Journal of Computer-Aided Molecular Design (2020) 34:659–669
https://doi.org/10.1007/s10822-020-00298-x

Towards the understanding of the activity of G9a inhibitors: an activity landscape and molecular modeling approach

Edgar López-López[1] · Obdulia Rabal[2] · Julen Oyarzabal[2] · José L. Medina-Franco[1]

Received: 6 August 2019 / Accepted: 7 February 2020 / Published online: 14 February 2020

Abstract

In this work, we analyze the structure–activity relationships (SAR) of epigenetic inhibitors (lysine mimetics) against lysine methyltransferase (G9a or EHMT2) using a combined activity landscape, molecular docking and molecular dynamics approach. The study was based on a set of 251 G9a inhibitors with reported experimental activity. The activity landscape analysis rapidly led to the identification of activity cliffs, scaffolds hops and other active an inactive molecules with distinct SAR. Structure-based analysis of activity cliffs, scaffold hops and other selected active and inactive G9a inhibitors by means of docking followed by molecular dynamics simulations led to the identification of interactions with key residues involved in activity against G9a, for instance with ASP 1083, LEU 1086, ASP 1088, TYR 1154 and PHE 1158. The outcome of this work is expected to further advance the development of G9a inhibitors.

Keywords Activity cliffs · Drug discovery · EHMT2 · Epi-pharmacology · Scaffold hops · Structure–activity relationships

Introduction

The study of diseases based on epigenetic approaches is crucial for the development of new therapeutic alternatives [1]. Protein lysine methyltransferases (PKMTs) have attracted much interest in the drug discovery field as their inhibition is believed to be specific at the functional level [2]. Particularly, G9a [also known as KMT1C (lysine methyltransferase 1C) or EHMT2 (euchromatic histone methyltransferase 2)] is a histone-lysine N-methyltransferase that catalyzes the transfer of one or two methyl groups to the ε-amino group of lysine 9 of histone H3 (H3K9me1 and H3K9me2), a hallmark associated with transcriptional gene silencing. Other protein targets of G9 include the tumor suppressor p53, whose methylation leads to its inactivation [3, 4]. G9a is upregulated in various cancers, and its overexpression has been associated with poor prognosis and metastasis [5–8]. In addition, a very recent study reveals that G9a expression is significantly associated with resistance to immunotherapy, programmed cell death protein 1 (PD1) inhibition, in a cohort of bladder cancer patients [9]. Moreover, G9a is involved in the maintenance of HIV-1 latency, cognitive disturbances (mental retardation, cocaine addiction, age-related cognitive decline), embryonic development, colitis, regulation of the cell cycle, and stem cell reprogramming, which has been used to produce inducible pluripotent stem cells (iPSCs) [10–15].

There is a large variety of compounds synthesized and evaluated against G9a [16]. Distinct structural classes are quinoline and indole derivatives (Fig. 1). Reports have shown the influence of substituents on the activity for each structural class. For instance, lysine mimetic substituents (positively charged) have been shown to have a higher affinity for the substrate pocket on G9a. Examples are the small molecules **23 m**, **CHEMBL3109631**, and **WO2017142947A1-4** (Fig. 1). Similarly, there are reports suggesting that the lysine mimetic substituents promote key interactions with residues involved in enzymatic inhibition, such as TYR 1067, LEU 1086, TYR 1154, and PHE 1152 [17–20]. However, it is necessary to compare the binding

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10822-020-00298-x) contains supplementary material, which is available to authorized users.

✉ José L. Medina-Franco
medinajl@unam.mx

[1] Department of Pharmacy, School of Chemistry, National Autonomous University of Mexico, 04510 Mexico City, Mexico

[2] Small Molecule Discovery Platform, Molecular Therapeutics Program, Center for Applied Medical Research, CIMA, University of Navarra, Pio XII, 55, 31008 Pamplona, Spain

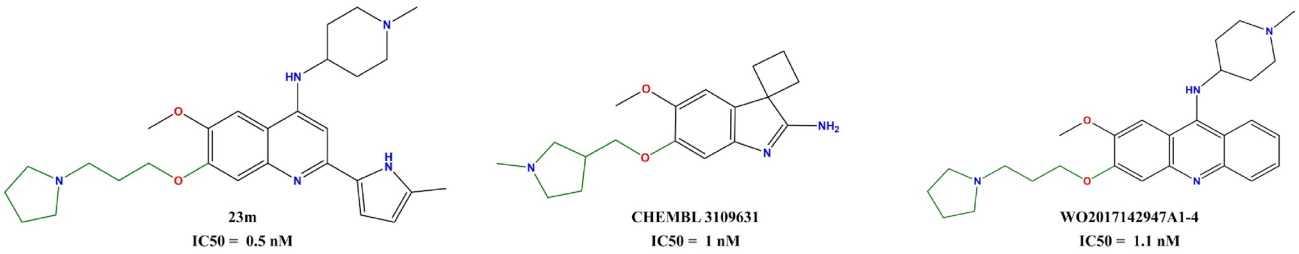

Fig. 1 Representative G9a inhibitors. The mimetic lysine substituents are colored green

mode and the molecular geometry of the inhibitors considering the different molecular scaffolds in the same study.

Recently, the structure–activity relationships (SAR) of G9a inhibitors has been explored using the concept of activity landscape modeling [21, 22]. In that work, the dual activity of inhibitors with a quinoline scaffold has been predominantly explored. These studies allowed the identification of key substituents in the selectivity and efficacy of compounds.

The objective of this paper was to further explore the SAR of G9a ligands using the established concept of the activity landscape model and molecular docking followed by molecular dynamics.

Materials and methods

Data set

This study is based on a data set of 251 compounds derived from aminoquinoline, indole and acridine core with reported activity against G9a (IC_{50}) [21, 23, 24]. All compounds were retrieved from the public database ChEMBL, except for compounds based on the acridine core, that are reported in patents [19, 23–25]. The SMILES representation of the structures and pIC_{50} ($-\log IC_{50}$) values are listed in Table S1 in the supplementary material. The pIC_{50} values range from 9.30 to 5.0 [26].

Software and online resources

The activity landscape analysis was carried out with Activity Landscape Plotter, an open web tool (https://www.difacquim.com/d-tools/) that enables the analysis of SAR of screening data sets [27]. This tool facilitates a first and rapid exploration of the SAR of compound data sets with a common scaffold and rapid decomposition of R-groups [28]. Molecular docking was performed with the software Yasara (YASARA Biosciences GmbH, Vienna, AUT). Molecular dynamics (MD) studies were done with Desmond (Schrödinger, New York, NY, USA) [29].

Activity landscape modeling

A structure–activity similarity (SAS) map is a tool for SAR analysis of compound data sets tested against a molecular target. SAS maps are based on the concept of activity landscape and are suited for the rapid identification of activity cliffs (AC), defined as compounds with a high structure similarity but unexpected large activity difference [30, 31]. SAS maps also enable the identification of scaffold hops (SH), defined as compounds with low structural similarity due to differences in their scaffold but similar biological activity. This concept has been useful in medicinal chemistry to identify and develop novel and diverse chemical entities [32].

Since SAS maps are based on systematic pairwise comparisons of the compounds in a data set, the SAS maps generated in this work represented all 31,375 pairwise comparisons between the 251 compounds of the set. To generate the map, the structural similarity was calculated with the ECFP4 fingerprint and the Tanimoto coefficient and was represented on the X-axis. The activity difference (in logarithmic scale) was plotted on the Y-axis. To differentiate the four major regions in the SAS map, two thresholds were set along the X- and Y-axis, respectively. The criteria to select the threshold along the X-axis was the mean of the similarity values of all compounds in the set (0.667) (calculated with Tanimoto and the ECFP4 fingerprint). The threshold of the activity difference (Y-axis) was set to two logarithmic units [33].

The data points in the SAS map were further colored by the corresponding Structure–Activity Landscape Index (SALI) value [34]. This index, as implemented in Activity Landscape Plotter, quantifies AC using the equation proposed by Guha and Van Drie:

$$SALI\ I,\ J = ((|Ai - Aj|)/(1 - sim(i,j)))$$

where Ai and Aj are the activities of the ith and the jth molecules, and sim (i, j) is the similarity coefficient between the two molecules (in this work computed with the ECFP4 fingerprint and the Tanimoto coefficient). The SALI values were mapped onto the SAS maps using a continuous color scale from the structurally most similar pairs (green) to the least similar pairs (red). For the quantitative analysis

of SAS maps, the structure similarity was also evaluated with MACCS keys (166-bits) and PubChem fingerprints as implemented in Activity Landscape Plotter [27].

Scaffold content analysis

For the study of scaffolds, we use the methodology implemented by Naveja et. al., generating the scaffolds of each compound based on Bemis and Murcko approach. Briefly, the method consists of a graph analysis for each compound, where a "scaffold" is defined as the union of ring systems and linkers in a molecule, and the side chains are removed (any non-ring, non-linker atoms). This was done with the RDKit Fragments node implemented in KNIME software v. 3.7.2. The chemical structures of the scaffolds are available in Table S1 of the supplementary material [35, 36].

Molecular docking

Protein preparation

The crystallographic structure of human G9a (PDB ID: 3RJW) was retrieved from the Protein Data Bank (https://www.rcsb.org/) [37]. The co-crystallized ligand was removed (UNC0638, a 4-aminoquinazoline). Missing loops and side-chains were added with YASARA software [38]. Finally, hydrogen atoms were added, followed by a minimization step with the AMBER99 forcefield in MOE software (Chemical Computing Group, Montreal, QC, Canada) [39].

Ligand preparation

The ligands were built and energy-minimized in MOE using the MMFF94x force field. The more stable protomers at physiological pH were identified [40].

Molecular docking

Yasara software was used to add the solvent model and assign the Gasteiger atomic charges to proteins and ligands [29]. The grid was centered on the binding site of the protein (PDB ID: 3RJW). Using the scoring function of AutoDock Vina, the binding compounds were subjected to 25 search steps and the default values for the other parameters. The clusters with an RMSD < 2 Å were visually explored. During the docking simulations the receptor was considered rigid and the ligands flexible. The conformations with the lowest binding energy and diverse scaffolds were selected for an additional MD analysis.

Molecular dynamics

MD studies of the protein–ligand complexes were performed using Desmond (version 2018-3, Schrödinger, New York, NY, USA) with the OPLS 2005 forcefield [41]. The most representative docking pose for each ligand was used as a starting point to initiate the MD simulations. The complexes were prepared with the System Builder Utility in a buffered orthomobic box ($10 \times 10 \times 10$ Å), using the transferable intermolecular potential with 3-point model for water (TIP3P). The complexes were neutralized and NaCl was added in a 0.15 M concentration.

Complexes were minimized using the steep-descent (SD) algorithm followed by the Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method in three stages. In the first stage water heavy atoms were restrained with a force constant of 1000 kcal mol^{-1} Å^{-2} for 5000 steps (1000 SD, 4000 LBFGS) with a convergence criterion of 50 kcal mol^{-1} Å^{-2}; for the second stage, backbones were constrained with a 10 kcal mol^{-1} Å^{-2} force constant using a convergence criterion of 10 kcal mol^{-1} Å^{-2} for 2000 steps (1000 SD, 1000 LBFGS); and for the third stage the systems were minimized with no restraints for 1000 steps (750 SD, 250 LBFGS) with a convergence criterion of 1 kcal mol^{-1} Å^{-2}.

Equilibration was carried out in several steps. Beginning with Brownian Dynamics for 250 ps with the Berendsen thermostat. Followed by simulation on the NVT ensemble, slowly heating from 10 to 300 K over 3000 ps. At this stage, constraints were enforced on solute heavy atoms, using a constant of 50 kcal/mol.

Finally, equilibration on NPT ensemble used the Berendsen thermostat and Langevin barostat for additional 250 ps. Subsequently, the system was submitted to 130 ns of production runs, under NPT ensemble at 1 bar using the Martyna-Tuckerman-Klein (MTK) barostat and 300 K using the Nose–Hoover thermostat. Electrostatic forces were calculated with the smooth PME method using a 9 Å cut-off, while constraints were enforced with the M-SHAKE algorithm. Integration was done every 1.2 fs, with a recording interval of 50 ps. Finally, the first 30 ns of production runs were removed, this is due to system stabilization after this period [22].

The quality of the simulation and trajectory analyses were carried out with the tools implemented in the Maestro-GUI (Schrödinger, New York, NY, USA) (Fig. S2).

Results and discussion

First, we present and discuss the results of the activity landscape analysis that led to the rapid identification of AC and other related pairs of active-inactive compounds with distinct SAR. Then, we discuss a structure-based analysis of

selected compounds from the activity landscape analysis (focused on the most active compounds of the cliffs in the data set). The structure-based analysis was performed with docking followed by MD simulations.

SAS map

Figure 2-A shows the SAS map for the 251 G9a inhibitors. The graph contains 31,375 data points, each of which represents a pairwise comparison (vide supra). For each pair of compound the maps show the relationship between the difference in activity and the molecular similarity. As detailed in the Materials and methods section, the molecular similarity was quantified with the Tanimoto coefficient using the ECFP4 fingerprints. The data points are further distinguished by the SALI value, using a continuous color scale from a low value (green) to a high value (red). As discussed in the Materials and methods section, AC will have a high

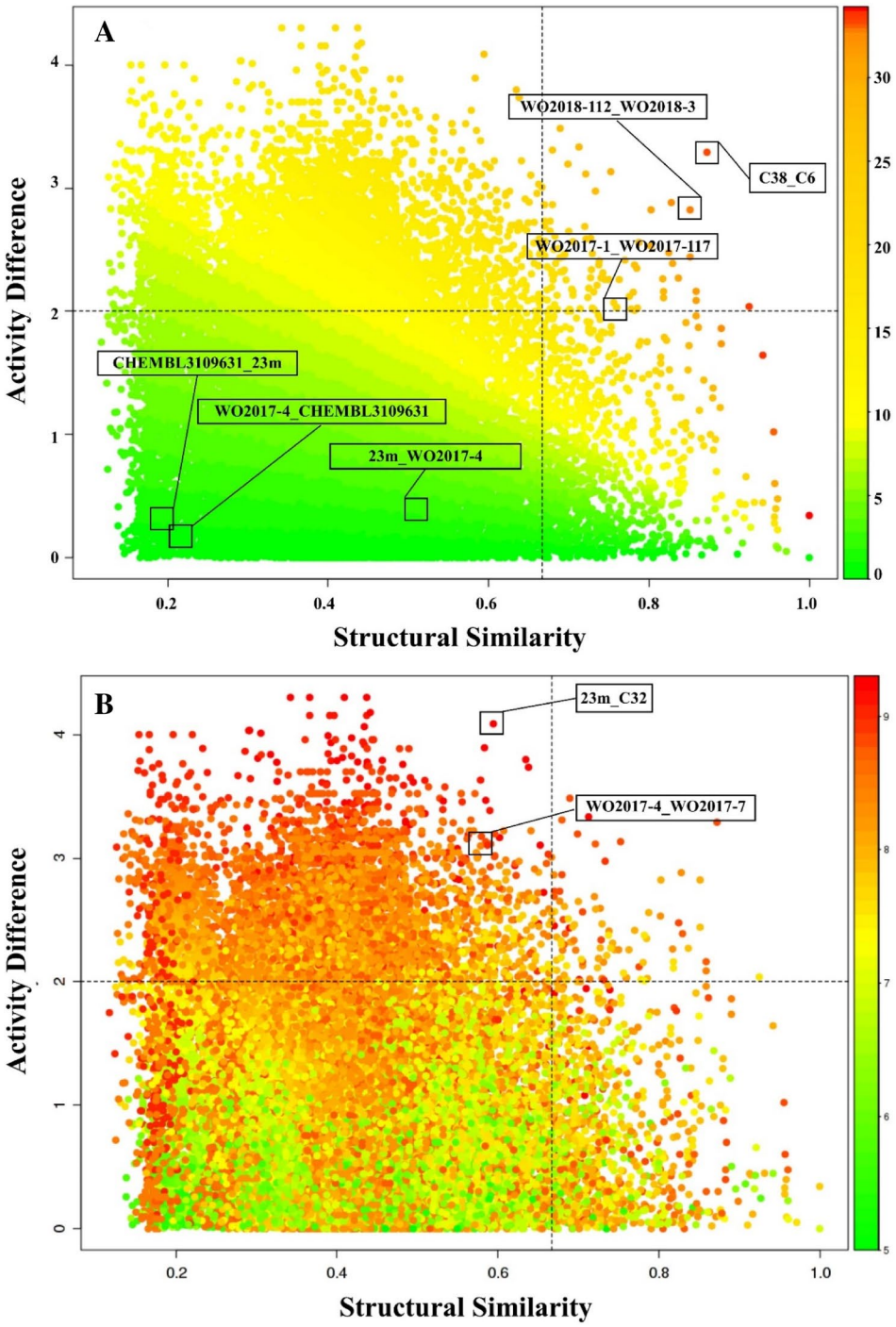

Fig. 2 SAS map of 251 compounds with activity against G9a (31,375 data points or pairwise comparison). **a** Data points are colored by the SALI value using a continuous scale from low (green) to high (red) values. **b** Data points are colored by the maximum activity of the compound in each pair (Max. Activity) value using a continuous scale from low (green) to high (red)

SALI value. In the SAS map in Fig. 2-A most pairs are green and yellow, indicating a more continuous SAR i.e., similar structures with similar activity. This result can be partially explained by the fact that the compounds in the data set come from a lead-optimization process [21, 23, 24]. Figure 2b shows the SAS map colored by "Max.Activity" value, i.e., the activity of the most active compound in the pair.

As shown in Fig. 2, a significant proportion of pairs of compounds have activity differences larger than two log units. This suggests that this set of compounds explores in detail the SAR for G9a. However, the AC zone (upper right corner) is not so coarse, suggesting that the compounds are structurally similar, although we know of the overall scaffold variety in the set of compounds (Fig. 1). In the SH zone of the SAS maps (lower left corner) is possible to identify compounds with small activity difference, but with different scaffolds (low structural similarity), as shown below. We want to emphasize that the SAS maps in Fig. 2 were generated using structural fingerprints (ECFP4, as described in the Method section) and the AC rapidly identified in the SAS maps may have actually different molecular scaffolds. In other words, the fingerprint-based similarity may not be highly associated with the scaffold or sub-structure-similarity and it is not a straightforward approach to identify molecular matched pairs.

Scaffold content analysis

Figure 3 shows the frequency of Bemis-Murcko scaffolds (138 different scaffolds), along with the distribution of their biological activity (pIC_{50}). It is noted that the set of compounds studied has several (more than 10) compounds with the same scaffold as **23 m** (**s17**) and **WO2017-7** (**s62**). However, there are compounds with unique scaffolds in the set such as **CHEMBL3109631** (**s59**).

Specific examples of AC and SH identified in the activity landscape analysis are labeled in Fig. 2 and 3, respectively. The chemical structures are shown in Fig. 4. An AC for G9a is the compound pair **23 m/C32**, differing in the methoxy group at the 6-position and the change of a pyrrole to furan substituent at the 2-position of the quinoline ring. Another example of a pair of compounds with similar structure

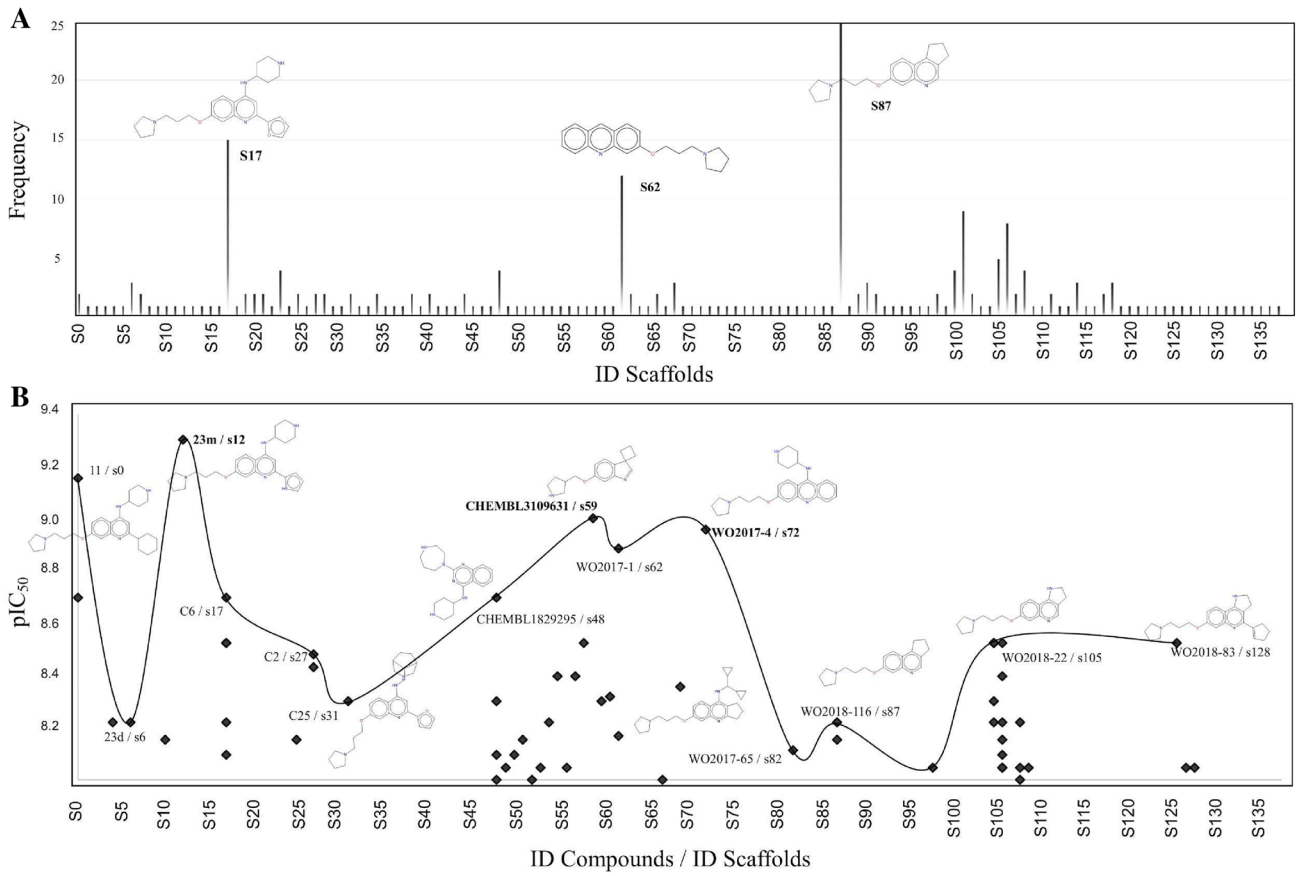

Fig. 3 Scaffold content analysis of compounds with activity against G9a. **a** Scaffold frequency. The chemical structures of the most frequent scaffolds are shown (frequency of at least 10 compounds). **b** Scaffolds of the most active compounds and activity differences. The ID of each compound and the ID of the corresponding scaffold are shown for each structure

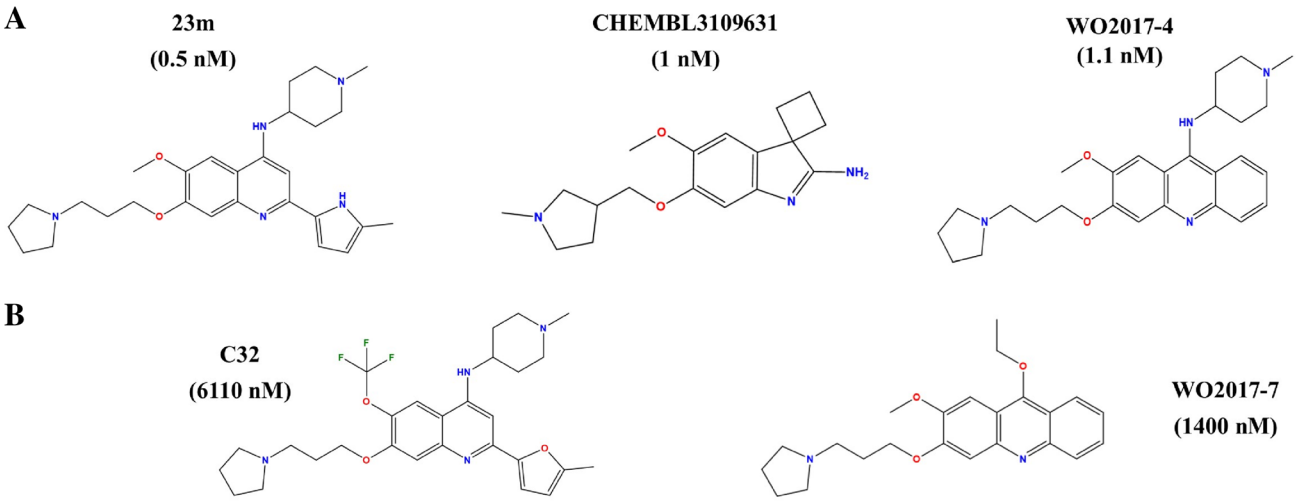

Fig. 4 Representative G9a inhibitors. **a** Most active compounds, **b** inactive compounds

(structural changes at the substituent group at 4-position) but large activity difference is **WO2017-4/WO2017-7** (Fig. 4). Examples of compounds associated with SH highlighted in Fig. 2a are **23 m**, **CHEMBL3109631** and **WO2017-4** (Fig. 4a).

Molecular docking

Based on the results of the activity landscape analysis, molecular docking was used to generate representative binding poses with G9a of selected compounds. Docking was conducted with Yasara software, as detailed in the Materials and methods section.

Figure 5 shows one of the best ranked binding poses of representative active and inactive compounds. The compounds were selected because the SAS map identified them as the most representative active and inactive compounds per scaffold. The docking model suggested that the three most active compounds (**23 m**, **CHEMBL3109631**, and **WO2017-4**) establish hydrogen-bond contacts with the side chains of ASP 1083, and ASP 1088, as well as hydrogen-bond contacts between the positively charged pyrrolidine group of **23 m** and **WO2017-4** with the backbone of LEU 1086.

Of note, in the three active compounds in Fig. 5a the substituents that mimic lysine (marked in green in Fig. 1) make polar and stacking interactions with amino acids such as TYR 1067, TYR 1154, ARG 1157, and PHE 1158.

We emphasize that the binding poses of the active and inactive compounds are similar, however we consider that the flexibility of the mimetic substituents of lysine (in the inactive compounds) leads to overfit their binding poses. This hypothesis was tested using MD simulations (vide infra).

The docking was done with the crystallographic structure of G9a PDB ID: 3RJW. Of note, the structure of G9a PDB ID: 4NVQ is co-crystallized with the compound A-366, which is structurally similar to **CHEMBL3109631**. Interestingly, Fig. S1 in the Supplementary Material shows the results of cross-docking of the co-crystal ligands with the two crystallographic structures (PDB ID: 3RJW and 4NVQ, respectively). In cross-docking, the binding poses change significantly suggesting the relevance of the flexibility of the binding pocket. In this work, such flexibility was addressed with MD simulations (discussed in the next section). Overall, the results of the cross-docking further emphasize the importance of MD to perform structure-based interpretation of inhibitors of G9a.

Molecular dynamics

Based on the insights from activity landscape modeling and docking calculations, we performed MD simulations (100 ns) for selected and structurally related active and inactive compounds.

Figure 6 summarizes the interactions between the most active compounds (**23 m**, **CHEMBL3109631**, and **WO2017-4**) and G9a, according to the MD simulations. It is noteworthy the well-conserved interactions with ASP 1083, ASP 1088 (except **CHEMBL3109631**), and TYR 1154. However, interactions with LEU 1086 were replaced throughout the dynamics by interactions with the TYR 1154. We highlight the interaction of compound 23 m with ASP 1088, since it is generated through the pyrrole substituent in 2-position. Pyrrole substituent is particularly interesting given that it is found in the most active compounds of this series of compounds (derivatives of 4 amino quinoline).

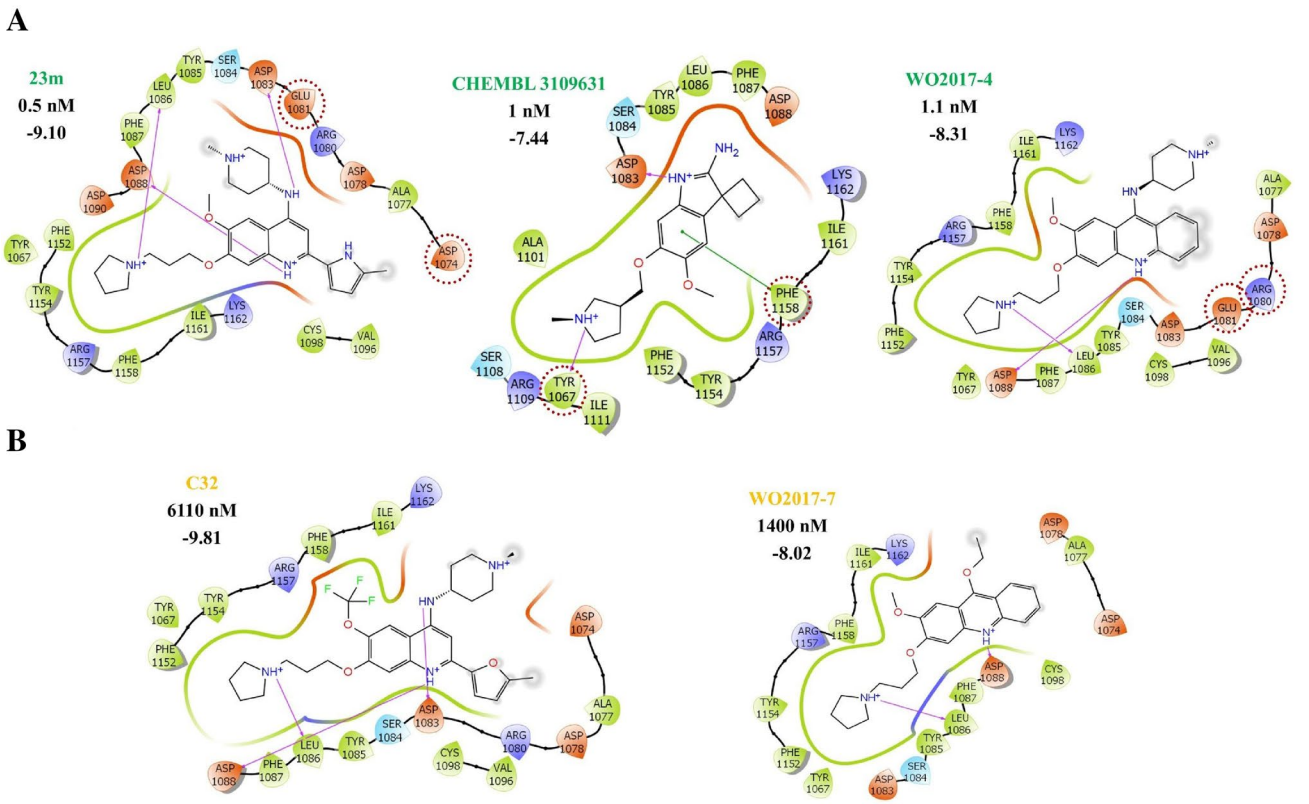

Fig. 5 Predicted binding poses of representative compounds. **a** Most active compounds and **b** inactive compounds. The interactions by hydrogen bonds (pink line) and Pi–Pi interactions (green line) are represented

Of the three compounds analyzed in MD, **WO2017-4** showed more stable direct interactions with G9a during the simulation time. In contrast, the interactions of **CHEMBL3109631** with the binding site are stabilized with structural waters. Of note, for the three most active compounds are predicted hydrogen-bond interactions between their positively charged scaffolds and ASP 1088. The lysine mimetic substituents in the most active compounds make interactions with PHE 1158 (Pi-cation interaction dependent of pyrrole). Also **23 m** and **WO2017-4** make interactions with LEU 1086 (hydrophobic interaction dependent of aliphatic portion), and TYR 1154 (Pi-cation).

Figure 7-A describes the interactions of inactive compounds against G9a (**C32** and **WO2017-7**), based on the MD simulations. As compared to the MD simulations of the active compounds it is remarkable that the inactive compounds do not make key interactions with the lysine-mimetic substituent: e.g. LEU 1086, TYR 1154, and PHE 1158 (Fig. 6). As per the interactions with their charged scaffolds, at least one of the key interactions observed for active compounds was lost (ASP 1083 and ASP 1088, Fig. 6). Moreover, as shown in the histograms of Fig. 6 it was also observed a decrease in the fraction of interaction with key residues during the MD simulations.

Figure 7 suggests non-specificity of interaction with respect to the most active compounds. This is further illustrated in Fig. S2 in the Supplementary Material, where the RMSD and RMSF values (per compound) are plotted, which show the conformational stability of G9a and the instability of the poses associated with inactive compounds (**C32** and **WO2017-7**).

At the structural level, we have identified two possible conformational states of pharmacology importance in the SET domain of G9a (domain involved in the function of histone methylation) [42]. The first (closed state) that is associated with active compounds, and the second (open state) that is associated with inactive compounds. The compounds favor one or another conformational state as shown in Fig. S3A and B (Supplementary Material), respectively. The distance between ASP 1088 (carboxyl group carbon in the side chain) and ARG 1157 (amino group carbon in the side chain) (plotted in Fig. S3C) seems crucial in the formation of an open or closed conformational states. Inactive compounds tend to generate an open conformational state (e.g., distances > 10 Å), in contrast to active compounds that favor the closed state (Fig. S3 and S4) stabilizing the interaction between ASP 1088 and ARG 1157 (e.g., distances < 10 Å). These results are consistent with the two crystallographic

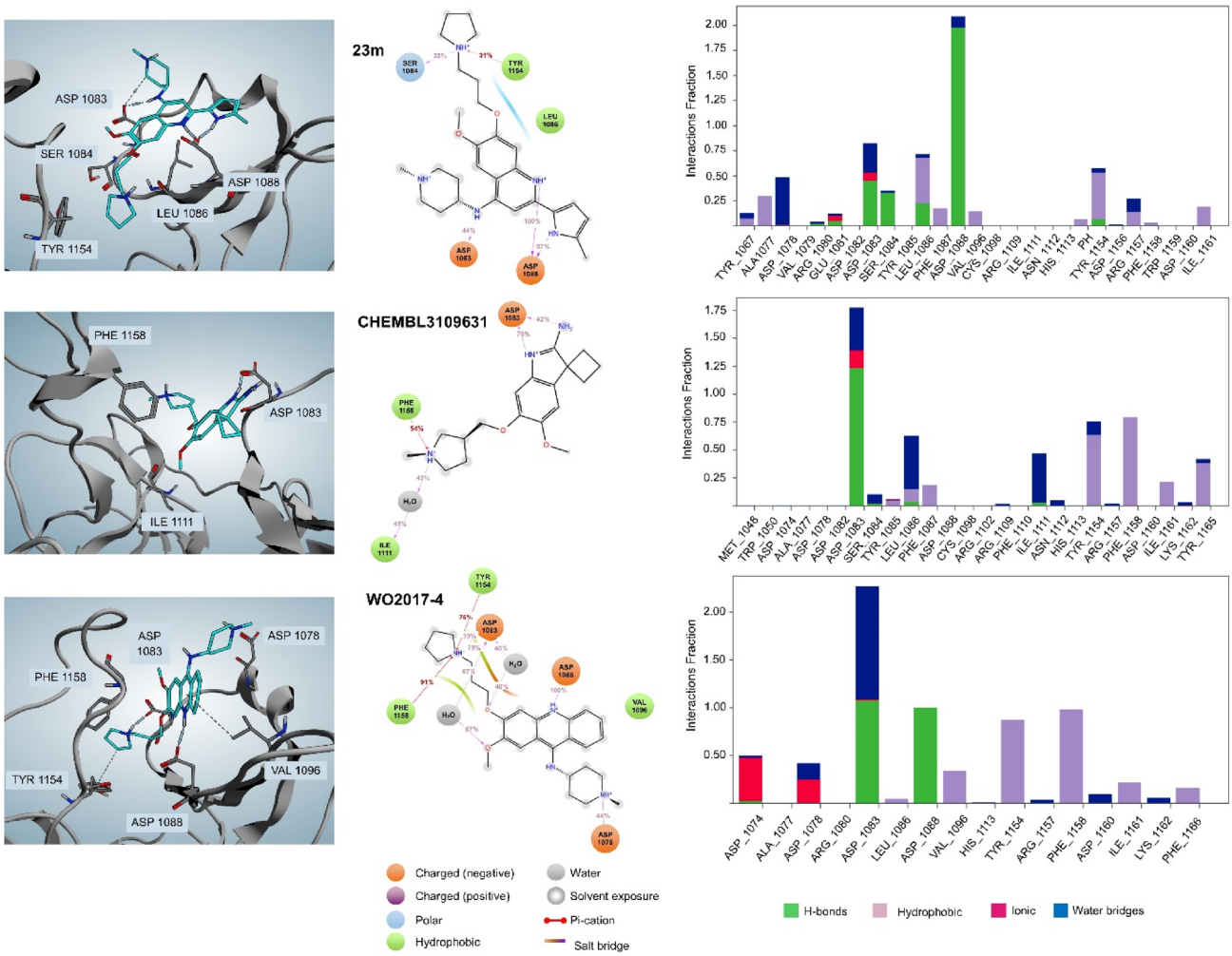

Fig. 6 G9a–most active compounds contacts analysis during the 100 ns MD simulations. The plots show the fraction and type of interactions with representative compounds during the simulation. Binding poses for each compound are shown at the end of the current MD simulation. The end-structure of G9a (for each case) is shown

states reported for G9a (PDB IDs: 3RJW and 4NVQ respectively). Conformational changes in the SET domain have been described for other lysine methyltransferases, using computational and experimental methods, breaking the paradigm of the existence of a unique conformational state associated with methylating action in this type of targets [43–45]. Recently, Shin C. et al. demonstrated the existence of specific conformational changes in the SET domain associated with the type of substrate recognized [46].

Conclusions and perspectives

The concept of epigenetic pharmacology (epi-pharmacology) is increasingly relevant [47, 48]. In parallel, activity landscape modeling has contributed to the characterization of the epigenetic relevant chemical space [49, 50]. Therefore, this study that combines activity landscape modeling with molecular modeling augmented the SAR information for G9a inhibitors. Activity landscape modeling allowed the identification of structurally different compounds, but with similar biological activity (scaffold hops) against G9a, which will facilitate the generation of more robust SAR studies. Based on the activity landscape, the set of G9a compounds analyzed in this study are suitable for the generation of QSAR models [51]. The results of this work further supported the convenience of exploring different scaffolds for the molecular recognition of small molecules with G9a. The scaffold analysis revealed that the new approaches should consider exploring in more detail the influence of more and diverse scaffolds on molecular recognition. Thus far, three major chemical scaffolds have been explored (scaffold IDs **S17**, **S62**, and **S87**), however, one of the most promising and least studied scaffolds is **s59**.

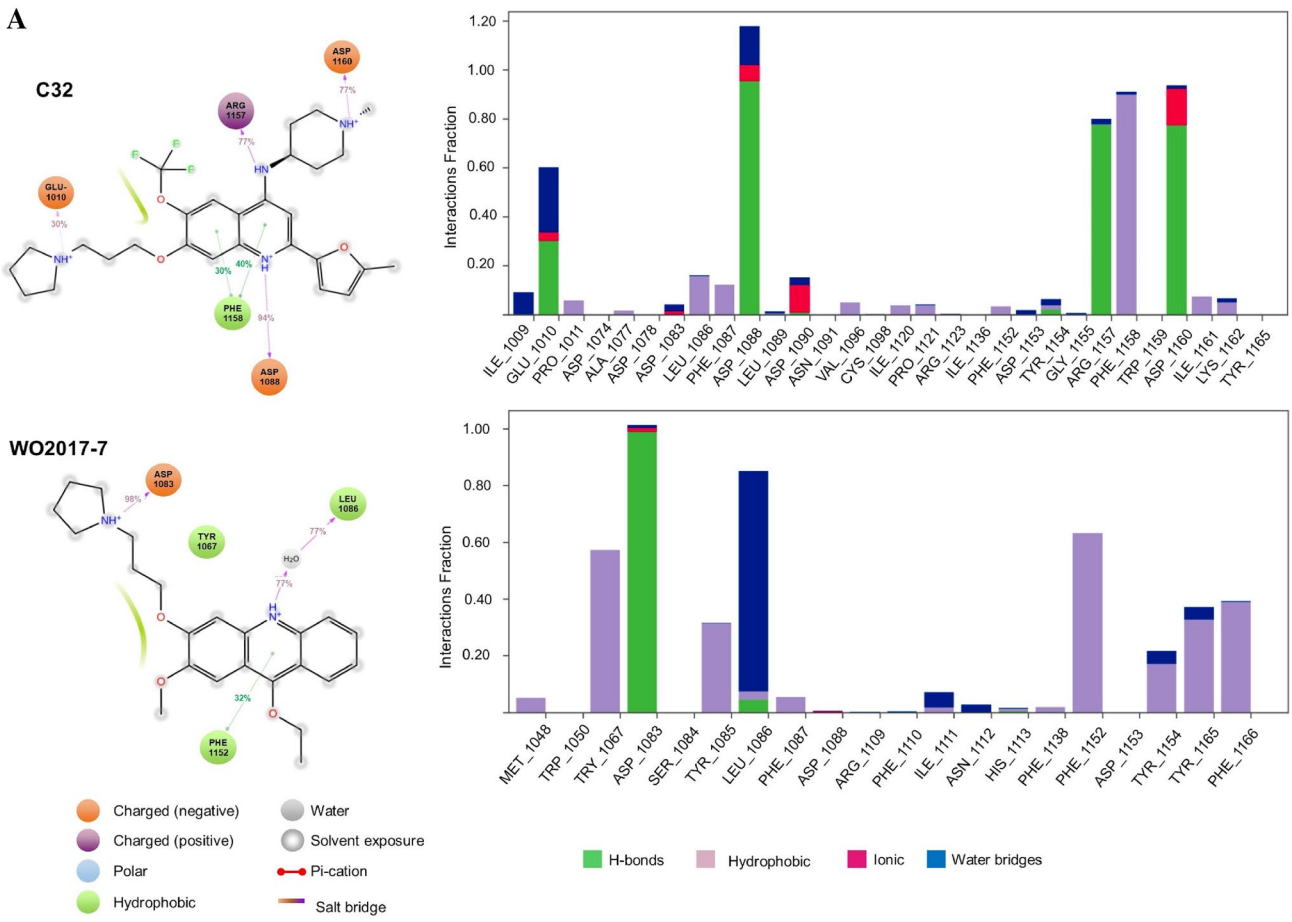

Fig. 7 G9a–inactive compounds contacts analysis during the 100 ns MD simulations. **a** The figures show the fraction and type of interactions with representative ligands during the simulation. **b** Binding poses at the beginning (orange) and at the end (cyan) of the MD simulation. The end-structure of G9a is shown

Docking and MD studies further emphasized the importance of the interactions between lysine mimetic substituents with G9a (LEU 1086, TYR 1154 and PHE 1158). From the MD simulations it was concluded that, despite the fact the chemical structures of the ligands is different, they maintain key interactions with ASP 1083 and ASP 1088. At the structural level, hypothetical conformational states (open and closed) on SET domain of G9a were identified. These

conformational states are apparently associated with interactions between the ligands and ASP 1088 and ARG 1157. A perspective of this work is conducting 3D-QSAR studies, as well as in silico scaffold replacement approaches with the long-term goal of identifying novel and alternative chemical entities with activity against G9a.

Authors contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by EL-L and OR. The first draft of the manuscript was written by EL-L and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was funded by the School of Chemistry of the Universidad Nacional Autónoma de México (UNAM), the *Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica* (PAPIIT) grant number IA203718, UNAM and the *Consejo Nacional de Ciencia y Tecnología* (CONACyT) grant number 282785. We also thank the program *Nuevas Alternativas para el Tratamiento de Enfermedades Infecciosas* NUATEI-UNAM for funding. We thank the Foundation for Applied Medical Research, University of Navarra (Pamplona, Spain), as well as Fundación Fuentes Dutor, for financial support.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

1. Miranda-Gonçalves V, Lameirinhas A, Henrique R, Jerónimo C (2018) Metabolism and epigenetic interplay in cancer: regulation and putative therapeutic targets. Front Genet 9:427. https://doi.org/10.3389/fgene.2018.00427
2. Rabal O, Castellar A, Oyarzabal J (2018) Novel pharmacological maps of protein lysine methyltransferases: key for target deorphanization. J Cheminform 10:32. https://doi.org/10.1186/s13321-018-0288-5
3. Rathert P, Dhayalan A, Murakami M et al (2008) Protein lysine methyltransferase G9a acts on non-histone targets. Nat Chem Biol 4:344–346. https://doi.org/10.1038/nchembio.88
4. Huang J, Dorsey J, Chuikov S et al (2010) G9a and Glp methylate lysine 373 in the tumor suppressor p53. J Biol Chem 285:9636–9641. https://doi.org/10.1074/jbc.M109.062588
5. Casciello F, Windloch K, Gannon F, Lee JS (2015) Functional role of g9a histone methyltransferase in cancer. Front Immunol 6:487. https://doi.org/10.3389/fimmu.2015.00487
6. Chen M-W, Hua K-T, Kao H-J et al (2010) H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. Cancer Res 70:7830–7840. https://doi.org/10.1158/0008-5472.CAN-10-0833
7. Hua K-T, Wang M-Y, Chen M-W et al (2014) The H3K9 methyltransferase G9a is a marker of aggressive ovarian cancer that promotes peritoneal metastasis. Mol Cancer 13:189. https://doi.org/10.1186/1476-4598-13-189
8. Bárcena-Varela M, Caruso S, Llerena S et al (2019) Dual targeting of histone methyltransferase G9a and DNA-methyltransferase 1 for the treatment of experimental hepatocellular carcinoma. Hepatology 69:587–603. https://doi.org/10.1002/hep.30168
9. Segovia C, San José-Enériz E, Munera-Maravilla E et al (2019) Inhibition of a G9a/DNMT network triggers immune-mediated bladder cancer regression. Nat Med 25:1073–1081. https://doi.org/10.1038/s41591-019-0499-y
10. Imai K, Togami H, Okamoto T (2010) Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294. J Biol Chem 285:16538–16545. https://doi.org/10.1074/jbc.M110.103531
11. Antignano F, Burrows K, Hughes MR et al (2014) Methyltransferase G9A regulates T cell differentiation during murine intestinal inflammation. J Clin Invest 124:1945–1955. https://doi.org/10.1172/JCI69592
12. Benevento M, van de Molengraft M, van Westen R et al (2015) The role of chromatin repressive marks in cognition and disease: A focus on the repressive complex GLP/G9a. Neurobiol Learn Mem 124:88–96. https://doi.org/10.1016/j.nlm.2015.06.013
13. Epsztejn-Litman S, Feldman N, Abu-Remaileh M et al (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat Struct Mol Biol 15:1176–1183. https://doi.org/10.1038/nsmb.1476
14. Fukuda M, Sakaue-Sawano A, Shimura C et al (2019) G9a-dependent histone methylation can be induced in G1 phase of cell cycle. Sci Rep 9:956. https://doi.org/10.1038/s41598-018-37507-5
15. Shi Y, Desponts C, Do JT et al (2008) Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell 3:568–574. https://doi.org/10.1016/j.stem.2008.10.004
16. Charles MRC, Dhayalan A, Hsieh H-P, Coumar MS (2019) Insights for the design of protein lysine methyltransferase G9a inhibitors. Future Med Chem 11:993–1014. https://doi.org/10.4155/fmc-2018-0396
17. Zang L, Kondengaden SM, Zhang Q et al (2017) Structure based design, synthesis and activity studies of small hybrid molecules as HDAC and G9a dual inhibitors. Oncotarget 8:63187–63207. https://doi.org/10.18632/oncotarget.18730
18. San José-Enériz E, Agirre X, Rabal O et al (2017) Discovery of first-in-class reversible dual small molecule inhibitors against G9a and DNMTs in hematological malignancies. Nat Commun 8:15424. https://doi.org/10.1038/ncomms15424
19. Rabal O, San José-Enériz E, Agirre X et al (2018) Discovery of reversible DNA methyltransferase and lysine methyltransferase G9a inhibitors with antitumoral in vivo efficacy. J Med Chem 61:6518–6545. https://doi.org/10.1021/acs.jmedchem.7b01926
20. Rodriguez-Madoz JR, San Jose-Eneriz E, Rabal O et al (2017) Reversible dual inhibitor against G9a and DNMT1 improves human iPSC derivation enhancing MET and facilitating transcription factor engagement to the genome. PLoS ONE 12:e0190275. https://doi.org/10.1371/journal.pone.0190275
21. Rabal O, Sánchez-Arias JA, San José-Enériz E et al (2018) Detailed exploration around 4-aminoquinolines chemical space to navigate the lysine methyltransferase G9a and DNA methyltransferase biological spaces. J Med Chem 61:6546–6573. https://doi.org/10.1021/acs.jmedchem.7b01925
22. López-López E, Prieto-Martínez FD, Medina-Franco JL (2018) Activity landscape and molecular modeling to explore the SAR of dual epigenetic inhibitors: A focus on G9a and DNMT1. Molecules. https://doi.org/10.3390/molecules23123282
23. Yu M, LI Z, Q L (2017) HISTONE METHYLTRANSFERASE INHIBITORS. 95.
24. Yu M, LI Z (2018) HISTONE METHYLTRANSFERASE INHIBITORS. 206.
25. Gaulton A, Hersey A, Nowotka M et al (2017) The ChEMBL database in 2017. Nucleic Acids Res 45:D945–D954. https://doi.org/10.1093/nar/gkw1074

26. Yongye AB, Byler K, Santos R et al (2011) Consensus models of activity landscapes with multiple chemical, conformer, and property representations. J Chem Inf Model 51:1259–1270. https://doi.org/10.1021/ci200081k
27. González-Medina M, Méndez-Lucio O, Medina-Franco JL (2017) Activity landscape plotter: a web-based application for the analysis of structure-activity relationships. J Chem Inf Model 57:397–402. https://doi.org/10.1021/acs.jcim.6b00776
28. López-López E, Naveja JJ, Medina-Franco JL (2019) DataWarrior: an evaluation of the open-source drug discovery tool. Expert Opin Drug Discov 14:335–341. https://doi.org/10.1080/17460441.2019.1581170
29. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461. https://doi.org/10.1002/jcc.21334
30. 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. https://doi.org/10.1021/ci300362x
31. Medina-Franco JL, Naveja JJ, López-López E (2019) Reaching for the bright StARs in chemical space. Drug Discov Today. https://doi.org/10.1016/j.drudis.2019.09.013
32. Bajorath J (2017) Computational scaffold hopping: cornerstone for the future of drug design? Future Med Chem 9:629–631. https://doi.org/10.4155/fmc-2017-0043
33. Maggiora GM (2006) On outliers and activity cliffs–why QSAR often disappoints. J Chem Inf Model 46:1535. https://doi.org/10.1021/ci060117s
34. Guha R, Van Drie JH (2008) Structure–activity landscape index: identifying and quantifying activity cliffs. J Chem Inf Model 48:646–658. https://doi.org/10.1021/ci7004093
35. Bemis GW, Murcko MA (1996) The properties of known drugs. 1. Molecular frameworks. J Med Chem 39:2887–2893. https://doi.org/10.1021/jm9602928
36. Naveja JJ, Norinder U, Mucs D et al (2018) Chemical space, diversity and activity landscape analysis of estrogen receptor binders. RSC Adv 8:38229–38237. https://doi.org/10.1039/C8RA07604A
37. Vedadi M, Barsyte-Lovejoy D, Liu F et al (2011) A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. Nat Chem Biol 7:566–574. https://doi.org/10.1038/nchembio.599
38. Krieger E, Joo K, Lee J et al (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. Proteins 77(Suppl 9):114–122. https://doi.org/10.1002/prot.22570
39. Chemical Computing Group Inc (2019) Molecular operating environment (MOE). Chemical Computing Group Inc., Montreal
40. Cho AE, Guallar V, Berne BJ, Friesner R (2005) Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach. J Comput Chem 26:915–931. https://doi.org/10.1002/jcc.20222
41. J. Bowers K, E. Chow D, Xu H, et al. (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters.
42. Rea S, Eisenhaber F, O'Carroll D et al (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406:593–599. https://doi.org/10.1038/35020506
43. Shi C, Rafal PW, Fanwang M et al (2019) The dynamic conformational landscape of the protein methyltransferase SETD8. eLIFE 8:e45403. https://doi.org/10.7554/eLife.45403
44. Matthieu S (2016) Chemical inhibition of protein methyltransferases. Cell Chem Biol 9:1067–1076. https://doi.org/10.1016/j.chembiol.2016.07.014
45. Po H, Shenglong W, Yingkai Z (2008) How do SET-domain protein lysine methyltransferases achieve the methylation state specificity? Revisited by ab initio QM/MM molecular dynamics simulations. J Am Chem Soc 12:3806–3813. https://doi.org/10.1021/ja075896n
46. Shi C, Kanishk K, Chamara S et al (2019) Substrate-differentiated transition states of SET7/9-catalyzed lysine methylation. J Am Chem Soc 20:8064–8067. https://doi.org/10.1021/jacs.9b02553
47. Medina-Franco JL (2016) Epi-Informatics. https://doi.org/10.1016/C2014-0-03789-6
48. Kulkarni RA, Montgomery DC, Meier JL (2019) Epigenetic regulation by endogenous metabolite pharmacology. Curr Opin Chem Biol 51:30–39. https://doi.org/10.1016/j.cbpa.2019.02.002
49. Naveja JJ, Medina-Franco JL (2018) Insights from pharmacological similarity of epigenetic targets in epipolypharmacology. Drug Discov Today 23:141–150. https://doi.org/10.1016/j.drudis.2017.10.006
50. Prieto-Martínez FD, López-López E, Eurídice Juárez-Mercado K, Medina-Franco JL (2019) Computational drug design methods—current and future perspectives. In: Roy K (ed) silico drug design. Elsevier, Amsterdam, pp 19–44. https://doi.org/10.1016/B978-0-12-816125-8.00002-X
51. Norinder U, Naveja JJ, López-López E et al (2019) Conformal prediction of HDAC inhibitors. SAR QSAR Environ Res 30:265–277. https://doi.org/10.1080/1062936X.2019.1591503