP3A4-mediated metabolism and a too-short half-life. (3) Start of LO to address caveats identified for compound [A]. (4) Compound [B] first synthesized. (5) Initial profiling of compound [B] indicating a more favorable profile compared with [A]. (6) Compound [B] failed to deliver effect in a preclinical PD model and, therefore, was stopped. (7) Preliminary reassessment of CD potential of [A]. Previously identified caveats now considered to be compliant with the TPP. Despite not being one of the most potent compounds, the predicted human dose was in a reasonable range and a label as a 'moderate sensitive CYP3A4 substrate' was considered compliant with the target product profile (TPP). It was decided to evaluate the drug–drug interaction potential in healthy volunteers. The short half-life could be addressed be developing a modified-release formulation. (8) CD profiling of [A] initiated. (9) Confirmed CD potential of [A] and selected as CD.

made during drug discovery [23]. The predicted human dose should be in focus throughout the drug discovery process to aid understanding of the overall quality of a compound series and to identify a potential clinical candidate as soon as data become available. The confidence in the human dose prediction will increase as more assay data are generated and by an evolving PKPD understanding during project progression.

Candidate drug target profile

The major breakthroughs in compound optimization often occur during the early drug discovery phase and we have seen that it is not uncommon that the CD is already synthesized in the LG phase. Line-of-sight is needed in early discovery for the rapid identification of the candidate drugs and we stress that a CDTP should be already available at the start of LG. Previously, the CDTP was not established until the start of LO and reflected CD criteria for the ideal compound with little concern about how to progress into clinical development. With the bar placed so high, the preclinical team would concentrate on meeting the CDTP criteria and to deliver the perfect molecule, given that it is in a scientist's nature to continue looking for opportunities for improvement. There will always be new compounds to synthesize that might be superior to current lead compounds, but there is an obvious risk that compound optimization, in the worst case, will continue for several years. At some point, a shortlist of the best compounds is assembled and the project team rank-orders these compounds. The top-ranked compound is then nominated for CD selection if considered to be of sufficient quality for clinical development. Big pharma has a reputation of being risk averse, often for good reasons, which could partly explain the approach described above aiming for the 'perfect' compound.

We challenge this way of working because it carries a substantial risk for unnecessarily delaying the delivery of the CD. In our opinion, the CDTP should reflect 'acceptable' criteria for the target and rely on the TPP [24] for the actual disease, including differentiation versus standard of care, patient endpoints for efficacy and safety, as well as a strategy for determining the optimal route of administration and dosing regimen. Examples of ideal criteria are a daily dose $<$50 mg because a high dose is associated with higher risk for adverse events [20] and an ideal compound does not carry any predicted risk of drug–drug interactions (DDI) or alerts for reactive metabolite formation. A higher predicted human dose and a label for DDI might be acceptable when the overall risk–benefit profile for the TPP is considered. Reactive metabolite formation is an alert for potential drug-induced liver injury (DILI), but needs to be contextualized by an overall risk assessment of DILI, including additional in vitro and in vivo assays as well as the predicted daily dose. When the CDTP has been established, it should normally not be changed. Exceptions to this could be in a rare situation where the TPP is updated or if the project change focus to a TPP for an alternative disease. Whether the aim is to develop a first-in-class or best-in-class compound should be considered (Fig. 2). As an example, for a drug project working on a novel target aiming for a first-in-class compound for treatment of a disease with a high unmet medical need, the focus must be to bring a compound into clinical development as quickly as possible. The compound must have an acceptable risk–benefit profile, but the PK profile does not necessarily need to be perfect. For a compound with a short half-life, it might be acceptable to administer the compound twice or even three times daily. For a best-in-class compound, the CDTP will typically carry criteria that are close to ideal. If there is a competitor compound within the same class already on the market or in late-stage development, the aim must be to be to deliver a compound of overall better quality, which should be reflected in the CDTP.

Building the CDTP with acceptable criteria can be cumbersome and requires engagement from all disciplines earlier in the process than is currently common practice. If only CD criteria for the ideal case are established, there is an obvious risk for excessive optimization of compound properties that already satisfy the TPP. A well-defined and high-quality CDTP with relevant endpoints is key to run a drug discovery program efficiently and to reduce the time between first synthesis and nomination of the CD. It is important that project teams in drug discovery focus their efforts and resources on the most relevant issues to address.

An effective screening cascade

Pharmaceutical companies have invested heavily developing high-throughput assays for the screening of DMPK and safety properties and to streamline compound synthesis and data flow. Therefore, the scope for even higher

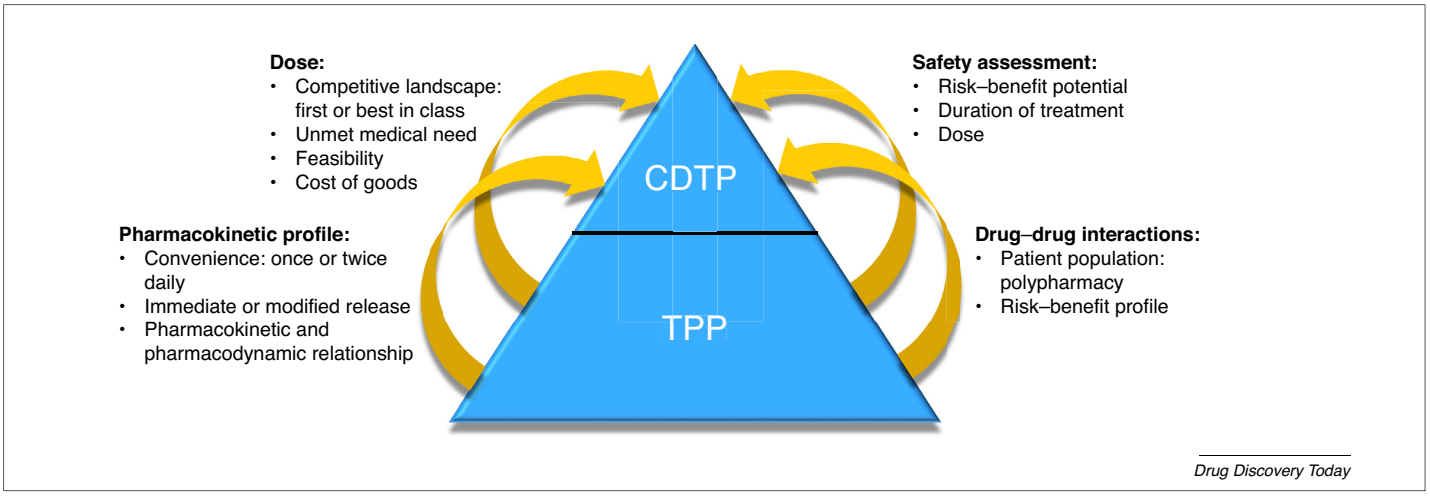

FIGURE 2

The disease area the target product profile (TPP) is the foundation for the candidate drug target profile (CDTP). Illustrated are aspects that should be considered for a few example CDTP criteria.

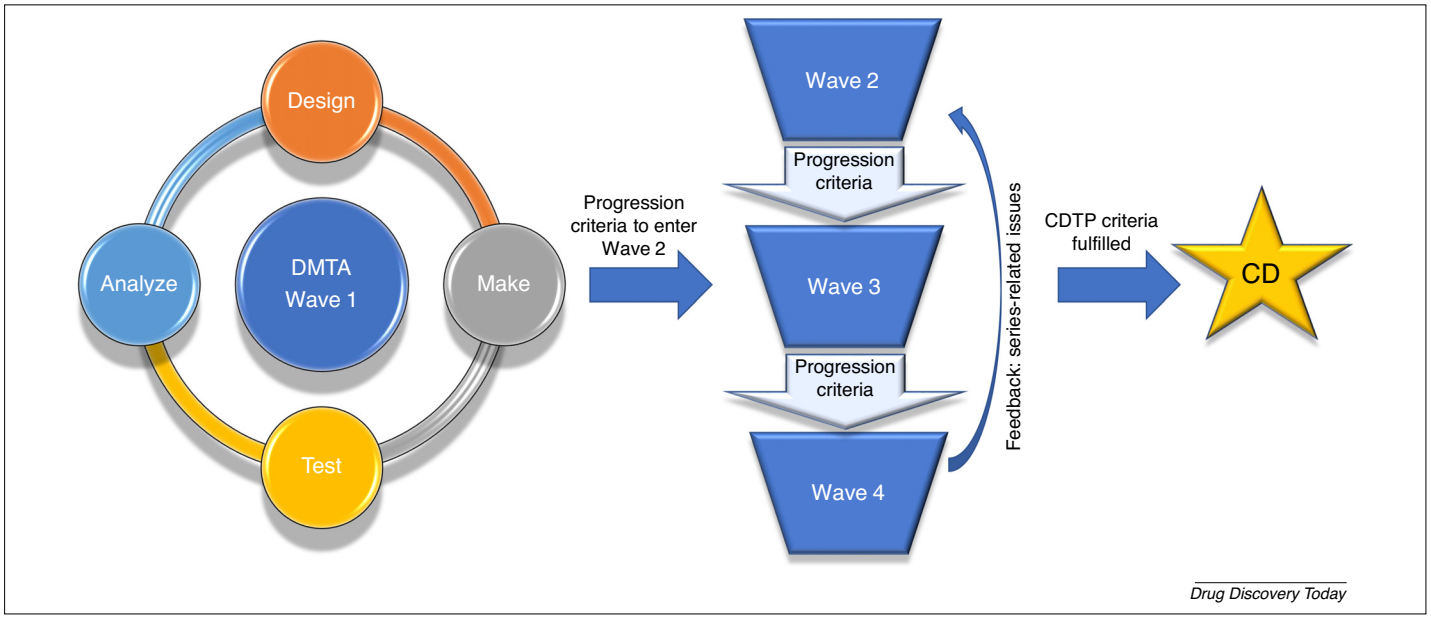

FIGURE 3

A schematic view of the drug discovery process illustrating how the design–make–test–analyze (DMTA) cycle and the screening cascade leads to the candidate drug (CD) nomination. In the DMTA cycle, compound design is driven by wave 1 screening data, including assays for target potency, off-target potency, metabolic clearance, plasma protein binding, solubility, and lipophilicity. A set of progression criteria based on wave 1 data, such as an early human dose prediction and selectivity profile, are defined and a compound fulfilling these criteria will enter wave 2 in the screening cascade and progress through the cascade until it fails to meet any criteria to enter the subsequent screening wave. The screening cascade should have a feed-back mechanism to ensure that, if compounds in later waves fall on the same criteria, the screening cascade can be adjusted by moving assays up in the cascade to identify and filter out compounds with these liabilities earlier. A compound that passes through all the waves in the cascade should fulfil the CDTP and be ready for CD nomination. The number of screening waves can be varied according to the priority of the project, with fewer waves running more assays in parallel for high-profile projects.

throughput in compound screening is limited. However, setting up the screening cascade in an optimal way for the rapid profiling and identification of compounds with the right properties is an area with potential for further improvement. Typically, a screening cascade works as a funnel where all compounds made are profiled in a first wave of assays, while fewer compounds are profiled in assays connected to later waves. This is a well-established approach in the pharmaceutical industry. A screening cascade should provide effective filtering of compounds to identify the best compounds and to filter out compounds with less-favorable properties that will never become a CD. The desired compound properties should be defined in the CDTP and, using that as a basis, the screening cascade should be set up with high-throughput assays that provide relevant information to enable the deselection of unwanted compounds in the first wave(s). The use of prudent predefined progression criteria enables compounds with the right properties to be submitted to the next wave of assays in a timely and efficient manner (Fig. 3). A screening cascade built with progression criteria of relevant composite parameters, such as human dose prediction and safety margin, is a necessity for the future as we consider the ever-increasing desire to reduce cost and time in all aspects of drug discovery.

It is important to adjust the screening cascade during the evolution of a compound series. When new series-related issues are discovered at a later stage, the screening cascade should be modified, and the related assay(s) should be moved to an earlier wave in the cascade to ensure that unwanted compound properties are filtered out as early as possible (Fig. 3).

There are ongoing initiatives within the pharmaceutical industry to integrate the synthesis and testing of new compounds and to even incorporate artificial intelligence (AI) in drug design [25–27]. If these approaches are successful, they should cut LG and LO timelines substantially [27–30]. At present, automated systems include one or a few assays, whereas in the future, more extensive screening could be incorporated. With all this in place, the demands for both a well-designed and efficient screening cascade and clear decision points in terms of progression criteria will be even higher.

Concluding remarks

As the outcome of a retrospective analysis of AstraZeneca drug discovery projects, we have identified opportunities to reduce timelines between LG and CD identification. We emphasize the need, at the start of LG phase, to have a clear line of sight to clinical development that should manifest in a CDTP based on the disease area TPP. The CDTP should be used to establish a screening cascade with meticulously selected progression criteria for each wave in the screening cascade. A compound that passes the criteria in an initial wave should be rapidly progressed to the next wave. With these improvements in both quality and efficiency in discovery, we anticipate more rapid identification of CDs and accelerated drug discovery.

Acknowledgments

We would like to thank Martin Hayes for reviewing and commenting on the manuscript.

References

1 Munos, B. (2009) Lessons from 60 years of pharmaceutical innovation. *Nat. Rev. Drug Discov.* 8, 959
2 Scannell, J.W. *et al.* (2012) Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* 11, 191
3 Hwang, T.J. *et al.* (2017) The FDA's expedited programs and clinical development times for novel therapeutics, 2012-2016. *JAMA* 318, 2137–2138
4 Kesselheim, A.S. *et al.* (2015) Trends in utilization of FDA expedited drug development and approval programs, 1987–2014: cohort study. *BMJ* 351, h4633

5 Raghavendra, M.S. *et al.* (2012) A study of decrease in R&D spending in the pharmaceutical industry during post-recession. *Int. J. Acad. Res. Part B* 4, 29–47
6 Cook, D. *et al.* (2014) Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nat. Rev. Drug Discov.* 13, 419
7 Elebring, T. *et al.* (2012) What is the most important approach in current drug discovery: doing the right things or doing things right? *Drug Discov. Today* 17, 1166–1169
8 Mignani, S. *et al.* (2016) Why and how have drug discovery strategies in pharma changed? What are the new mindsets?. *Drug Discov. Today* 21, 239–249
9 Plowright, A.T. *et al.* (2012) Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle. *Drug Discov. Today* 17, 56–62
10 Andersson, S. *et al.* (2009) Making medicinal chemistry more effective: application of Lean Sigma to improve processes, speed and quality. *Drug Discov. Today* 14, 598–604
11 Segall, M.D. (2012) Multi-parameter optimization: identifying high quality compounds with a balance of properties. *Curr. Pharm. Des.* 18, 1292–1310
12 Abad-Zapatero, C. and Metz, J.T. (2005) Ligand efficiency indices as guideposts for drug discovery. *Drug Discov. Today* 10, 464–469
13 Hopkins, A.L. *et al.* (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* 9, 430–431
14 Leeson, P.D. and Springthorpe, B. (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 6, 881
15 Mortenson, P.N. and Murray, C.W. (2011) Assessing the lipophilicity of fragments and early hits. *J. Comput. Aided Mol. Des.* 25, 663–667
16 Reynolds, C.H. *et al.* (2008) Ligand binding efficiency: trends, physical basis, and implications. *J. Med. Chem.* 51, 2432–2438
17 Mignani, S. *et al.* (2016) Compound high-quality criteria: a new vision to guide the development of drugs, current situation. *Drug Discov. Today* 21, 573–584
18 Leeson, P.D. and Young, R.J. (2015) Molecular property design: does everyone get it? *ACS Med. Chem. Lett.* 6, 722–725
19 Wager, T.T. *et al.* (2013) Improving the odds of success in drug discovery: choosing the best compounds for in vivo toxicology studies. *J. Med. Chem.* 56, 9771–9779
20 Stepan, A.F. *et al.* (2011) Structural alert/reactive metabolite concept as applied in medicinal chemistry to mitigate the risk of idiosyncratic drug toxicity: a perspective based on the critical examination of trends in the top 200 drugs marketed in the United States. *Chem. Res. Toxicol.* 24, 1345–1410
21 Thompson, R.A. *et al.* (2012) In vitro approach to assess the potential for risk of idiosyncratic adverse reactions caused by candidate drugs. *Chem. Res. Toxicol.* 25, 1616–1632
22 Grime, K.H. *et al.* (2013) Application of in silico, in vitro and preclinical pharmacokinetic data for the effective and efficient prediction of human pharmacokinetics. *Mol. Pharm.* 10, 1191–1206
23 Page, K.M. (2016) Validation of early human dose prediction: a key metric for compound progression in drug discovery. *Mol. Pharm.* 13, 609–620
24 Curry, S. and Brown, R. (2003) The Target Product Profile as a planning tool in drug discovery research. *Pharmatech* 2003, 67–71
25 Blaschke, T. *et al.* (2018) Application of generative autoencoder in de novo molecular design. *Mol. Informat.* 37, 1700123
26 Olivecrona, M. *et al.* (2017) Molecular de-novo design through deep reinforcement learning. *J. Cheminformat.* 9, 48
27 Segler, M.H.S. *et al.* (2018) Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Central Sci.* 4, 120–131
28 Baranczak, A. *et al.* (2017) Integrated platform for expedited synthesis-purification-testing of small molecule libraries. *ACS Med. Chem. Lett.* 8, 461–465
29 Desai, B. *et al.* (2013) Rapid discovery of a novel series of abl kinase inhibitors by application of an integrated microfluidic synthesis and screening platform. *J. Med. Chem.* 56, 3033–3047
30 Schneider, G. (2017) Automating drug discovery. *Nat. Rev. Drug Discov.* 17, 97

Fredrik Bergström*
Bo Lindmark

Drug Metabolism and Pharmacokinetics, Cardiovascular, Renal and Metabolism, IMED Biotech Unit, AstraZeneca, Gothenburg, Sweden

*Corresponding author.