Advances in the Exploration of the Epigenetic Relevant Chemical Space

Diana L. Prado-Romero and José L. Medina-Franco*

ABSTRACT: Epigenetic drug discovery is a promising avenue to find therapeutic agents for treating several diseases and developing novel chemical probes for research. In order to identify hit and lead compounds, the chemical space has been explored and screened, generating valuable bioactivity information that can be used for multiple purposes such as prediction of the activity of existing chemicals, e.g., small molecules, guiding the design or optimization of compounds, and expanding the epigenetic relevant chemical space. Herein, we review the chemical spaces explored for epigenetic drug discovery and discuss the advances in using structure–activity relationships stored in public chemogenomic databases. We also review current efforts to chart and identify novel regions of the epigenetic relevant chemical space. In particular, we discuss the development and accessibility of two significant types of compound libraries focused on epigenetic targets: commercially available libraries for screening and targeted chemical libraries using de novo design. In this mini-review, we emphasize inhibitors of DNA methyltransferases.

1. INTRODUCTION

Epigenetics has a central role in understanding the inheritance, development, and progression of diseases. Modulation and, in particular, inhibition of epigenetic targets was considered as an approach for cancer treatment. Several conditions are currently associated with the misregulation of epigenetic targets, such as depressive disorders, multiple sclerosis, diabetes, or Alzheimer’s disease. Epigenetic target inhibitors are attractive not only for drug development but also as chemical tools to understand the underlying mechanisms of epigenetic regulation. Currently, there are 10 Food and Drug Administration (FDA)-approved drugs related to epigenetic targets, and there are several others under clinical development. Likewise, there are several chemical probes focused on epigenetic targets. Figure 1 shows the chemical structures of representative epigenetic drugs.

DNA methyltransferases (DNMTs) are major epigenetic targets with therapeutic relevance (Figure 1). The enzyme family of DNMTs promotes the covalent addition of a methyl group from S-adenosyl-L-methionine (SAM) to the 5-carbon of cytosine, mainly within CpG dinucleotides, yielding S-adenosyl-L-homocysteine (SAH). Alterations in the functions of DNMT1, DNMT3A, and DNMT3B are related to tumorigenesis and other diseases. Several reviews have been published regarding the status of the DNMTs inhibitors proposed so far. Figure 2 shows chemical structures of representative DNMT inhibitors and compounds associated with a demethylating activity.

Identifying inhibitors of epigenetic targets, including DNMTs, is an active area of research. Screening compound libraries and optimization of hit and lead compounds, from either synthetic or natural sources, has led to the population of the so-called epigenetic relevant chemical space (ERCS). More and more chemical libraries have been tested biologically, and the information has been deposited in public libraries such as ChEMBL and other chemogenomic databases as reviewed recently. Consequently, the structure–activity relationships (SAR), or, more specifically, the structure–epigenetic activity relationships (SEAR), have increased, paving the way for developing predictive models. Although several drug discovery strategies are being successfully implemented and developed to augment the epigenetic relevant chemical space, some approaches have been used on a limited basis. Examples are de novo design and assembly of focused libraries from commercial sources for screening and experimental screening.

Herein, we review advances on the application of SEAR available in public chemogenomics databases. We also discuss techniques to chart novel and unexplored regions of chemical space to identify potential hits, e.g., augment the ERCS. In particular, we discuss avenues to design or test focused chemical libraries and optimization of hit and lead compounds, from either synthetic or natural sources, has led to the population of the so-called epigenetic relevant chemical space (ERCS). More and more chemical libraries have been tested biologically, and the information has been deposited in public libraries such as ChEMBL and other chemogenomic databases as reviewed recently. Consequently, the structure–activity relationships (SAR), or, more specifically, the structure–epigenetic activity relationships (SEAR), have increased, paving the way for developing predictive models. Although several drug discovery strategies are being successfully implemented and developed to augment the epigenetic relevant chemical space, some approaches have been used on a limited basis. Examples are de novo design and assembly of focused libraries from commercial sources for acquisition and experimental screening.

Received: June 28, 2021
Accepted: August 10, 2021
Published: August 23, 2021

© 2021 The Authors. Published by American Chemical Society

http://pubs.acs.org/journal/acsono
than 5 years. Current trend in epi-informatics includes the development of predictive models using machine learning, and the systematic generation and maintenance of target-compound databases, the implementation of ETP are described in detail elsewhere. It is anticipated that ETP will help guide the identification of compounds with epigenetic activity.

4. OPPORTUNITIES TO EXPAND THE EPGENETIC RELEVANT CHEMICAL SPACE

The number of synthetically viable organic compounds exceeds 166 billion molecules. However, as for many other targets of therapeutic relevance, just a tiny fraction of that chemical space has been screened, and there is a need to keep expanding the ERCS, balancing novelty with relevance in medicinal chemistry.

4.1. Focused Libraries Commercially Available for Screening. Recently, commercially available compound libraries focused on epigenetic targets have emerged. Table 2 summarizes representative commercial libraries focused on epigenetic targets, including the number of compounds and the main targets. In total, there are over 53,000 compounds. Most of them are commercialized as epigenetic-focused libraries in different designs. Full details of the development and implementation of ETP are described in detail elsewhere. It is anticipated that ETP will help guide the identification of compounds with epigenetic activity.

Figure 1. Chemical structures of representative epigenetic drugs.
Table 1. Design and Discovery of DNMT Inhibitors

| strategy                                      | source                      | target              | results                          | ref |
|-----------------------------------------------|-----------------------------|---------------------|----------------------------------|-----|
| pharmacophore-based virtual screening and     | specs                       | DNMT3A (PDB ID: 4U7T) | compounds 40 and 40_3 with inhibitory activity (IC50 = 46.5 and 41 μM) | 8   |
| similarity searching                           |                             |                     |                                  |     |
| pharmacophore-based virtual screening         | Maybridge                   | DNMT1 (PDB ID: 3PTA) | identification of three novel hits: JFD01881, RJC02836, and RJC02837 | 9   |
| pharmacophore and docking-based virtual       | in-house database           | DNMT1 (PDB ID: 3SWR) | one derivative from an identified compound: 4b (IC50 = 4.1 μM) | 10  |
| screening                                      |                             |                     |                                  |     |
| fragment-based design                          | natural products from       | DNMT1 (PDB ID: 4WXX) | two proposals of lead compounds: HAMI 9 and HAMI 14 | 11  |
| fragment-based design and fragment merging    | natural products’ fragments | DNMT1 (PDB ID: 3AV5, 3AV6, 3PTA, 3SWR, 4WXX) | most promising hit: MAH11 | 11  |
| with SAH                                        | from PubChem                |                     |                                  |     |
| fragment-based design                          | natural products’ fragments | DNMT1 (PDB ID: 3AVS) | potential drug lead: C-7756 | 12  |
| from PubChem                                   |                             |                     |                                  |     |

“The chemical structures of selected compounds are shown in Figure 2.”
| source                        | size (initial) | size (curated) | brief description                                                                                                                                                                                                 | website accessed June 2021                                                                                     |
|------------------------------|----------------|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| ApeXBio (DiscoveryProbe Epigenetics Compound Library) | 328            | 310            | small molecules with activity against different epigenetic targets; design approach not disclaimed                                                                                                                | https://www.apexbt.com/discoveryprobetm-epigenetics-compound-library.html                                      |
| Asinex (Epigenetic Library)  | 5391           | 5313           | library focused on bromodomains and histone methyltransferase inhibitors; compounds designed by a combination of structure and ligand-based methods                                                                               | http://www.asinex.com/                                                                                                                                               |
| ChemDiv (Epigenetic Set)     | 30,431         | 27,543         | library designed for modulation for all three classes of epi-targets; molecules were designed based on a combination of structure-based (based on X-ray and nuclear magnetic resonance) and ligand-based approaches | https://www.chemdiv.com/epigenetics-library/                                                                                                                         |
| Enamine (Epigenetics Library)| 9352           | 9352           | library designed with a combination of ligand and structure-based methods; compound optimization was made via pharmacophore profiling                                                                                   | https://enamine.net/hit-finding/focused-libraries/view-all/epigenetics-libraries                                                                               |
| Life Chemicals (Epigenetics Screening Library) | 7019           | 7011           | derived with the use of similarity methods, this library is focused on methylation-related epi-enzymes                                                                                                          | https://lifechemicals.com/screening-libraries/targeted-and-focused-screening-libraries/epigenetic-screening-libraries |
| MedChemExpress (Epigenetics Library) | 700            | 650            | library designed for modulation of several epi-targets; the targets and design approaches are not disclaimed                                                                                                     | https://www.medchemexpress.com/virtual-screening/epigenetics-library.html                                                                                           |
| OTAVA DNMT1 (DNMTs Targeted Libraries) | 466            | 399            | drug-like compounds selected from virtual screening using docking and pharmacophore modeling                                                                                                                      | https://www.otavachemicals.com/targets/dnmt1-and-dnmt3b-targeted-libraries                                                                                         |
| OTAVA DNMT3b (DNMTs Targeted Libraries) | 1261           | 1230           | drug-like compounds selected from virtual screening using docking and pharmacophore modeling                                                                                                                     | https://www.otavachemicals.com/targets/dnmt1-and-dnmt3b-targeted-libraries                                                                                         |
| Targetmol (Epigenetics Compound Library) | 932            | 859            | a set of epi-regulators whose primary focus is lead optimization and HTS; design approach not disclaimed                                                                                                         | https://www.targetmol.com/                                                                                                                                            |
| TocrisScreen Epigenetics     | 101            | 99             | collection of small molecules covering more than 40 epigenetic targets (including readers, writers, erasers, and transcriptional modulators); design approach not disclaimed                                                                 | https://www.tocris.com/products/tocriscreen-epigenetics-library_6801                                                                                               |
| SelleckChem (Epigenetics Compound Library) | 699            | 677            | focused on experimental tests and for HTS validation, this library contains inhibitors for several epi-targets; design approach not disclaimed                                                                           | https://www.selleckchem.com/screening/epigenetics-compound-library.html                                                                                             |

*a Data curation was done with a standard protocol described in ref 16 to conduct a comparative chemoinformatic profile of the compound libraries.
rotatable bonds, number of hydrogen bond donors and acceptors atoms, topological polar surface area (TPSA), and octanol–water partition coefficient (LogP). The total variance captured by the first two principal components is 72%. The properties that contributed most to the first two principal components were TPSA and LogP. The visualization indicates that most libraries have comparable and drug-like properties, as designed and prefiltered by the chemical companies selling the

Figure 3. Visual representation of the chemical space of 11 compound libraries focused on epigenetic targets in Table 2. The visualization was done with a principal component analysis of six autoscaled properties of pharmaceutical relevance. Each compound library is plotted using the same coordinates. The plot on the bottom-right corner shows all 11 compound libraries plotted on the same graph. The percentage of variance recovered by each principal component is indicated along the X- and Y-axis.
libraries. Compounds in the Enamine collection populate a more constrained region of the chemical space compared with other libraries such as APExBIO and SelleckChem, covering broader areas of the property-based chemical space.

Recently, the need to include privileged substructures present in the collection of approved drugs for clinical use in the focused libraries for epigenetic drug discovery has been highlighted.15

4.2. De Novo Design.

The main goal of de novo design is to generate new molecules with desired properties. Herein, we focus on the design of bioactive hits. A schematic summary of de novo design is presented in Figure 4. The method has four general stages. First of all, primary target constraints are established. These refer to the information related to ligand−receptor interaction.17 Three-dimensional (3D) coordinates are needed to determine the constraints. The information could be obtained from the 3D structure of the target (X-ray or nuclear magnetic resonance), a homology model, or a pharmacophore model. The last one enables the use of known ligands to define primary target constraints, even if the 3D structure of the receptor is not available. Second, atoms or fragments could be the building blocks of the new molecules. Atom-based construction will likely give more diversity to the chemical space. However, proposed structures usually are synthetically challenging or unfeasible. In that sense, fragment-based approaches could overcome this disadvantage as they are typically larger and retain more chemical information. Fragments are drawn from sources like general screening databases, natural products, or focused libraries (vide supra). Once the building blocks are established, the algorithm continues to the ligand construction. This stage requires a seed atom or fragment. Techniques for molecular assembly include growing, linking, alignment-based, lattice-based, reaction-based, and graph-based methods.18 Some software programs or algorithms also have strategies to address the issue of combinatorial search explosion, like breadth-first and depth-first search, Monte Carlo search combined with Metropolis criterion or evolutionary algorithms. Finally, the scoring function guides the ligand construction and defines the best candidate compounds. Functions could also consider the 3D structure of the target (structure-based) or information from known ligands (ligand-based). In general, structure-based methods are classified into force-field, knowledge-based, and empirical functions. Table 3 lists examples of de novo design programs. The description of the algorithm includes the selected building blocks, the technique for the ligand construction, the search strategy for the combinatorial problem, and the scoring function. Earlier de novo design programs considered atoms preferably as building blocks until synthetic tractability became an issue. Consequently, most recent programs typically consider fragments. Synthetic feasibility quantified with different approaches is also addressed. Current programs include fragment databases generated from already known drug-like compounds (LEA3D) or incorporate a new score to evaluate synthetic accessibility (PhDD). Current programs also incorporate genetic algorithms to optimize the searching (Table 3).

**Table 3. De Novo Design Programs**

| Program   | Year | Building Blocks | Ligand Construction | Search Strategy | Scoring Function               | Refs     |
|-----------|------|-----------------|---------------------|-----------------|--------------------------------|----------|
| LUDI      | 1992 | fragments       | growing, linking    | Breadth-first search | empirical                     | 19,20    |
| LEA3D     | 2005 | fragments       | growing, linking    | genetic algorithm | user-defined fitness function  | 21       |
| PhDD      | 2010 | fragments       | linking             | random           | fit value (alignment with pharmacophore model) | 22       |
| eLEA3D    | 2010 | fragments       | growing, linking    | genetic algorithm | user-defined fitness function  | 23       |
| DOGS      | 2012 | fragments       | reaction-based      | deterministic process | similarity with reference ligand | 24       |
| DENOPTIM  | 2019 | fragments       | graph-based         | genetic algorithm | customizable fitness function | 25       |
| LigBuilder V3 | 2020 | fragments       | growing, linking    | genetic algorithm | empirical                     | 26       |

*Free software for academic use.*
In epigenetic drug discovery, de novo design has been applied to find new inhibitors of the (CREB (cAMP responsive element binding protein) binding protein) bromodomain, using the program LUDI (a structure-based approach). In this example,

| target | BIOFACQUIM (510) | MEGx (3707) | NuBBE (1995) | Fungi (178) | Marines (5371) | Cyanobacteria (225) | COCONUT (350,070) |
|--------|------------------|-------------|--------------|-------------|---------------|-------------------|--------------------|
| APEX1  | 433              | 3405        | 1614         | 139         | 4109          | 151               | 237,341           |
| ATM    | 322              | 2812        | 1255         | 116         | 4233          | 203               | 258,867           |
| AURKA  | 75               | 712         | 338          | 21          | 717           | 7                 | 60,780            |
| AURKB  | 137              | 850         | 533          | 44          | 1301          | 68                | 97,705            |
| BRD2   | 310              | 2414        | 1143         | 119         | 3407          | 141               | 209,233           |
| BRD4   | 117              | 1129        | 416          | 63          | 1515          | 141               | 100,391           |
| BRPF1  | 0                | 15          | 19           | 0           | 12            | 0                 | 2113              |
| CARM1  | 38               | 294         | 171          | 4           | 403           | 31                | 64,177            |
| CDK1   | 264              | 1539        | 871          | 81          | 2071          | 103               | 127,126           |
| CDK2   | 44               | 401         | 177          | 20          | 894           | 37                | 40,350            |
| CDK5   | 32               | 170         | 87           | 18          | 242           | 3                 | 11,396            |
| CDK7   | 3                | 16          | 5            | 1           | 90            | 3                 | 7850              |
| CHEK1  | 3                | 49          | 15           | 6           | 144           | 3                 | 8989              |
| CHUK   | 0                | 0           | 1            | 0           | 4             | 0                 | 336               |
| CREBBP | 26               | 181         | 83           | 18          | 415           | 59                | 50,678            |
| DAPK3  | 1                | 6           | 16           | 0           | 17            | 1                 | 1559              |
| DNMT1  | 0                | 0           | 0            | 0           | 0             | 0                 | 3 (KNOWN)         |
| DOT1L  | 0                | 4           | 0            | 0           | 5             | 15                | 3199              |
| EHMT2  | 4                | 17          | 9            | 9           | 40            | 13                | 6828              |
| EP100  | 132              | 1009        | 451          | 88          | 2080          | 98                | 119,920           |
| EZH2   | 24               | 200         | 84           | 19          | 396           | 61                | 34,732            |
| HDAC1  | 100              | 1236        | 352          | 57          | 1365          | 113               | 98,671            |
| HDAC10 | 8                | 151         | 39           | 9           | 302           | 85                | 21,772            |
| HDAC11 | 27               | 281         | 130          | 17          | 483           | 60                | 28,158            |
| HDAC2  | 47               | 302         | 212          | 37          | 661           | 72                | 47,019            |
| HDAC3  | 27               | 184         | 132          | 21          | 516           | 54                | 48,494            |
| HDAC4  | 28               | 178         | 76           | 15          | 435           | 44                | 40,606            |
| HDAC5  | 7                | 86          | 30           | 7           | 163           | 20                | 21,118            |
| HDAC6  | 51               | 453         | 202          | 32          | 812           | 66                | 49,533            |
| HDAC7  | 44               | 252         | 123          | 25          | 490           | 25                | 38,319            |
| HDAC8  | 103              | 453         | 405          | 37          | 874           | 22                | 53,706            |
| HDAC9  | 0                | 20          | 5            | 0           | 48            | 0                 | 5223              |
| JAK2   | 14               | 112         | 95           | 13          | 355           | 10                | 28,058            |
| KAT2B  | 4                | 20          | 7            | 3           | 71            | 10                | 8507              |
| KDM1A  | 55               | 244         | 223          | 16          | 410           | 72                | 47,451            |
| KDM4A  | 1                | 15          | 9            | 0           | 13            | 0                 | 2343              |
| KDM4C  | 54               | 549         | 248          | 28          | 855           | 51                | 85,910            |
| KDM4E  | 266              | 1654        | 818          | 77          | 1994          | 40                | 115,675           |
| KDM5A  | 64               | 409         | 160          | 21          | 770           | 79                | 64,053            |
| KDM6B  | 0                | 0           | 0            | 0           | 0             | 0                 | 9                 |
| L3MBTL1 | 0             | 0           | 0            | 0           | 0             | 0                 | 336               |
| PARG   | 0                | 0           | 0            | 0           | 0             | 0                 | 197               |
| PARP1  | 144              | 1434        | 610          | 85          | 2469          | 171               | 168,852           |
| PKN1   | 50               | 267         | 129          | 12          | 220           | 27                | 28,310            |
| PRKAA1 | 15               | 11          | 16           | 0           | 28            | 0                 | 3600              |
| PRKCB  | 77               | 570         | 218          | 63          | 1432          | 51                | 56,473            |
| PRKCD  | 182              | 1799        | 685          | 106         | 2700          | 87                | 112,135           |
| PRKDC  | 107              | 820         | 352          | 39          | 1193          | 65                | 90,455            |
| PRMT3  | 0                | 2           | 0            | 0           | 0             | 0                 | 797               |
| RPS6KA5 | 17            | 87          | 86           | 1           | 105           | 1                 | 6038              |
| SIRT1  | 5                | 52          | 27           | 4           | 108           | 13                | 8878              |
| SIRT2  | 13               | 13          | 9            | 0           | 20            | 11                | 2190              |
| SIRT3  | 0                | 0           | 0            | 0           | 0             | 0                 | 72                |
| TOP2A  | 3                | 2           | 5            | 0           | 16            | 0                 | 1026              |
| USP7   | 30               | 350         | 114          | 11          | 592           | 17                | 29,020            |

a Compounds classified as potentially active with the free server Epigenetic Target Profiler. b Number of compounds in each library.
the authors added new fragments to the core of the compound CPI-637. The search was guided with the information on crucial residues. As a result, structures with similar binding affinities to the original molecule were obtained. These results state the applicability of de novo design to aim the discovery of novel epigenetic drugs.

4.3. Profiling of Large Libraries of Natural Products. As with several other therapeutic targets, natural products are sources of compounds or starting points to inspire the development of active compounds with epigenetic targets. Natural products have been the sources of inhibitors of DNMTs and molecules with demethylating activity (Figure 2). One of the most recent DNMT1 inhibitors identified from natural sources is theaflavin, a polyphenol compound in black tea and previously identified as an inhibitor of DNMT3B. The importance of natural products as sources of epigenetic compounds, the increasing availability of large natural product databases in the public domain, and overall significance of the emerging “natural products informatics” (e.g., rational application of informatics approaches to enhance natural product-based drug discovery) encourages the continued systematic virtual screening of natural products databases to identify compounds with potential epigenetic inhibitory activity, including DNMT inhibitors.

To illustrate this point, Table 4 summarizes the predicted classification of seven natural product libraries in the public domain. The predictions were made with the algorithms implemented in the free server ETP described in section 3. The natural product databases used in the epigenetic target prediction (e.g., epigenetic target fishing) were the same files previously curated and used in diversity profiling reported in detail elsewhere. As discussed in those studies, the databases include a collection of ≤10 natural products isolated and characterized in Mexico (BIOFACQUIM), 3707 compounds from a screening library (MEGs), 1995 natural products from Brazil (NUBBEDB), 178 fungi metabolites (Fungi), 5371 marine natural products, 235 cyanobacterial metabolites, and 350 070 molecules from the Collection of Open Natural Products (COCONUT). As discussed elsewhere, COCONUT is one of the largest public collections of natural products available today. Table 4 reports the number of compounds predicted as “active” (defined in ETP as molecules with at least IC_{50} = 10 μM) with confidence in the range quartiles 1-4. The predictions were made with the consensus model Morgan::SVM–RDKit::SVM as described elsewhere. Results of the profiling indicated that three compounds (nucleoside analogues) in COCONUT had reported epigenetic activity with DNMT1 (including Vidaza and S-(S'-adenosyl)-L-homocysteine, AdoHcy). However, there are no other compounds with predicted activity against DNMT1.

5. CONCLUSIONS AND FUTURE DIRECTIONS

The ERCS is expanding, as reflected by the increase in the number of SEAR. The ERCS’ expansion is being driven by (1) novel and multiple compound libraries focused on epigenetic targets, ready for experimental testing, and (2) the growing interest to build de novo chemical libraries focused on epigenetic targets. The concurrent growth in the experimental screening information has encouraged the development of machine learning models, now implemented into a free Web server to predict small molecules epigenetic activity, including natural products, and guide the design of novel compounds. It is anticipated that as the increase of screening data continues, more predictive models will be developed. Perspectives in the field include augmenting the implementation of de novo design of compounds as candidate epi-drugs, experimental screening of the focused libraries already available for testing, and continuing epigenetic target profiling (e.g., inverse virtual screening) of natural product databases followed by testing of the computational hits and further refinement using other approaches such as structure-based virtual screening.

AUTHOR INFORMATION

Corresponding Author
José L. Medina-Franco — DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, México City 04510, Mexico; orcid.org/0000-0003-4940-1107; Phone: +52-55-5622-3899; Email: medinajl@unam.mx, jose.medina.franco@gmail.com

Author
Diana L. Prado-Romero — DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c03389

Notes
The authors declare no competing financial interest.

Biographies
Diana L. Prado-Romero received her M.Sc. degree from the National Autonomous University of Mexico (UNAM) in 2020. She began her research at the School of Chemistry, UNAM, in 2017, focused on medicinal chemistry under the supervision of Dr. Rafael Castillo and Dr. Alicia Hernández-Campos. She is currently working to obtain her Ph.D. with Dr. Medina-Franco at DIFACQUIM. Her research interest is the design of novel inhibitors of DNA methyltransferases.

José L. Medina-Franco, Ph.D., FRSC, is a full professor in the Department of Pharmacy of the School of Chemistry at the National Autonomous University of Mexico (UNAM). He received his Ph.D. degree at UNAM. In 2005, Dr. Medina-Franco joined the University of Arizona as a postdoctoral fellow under the supervision of Prof. Gerald Maggiora, and he was named Assistant Member at the Torrey Pines Institute for Molecular Studies in Florida in August 2007. In 2013, he conducted research at the Mayo Clinic. In 2014, he joined UNAM and now is a Full-Time Research Professor. He leads the DIFACQUIM research group at UNAM. The research focus is on computer-aided drug design using chemoinformatics, machine learning, and molecular modeling with applications on epigenetic targets and natural products. In 2017, he was named Fellow of the Royal Society of Chemistry.
Acknowledgments

We thank the support of DGAPA, UNAM, Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), Grant No. IN201321. D.L.P.-R. is thankful to Consejo Nacional de Ciencia y Tecnología (CONACyT), Mexico for the Ph.D. scholarship number, 888207. The technical support of Norberto Sánchez-Cruz, K. Euridice Juárez-Mercado, and Ana L. Chávez-Hernández is greatly acknowledged. We also thank Jocelyn Salazar and Alexis Padilla for providing the data for Figure 1 and Table 1, respectively.

Abbreviations

3D, three-dimensional; COCONUT, Collection of Open Natural Products; DNMTs, DNA methyltransferases; ERCS, epigenetic relevant chemical space; ETP, Epigenetic Target Profiler; FDA, Food and Drug Administration; HTS, high-throughput screening; LogP, octanol–water partition coefficient; SAH, S-adenosyl-L-homocysteine; SAM, S-adenosyl-L-methionine; SAR, structure–activity relationships; SEAR, structure–epigenetic activity relationships; TPSA, topological polar surface area

References

1. Sessions, Z.; Sánchez-Cruz, N.; Prieto-Martínez, F. D.; Alves, V. M.; Santos, H. P.; Muratov, E.; Troshpa, A.; Medina-Franco, J. L. Recent progress on cheminformatics approaches to epigenetic drug discovery. Drug Discovery Today 2020, 25, 2268–2276.
2. Quinlan, R. B. A.; Brennan, P. E. Chemogenomics for drug discovery: clinical molecules from open access chemical probes. RSC Chem. Biol. 2021, 2, 759–795.
3. Ganesan, A.; Arimondo, P. B.; Rotis, M. G.; Jeronimo, C.; Berdasco, M. The timeline of epigenetic drug discovery: From reality to dreams. Clin. Epigenet. 2019, 11, 174.
4. Lyko, F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. Nat. Rev. Genet. 2018, 19, 81.
5. Juárez-Mercado, K. E.; Prieto-Martínez, F. D.; Sánchez-Cruz, N.; Peña-Castillo, A.; Prada-Gracia, D.; Medina-Franco, J. L. Expanding the structural diversity of dna methyltransferase inhibitors. Pharmaceuticals 2021, 14, 17.
6. Gortari, E. F.-d.; Medina-Franco, J. L. Epigenetic relevant chemical space: A chemoinformatic characterization of inhibitors of DNA methyltransferases. RSC Adv. 2015, 5, 87465–87476.
7. Yu, J.; Xie, T.; Wang, Z.; Wang, X.; Zeng, S.; Kang, Y.; Hou, T. DNA methyltransferases: Emerging targets for the discovery of inhibitors as potent anticancer drugs. Drug Discovery Today 2019, 24, 2323–2331.
8. Shao, Z.; Xu, P.; Xu, W.; Li, L.; Liu, S.; Zhang, R.; Liu, Y.-C.; Zhang, C.; Chen, S.; Luo, C. Discovery of novel DNA methyltransferase 3a inhibitors via structure-based virtual screening and biological assays. Bioorg. Med. Chem. Lett. 2017, 27, 342–346.
9. Krishna, S.; Shukla, S.; Lakra, A. D.; Meeran, S. M.; Siddiqui, M. I. Identification of potential inhibitors of DNA methyltransferase 1 (DNMT1) through a pharmacoaphore-based virtual screening approach. J. Mol. Graphics Modell. 2017, 75, 174–188.
10. Hassanazadeh, M.; Kasymov, R.; Mahernia, S.; Adib, M.; Emperle, M.; Dukatz, M.; Bashtrykov, P.; Jeltsch, A.; Amanlou, M. Discovery of novel and selective DNA Methyltransferase 1 inhibitors by pharmacoaphore and docking-based virtual screening. ChemistrySelect 2017, 2, 8383–8392.
11. Saragih, M. Discovery of novel DNMT-1 inhibitor by fragment-based drug design as a potential breast cancer treatment. GEOMATE 2019, 17, 41–46.
12. Alkaff, A. H.; Saragih, M.; Imana, S. N.; Nasution, M. A. F.; Tambunan, U. S. F. Identification of DNA methyltransferase-1 inhibitor for breast cancer therapy through computational fragment-based drug design. Molecules 2021, 26, 375.