



editorial



Celerino Abad-Zapatero

Are SAR tables obsolete?

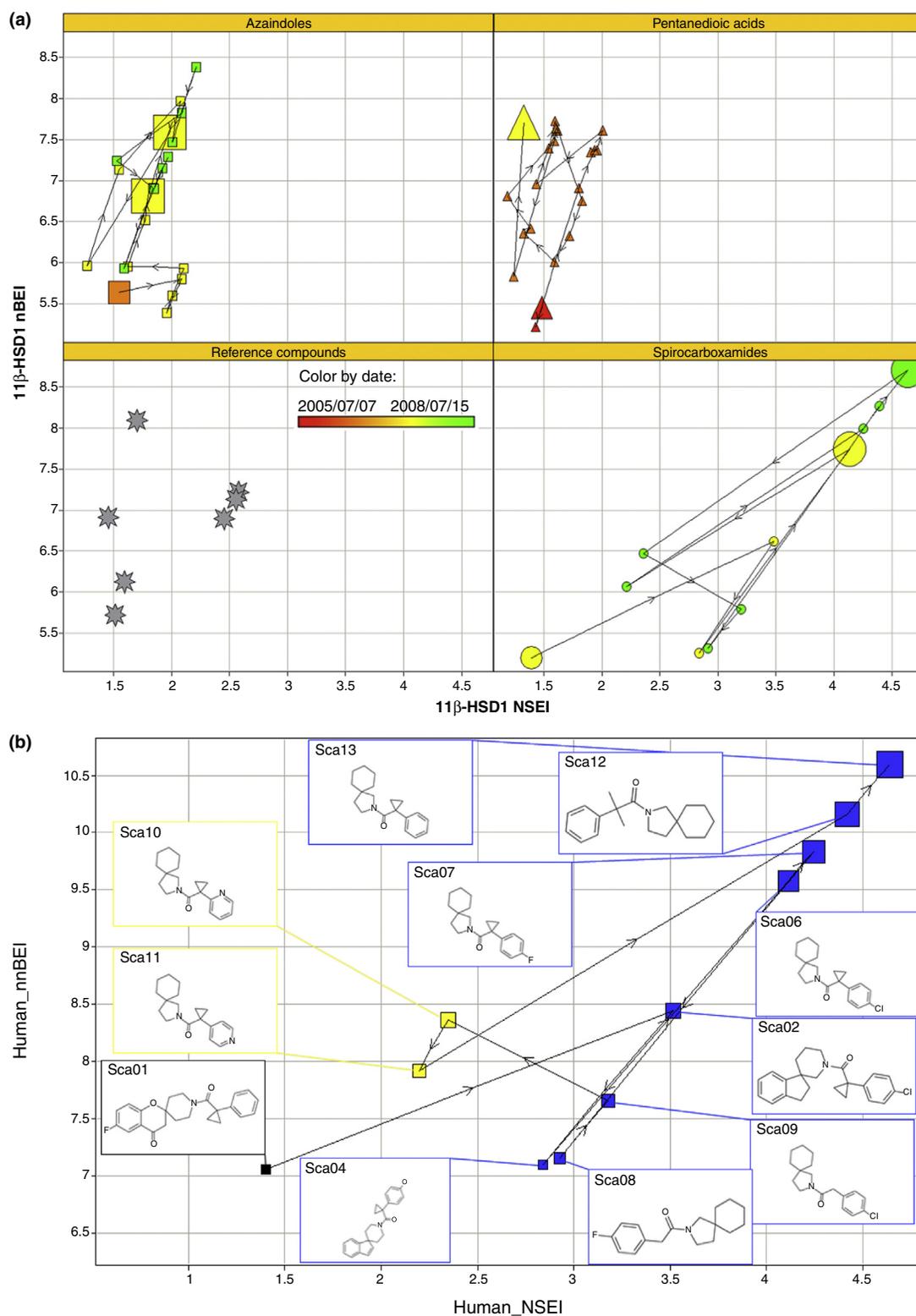
Perhaps it is because I am not a professional medicinal chemist, or possibly it is because I am a person that prefers the visual presentation of data as opposed to a long list of numbers in parallel vertical columns; either way, the fact is that the conventional structure–activity relation tables (SAR-tables), which have been and still are the essential item in most medicinal chemistry papers, are difficult for me to assimilate and might not be the best way to follow the path of a drug discovery project.

I can understand that nearly 100 years ago (Paul Ehrlich's time at the beginning of the 20th century), when the main and possibly the only parameter to guide drug discovery was a measure of the activity in a biological assay, SAR tables were an effective way of condensing the two main variables: activity and full chemical structures. Those were simpler times when there was a weaker understanding of the role that the physicochemical parameters had in making a compound a successful drug. The driving force was the *in vivo* activity because it related to the limited synthetic compounds available for testing.

Nowadays, there is enormous amount of medicinal chemistry effort that goes into producing different chemical cores and optimization of various series, without even mentioning the more extensive efforts involved in parallel synthesis of a relatively large

number of related compounds to a common scaffold. In addition, what we have learned since the publication of Lipinski's 'Rule of 5' (Ro5) guidelines and other related *ad hoc* rules, is that many of those compounds listed in myriad SAR tables do not have a chance of passing several basic filters on their way to even being selected for further development. I challenge even the savviest of medicinal chemists to examine the results of any conventional SAR table for a project to follow the trends in the affinity (typically K_i or IC_{50}), and simultaneously consider mentally the size and polarity of the compounds that are being listed. Yet, we do know now that those properties are crucial parameters in the selection and optimization of any compound series. Why do we not consider the possibility of condensing all that information in a more convenient 2D diagram, one that is easy to examine and follow and that will make grasping all the essential information essentially instantaneous. Would this not be better? Is this possible? Are these tables not obsolete?

During the first few years of the 21st century, beginning with the original paper by Hopkins and colleagues in 2004 defining 'ligand efficiency' ($LE = \Delta G/NHA$), a succession of 'combined variables' have been introduced that condense in various ways three crucial variables for drug discovery into a more manageable set. They are normally referred to as 'Ligand Efficiency Indices' (LEIs) (see, for instance, Tables 2.1 and 2.2 in [1]). Their merit is that they merge affinity, size, and polarity (most commonly) into a set of two variables that can be easily represented in Cartesian planes that had been named 'efficiency planes' [2]. There have been various formulations following the initial Hopkins definition and their use is now sprinkled throughout the literature, mostly using only one efficiency index at the time, either size related or polarity related. For a while, there was controversy as to the mathematical rigor of their formulation, which I do not discuss here. I believe that these 'combined variables' are here to stay in way or another. Their use to drive and optimize drug discovery rigorously in a robust multiparameter optimization methodology is still open to question. However, I would argue that they could have genuine value in helping us to visualize the drug discovery process in terms of mapping the discovery and optimization process in a convenient, easy-to-read, Cartesian plane. A basic background about the concepts should help the reader to easily



Drug Discovery Today

FIGURE 1

Overall trajectories of the Merck-Serono compounds for 11β-hydroxysteroid-dehydrogenase type I (11β-HSD1) described in [3,4] as illustrated in [5]. **(a)** Clockwise from upper left: azaindoles, pentanedioic acids, spirocarboxamides, and reference compounds. The ranges on the x- and y-axes are the same, showing clearly that the spirocarboxamide series is superior to the other three (including the reference compounds) in the polar efficiency direction (horizontal). Size of the markers is proportional to potency and color relates to the date as indicated (2005/05/07 red to 2008/07/15 green). The line connecting the starting and the end compounds marks the trajectory of the drug discovery process for each series and would be the vectorial addition of the different steps. **(b)** Chemically annotated 'efficiency plane' of the spirocarboxamide series of compounds from the lower right panel in (a). The numbers refer to the actual sequence of chemical compounds and can be followed by the arrowheads. In this plane, it is easier to follow the movements of the compounds on the plane by just counting the N + O atoms (slope of the

understand the general idea and then I will illustrate the concept with one example.

The affinity of a small ligand for a certain macromolecular target is a natural link between the two domains, the chemical and the biological, and there are several ways of combining these. I have suggested previously that an effective way to summarize and map the optimization of affinity, size, and polarity is to use pairs of LEIs, typically 'efficiency per polarity' on the x -axis [e.g., surface EI (SEI) and number SEI (pK_i /NPOL)] and 'efficiency per size' [e.g., binding EI (BEI) and number BEI (pK_i /NHA)] on the y -axis. Taking the first pair [SEI = pK_i /(PSA/100) and BEI = pK_i /(MW/1000) (where PSA is polar surface area and MW is molecular weight)], these simple definitions connect the biological affinity ($pK_i = -\log K_i$) in the numerator with a physicochemical parameter of the ligand in the denominator (PSA and MW, respectively). The scaling numbers 100 and 1000 are chosen for convenience, to put them on approximately the same scale. From the definition, it is clear that BEI/SEI [i.e., the slope of the lines in the SEI, BEI (x,y) plane] depends only on the two physicochemical properties and is equal to 10 PSA/MW. Thus, any compound analyzed can be placed along a line just based on MW and PSA; where along the line would depend on the pK_i value, that is, the affinity. The polarity of the compounds increases counterclockwise on the plane as the PSA:MW ratio increases, allowing a simple 'decoupling' of the chemistry and biology.

Similar considerations can be used to understand the basic properties of the NSEI [$=pK_i$ /number of polar atoms (NPOL)], NBEI [$=pK_i$ /number of non-hydrogen atoms (NHA)] (x,y) planes. In the case of NSEI, NBEI (x,y) planes, the slope of the lines increases counterclockwise as the ratio NPOL:NHA. These simple concepts permit mapping the drug discovery effort, whether it is optimization, series comparisons, or any other endeavor, in simple and easy-to-grasp 2D plots that could be called 'efficiency planes'. The corresponding values of the variables can be read from the axis and, using conventional software, it is now possible to hover the points on the plane and follow the chemical changes that have resulted on the various paths through the plane. An 'efficiency plane' that is particularly appealing and easy to interpret and follow is that defined by NSEI, nBEI {nBEI = $-\log[(K_i/NHA)]$ } (x,y). In this plane, the slope of the lines is given by NPOL (i.e., N + O count) and increases counterclockwise. Is this so complicated? Could we not just do away with SAR tables?

There are several examples of this concept in the original publication describing the use of LEIs to map chemico-biological space (CBS) [2]. More examples can be prepared using the AtlasCBS server (www.ebi.ac.uk/chembl/atlasCBS/intro.jsp). In these plots, the optimum candidates can be selected by inspection of the compounds lying in the most northeastern corner of the diagram. Alternatively, a single efficiency index, namely LELP = $c\text{LogP}/LE$, is becoming more widely used as a single parameter to monitor the optimization process, possibly because it combines polarity, size, and affinity. One single variable could also be used but I prefer the

2D separation of size (y -axis) and polarity (x -axis). In addition, thinking in polar coordinates, one can decouple the chemistry (with polarity on the angular variable) and affinity (biology) on the radial coordinate, as suggested above.

I wish to emphasize that this is not, by any means, the only way to present the three critical variables of drug discovery (affinity, size, and polarity) in a Cartesian plane to illustrate and guide the drug discovery process. I am certain that other ways to present the same data relating affinity to size and polarity can be found that can help to present the limited data presented in the SAR tables in a more effective way. The existence of several other combined variables related to efficiency indices (LLE, LELP, FQ, SILE, EE, SIHE, among others [1]) should inspire researchers to find other ways to map the drug discovery process effectively in 2D or 3D to make it easier to follow.

What would the community gain by replacing conventional SAR tables with a pictorial representation of the data? A picture is worth a thousand words! Finding graphical ways to summarize the data (chemical and biological) contained in conventional SAR tables would enable the progress of medicinal chemistry and drug discovery efforts to be monitored more effectively, making them easy to follow even for the nonprofessional medicinal chemist, facilitating the communication of results and strategies among the project team members and management to optimize approaches and paths making drug discovery more effective. Comparing trajectories, time lines, and even in the future, device optimal trajectories could help expedite the drug discovery pipeline.

In Fig. 1, I provide a simple but complete example of the idea that the data contained in SAR tables can be presented more comprehensively and effectively using graphical means rather than by the listing of detailed numbers in several columns in a table form. Fig. 1 illustrates the data relating to the structure-guided drug design of several series of inhibitors of 11 β -hydroxysteroid-dehydrogenase type I published by Lepifre *et al.* [3,4]. The four panels in Fig. 1 (clockwise from top left) illustrate the overall trajectories of three series of compounds (azaindoles, pentanedioic acids, and spirocarboxamides) plus the reference compounds. The authors originally quoted in their published SAR tables the LEIs related to size. Adding the efficiency related to polarity permitted the complete mapping in efficiency planes [5] for a full comparison in terms of compound quality, series progression, and determination of the optimum series the Spirocarboxamides.

I am certain that other ways to present the same data relating affinity to size and polarity (or even additional variables) can be found that can help to present the limited data presented in the SAR tables in a more effective way. Select the one that fits you and your team best, but please provide a visual summary of your drug discovery efforts in chemico-biological space in a concise and graphical manner. Detailed SAR tables could well be part of the supplementary material for those researchers looking for the finer details of the project.

lines and color of squares) in the molecules: increasing to the left and decreasing to the right. The general direction of optimization is north-east and the optimum compounds in terms of affinity, size, and polarity are at the farthest upper right-hand corner of the diagram. The size of the squares is proportional to affinity. Source: Adapted and reproduced, with permission, from Ref. [5] (a, b). Figures prepared with Spotfire™.

References

- 1 Abad-Zapatero, C. (2013) *Ligand efficiency indices for drug discovery: towards an atlas-guided paradigm*. Academic Press
- 2 Abad-Zapatero, C. *et al.* (2010) Ligand efficiency indices for an effective mapping of chemico-biological space: the concept of an atlas-like representation. *Drug Discov. Today* 15, 804–811
- 3 Roche, D. *et al.* (2009) Discovery and structure-guided relationships of pentanedioic acid diamides as potent inhibitors of 11beta-hydroxysteroid dehydrogenase type I. *Bioorg. Med. Chem. Lett.* 19, 2674–2678
- 4 Lepifre, F. *et al.* (2009) Discovery and structure-guided drug design of inhibitors of 11beta-hydroxysteroid dehydrogenase type I based on spiro-carboxamide scaffold. *Bioorg. Med. Chem. Lett.* 219, 3682–3685
- 5 Christmann-Franck, S. *et al.* (2011) Time trajectories in efficiency maps as effective guides for drug discovery efforts. *Mol. Inf.* 30, 137–144

Celerino Abad-Zapatero

Center for Biomolecular Sciences, Chicago, IL, USA

email: caz@uic.edu