Mol Divers (2016) 20:771–780
DOI 10.1007/s11030-016-9659-x

SHORT COMMUNICATION

Activity landscape analysis of novel 5α-reductase inhibitors

J. Jesús Naveja[1,2] · Francisco Cortés-Benítez[1] · Eugene Bratoeff[1] · José L. Medina-Franco[1]

Received: 31 August 2015 / Accepted: 12 January 2016 / Published online: 1 February 2016

Abstract Inhibitors of the enzyme 5α-reductase (5aR) are promising therapeutic agents for the treatment of benign prostatic hyperplasia (BPH) and prostate cancer. The lack of structural data of the enzyme 5aR prompts the application of ligand-based approaches to systematically explore the activity landscape of 5aR inhibitors. As part of an effort to develop inhibitors of this enzyme for the treatment of BPH, herein we discuss a chemoinformatic-based analysis of the activity landscape of a novel set of 53 novel pregnane and androstene compounds. It was found that, in general, for each pair of compounds in the set, as the structure similarity of the compounds increases the corresponding potency difference decreases. These results are in agreement with an overall smooth activity landscape. However, two potent activity cliff generators were identified pointing to specific small structural changes that have a large impact on the inhibition of 5aR.

Keywords Activity cliff generators · Benign prostatic hyperplasia · Chemical space · Chemoinformatics · Prostatic 5α-reductase · Structure–activity relationships

Abbreviations

5aR	5α-Reductase
ALM	Activity landscape modeling
AR	Androgen receptor
BPH	Benign prostatic hyperplasia
DHT	5α-Dihydrotestosterone
ECFPs	Extended connectivity fingerprints
FIDE	Finasteride
PCa	Prostate cancer
PCA	Principal component analysis
SARs	Structure–activity relationships
SAS	Structure–activity similarity
T	Testosterone
Tc	Tanimoto coefficient

Introduction

Approximately, 3 % of testosterone (T) is free in serum and biological activity making it in the most abundant androgen [1]. In prostatic stromal and basal cells, T is converted to 5α-dihydrotestosterone (DHT) by the 5α-reductase (5aR) enzyme through the irreversible reduction of the Δ^4 bond [2]. DHT is the preferred ligand for the androgen receptor (AR) [3]. Once DHT–AR has been formed, in its activated form, AR undergoes dimerization, phosphorylation, and translocation to the nucleus. Then, it activates the transcription of certain genes that include transcription co-regulators and transcriptional machinery which triggers the synthesis of spe-

Electronic supplementary material The online version of this article (doi:10.1007/s11030-016-9659-x) contains supplementary material, which is available to authorized users.

J. Jesús Naveja and Francisco Cortés-Benítez have contributed equally to the work.

Eugene Bratoeff: deceased.

✉ José L. Medina-Franco
jose.medina.franco@gmail.com; medinajl@unam.mx

[1] Departamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, Avenida Universidad 3000, 04510 Mexico, DF, Mexico

[2] Facultad de Medicina, PECEM, Universidad Nacional Autónoma de México, Avenida Universidad 3000, 04510 Mexico, DF, Mexico

cific proteins and also cell proliferation [4]. Elevated levels of DHT are related to benign prostatic hyperplasia (BPH) and prostate cancer (PCa) which are two of the most common diseases in men [5,6].

Due to the crucial role of 5aR in the progress of the prostate pathogenesis, various prostatic 5aR inhibitors have been developed in an effort to block the T conversion and inhibit the effects of DHT in prostate disease. Currently, there is a wide range of azasteroids reported such as finasteride (FIDE), which is a drug broadly used for BPH treatment [7]. However, side effects of FIDE [8,9] have encouraged the development of new steroidal compounds with improved selectivity toward 5aR.

Over the past several years, the research group of Bratoeff has reported a large set of pregnane and androstane molecules as inhibitors of 5aR. Most of these steroidal compounds have shown potent inhibitory effect in prostatic 5aR enzyme. Moreover, several of these compounds had been able to reduce the weight of prostate in gonadectomized hamster more than FIDE in the in vivo experiments [10–24]. However, limited studies have been reported for the computational analysis of the structure–activity relationships (SARs) of pregnane and androstane molecules. This is in part due to the lack of experimental three-dimensional structures of the 5aR enzyme. Recently, a self-organizing molecular field analysis (SOMFA) [25], a 3D-QSAR technique, was employed to explore the SAR of a set of pregnane and androstane derivatives [26] and different azasteroids [27]. In those studies, the master grid for various SOMFA models indicated that bulky groups around C-3, C-6, and C-17 of the steroidal skeleton suggest favorable interactions, whereas electronegative groups around C-3, C-6, and C-17 as well as electropositive groups at C-4 are responsible for the observed variations in the 5aR inhibition. However, drawbacks of 3D-QSAR methods such as molecular alignment and selection of bioactive conformations are well known, making this technique challenging to implement [28]. Therefore, it is convenient to apply additional computational strategies using a fast and robust method to further advance the understanding of the SAR and structure–property relationships of inhibitors of 5aR.

Herein, we report a comprehensive SAR study of 54 5aR inhibitors (53 novel compounds plus FIDE) using the concept of activity landscape modeling (ALM) [29,30]. ALM has emerged as an approach to rapidly navigate through the SAR of datasets [29,31]. Of note, a number of ALM methodologies are ligand-based only, so no data on the three-dimensional structure of the receptor or prior knowledge of the putative binding site are required. Indeed, ligand-based methods are particularly useful when no detailed structural information of the putative ligand–target interactions is known or when there is a high uncertainty in the bioactive conformations [32]. ALM emphasizes on the identification of compounds with very similar chemical structures but different potency differences, i.e., activity cliffs [33]. Of note, ALM is part of the broader concept 'property landscape modeling' used in chemistry [34]. Here, 'activity' (typically referring to biological activity) is a specific case of 'property.' In this context, ALM does not necessarily assume that all compounds have exactly the same mechanism of action such as other typical quantitative SAR methods such as QSAR. Indeed, ALM has been proved to be useful to gain insights from datasets where the biological activity has been tested in cell-based assays (where the precise mechanism of action is actually unknown) [35]. This is because, in contrast to traditional QSAR analysis such as 3D-QSAR, ALM does not assume continuous SARs [36,37].

Methods

Dataset

The chemical structures and biological activity of 53 in-house pregnane and androstene molecules were retrieved from the literature. The activity data for all compounds were obtained in vitro in enzymatic inhibition assays using the same experimental conditions. The experimental IC_{50} values against the prostatic 5aR were converted to pIC_{50} ($-\log IC_{50}$) values. The pIC_{50} values ranged from 5.00 to 10.60. The chemical structure of FIDE was retrieved from the ZINC database [38].

Chemical space

A visual representation of the chemical space [39] was generated by conducting a principal components analysis (PCA) of the similarity matrix of the dataset with 54 molecules. This is a well-established method to generate visual representation of chemical spaces [40]. The similarity matrix was calculated using the extended connectivity fingerprints (ECFPs) available in MayaChemTools (http://www.mayachemtools.org) using a neighborhood radius of two [41]. ECFPs have been used to perform activity landscape analysis with interpretable results [32,42]. The structural similarity was computed with the Tanimoto coefficient (Tc) [43]. The PCA analysis was performed with the FactoMineR R package version 1.29. PCA of the similarity matrix has been used as a strategy to visualize the chemical space of several datasets [43,44].

Activity landscape

Several methods have been developed for ALM [29]. In this work, the activity landscape of the 5aR inhibitors was studied by means of the structure–activity similarity (SAS) maps [45], which have been extensively employed to characterize the SAR of datasets with activity data for one or more biological endpoints [34,35,46,47]. A detailed description of

the construction of SAS maps is provided elsewhere [48]. In summary, the relationship between structural similarity and potency difference (or potency similarity) can be represented in a two-dimensional graph called a SAS map. Usually, the structural similarity (that can be obtained with any single or combination of similarity approaches) is plotted on the x axis and the potency difference is plotted on the y axis [49]. SAS maps can be roughly divided in four zones. The most significant one explored in this work was the upper-right region of the plot that identifies pairs of molecules with high structure similarity but large potency difference. Therefore, this zone identified activity cliffs and was considered the 'activity cliff region' of the SAS maps. In order to define the four zones of the SAS maps, we used two thresholds following established criteria [48]: one standard deviation above the mean structure similarity (ECFPs/Tanimoto = 0.43) along the x-axis, and one standard deviation above the activity difference ($\Delta pIC_{50} = 2.75$) along the y-axis. Since the visual interpretation of the SAS maps can be difficult if many data points are present, we implemented the recently developed density SAS maps [32]. In a density SAS map, the frequency of the data points is represented with a continuous color scale as detailed elsewhere [32].

Activity cliff generators

An activity cliff generator is a molecule frequently found in activity cliffs [50]. In medicinal chemistry, the most attractive activity cliff generators are compounds with high biological activity. Identification of activity cliff generators and compounds with similar chemical structure but large potency difference is particular relevant because it points to specific structural data with significant impact on the biological activity. Herein, the activity cliff generators were identified as molecules with high frequency in the 'activity cliff region' of the SAS maps.

Results and discussion

Since the activity landscape of a dataset is the association between the chemical space with the structure representation [42], the "Results and discussion" section of the paper are further organized in two parts: (1) exploration of the chemical space of the dataset followed by (2) activity landscape analysis. Analysis of the chemical space provides information of the chemical diversity of the dataset, while the second part discusses the pairwise structural relationships between the chemical structures with the biological activity.

Chemical space

Figure 1 shows a visual representation of the chemical space of the 53 pregnane and androstene molecules and FIDE. In this plot, each data point represents one compound. Data points are colored by the corresponding pIC_{50} value using a continuous scale indicated in the figure. The relative position of FIDE and two representative compounds (activity cliff generators EB-13 and EB-37, discussed below) is shown. The visual representation of chemical space in Fig. 1 is an approach to rapidly classify the compounds in the dataset. Such visual classification provides a first and rapid idea of the different groups (or clusters) of compounds that can be found in the set of 53 pregnane and androstene molecules. Furthermore, adding activity data to the plot enables the identification of regions in chemical space with, for example, smooth SAR or groups of compounds enriched with activity. The distribution in chemical space distinguishes several groups, both by visual inspection and k-means analysis (data not shown). The most significant activity cliffs generators, i.e., EB-13 and EB-37 (cf. chemical structures in Figs. 5, 6, respectively), are included in two different groups. Table 1 contains the main clusters that can be distinguished and the compounds present in each group. The chemical structures are given in the Supporting Information. Cluster A is composed by 6-halopregna-4,6-diene-3,20-dione, pregna-4-en-3,20-dione, and methylenepregna-4,6-diene-3,20-dione skeletons having different aliphatic and aromatic esters as well as aromatic carbamates at C-17 α. Some of these compounds have an alpha epoxide group at C-4, C-5 or

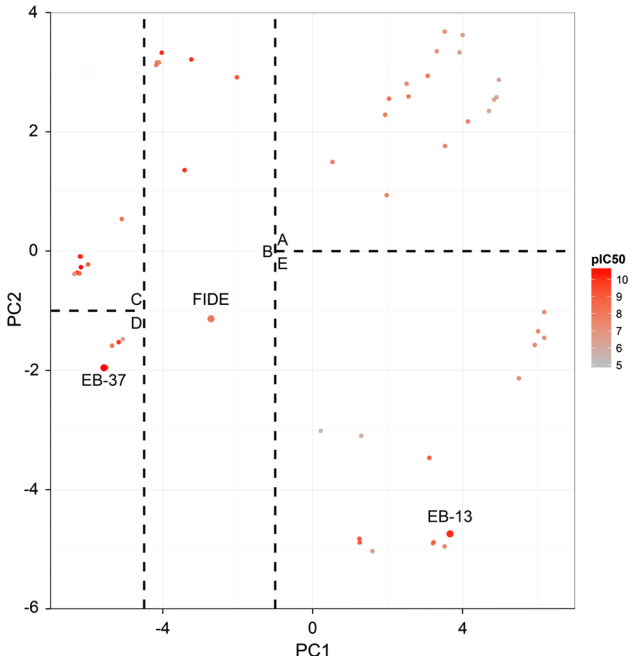

Fig. 1 Visual representation of the chemical space of the 54 molecules in the dataset. The PCA shows the relative position of finasteride (FIDE). The visualization was obtained by principal component analysis of the similarity matrix computed with ECFPs. *Data points* are *colored* by the pIC_{50} values in a *continuous scale*. Five major groups or clusters (*A*–*E*) are readily distinguished (see also Table 1)

Table 1 Major clusters identified in the visual representation of the chemical space in Fig. 1

Cluster A	Cluster B	Cluster C	Cluster D	Cluster E
EB-1	EB-19	EB-31	EB-36	EB-5
EB-2	EB-20	EB-32	EB-37	EB-6
EB-3	EB-21	EB-33	EB-38	EB-7
EB-4	EB-22	EB-34	EB-41	EB-8
EB-10	EB-23	EB-35	EB-43	EB-9
EB-15	EB-24	EB-39	EB-44	EB-11
EB-16	EB-25	EB-40		EB-12
EB-17	FIDE	EB-42		EB-13
EB-18				EB-14
EB-29				EB-26
EB-30				EB-27
EB-45				EB-28
EB-47				EB-46
EB-48				EB-52
EB-49				EB-53
EB-50				
EB-51				

C-6. Cluster B includes pregna-4,16-diene-6,20-dione derivatives with aliphatic and aromatic esters at C-3 β position, the 4-azasteroid FIDE are present in this cluster. Cluster C contains the modified D-homo lactone androstane and androstane skeletons having different aliphatic esters at C-3 β . Interestingly, Cluster D has the same skeletons than C, but it differs in the presence of 5α,6 β-dibromo group. Finally, Cluster E contains the 17α-methylpregna-4,6-diene-3,20-dione, 17α-methylpregna-1,4,6-triene-3,20-dione, pregna-4,6-diene-3,20-dione, and 17β-methyl-16β-phenyl-D-homoandrost-4,6-diene-3,17a-dione skeletons having different ester groups at the C-17 α position. Of note, the ECPFs were able to distinguish the structures, despite the fact that most of the compounds in the dataset have a steroid skeleton. This is due to the high resolution of this fingerprint-based structure representation [51].

Activity landscape analysis

Figure 2a depicts a typical SAS map for this dataset. Each point in the SAS map represents a paired comparison (there are 1431 possible pairs for this dataset of 54 compounds). Importantly, it can be seen that at higher similarity values (*x* axis), the range of the difference in activity decreases (*y* axis). This feature is in overall agreement with a continuous SAR, i.e., as two compounds are structurally more similar, the activity values are also similar. This result is further illustrated in panel b. In the SAS map of Fig. 2b, points are distinguished using a categorical classification

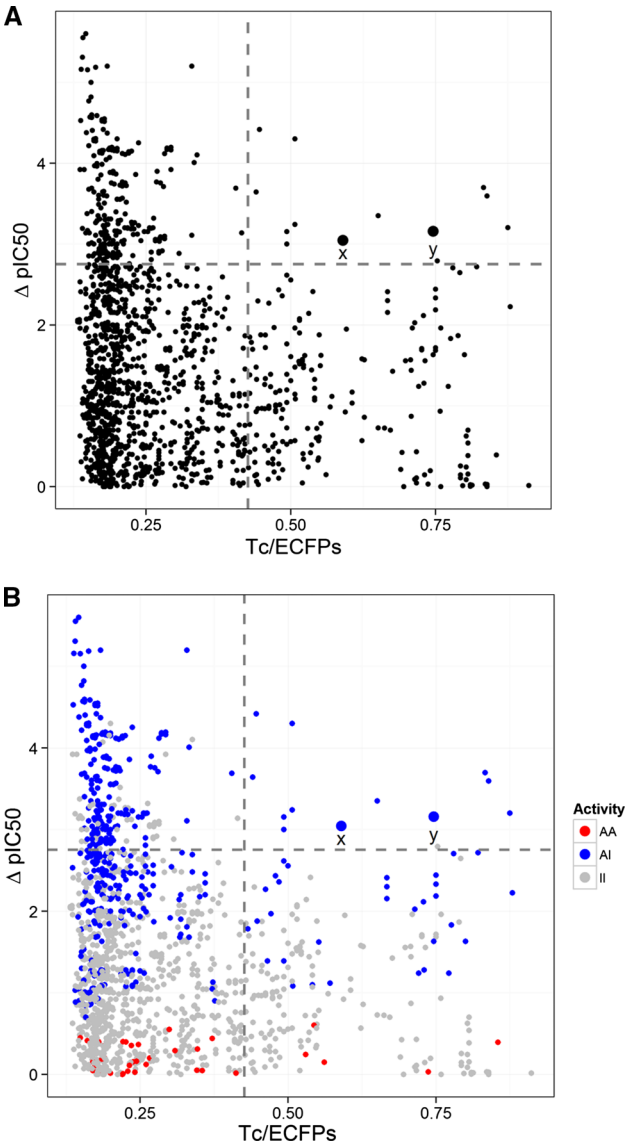

Fig. 2 SAS map for the 54 molecules in the dataset. Each *data point* (1431 total) represents a pairwise comparison. **a** Relative position of the 1431 pairwise structure and potency difference comparisons. **b** Activity SAS map identifying the *data points* by the categorical activity of the molecules in each pair. 'AA' represents pairs of active molecules, 'AI' pairs with one active and one inactive, and 'II' pairs in which both compounds are inactive (see text for details). Activity cliffs are located in the *upper-right region* of the *plot* herein defined using established criteria [45]. Representative activity cliffs are labeled EB-13/EB-46 ('x') and EB-37/EB-38 ('y')

of the biological activity taking as reference the pIC_{50} of FIDE: a compound was defined as 'active' in this figure if $pIC_{50} \geq 8.0$. Thus, in Fig. 2 data points labeled as 'AA,' are pairs where both compounds are active. Data points labeled as 'AI' contain one active and one inactive molecules. Finally, data points in gray (II pairs), contain two inactive compounds. Figure 2b clearly shows that the proportion of red (AA) points increases as the structure similarity also increases. This result

highlights the proficiency of the chemical representation used and also suggests a smooth SAR of the dataset.

In the SAS map of Fig. 2a, there are several pairs of compounds with both, low structure similarities and low activity differences (lower left quadrant). These are called 'similarity cliffs' [52] or scaffold hops. The relative amount of data points in the similarity cliff region can be visually analyzed in the density SAS map of Fig. 3. Note that the density SAS map can be clearly differentiated from the SAS maps in Fig. 2 in that Fig. 3 is focused on the counts of data points in each different regions of the map. Indeed, density SAS maps have been recently developed as complementary tools to alleviate the issue of crowded SAS maps that may be difficult to interpret [53].

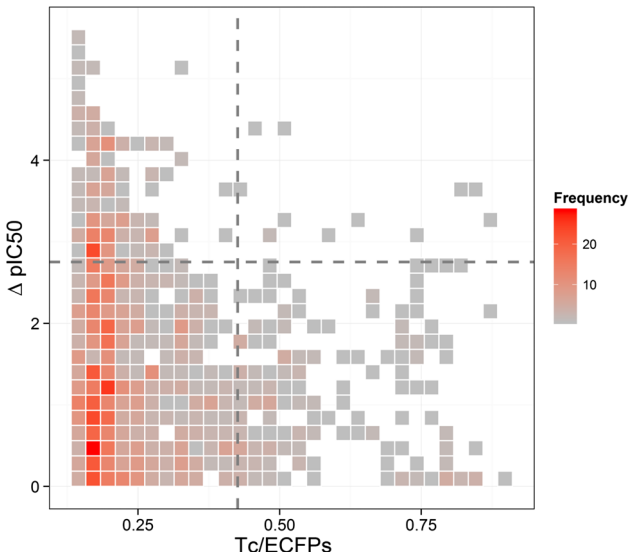

Fig. 3 Density SAS map of paired comparisons plotting structure similarity versus potency difference. The map is *colored* by the frequency of *data points* in the coordinates given

Figure 2a, b shows that there are few but notable data points in the upper-right zone of the plot, this is, pairs if compounds with high structure similarity but have large potency difference. These exceptions to the continuous SAR are the activity cliffs and are discussed in the next section. Of note, the activity cliffs are not systematically identified in 3D-QSAR studies such as SOMFA (*vide supra*). The identification and interpretation of the activity cliffs is an outcome of particular significance of the ALM study reported in this work.

Activity cliffs generators: identification

Figure 4 shows the distribution of molecules in the activity cliff region of the SAS map. The most frequent compounds with high activity were EB-13 (pIC_{50} = 10.20; IC_{50} = 0.063 nM) and EB-37 (pIC_{50} = 10.60; IC_{50} = 0.025 nM) with frequencies of five and three, respectively. In other words, EB-13 forms five activity cliffs and EB-37 forms three activity cliffs (see discussion below). Figure 4 also shows that there were additional activity cliff generators with high activity (red bars) but with lower frequencies such as EB-32, EB-33, and EB-42 (frequency of one). Finally, there were compounds with relative low pIC_{50} values (lower than the median pIC_{50} values of the entire dataset) that form several activity cliffs. One example was EB-34 (pIC_{50} = 6.96; IC_{50} = 110 nM) that forms four activity cliffs (Fig. 4). The SAR of the two most prominent activity cliff generators is discussed in the next section.

Activity cliffs generators: interpretation of the SAR

As discussed above, analysis of the chemical space and activity landscape of the dataset of inhibitors of 5aR led to the rapid identification of two activity cliff generators (EB-13 and EB-37) that occupy different regions in the chemical space of

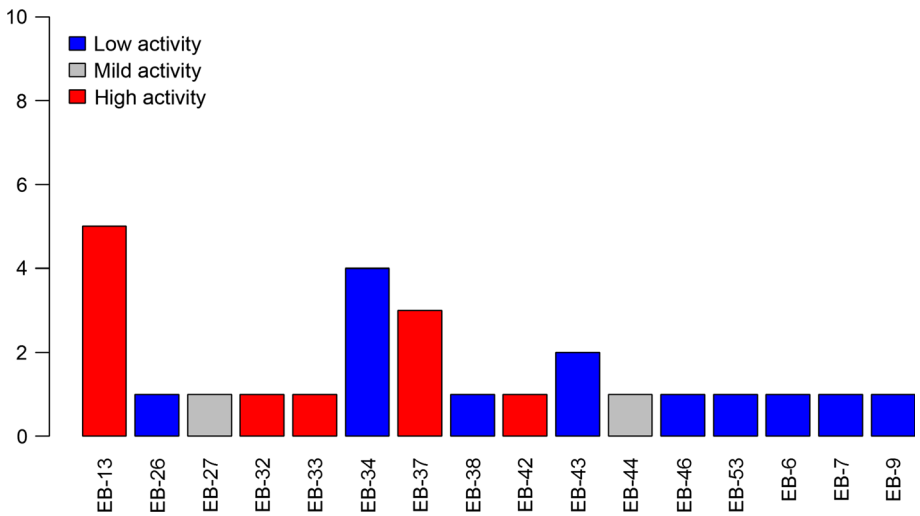

Fig. 4 Activity cliff generators: *bar graph* of compounds in the cliff region of the SAS map. The *bars* are *colored* by the relative biological activity (pIC_{50} values) of the compounds with respect to the entire set, *red* compounds with activities above one standard deviation from the mean, *blue* compounds with an activity below the mean, *gray* if the activity is between the mean and one standard deviation. Compounds EB-13 and EB-37 are the most active activity cliff generators

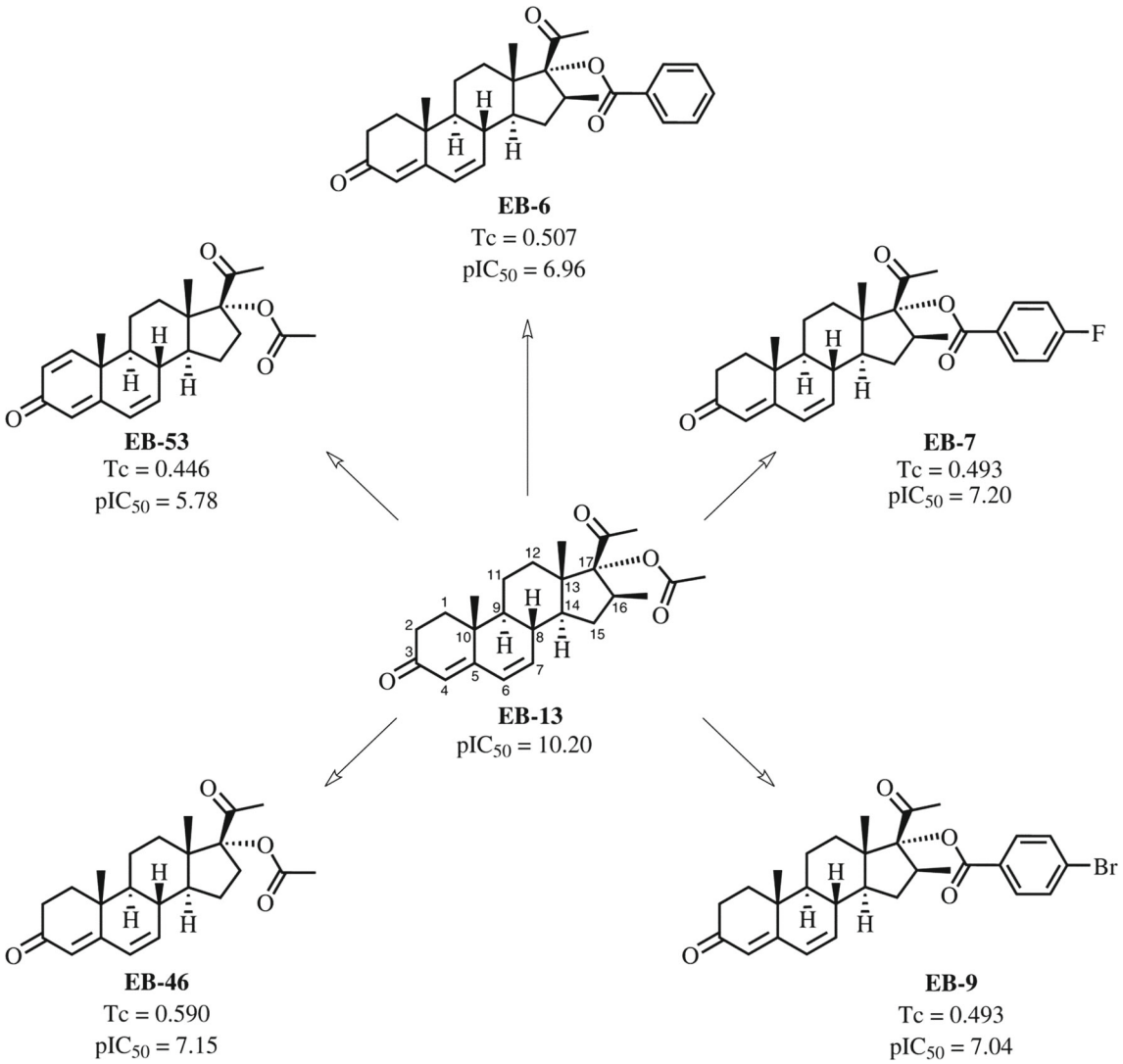

Fig. 5 SAR of the activity cliff generator EB-13

this dataset. In this section, we discuss an interpretation of the SAR of each of the two generators.

SAR of activity cliff generator EB-13

Figure 5 shows the chemical structure, pIC_{50} value, and structural similarity (Tc/ECFPs) values of the five compounds forming activity cliffs with EB-13, which is the 17α-acetoxy-16-methylpregna-4,6-diene-3,20-dione derivative (pIC_{50} = 10.20). In this figure, the standard numbering for the molecular scaffold is shown for EB-13. All compounds in this figure have a potency difference $\Delta pIC_{50} > 2$ with EB-13. The position in the SAS maps of a representative activity cliff, compound pair EB-13/EB-46, is shown in Fig. 2 (e.g., data point labeled with an 'x'). Comparison of the structure and activity of the activity cliff EB-13/EB-46 (Fig. 5) led to the remarkable finding that the 17β-methyl group in EB-13 can drastically change the inhibitory potency in three log units.

Overall, the molecules structurally related to the activity cliff generator EB-13 are pregnane derivatives with three conjugated double bonds and different esters attached to C-17 α position (Fig. 5). The inhibitory effect of these compounds has been explained with a plausible mechanism proposed wherein the inhibition of the 5aR enzyme is based on the Michael type addition reaction of the 5aR enzyme to the steroidal enone, dienone, or trienone to form irreversible adducts [54]. For EB-46 and EB-53, the different pIC_{50} values could be explained by the number of unsaturated bonds in the steroidal skeleton. Therefore, we can infer that the trienone in EB-53 has more resonance stabilization than EB-46, thus it is less susceptible to Michael addition and therefore EB-53 is the less active molecule in Fig. 5.

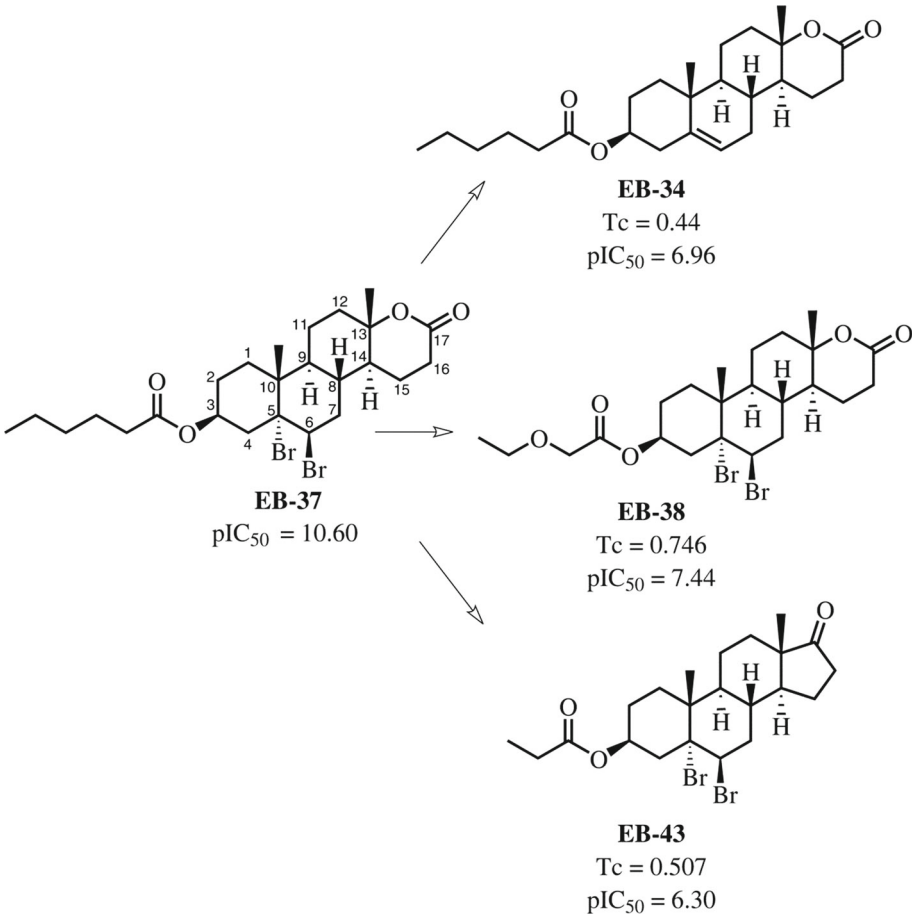

Fig. 6 SAR of the activity cliff generator EB-37

An additional intriguing finding that emerged from the analysis of the activity cliff generator in Fig. 5 is that the most similar steroidal compounds to EB-13 (i.e., with Tc >0.40) have in common two double bonds between C-4 and C-5, and C-6 and C-7, respectively. Consequently, the changes in activity can be attributed to the C-17α substituent. In general, a bulky ester group at C-17α decreases the activity for this series. However, comparing the pIC_{50} values of EB-6 with EB-7, and EB-9, it can be concluded that the *p*-substituted aromatic esters slightly increase the potency.

SAR of activity cliff generator EB-37

Figure 6 illustrates the three compounds forming activity cliffs with EB-37 (the standard numbering for the molecular scaffold is shown for EB-37). All compounds in this figure have a potency difference $\Delta pIC_{50} > 3$ with EB-37, which is 5α, 6β-dibromo-17a-oxa-D-homoandrostane-3β-yl-3′-oxahexanoate ($pIC_{50} = 10.60$). The position in the SAS map of the representative activity cliff EB-37/EB-38 is shown in Fig. 2 (labeled with a 'y'). The structural difference in this activity cliff is the side chain at C-3β indicating that the more lipophilic character of the hexanoate moiety in EB-37 versus the ethoxyacetate in EB-38 seems to favor the interaction with 5aR.

All molecules in Fig. 6 lack of an unsaturated ketone (in contrast to EB-13 and all compounds in Fig. 5). This observation is consistent with the different relative positions of EB-13 and EB-37 in chemical space (Fig. 1). Such structural difference can be related to a different mechanism of action. For example, the electrophilic C-5 and C-6 positions form irreversible adducts with nucleophilic residues and therefore promotes the inhibition of catalytic activity of the 5aR enzyme. Another mechanism suggested is that these compounds could have an alternative binding mode due to the pseudosymmetry of C-19 steroids, causing the C-3 end to be in the usual position of C-17. This finding was previously observed for DHT steroid in the 17β-HSD1 enzyme [55] and provides support for the hypothesis that the interchange of D-ring with the A-ring in androstene series illustrated above makes the C-3β (or pseudo C-17β position) ester moiety responsible for the lipophilic interaction at the hypothetical pocket in 5aR in a similar way found by the 4-azasteroids, wherein lipophilic ketones and amides at C-17β increase the inhibitory effect in the 5aR type II enzyme [7].

Conclusions and perspectives

Systematic characterization of the chemical space of 53 5aR inhibitors here presented demonstrated the structural uniqueness with respect to the approved drug FIDE and was able to distinguish different chemical classes, for example, those represented by the activity cliff generators EB-13 and EB-37. Systematic comparison of the structural activity and potency difference of each pair of molecules in the dataset showed that, in general, as the structural similarity increases, the potency difference decreases, in overall agreement with the similarity principle. However, the dataset analyzed in this work has two chemically distinct and potent activity cliff generators, EB-13 ($pIC_{50} = 10.20$; $IC_{50} = 0.063$ nM) and EB-37 ($pIC_{50} = 10.60$; $IC_{50} = 0.025$ nM), this is, compounds with high affinity to 5aR that that are very similar to compounds analogues but have large potency differences. Although the current study is based on a relatively small set of steroidal derivatives, the findings of this investigation complement those of previous studies. The outcome of the activity landscape analysis also provided hypothesis that compounds may be inhibiting 5aR in different forms, for example, forming irreversible adducts through a Michael type addition or having binding modes different from FIDE. The application of activity landscape analysis to identify possible different mechanisms of action or alternative binding modes is related to the concept 'activity landscape sweeping' introduced recently [53]. Of note, in contrast to other quantitative approaches to analyze SARs (such as QSAR), ALM does not require that the compounds analyzed have exactly the same mechanism of action. This is because activity landscape studies do not assume continuous SARs [36,37]. Despite the fact that further experimental investigations are needed to assess this hypothesis, novel non-4-azasteroidal inhibitors can be developed in order to improve the potency and selectivity in the prostatic 5aR enzyme and then be used for the treatment of BPH and PCa. Indeed, once the underlying SAR of the set of 53 5aR inhibitors has been explored, what was the focus of this work; the next logical steps are design new compounds and predict their activity. As previously discussed in the literature, understating the SAR of a data set become before prediction [37].

Supporting information

List of the 54 compounds used in this study (53 pregnane and androstene compounds and FIDE).

Acknowledgments FCB thanks the Consejo Nacional de Ciencia y Tecnología (CONACyT) by the Fellowship Awarded (Number 255249) to carry out the PhD studies. JLMF acknowledges the National Autonomous University of Mexico (UNAM) for Grant PAIP 5000-9163.

References

1. Azzouni F, Godoy A, Li Y, Mohler J (2012) The 5 alpha-reductase isozyme family: a review of basic biology and their role in human diseases. Adv Urol 2012:18. doi:10.1155/2012/530121
2. Bull HG, Garcia-Calvo M, Andersson S, Baginsky WF, Chan HK, Ellsworth DE, Miller RR, Stearns RA, Bakshi RK, Rasmusson GH, Tolman RL, Myers RW, Kozarich JW, Harris GS (1996) Mechanism-based inhibition of human steroid 5α-reductase by finasteride: enzyme-catalyzed formation of NADP–dihydrofinasteride, a potent bisubstrate analog inhibitor. J Am Chem Soc 118:2359–2365. doi:10.1021/ja953069t
3. Saartok T, Dahlberg E, Gustafsson J (1984) Relative binding affinity of anabolic–androgenic steroids: comparison of the binding to the androgen receptors in skeletal muscle and in prostate, as well as to sex hormone-binding globulin. Endocrinology 114:2100–2106. doi:10.1210/endo-114-6-2100
4. Gao W, Bohl CE, Dalton JT (2005) Chemistry and structural biology of androgen receptor. Chem Rev 105:3352–3370. doi:10.1021/cr020456u
5. Schauer IG, Rowley DR (2011) The functional role of reactive stroma in benign prostatic hyperplasia. Differentiation 82:200–210. doi:10.1016/j.diff.2011.05.007
6. Salvador JAR, Pinto RMA, Silvestre SM (2013) Steroidal 5α-reductase and 17α-hydroxylase/17,20-lyase (cyp17) inhibitors useful in the treatment of prostatic diseases. J Steroid Biochem Mol Biol 137:199–222. doi:10.1016/j.jsbmb.2013.04.006
7. Aggarwal S, Thareja S, Verma A, Bhardwaj TR, Kumar M (2010) An overview on 5α-reductase inhibitors. Steroids 75:109–153. doi:10.1016/j.steroids.2009.10.005
8. Thompson IM, Goodman PJ, Tangen CM, Lucia MS, Miller GJ, Ford LG, Lieber MM, Cespedes RD, Atkins JN, Lippman SM, Carlin SM, Ryan A, Szczepanek CM, Crowley JJ, Coltman CA (2003) The influence of finasteride on the development of prostate cancer. N Engl J Med 349:215–224. doi:10.1056/NEJMoa030660
9. Volpi R, Maccarini PA, Boni S, Chiodera P, Coiro V (1995) Case report: finasteride-induced gynecomastia in a 62-year-old man. Am J Med Sci 309:322–325
10. Cabeza M, Heuze I, Bratoeff E, Murillo E, Ramirez E, Lira A (2001) New progesterone esters as 5alpha-reductase inhibitors. Chem Pharm Bull 49:1081–1084. doi:10.1248/cpb.49.1081
11. Bratoeff E, Ramírez E, Flores E, Valencia N, Sánchez M, Heuze I, Cabeza M (2003) Molecular interactions of new pregnenedione derivatives. Chem Pharm Bull 51:1132–1136. doi:10.1248/cpb.51.1132
12. Ramírez E, Cabeza M, Bratoeff E, Heuze I, Pérez V, Valdez D, Ochoa M, Teran N, Jimenez G, Ramírez T (2005) Synthesis and pharmacological evaluation of new progesterone esters as 5alpha-reductase inhibitors. Chem Pharm Bull 53:1515–1518. doi:10.1248/cpb.53.1515
13. Cabeza M, Flores E, Heuze I, Sánchez M, Bratoeff E, Ramírez E, Francolugo VA (2004) Novel 17 substituted pregnadiene derivatives as 5α-reductase inhibitors and their binding affinity for the androgen receptor. Chem Pharm Bull 52:535–539. doi:10.1248/cpb.52.535
14. Bratoeff E, Sainz T, Cabeza M, Heuze I, Recillas S, Pérez V, Rodríguez C, Segura T, Gonzáles J, Ramírez E (2007) Steroids with a carbamate function at C-17, a novel class of inhibitors for human and hamster steroid 5α-reductase. J Steroid Biochem Mol Biol 107:48–56. doi:10.1016/j.jsbmb.2007.03.038
15. Bratoeff E, Cabeza M, Pérez-Ornelas V, Recillas S, Heuze I (2008) In vivo and in vitro effect of novel 4,16-pregnadiene-6,20-dione derivatives, as 5α-reductase inhibitors. J Steroid Biochem Mol Biol 111:275–281. doi:10.1016/j.jsbmb.2008.06.014

16. Cabeza M, Bratoeff E, Gómez G, Heuze I, Rojas A, Ochoa M, Palomino MA, Revilla C (2008) Synthesis and biological effect of halogen substituted phenyl acetic acid derivatives of progesterone as potent progesterone receptor antagonists. J Steroid Biochem Mol Biol 111:232–239. doi:10.1016/j.jsbmb.2008.06.011
17. Pérez-Ornelas V, Cabeza M, Bratoeff E, Heuze I, Sánchez M, Ramírez E, Naranjo-Rodríguez E (2005) New 5α-reductase inhibitors: in vitro and in vivo effects. Steroids 70:217–224. doi:10.1016/j.steroids.2004.11.008
18. Cabeza M, Zambrano A, Heuze I, Carrizales E, Palacios A, Segura T, Valencia N, Bratoeff E (2009) Novel C-6 substituted and unsubstituted pregnane derivatives as 5α-reductase inhibitors and their effect on hamster flank organs diameter size. Steroids 74:793–802. doi:10.1016/j.steroids.2009.04.009
19. Bratoeff E, Zambrano A, Heuze I, Palacios A, Ramírez D, Cabeza M (2009) Synthesis and biological activity of progesterone derivatives as 5α-reductase inhibitors, and their effect on hamster prostate weight. J Enzyme Inhib Med Chem 25:306–311. doi:10.3109/14756360903179401
20. Bratoeff E, García P, Heuze Y, Soriano J, Mejía A, Labastida AM, Valencia N, Cabeza M (2010) Molecular interactions of progesterone derivatives with 5α-reductase types 1 and 2 and androgen receptors. Steroids 75:499–505. doi:10.1016/j.steroids.2010.03.006
21. Garrido M, Bratoeff E, Bonilla D, Soriano J, Heuze Y, Cabeza M (2011) New steroidal lactones as 5α-reductase inhibitors and antagonists for the androgen receptor. J Steroid Biochem Mol Biol 127:367–373. doi:10.1016/j.jsbmb.2011.07.001
22. Arellano Y, Bratoeff E, Garrido M, Soriano J, Heuze Y, Cabeza M (2011) New ester derivatives of dehydroepiandrosterone as 5α-reductase inhibitors. Steroids 76:1241–1246. doi:10.1016/j.steroids.2011.05.015
23. Bratoeff E, Segura T, Recillas S, Carrizales E, Palacios A, Heuze I, Cabeza M (2009) Aromatic esters of progesterone as 5α-reductase and prostate growth inhibitors. J Enzyme Inhib Med Chem 24:655–662. doi:10.1080/14756360802323720
24. Cabeza M, Bratoeff E, Heuze I, Rojas A, Terán N, Ochoa M, Teresa Ramírez-Apan M, Ramírez E, Pérez V, Gracia I (2006) New progesterone derivatives as inhibitors of 5α-reductase enzyme and prostate cancer cell growth. J Enzyme Inhib Med Chem 21:371–378. doi:10.1080/14756360600748474
25. Robinson DD, Winn PJ, Lyne PD, Richards WG (1999) Self-organizing molecular field analysis: a tool for structure–activity studies. J Med Chem 42:573–583. doi:10.1021/jm9810607
26. Aggarwal S, Thareja S, Bhardwaj TR, Kumar M (2010) Self-organizing molecular field analysis on pregnane derivatives as human steroidal 5α-reductase inhibitors. Steroids 75:411–418. doi:10.1016/j.steroids.2010.02.005
27. Aggarwal S, Thareja S, Bhardwaj TR, Kumar M (2010) 3D-QSAR studies on unsaturated 4-azasteroids as human 5α-reductase inhibitors: a self organizing molecular field analysis approach. Eur J Med Chem 45:476–481. doi:10.1016/j.ejmech.2009.10.030
28. Myint KZ, Xie X-Q (2010) Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. Int J Mol Sci 11:3846. doi:10.3390/ijms11103846
29. Bajorath J (2012) Modeling of activity landscapes for drug discovery. Expert Opin Drug Discov 7:463–473. doi:10.1517/17460441.2012.679616
30. Stumpfe D, Hu Y, Dimova D, Bajorath J (2014) Recent progress in understanding activity cliffs and their utility in medicinal chemistry. J Med Chem 57:18–28. doi:10.1021/jm401120g
31. Guha R (2012) Exploring structure–activity data using the landscape paradigm. Wiley Interdiscip Rev Comput Mol Sci 2:829–841. doi:10.1002/wcms.1087
32. Naveja JJ, Medina-Franco JL (2015) Activity landscape of DNA methyltransferase inhibitors bridges chemoinformatics with epigenetic drug discovery. Expert Opin Drug Discov 10:1059–1070. doi:10.1517/17460441.2015.1073257
33. Maggiora GM (2006) On outliers and activity cliffs—why QSAR often disappoints. J Chem Inf Model 46:1535. doi:10.1021/ci060117s
34. Rojas-Aguirre Y, Medina-Franco J (2014) Analysis of structure–caco-2 permeability relationships using a property landscape approach. Mol Divers 18:599–610. doi:10.1007/s11030-014-9514-x
35. Pérez-Villanueva J, Santos R, Hernández-Campos A, Giulianotti MA, Castillo R, Medina-Franco JL (2011) Structure–activity relationships of benzimidazole derivatives as antiparasitic agents: dual activity-difference (DAD) maps. Med Chem Commun 2:44–49. doi:10.1039/C0MD00159G
36. Guha R, VanDrie JH (2008) Structure–activity landscape index: identifying and quantifying activity cliffs. J Chem Inf Model 48:646–658. doi:10.1021/ci7004093
37. Medina-Franco J, Navarrete-Vázquez G, Méndez-Lucio O (2015) Property landscape modeling is at the interface of chemoinformatics and experimental sciences. Future Med Chem 7:1197–1211. doi:10.4155/fmc.15.51
38. Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. J Chem Inf Model 45:177–182. doi:10.1021/ci049714+
39. Medina-Franco JL, Martínez-Mayorga K, Giulianotti MA, Houghten RA, Pinilla C (2008) Visualization of the chemical space in drug discovery. Curr Comput Aided Drug Des 4:322–333. doi:10.2174/157340908786786010
40. Medina-Franco JL, Maggiora GM, Giulianotti MA, Pinilla C, Houghten RA (2007) A similarity-based data-fusion approach to the visual characterization and comparison of compound databases. Chem Biol Drug Des 70:393–412. doi:10.1111/j.1747-0285.2007.00579.x
41. Rogers D, Brown RD, Hahn M (2005) Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up. J Biomol Screen 10:682–686. doi:10.1177/1087057105281365
42. Stumpfe D, Bajorath J (2012) Exploring activity cliffs in medicinal chemistry. J Med Chem 55:2932–2942. doi:10.1021/jm201706b
43. Medina-Franco JL, Maggiora GM (2014) Molecular similarity analysis. In: Bajorath J (ed) Chemoinformatics for drug discovery. Wiley, New York, pp 343–399. doi:10.1002/9781118742785.ch15
44. López-Vallejo F, Castillo R, Yépez-Mulia L, Medina-Franco JL (2011) Benzotriazoles and indazoles are scaffolds with biological activity against *Entamoeba histolytica*. J Biomol Screen 16:862–868. doi:10.1177/1087057111414902
45. Shanmugasundaram V, Maggiora GM (2001) Characterizing property and activity landscapes using an information-theoretic approach. Cinf-032. In: Paper presented at the 222nd ACS national meeting, Chicago, IL, 26–30 August
46. Medina-Franco JL, Waddell J (2012) Towards the bioassay activity landscape modeling in compound databases. J Mex Chem Soc 56:163–168
47. Yongye AB, Medina-Franco JL (2013) Systematic characterization of structure–activity relationships and ADMET compliance: a case study. Drug Discov Today 18:732–739. doi:10.1016/j.drudis.2013.04.002
48. Medina-Franco JL (2012) Scanning structure–activity relationships with structure–activity similarity and related maps: from consensus activity cliffs to selectivity switches. J Chem Inf Model 52:2485–2493. doi:10.1021/ci300362x
49. Pérez-Villanueva J, Santos R, Hernández-Campos A, Giulianotti MA, Castillo R, Medina-Franco JL (2010) Towards a systematic characterization of the antiprotozoal activity landscape of benz-

imidazole derivatives. Bioorg Med Chem 18:7380–7391. doi:10.1016/j.bmc.2010.09.019
50. Mendez-Lucio O, Perez-Villanueva J, Castillo R, Medina-Franco JL (2012) Identifying activity cliff generators of PPAR ligands using SAS maps. Mol Inform 31:837–846. doi:10.1002/minf.201200078
51. Rogers D, Hahn M (2010) Extended-connectivity fingerprints. J Chem Inf Model 50:742–754. doi:10.1021/ci100050t
52. Iyer P, Stumpfe D, Vogt M, Bajorath J, Maggiora GM (2013) Activity landscapes, information theory, and structure–activity relationships. Mol Inform 32:421–430. doi:10.1002/minf.201200120
53. Naveja JJ, Medina-Franco JL (2015) Activity landscape sweeping: insights into the mechanism of inhibition and optimization of DNMT1 inhibitors. RSC Adv 5:63882–63895. doi:10.1039/C5RA12339A
54. Flores E, Bratoeff E, Cabeza M, Ramirez E, Quiroz A, Heuze I (2003) Steroid 5α-reductase inhibitors. Minirev Med Chem 3:225–237. doi:10.2174/1389557033488196
55. Gangloff A, Shi R, Nahoum V, Lin S-X (2002) Pseudo-symmetry of C19-steroids, alternative binding orientations and multispecificity in human estrogenic 17β-hydroxysteroid dehydrogenase. FASEB J. doi:10.1096/fj.02-0397fje