

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Inhibition of Human Enhancer of Zeste Homolog 2 (EZH2) with tambjamine analogs

Journal:	<i>Journal of Chemical Information and Modeling</i>
Manuscript ID	ci-2017-00178g.R1
Manuscript Type:	Article
Date Submitted by the Author:	25-May-2017
Complete List of Authors:	Kotev, Martin; Barcelona Supercomputing Center, Life Science Manuel-Manresa, Pilar ; University of Barcelona, Cancer Cell Biology Research Group, Department of Pathology and Experimental Therapeutics Hernando , Elsa ; Universidad de Burgos, Química Soto-Cerrato, Vanessa; University of Barcelona, Cancer Cell Biology Research Group, Department of Pathology and Experimental Therapeutics Orozco, Modesto; Institute for Research in Biomedicine, Quesada, Roberto; Universidad de Burgos, Química Perez-Tomas, Ricardo; Universidad de Barcelona, Department of Pathology and Experimental Therapeutics Guallar, Victor; Barcelona Supercomputing Center, Life Science

SCHOLARONE™
Manuscripts

Inhibition of Human Enhancer of Zeste Homolog 2 (EZH2) with tambjamine analogs

*Martin Kotev^{†, ‡, †}, Pilar Manuel-Manresa[§], Elsa Hernando^{¶, †}, Vanessa Soto-Cerrato^{§, *},
Modesto Orozco[‡], Roberto Quesada[¶], Ricardo Pérez-Tomás[§] and Victor Guallar^{†, †, *}*

[†]Joint BSC-CRG-IRB Research Program in Computational Biology. Barcelona Supercomputing Center, c/ Jordi Girona 29, 08034 Barcelona, Spain

[‡]Institute for Research in Biomedicine (IRB Barcelona), Baldori i Reixac 8, Barcelona 08028, Spain

[§]Cancer Cell Biology Research Group, Department of Pathology and Experimental Therapeutics. University of Barcelona. E-08907 Barcelona, Spain

[¶]Department of Chemistry, Faculty of Science, University of Burgos, 09001 Burgos, Spain

[†]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Spain

Abstract

Combining computational modeling, de novo compound synthesis and *in vitro* and cellular assays, we have performed an inhibition study against the enhancer of zeste homolog 2 (EZH2) histone-lysine N-methyltransferase. This enzyme is an important catalytic component of the PRC2 complex whose alterations have been associated to different cancers. We introduce here several tambjamine-inspired derivatives with low micromolar *in vitro* activity that produce a significant decrease in histone 3 trimethylation levels in cancer cells. We demonstrate binding at the methyl transfer active

1
2
3 site, showing, in addition, that the EZH2 isolated crystal structure is capable of being
4
5 used in molecular screening studies. Altogether, this work provides a succesful
6
7 molecular model that will help in the identification of new specific EZH2 inhibitors and
8
9 identify a novel class of tambjamine-derived EZH2 inhibitors with promising activities
10
11 for their use in cancer treatment.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1. Introduction

Enhancer of zeste homolog 2 (EZH2) is a histone-lysine N-methyltransferase enzyme, the catalytic component of the Polycomb Repressive Complex 2 (PRC2) that acts as a transcriptional repressor. It catalyzes the addition of three methyl groups, from the *S*-adenosyl-L-methionine (SAM) cofactor, to lysine 27 of Histone H3, one of the five main histone proteins involved in the structure of chromatin in eukaryotic cells. This modification facilitates chromatin compaction and gene silencing of tumor suppressor genes in cancer cells¹. Since the early observation that EZH2 overexpression was associated with progression of prostate cancer², several EZH2 alterations have been associated to other different cancers, including breast, colorectal and lung³⁻⁵. Moreover, EZH2 overexpression has also been associated with metastasis and poor clinical outcome⁶. Due to these evidences, we find several efforts in developing specific inhibitors for EZH2, including various preclinical studies and several human phase 1 and 2 clinical trials^{3,7,8}.

Such biomedical interest has been translated in recent efforts to obtain crystal structures. Since 2013, there is public access to two human EZH2 crystal structures, protein data bank (PDB) entries 4MI0 and 4MI5. Since EZH2 without PRC2 components (SUZ12 and EED) is inactive, there is some controversy about the utility of these structures for drug design. In addition, during 2016 several structures of the human PRC2 complex were solved, including one with a pyridone inhibitor (pdb entries 5IJ7, 5IJ8) and a *S*-adenosyl homocysteine (SAH) cofactor (pdb entry 5HYN, clearly indicating the methyl transfer active site)⁹. Interestingly, these structures reveal an unusual binding mode. The inhibitor pocket is formed on the interface of the I-SET domain of EZH2, a stretch of 17 residues called by the authors activation loop, and the EED protein part of PRC2 complex. Comparing the EZH2/pyridone and the EZH2/SAH

1
2
3 structures, we can observe that the only overlapping part of the inhibitor with the SAH
4 molecule is the pyridone moiety, participating in a similar H-bond networking with
5 Trp624.
6
7

8
9 We are involved in an anticancer drug discovery program focused at small
10 molecules capable of facilitating the transmembrane transport of anions¹⁰, some of them
11 (which we address here) inspired in the structure of marine secondary metabolites
12 tambjamines¹¹. We have synthesized and studied synthetic analogs of these natural
13 products, which proved to possess potent anticancer activity *in vitro*, including activity
14 against cancer stem cells¹². We are also studying in depth the molecular mechanisms
15 responsible for this activity¹³. A series of these derivatives were submitted to the Target
16 Drug Discovery Program, within the Open innovation drug discovery panel, operated by
17 Lilly^a. Preliminary results from this program indicated important activity towards their
18 oncology-EZH2 target. Moreover, our cellular assay showed an important reduction of
19 histone H3 tri-methylation levels using anti-trimethyl-Histone H3 (Lys27) antibody in
20 two selected compounds. These encouraging results prompted us to carry out a study to
21 design optimized candidates for this target using computational modeling, to synthesize
22 these molecules and to test their *in vitro* efficacy.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 We report here the results of this broad analysis. Our study indicates for the first
41 time that our synthetic tambjamine molecules bind in the methyl transfer active site
42 region of the enzyme. These results suggest that the inhibition of EZH2 may be
43 involved in the cytotoxic properties of these molecules. Moreover, the knowledge of the
44 binding mode, coupled to extensive state-of-the-art induced fit simulations, allowed us
45 to design and synthesize 4 new molecules, all presenting low micromolar activity,
46
47
48
49
50
51
52
53
54
55

56
57
58 ^a <https://openinnovation.lilly.com/dd/>
59
60

1
2
3 supporting that the *in silico* model used is valid for the identification of novel effective
4
5 EZH2 inhibitors.
6
7
8
9

10 **2. Results and Discussion**

13 **Computational Results**

14
15 The exploration of possible binding modes between the studied tambjamine analogs and
16
17 EZH2 started with the standard practice of finding possible binding sites followed by a
18
19 docking procedure. From the apo EZH2 crystal (PDB entry 4MI0), we collected the first
20
21 nine sites identified with SiteMap, which included the SAM catalytic binding site, and
22
23 dock in each of them all the tambjamine analogs (Figure 1). Compounds **1**, **2**, **3** and **4**
24
25 were selected as negative controls (were found negative in the biological assays);
26
27 whereas all other 7 molecules (compounds **5-11**) showed experimental low μM inhibitor
28
29 activities against EZH2 in the target drug discovery program performed at Lilly (Table
30
31 1). The complete results of the Glide XP docking scores for the nine discovered binding
32
33 sites are shown in supporting information Table S1. All of them indicate similar (poor)
34
35 scoring values between -3.0 to -6.5 kcal/mol. While absolute docking scores are often
36
37 ligand dependent, low micromolar activities tend to correlate with values < -10 . Thus,
38
39 based on the scores obtained from docking, significant inhibition activity for EZH2
40
41 could not be predicted for the studied compounds. Moreover, there seems to be no
42
43 preferential binding of the ligands into the methyl transfer catalytic site. Nevertheless,
44
45 the negative controls presented overall (slightly) lower scores than the rest of molecules,
46
47 introducing some differentiation between the negative controls **1-4** and the other active
48
49 ligands.
50
51
52
53
54
55
56
57
58
59
60

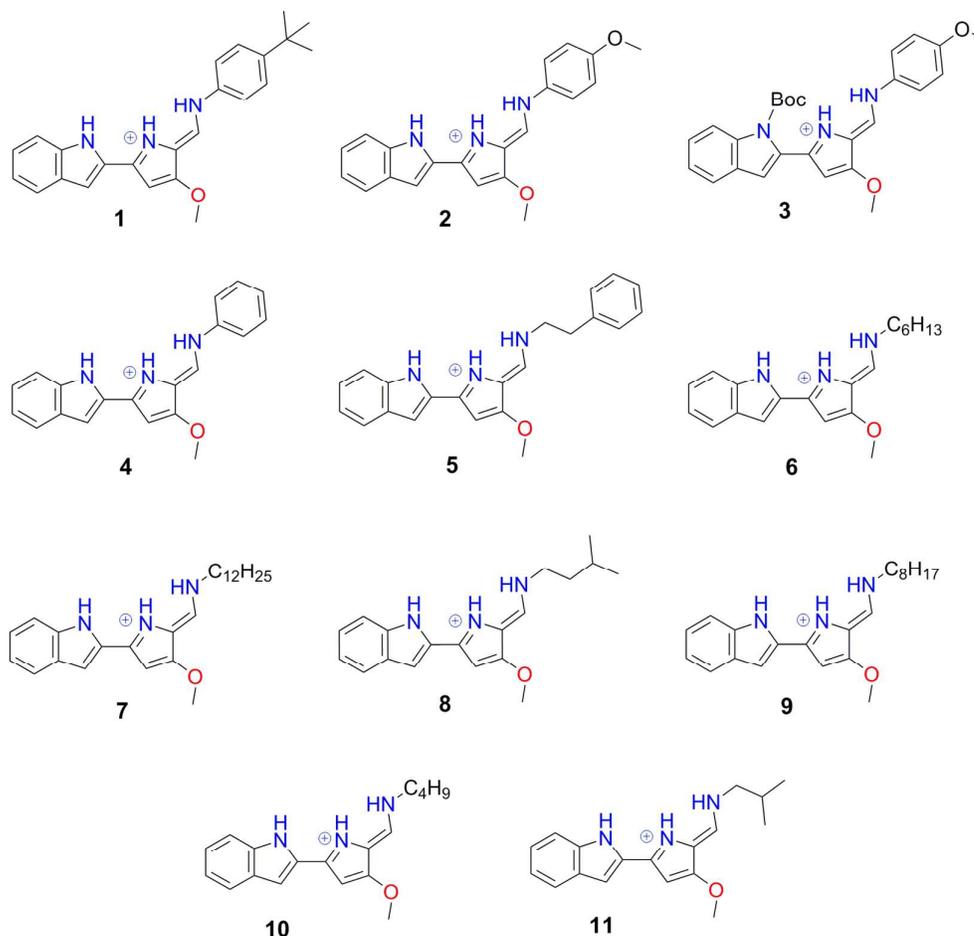
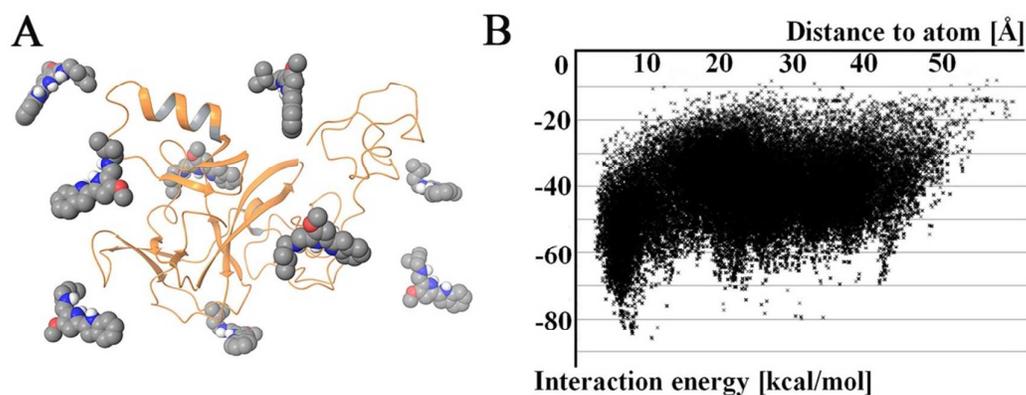


Figure 1. Structure of tambjamine analogs used in the initial docking studies. Negative control systems (no detectable inhibition on biological assays) involve molecules **1-4**.

Since standard docking techniques do not sample the receptor conformational changes, both at the induced fit or conformational selection level (only a limited ligand flexibility is introduced), we turned into PELE for a more robust exploration of the protein-ligand energy landscape. In recent studies in mTOR²⁵ and BCL²⁶, for example, we have shown the necessity of introducing such type of analysis in order to correlate with experimental binding affinities. The first simulation involved a non-biased exploration of the entire EZH2 surface (global exploration). In such a procedure, the ligand is placed in the bulk solvent and is allowed to explore all the space without any restriction or predefined

1
2
3 goal. At each PELE Monte Carlo step, the system is perturbed and relaxed, introducing
4 significant conformational changes, and the protein-ligand interaction energy is
5 computed. The global exploration, aimed at identifying the correct binding site, was
6 performed only for compound **8**, which introduces an average molecular structure
7 among all ligands. Figure 2A shows the nine different ligand starting positions around
8 EZH2 (at a distances of ~ 15 Å from the enzyme's surface). The 2B panel of the same
9 figure shows the results from the exploration (64 processors and 48 hours each) where
10 we depict the protein-ligand interaction energies with respect to the (ligand) distance to
11 an active site atom: the alpha carbon from Tyr726. By introducing conformational
12 sampling in both the protein and the ligand, we observe now a funnel shaped energy
13 landscape towards the methyl transfer active site; this site is now clearly preferred to the
14 rest of the protein surface and other cavities.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31



32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47 **Figure 2.** (A) Different initial positions for tambjamine analog **8** around EZH2 for the non-
48 biased global PELE exploration. (B) Binding (interaction) energy between compound **8** and
49 EZH2 with respect to the distance of the center of mass of **8** and the CG atom from Tyr726 (a
50 tyrosine residue located at the SAM binding pocket).
51
52
53
54

55
56 Following the global exploration, we performed a locally restricted simulation where,
57 starting from the best position (lowest interaction energy) obtained in the global search,
58
59
60

we limited the ligand to explore only the vicinity of the active site. Here, we expanded a local sampling for all the ligands shown in Figure 1, where we manually edited the changes in chemical substituents with Maestro. As an additional positive control, we also performed the local exploration for the SAH co-factor, using again the apo EZH2 crystal, and compared the best interaction energy to the recent human PRC2 complex crystal (PDB entry 5HYN), including a SAH cofactor. Figure 3 shows an excellent level of agreement between our model and the crystal, where we reproduce the overall position and main interactions.

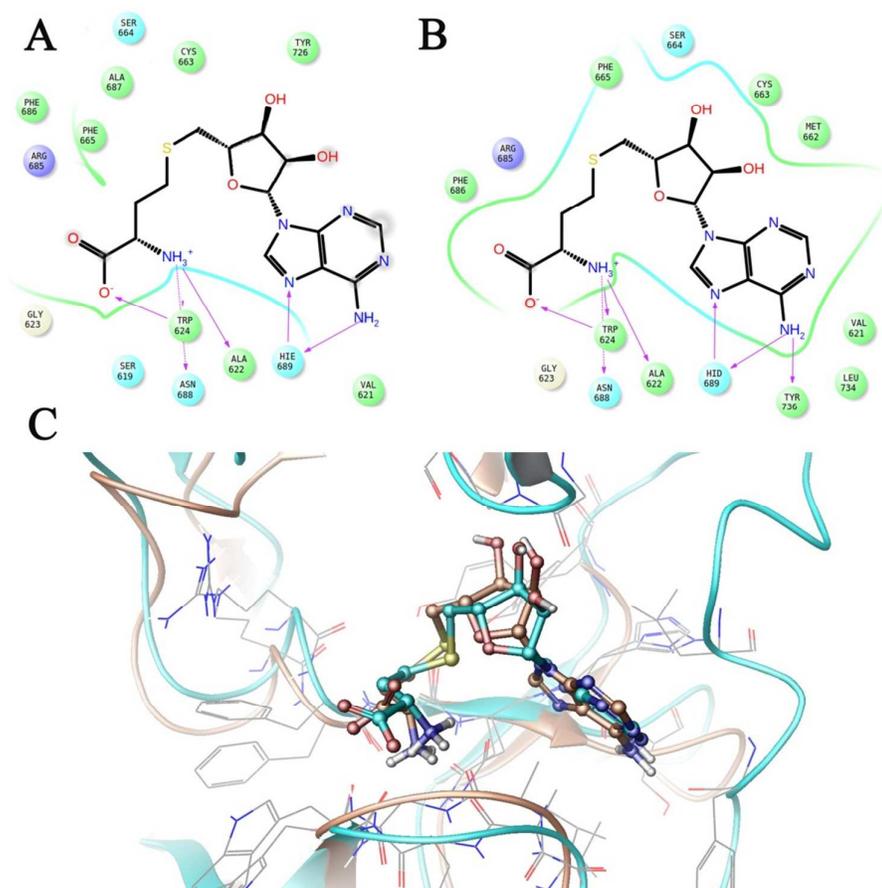


Figure 3. SAH co-factor binding comparison between the PELE simulation using the EZH2 apo structure (pdb entry 4MI0, panel A) and the recent human PRC2 complex (pdb entry 5HYN, panel B). Panel C shows our model and the crystal overlapped.

From the local sampling we extracted the best 10 structures (with lowest interaction energy) for each compound and performed additional scoring calculations with Glide. Thus, we were able to compare the protein-ligand affinity score before and after the induced fit process. The results clearly indicate a significant improvement in the ligands scores, reaching now values in agreement to their level of activity (~ -10 kcal/mol) (Table 1). Importantly, all the negative controls still show significant lower values than the active ligands. The binding mode shows the ability of the molecule to enter in the “end” of the channel where Lys27, the histone’s lysine, is methylated. We should emphasize here that this channel is collapsed in the original apo crystal, which severely restricts the model to be used in standard docking calculations, thus requiring induced fit techniques.

Compound	Initial docking	PELE + docking	IC ₅₀ (μM)
1	-3.7	-7.3 (-6.5)	-
2	-3.9	-8.0 (-7.3)	-
3	-3.7	-5.8 (-5.3)	-
4	-4.0	-8.1 (-7.6)	-
5	-4.6	-10.5	11.7
6	-4.3	-10.4	8.4
7	-5.1	-9.4	11.0
8	-5.2	-10.5	12.9
9	-4.2	-9.7	8.7
10	-4.6	-9.3	23.5
11	-4.7	-9.3	26.9

1
2
3 **Table 1.** Average docking scores (in kcal/mol) before and after PELE and experimental
4 affinities obtained from the target drug discovery program performed at Lilly (half inhibitory
5 concentrations (IC_{50})). Negative control ligands (in grey background color) show both the
6 protonated and deprotonated (in parenthesis) docking scores.
7
8
9

10
11
12
13
14
15 The flexibility of the binding pocket and ligands produced several binding modes of
16 similar (low) interaction energies. Nevertheless, in most of the best (lowest energy)
17 poses, the more flexible moiety of the ligand occupies the hydrophobic pocket around
18 Leu666, at the I-SET domain (Figure 4) observed by Hong Wu et al.²⁷, directly
19 distressing the cognate substrate binding. In addition, the characteristic methoxy group
20 of these derivatives binds approximately at the methyl transfer area of EZH2, as
21 observed when comparing our models with the recent 5HYN crystallographic structure.
22 As seen in Figure 4, the indole moiety participates in a hydrogen bond with Asn688,
23 mimicking the role observed for the SAH co-factor. Moreover, recently deposited PDB
24 entries 5T5G and 5TH7 of the SETD8 protein (with a catalytic pocket structurally
25 similar to EZH2) showed methoxy group containing inhibitors whose positions overlap
26 well with that of our tambjamine analogs methoxy group. All these results underscore
27 the ability of these derivatives inhibiting the methyl transfer from SAM to Lys27.
28
29

30
31
32
33
34
35
36
37
38
39
40
41
42
43
44 All active compounds bear a positive charge, which seems to benefits from binding in
45 the methyl transfer cavity, which has evolved to accommodate the positively charged
46 histone's lysine and SAM cofactor. The increase in the binding score for the protonated
47 form of control ligands (with respect to their deprotonated forms) seems to confirm this
48 point. Nevertheless, the increase is moderate and we still observe a clear difference
49 between active and inactive compounds, indicating that additional aspects, such as the
50
51
52
53
54
55
56
57
58
59
60

imine substituent flexibility, are necessary to accommodate the molecule in the catalytic pocket.

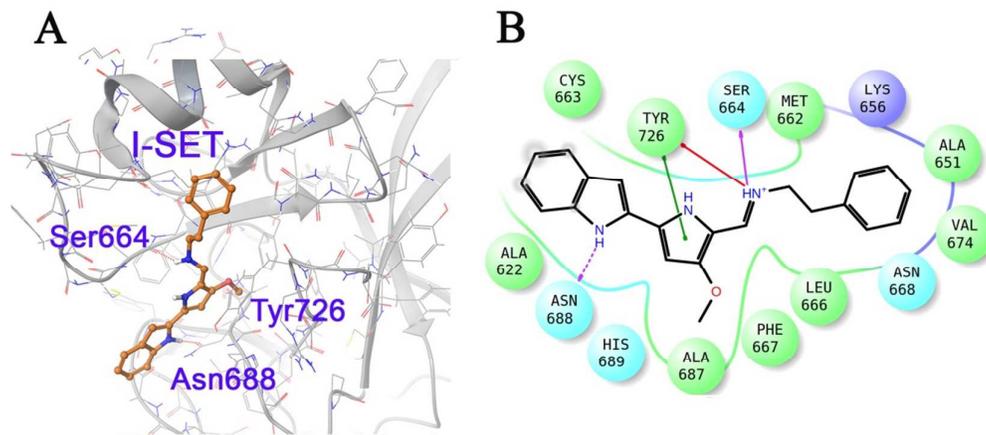


Figure 4. (A) Induced fit docking derived structure for tambjamine **5** (important residues and I-SET domain marked in blue). (B) Ligand 2D interaction diagram for tambjamine **5** with H-bonds underlined in purple, π - π interactions in green and π -cation ones in red.

The molecular knowledge of the binding mode was further challenged by designing a second generation set of inhibitors. Based on the above local exploration, eight different snapshots were selected for molecular docking. A small library varying the imine substituent of the tambjamine analogs was created. The library contained 150 fragments selected by molecular complementarity based on the docking grid (available volume). A further selection of 19 molecules was done based on their docking scores (< 11) (see Figure S2 for details) and from this set of candidates four of them were successfully synthesized to experimentally test the validity of our calculations (compounds **12-15**, Figure 5). The four molecules were tested again in the OIDD Lilly assay and all of them were found to be positive hits, displaying activities in the low micromolar range (Table

2). The fact that all the designed molecules were found active even when significant structural changes were introduced (at the hydrogen bond acceptor/donor level), underscore the capabilities of the molecular model and overall procedure. Although the IC_{50} values obtained did not improve those found in the first generation of molecules, the scaffold is useful for the identification of EZH2 inhibitors and further studies with more diverse compounds should be performed.

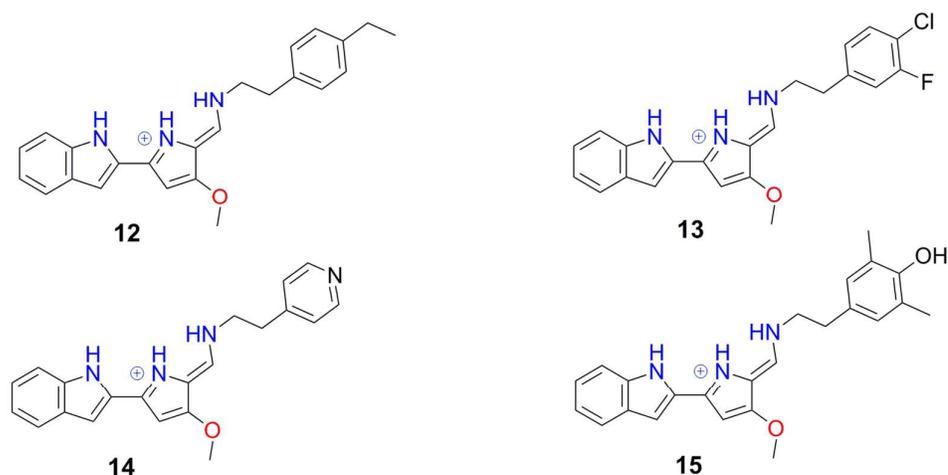


Figure 5. Second generation of synthesized tambjamine analogs after computational directed design.

Compound	Docking	IC_{50} (μ M)
12	-11.9	12.56
13	-11.4	11.77
14	-11.8	36.3
15	-11.8	15.81

Table 2. Docking scores and experimental affinities for the second generation ligand molecules.

Cell assays

EZH2 is a well-established therapeutic target for several cancers, so we wanted to analyze whether our tambjamine-inspired molecules are potential chemotherapeutic compounds. Thus, in order to complement and validate the *in vitro* enzymatic assay, we selected two compounds and their activity in cancer cells was assessed. Lung adenocarcinoma A549 cells were treated with 2 μ M of compound **5** and compound **6** for 96 h and tri-methylation levels in Lys27 of histone 3 were assessed by immunofluorescence. As observed in Figure 6, both compounds induced a very significant decrease in histone 3 tri-methylation levels in cells. After fluorescence intensity quantification and normalization, cells treated with tambjamine analogs **5** and **6** showed a relative fluorescence (in percentage) of 11.37 ± 8.5 and 12.03 ± 6.5 respectively, compared to non-treated cells. These results indicate a statistically significant effect and corroborate in cancer cells the potent EZH2 inhibition by these compounds, formerly detected in *in vitro* experiments (Figure 7). Therefore, these compounds effectively inhibit the cancer therapeutic target EZH2, suggesting that they may be good candidate agents for the treatment of different cancers.

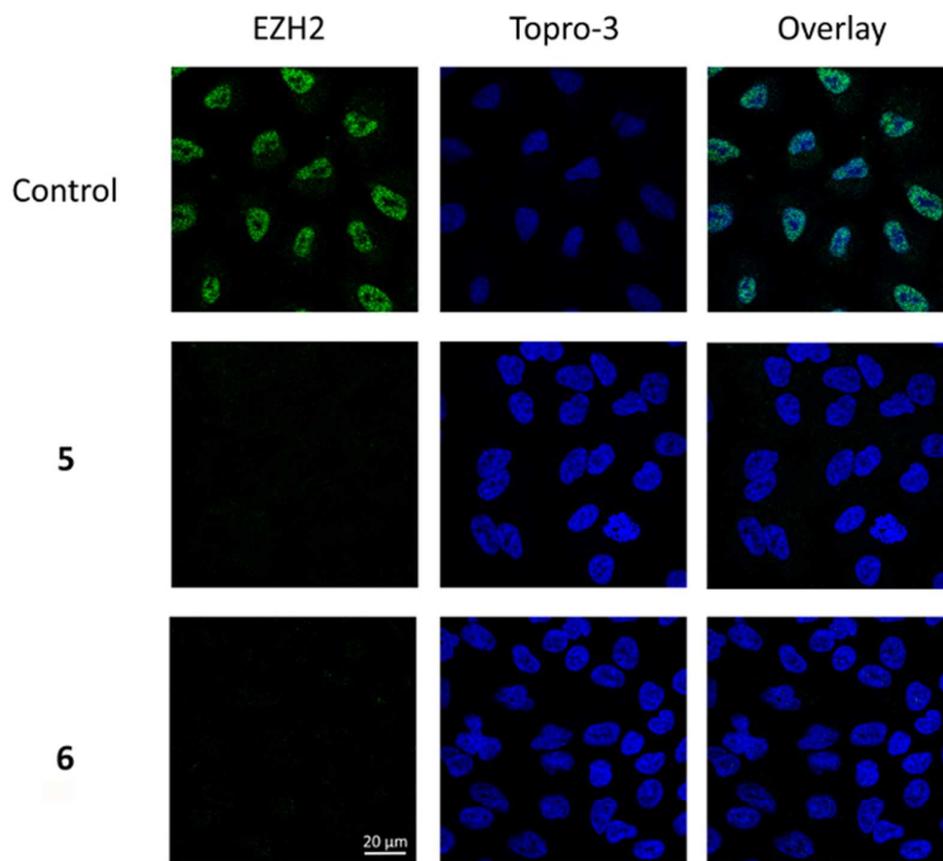


Figure 6. Histone H3 Lys 27 tri-methylation levels. A549 cells were treated with 2 μ M of each compound (compounds **5** and **6**) or with 1% DMSO (CT, control cells) for 96 h and immunostaining was performed using anti-trimethyl-Histone H3 (Lys27) antibody. Results shown are representative images of three immunostaining experiments.

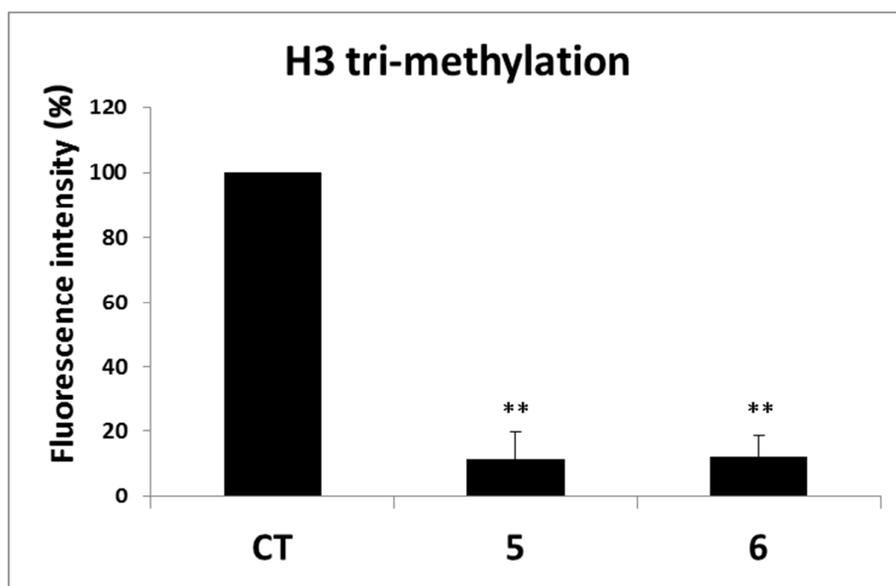


Figure 7. Histone H3 Lys 27 tri-methylation quantification. Histone H3 tri-methylation fluorescence was quantified in treated (compounds **5** and **6**) or non-treated (CT) A549 cells; average normalized intensities are shown (percentage). Statistical significance was assessed by One-way ANOVA with post-hoc Tukey HSD test (**, p value<0.01).

3. Conclusion

Molecular modeling together with *in vitro* and cell assays indicate the capabilities of several tambjamine-inspired derivatives to inhibit EZH2, an important catalytic component of the PRC2 complex whose alterations have been associated to different cancers. These molecules are predicted to bind in the methyl transfer active site, which has evolved to accommodate the positively charged histone's lysine and SAM cofactor. In agreement with this, the most active compounds present a positive charge, although our simulations indicate that introducing a positive charge in the control (neutral) ligands only slightly improve the docking scores; a flexible hydrophobic group is necessary for an effective inhibition of the catalytic pocket. While we show here that the

1
2
3 EZH2 isolated apo crystal is capable of being used in molecular modeling studies, these
4
5 require an important level of conformational sampling. These results represents another
6
7 (yet) example of how rigid docking studies are not capable of predicting potency
8
9 (screening capabilities) nor the correct binding pose (lead optimization capabilities),
10
11 being necessary to carry on conformational landscape exploration simulations, such as
12
13 PELE.
14

15
16
17 EZH2 role in tumorigenesis has been widely validated in many different tumors, being a
18
19 relevant therapeutic target for the treatment of this malignancy²⁸. Many pre-clinical data
20
21 support that EZH2 inhibitors effectively block the growth of cancer cells from different
22
23 tumors, and preliminary clinical results with the most advanced EZH2 inhibitor EPZ-
24
25 6438 (Tazemetostat) have shown durable tumor responses, especially in patients with
26
27 B-cell lymphomas³. Altogether, these data suggest that synthetic tambjamine analogs
28
29 may be promising novel EZH2 inhibitors for cancer treatment.
30
31
32
33
34
35

36 **4. Material and Methods**

37 38 39 **Computational Methods**

40
41
42 **System preparation and docking.** The crystal structure, PDBID 4MI0, was optimized
43
44 with the protein preparation wizard tool from Schrodinger.¹⁴ The protonation state of
45
46 titrable residues were assigned at physiological pH with PROPKA and double-checked
47
48 with the H++ server.¹⁵ Eleven tambjamine analogs, including four negative controls
49
50 (molecules **1-4**), were studied in the initial docking calculations (Figure 1). Selected
51
52 tambjamines were optimized with the LigPrep default protocol, which indicated the
53
54 ambiguous protonation nature of all control ligands, the remaining molecules showing a
55
56 positive charge at physiological pH. Charge assignments were confirmed with
57
58
59
60

1
2
3 additional pKa calculations at the B3LYP and 6-31G (d, p) basis set level of quantum
4
5 chemistry theory by using Jaguar.¹⁶ All negative controls gave a pKa range between 5-
6
7 7, possibly due to the delocalization of electron density into the aromatic substituent,
8
9 while the rest of the systems presented pKa values >8. Thus, the negative controls were
10
11 prepared in a protonated and deprotonated forms. Schrodinger ConfGen tool was then
12
13 used to generate possible conformations for the tambjamins, and their electrostatic
14
15 potential charges extracted from the Jaguar calculations described above. SiteMap¹⁷ was
16
17 then used for selecting probable docking sites and nine grids were calculated and a
18
19 library of possible tambjamins conformations was docked to all nine sites. Docking
20
21 was performed with Glide, using the XP scoring function.¹⁸
22
23
24

25
26 **Induced fit simulations.** PELE,¹⁹ our in-house software combining a Monte Carlo
27
28 (MC) stochastic approach with protein structure prediction algorithms, was used to
29
30 refine the initial docking poses. PELE is capable of accurately reproducing long-
31
32 timescale processes in a 1–2 order of magnitude faster manner than molecular
33
34 dynamics. It has been underlined, for example, as one of the most promising techniques
35
36 in the latest CSAR blind docking competition²⁰. When compared to docking techniques,
37
38 it provides a good induced fit description, allowing the docking in difficult cases (apo,
39
40 cross docking, etc.)²¹. Briefly, the sampling algorithm is based on a consecutive
41
42 iteration of three main steps: a ligand and protein (backbone) perturbation, a side-chain
43
44 sampling, and a minimization.¹⁹ Thus, the procedure begins by a ligand perturbation
45
46 involving a random translation and rotation of the ligand. In the case of the protein, the
47
48 perturbation is based on the α -carbons anisotropic network model²². Following the
49
50 perturbation, a relaxation step is introduced, involving a side chain prediction for all
51
52 residues with 6 Å from the ligand and an overall minimization. These steps compose a
53
54 move that is accepted (defining a new local minimum) or rejected based on a Metropolis
55
56
57
58
59
60

1
2
3 criterion, forming a stochastic trajectory. PELE uses an OPLS (Optimized Potentials for
4 Liquid Simulations) all-atom force field²³ with an implicit surface-generalized Born
5 continuum solvent model²⁴.
6
7
8

9
10 Two different types of PELE simulations were performed: a global and a local
11 sampling. Global sampling, using large ligand translations (up to 5 Å) and rotations (up
12 to 90°) without any restriction of the space explored, aims at finding ligand binding
13 sites. Nine different positions, with the ligand in the protein surface, were used as the
14 initial protein-ligand structures (Figure 2). To enhance the exploration, the ligand
15 perturbation random search direction is also kept for 3 steps. In order to map possible
16 conformational selection on the receptor we used larger backbone perturbations, with
17 ~1 Å alpha carbon displacements in each MC step, where we keep the normal mode for
18 6 consecutive steps. The global exploration used 64 processors and 48 hours each,
19 giving a total of approximately 70,000 MC new minima that explored the entire surface
20 (and inner cavities).
21
22
23
24
25
26
27
28
29
30
31
32
33

34
35 The local sampling focuses the sampling in one region by reducing considerably the
36 ligand translation (up to 1.5 Å), and by keeping the ligand's perturbation direction for
37 only one step. The backbone perturbation direction is also reduced to 1 step. The
38 sampling was performed using 12 processors and 12 hours, producing approximately
39 500-1000 MC minima (notice that, due to steric clashes, the acceptance in local
40 sampling is reduced from that of global sampling). Disregarding duplicates (RMSD <
41 0.2 Å), the average ligand heavy atom RMSD along the different poses obtained is
42 significantly large, ~2.5 Å. Thus the local sampling effectively maps the different
43 conformations adopted by the ligands in an active site.
44
45
46
47
48
49
50
51
52
53
54
55

56 **Experimental Methods**

57
58
59
60

1
2
3 **EZH2 SPA Enzymatic Assay.** Human EZH2 was co-expressed as a 5-member
4 enzymatic complex with SUZ12, EED, AEBP2 and RbAp48 using a baculovirus/Sf9
5 insect cell system. The EZH2 complex was purified using FLAG affinity
6 chromatography. A biotinylated peptide corresponding to human Histone 3 (H3)
7 residues 21-44 was used as the substrate (CPC Scientific Inc., San Jose, CA, Cat#
8 811115). A peptide corresponding to H3 residues 21-44 that was trimethylated on
9 Lys27 was used as a coactivating factor (CPC Scientific Cat # 869799). Adenosyl-L-
10 methionine,S-(methyl)-3H (Adomet) (15 Ci/mmol or 0.55 mCi/ml) was used as the
11 radiolabelled tracer (Perkin Elmer NET155). Streptavidin-coated yttrium silicate SPA
12 beads were used to capture and quantify substrate labelled with tritiated methyl groups
13 (Perkin Elmer Cat# RPNQ0012). Working solutions of any of the above reagents were
14 diluted in Assay buffer (50 mM Tris-HCl pH 8.5, 10 mM DTT, 0.005% Triton X-100).
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 Compounds were serially diluted 3-fold (100% DMSO) for initial top concentrations of
31 either 1 or 100 μ M. The reaction was assembled at room temp in a 384-well white clear
32 bottom assay plate (Corning, NY, USA, Cat# 3706); 15 μ L of a premixed solution of
33 EZH2 5 mER/trimethylated coactivator peptide; 100-200 nL of serially-diluted
34 compound; 5 μ L of a premixed solution of 3H-Adomet/peptide substrate to yield final
35 reaction conditions of 2.5 nM EZH2, 10 nM trimethyl coactivator peptide, 1 μ M 3H-
36 Adomet and 1 μ M peptide substrate. The plate was sealed and shaken for 2 hours at
37 room temperature (RT). Streptavidin-coated yttrium silicate SPA beads were
38 reconstituted at 1mg/mL in 3 M Guanidine HCL and briefly sonicated prior to a 20 μ L
39 dispense in the assay plate (20 μ g beads per well) to stop the reaction. The plates were
40 sealed and then shaken for 10 minutes. The beads were then allowed to settle for 60
41 minutes prior to counting on a Microbeta Trilux (Perkin Elmer).
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Immunofluorescence staining.** A549 cells (8×10^4 cells/mL) were cultured in a 12-well
4 plate containing FBS-coated glass coverslips. After 24 h, they were incubated 96 h with
5 $2 \mu\text{M}$ of two of the studied compounds (**5** and **6**, as shown in Figure 1). Cells were then
6 washed twice with PBS and fixed with PBS-2% paraformaldehyde for 10 min at RT.
7 Fixed cells were washed twice with PBS and permeabilized with PBS-0.1% Triton X-
8 100 during 10 min. They were then washed twice with washing solution (WS, PBS-
9 0.1% Tween 20) for 5 min and treated 30 min at RT with blocking solution (PBS-
10 Tween-20 0.1%, 0.1% Bovine Serum Albumin and 10% normal goat serum). Cells were
11 incubated overnight at 4°C with primary antibody (1:500, Anti-trimethyl-Histone H3
12 (Lys27) Antibody, EMD Millipore Corporation, Billerica, MA, USA) diluted in
13 blocking solution. Cells were then washed with WS three times and incubated with
14 Alexa Fluor 488-conjugated donkey anti rabbit (Molecular Probes, Invitrogen,
15 Carlsbad, CA, USA) at 1:400 dilution for 1 h at RT. Afterwards, coverslips were
16 washed with WS and placed on the slides with Mowiol (Sigma-Aldrich, St. Louis, MO,
17 USA). Images were captured using a Leica TCS-SL filter-free spectral confocal
18 microscope (Leica Microsystems, Wetzlar, Germany). Three independent experiments
19 were performed and fluorescence intensities ($n=30/\text{condition}$) were normalized and
20 quantified by Image J software. Statistical analysis was performed with Statgraphics
21 Centurion XVI software (Manugistics, Rockville, MD, USA) using One-way ANOVA
22 with post-hoc Tukey HSD test.

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48 **General procedures.** Purchased reagents were used as received without any further
49 purification. The ^1H and ^{13}C NMR spectra were recorded on Varian Mercury-300 MHz
50 and Varian Unity Inova-400 MHz spectrometers, using CDCl_3 or $\text{DMSO}-d_6$ as solvents.
51 Chemical shifts are reported as δ values in parts per million (ppm) using the residual
52 solvent peak as a reference and coupling constants are reported in Hz. Data are reported
53
54
55
56
57
58
59
60

1
2
3 as follows: chemical shift, multiplicity (splitting pattern abbreviations are: s: singlet, d:
4 doublet, dd: doublet of doublets, dt: doublet of triplets, t: triplet, td: triplet of doublets,
5 q: quartet, m: multiplet, app t: apparent triplet, app q: apparent quartet). Mass spectra
6 (HRMS) were obtained with a Micromass Autospec S-2 spectrometer using EI at 70 eV.
7
8

9
10
11
12 **Synthesis of compounds (1-15).** 4-methoxy-1H,1'H-[2,2'-bipyrrole]-5-carbaldehyde
13 was synthesized as described.²⁹ This set of compounds was synthesized applying
14 modifications to a previously reported method.³⁰ Briefly, the corresponding amine (1.7–
15 4 equivalents) was added, at room temperature, over a suspension of the carboxaldehyde
16 in chloroform (80–100 mM aldehyde) or 1,2-dichloroethane (for **2**, **3**, **7** and **15**). Then,
17 acetic acid (40–400 μ L) was added. The reaction mixture was stirred under reflux
18 overnight or until TLC (hexane:AcOEt 1:1) analysis indicated complete consumption of
19 the starting material. The mixture was then cooled to room temperature. It was washed
20 three times (x 30 mL) with HCl 1M. The organic phase was dried on Na₂SO₄ and the
21 solvent was removed under vacuum to yield the desired products, as yellow-orange
22 solids in good yields. In the case of compound **14** the work up is slightly different. The
23 organic layer was washed with a buffer solution (20 mM NaH₂PO₄ pH 7.2, ionic
24 strength 1 M NaCl). Compounds **5**, **6**, were previously reported by us.¹² See Supporting
25 information for a reaction scheme and the NMR and HRMS spectra of each compound.
26 The full list of logP and LLE (pIC₅₀ – logD) values for all compounds is provided in
27 Supporting Information.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48
49 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-4-(tert-**
50 **butyl)benzenaminium chloride (1.HCl).** Yield: 59%. ¹H NMR (300 MHz, CDCl₃) δ =
51 13.56 (s, 1H), 11.48 (d, *J* = 12.8 Hz, 1H), 10.19 (s, 1H), 7.40–7.28 (m, 5H), 7.22–7.13
52 (m, 3H), 6.97 (t, *J* = 7.5 Hz, 1H), 6.92 (s, 1H), 6.03 (s, 1H), 3.75 (s, 3H, OCH₃), 1.29 (s,
53 9H). ¹³C NMR (75 MHz, CDCl₃) δ = 164.33 (C), 149.68 (C), 142.53 (C), 138.04 (C),
54
55
56
57
58
59
60

1
2
3 135.58 (C), 131.66 (CH), 127.88 (C), 127.32 (C), 126.67 (2CH), 125.01 (CH), 121.46
4 (CH), 120.71 (CH), 117.12 (2CH), 113.22 (C), 112.09 (CH), 107.11 (CH), 93.49 (CH),
5 58.74 (OCH₃), 34.63 (C), 31.30 (3CH₃).). HRMS (EI) m/z calc for [C₂₄H₂₅N₃O]
6
7 371.1998; found 371.2000.
8
9
10

11
12 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl) methylene)-4-**
13 **methoxybenzenaminium chloride (2.HCl).** Yield: 95%. ¹H NMR (300 MHz, DMSO-
14
15 *d*₆) δ = 13.52 (s, 1H), 13.09 (d, *J* = 14.3 Hz, 1H), 12.00 (s, 1H), 8.41 (d, *J* = 14.0 Hz,
16 1H), 7.74–7.59 (m, 3H), 7.52–7.42 (m, 2H), 7.24 (t, *J* = 7.6 Hz, 1H), 7.15–7.00 (m,
17 3H), 6.77 (s, 1H), 4.03 (s, 3H, OCH₃), 3.78 (s, 3H, OCH₃). ¹³C NMR (75 MHz, DMSO-
18 *d*₆) δ = 164.04 (C), 157.92 (C), 141.03 (C), 137.92 (C), 134.84, (CH), 131.49 (C),
19 127.79 (C), 124.12 (CH), 121.27 (CH), 120.50 (CH), 119.47 (2CH), 114.98 (2CH, C),
20 112.83 (C), 111.84 (CH), 103.73 (CH), 94.06 (CH), 58.87 (OCH₃), 55.50 (OCH₃).
21
22 HRMS (EI) m/z calc for [C₂₁H₁₉N₃O₂] 345.1477; found: 345.1469.
23
24
25
26
27
28
29
30
31
32

33
34 **N-((5-(1-(tert-butoxycarbonyl)-1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-**
35 **yl)methylene)-4-methoxybenzenaminium chloride (3.HCl).** Yield: 64%. ¹H NMR
36 (300 MHz, CDCl₃) δ = 13.73 (d, *J* = 13.7 Hz, 1H), 13.45 (s, 1H), 8.08 (d, *J* = 8.4 Hz,
37 1H), 7.96 (d, *J* = 14.8 Hz, 1H), 7.65 (d, *J* = 7.4 Hz, 1H), 7.54–7.48 (m, 2H), 7.42 (s,
38 1H), 7.37 (d, *J* = 7.3 Hz, 1H), 7.27 (t, *J* = 7.5, 1H), 7.00–6.90 (m, 2H), 6.06 (d, *J* = 1.8,
39 1H), 3.97 (s, 3H, OCH₃), 3.82 (s, 3H, OCH₃), 1.63 (s, 9H, 3 CH₃). ¹³C NMR (75 MHz,
40 CDCl₃) δ = 162.94 (C), 158.80 (C), 149.86 (C), 142.36 (C), 138.51 (C), 134.37 (CH),
41 131.77 (C), 129.56 (C), 128.83 (C), 126.63 (CH), 123.74 (CH), 122.35 (CH), 119.76
42 (2CH), 115.95 (CH), 115.72 (CH), 115.21 (2CH), 112.61 (C), 97.17 (CH), 85.12 (C),
43 58.66 (OCH₃), 55.76 (OCH₃), 28.25 (3CH₃). HRMS (EI) m/z calc for [C₂₆H₂₇N₃O₄]
44
45 445.2002; found: 445.2011.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **(E)-N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)benzenaminium**
4
5 **chloride (4.HCl)**. Yield: 53%. ^1H NMR (300 MHz, CDCl_3) δ = 13.87 (s, 1H), 11.63 (s,
6
7 1H), 10.31 (s, 1H), 7.57 (s, 1H), 7.49 (d, J = 3.0 Hz, 1H), 7.46 (d, J = 2.7 Hz, 1H),
8
9 7.41–7.35 (m, 4H), 7.28–7.18 (m, 2H), 7.10–7.00 (m, 2H), 6.15 (s, 1H), 3.91 (s, 3H,
10
11 OCH_3). ^{13}C NMR (75 MHz, CDCl_3) δ = 164.89 (C), 143.54 (C), 138.39 (C), 138.28 (C),
12
13 131.84 (CH), 130.02 (2CH), 128.03 (C), 127.35 (C), 126.47 (CH), 125.46 (CH), 121.63
14
15 (CH), 121.00 (CH), 117.55 (2CH), 113.74 (C), 112.46 (CH), 107.74 (CH), 93.63 (CH),
16
17 58.95 (OCH_3). HRMS (EI) m/z calc for $[\text{C}_{20}\text{H}_{17}\text{N}_3\text{O}]$ 315.1363; found: 316.1447.
18
19

20
21 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene) dodecan-1-aminium**
22
23 **acetate (7.HOAc)**. Yield: 38%. ^1H NMR (300 MHz, CDCl_3) δ = 12.21 (s, 1H), 7.59 (d,
24
25 J = 8.0 Hz, 1H), 7.52 (dd, J = 8.2, 0.8 Hz, 1H), 7.41 (s, 1H), 7.27–7.21 (m, 1H), 7.12–
26
27 7.04 (m, 1H), 6.94 (d, J = 1.1 Hz, 1H), 6.15 (s, 1H), 3.90 (s, 3H, OCH_3), 3.46 (app t, J =
28
29 7.1 Hz, 2H), 2.13 (s, 3H, CH_3CO_2^-), 1.77–1.66 (m, 2H), 1.35–1.21 (m, 18H), 0.87 (t, J =
30
31 6.7 Hz, 3H). ^{13}C NMR (75 MHz, CDCl_3) δ = 178.34 (C, CH_3CO_2^-), 163.09 (C), 142.53
32
33 (CH), 141.90 (C), 138.40 (C), 128.61 (C), 128.14 (C), 124.01 (CH), 120.96 (CH),
34
35 120.29 (CH), 112.44 (CH), 111.63 (C), 105.09 (CH), 92.82 (CH), 58.38 (OCH_3), 51.54
36
37 (CH_2), 32.03 (CH_2), 30.46 (CH_2), 29.74 (2 CH_2), 29.69 (CH_2), 29.60 (CH_2), 29.47
38
39 (CH_2), 29.29 (CH_2), 26.65 (CH_2), 23.87 (CH_3 , CH_3CO_2^-), 22.81 (CH_2), 14.25 (CH_3).
40
41
42 HRMS (EI) m/z calc for $[\text{C}_{26}\text{H}_{37}\text{N}_3\text{O}]$ 407.2937; found: 407.2953.
43
44
45

46
47 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-3-methylbutan-1-**
48
49 **aminium chloride (8.HCl)**. Yield: 90%. ^1H NMR (300 MHz, CDCl_3) δ = 13.89 (s, 1H),
50
51 10.32 (s, 1H), 10.02 (d, J = 14.8 Hz, 1H), 7.59 (d, J = 8.0 Hz, 1H), 7.52 (d, J = 8.3 Hz,
52
53 1H), 7.36 (d, J = 15.2 Hz, 1H), 7.26 (t, J = 7.7 Hz, 1H), 7.10 (t, J = 7.5 Hz, 1H), 7.01 (d,
54
55 J = 2.0 Hz, 1H), 6.14 (d, J = 2.3 Hz, 1H), 3.90 (s, 3H, OCH_3), 3.52 (app q, 2H), 1.79–
56
57 1.63 (m, 3H), 0.96 (d, J = 6.3 Hz, 6H). ^{13}C NMR (75 MHz, CDCl_3) δ = 163.30 (C),
58
59
60

1
2
3 142.05 (CH), 141.31 (C), 138.03 (C), 128.05 (C), 127.80 (C), 124.88 (CH), 121.29
4 (CH), 120.81 (CH), 112.45 (CH), 111.45 (C), 106.03 (CH), 92.69 (CH), 58.66 (OCH₃),
5
6 49.65 (CH₂), 38.88 (CH₂), 25.59 (CH), 22.40 (2CH₃). HRMS (EI) m/z calc for
7
8 [C₁₉H₂₃N₃O] 309.1841; found: 309.1850.
9

10
11
12 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)octan -1-aminium**
13 **chloride (9.HCl)**. Yield 98%. ¹H NMR (300 MHz, CDCl₃) δ = 13.44 (s, 1H), 10.36 (s,
14 1H), 9.93 (d, *J* = 14.3 Hz, 1H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.45 (d, *J* = 8.1 Hz, 1H), 7.23–
15 7.09 (m, 2H), 7.04 (t, *J* = 7.5 Hz, 1H), 6.98 (s, 1H), 6.11 (s, 1H), 3.72 (s, 3H, OCH₃),
16 3.36 (app q, 2H), 1.75–1.60 (m, 2H), 1.36–1.16 (m, 10H), 0.83 (t, *J* = 6.5 Hz, 3H). ¹³C
17 NMR (75 MHz, CDCl₃) δ = 163.09 (C), 141.84 (CH), 140.71 (C), 137.72 (C), 127.90
18 (C), 127.65 (C), 124.37 (CH), 121.11 (CH), 120.49 (CH), 112.00 (CH), 111.04 (C),
19 105.47 (CH), 92.74 (CH), 58.41 (OCH₃), 51.14 (CH₂), 31.68 (CH₂), 29.99 (CH₂), 29.04
20 (CH₂), 29.00 (CH₂), 26.44 (CH₂), 22.55 (CH₂), 14.05 (CH₃). HRMS (EI) m/z calc for
21 [C₂₂H₂₉N₃O] 351.2311; found: 351.2306.
22
23
24
25
26
27
28
29
30
31
32
33

34
35
36 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-4-**
37 **butylbenzenaminium chloride (10.HCl)**. Yield: 89%. ¹H NMR (300 MHz, CDCl₃) δ =
38 13.46 (s, 1H), 10.35 (s, 1H), 9.93 (d, *J* = 14.8 Hz, 1H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.45 (d,
39 *J* = 8.2 Hz, 1H), 7.22–7.10 (m, 2H), 7.04 (t, *J* = 7.5 Hz, 1H), 6.98 (s, 1H), 6.11 (s, 1H),
40 3.73 (s, 3H, OCH₃), 3.38 (app q, 2H), 1.74–1.59 (m, 2H), 1.45–1.30 (m, 2H), 0.90 (t, *J*
41 = 7.3 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ = 163.07 (C), 141.85 (CH), 140.67 (C),
42 137.67 (C), 127.86 (C), 127.61 (C), 124.30 (CH), 121.09 (CH), 120.44 (CH), 111.91
43 (CH), 110.99 (C), 105.39 (CH), 92.73 (CH), 58.35 (OCH₃), 50.78 (CH₂), 31.86 (CH₂),
44 19.60 (CH₂), 13.46 (CH₃). HRMS (EI) m/z calc for [C₁₈H₂₁N₃O] 295.1685; found:
45 295.1683.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-2-methylpropan-1-**
4 **aminium chloride (11.HCl).** Yield: 43%. ¹H NMR (300 MHz, CDCl₃) δ = 13.85 (s,
5 1H), 10.36 (s, 1H), 9.97 (d, *J* = 14.8 Hz, 1H), 7.58 (d, *J* = 8.0 Hz, 1H), 7.51 (dd, *J* = 8.3,
6 0.8, 1H), 7.32–7.21 (m, 2H), 7.12–7.09 (*J* = 8.0 Hz, 1H), 7.01 (d, *J* = 1.3 Hz, 1H), 6.14
7 (d, *J* = 2.2 Hz, 1H), 3.87 (s, 3H, OCH₃), 3.29 (app t, 2H), 2.10–1.97 (m, 1H), 1.03 (d, *J*
8 = 6.7 Hz, 6H). ¹³C NMR (75 MHz, CDCl₃) δ = 163.44 (C), 142.43 (CH), 141.34 (C),
9 137.99 (C), 128.02 (C), 127.76 (C), 124.80 (CH), 121.28 (CH), 120.75 (CH), 112.38
10 (CH), 111.28 (C), 106.02 (CH), 92.75 (CH), 59.00 (CH₂), 58.64 (OCH₃), 29.55 (CH),
11 19.94 (2CH₃). HRMS (EI) *m/z* calc for [C₁₈H₂₁N₃O] 295.1685; found: 295.1685.
12
13
14
15
16
17
18
19
20
21
22
23

24 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-2-(4-**
25 **ethylphenyl)ethanaminium chloride (12.HCl).** Yield: 93%. ¹H NMR (300 MHz,
26 CDCl₃) δ = 13.94 (s, 1H), 10.31 (s, 1H), 10.16 (d, *J* = 14.5 Hz, 1H), 7.60 (d, *J* = 8.1 Hz,
27 1H), 7.55 (dd, *J* = 8.3, 0.8 Hz, 1H), 7.32–7.24 (m, 1H), 7.20–7.09 (m, 6H), 7.01 (dd, *J* =
28 2.1, 0.8 Hz, 1H), 6.13 (d, *J* = 2.3 Hz, 1H), 3.89 (s, 3H, OCH₃), 3.74 (app q, 2H), 3.07 (t,
29 *J* = 7.2 Hz, 2H), 2.62 (q, *J* = 7.6 Hz, 2H), 1.21 (t, *J* = 7.6 Hz, 3H). ¹³C NMR (75 MHz,
30 CDCl₃) δ = 163.49 (C), 143.19 (C), 142.14 (CH), 141.51 (C), 138.07 (C), 134.18 (C),
31 129.09 (2CH), 128.46 (2CH), 128.06 (C), 127.73 (C), 124.93 (CH), 121.32 (CH),
32 120.84 (CH), 112.47 (CH), 111.42 (C), 106.19 (CH), 92.74 (CH), 58.62 (OCH₃), 52.82
33 (CH₂), 36.12 (CH₂), 28.59 (CH₂), 15.74 (CH₃). HRMS (EI) *m/z* calc for [C₂₄H₂₅N₃O]
34 371.1998; found: 371.2003.
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-2-(3-chloro-4-**
50 **fluorophenyl)ethanaminium chloride (13.HCl).** Yield: 80%. ¹H NMR (300 MHz,
51 DMSO-*d*₆) δ = 13.27 (s, 1H), 11.92 (s, 1H), 11.27 (s, 1H), 8.12 (d, *J* = 12.9 Hz, 1H),
52 7.64 (d, *J* = 7.9 Hz, 1H), 7.53 (t, *J* = 8.1 Hz, 1H), 7.48–7.42 (m, 2H), 7.43 (s, 1H), 7.26–
53 7.18 (m, 2H), 7.08 (app t, *J* = 7.5 Hz, 1H), 6.69 (s, 1H), 3.96 (s, 3H, OCH₃), 3.91 (s,
54
55
56
57
58
59
60

1
2
3 2H), 3.02 (t, $J = 6.7$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) $\delta = 162.79$ (C), 157.07 (d,
4 $J_{\text{C-F}} = 244.5$ Hz, C), 145.24 (CH), 139.67 (d, $J_{\text{C-F}} = 6.7$ Hz, C), 139.41 (C), 137.65 (C),
5 130.52 (CH), 127.99 (C), 127.76 (C), 126.51 (d, $J_{\text{C-F}} = 3.2$ Hz, CH) 123.80 (CH),
6 121.12 (CH), 120.38 (CH), 117.56 (d, $J_{\text{C-F}} = 17.25$ Hz, C), 117.45 (d, $J_{\text{C-F}} = 20.25$ Hz,
7 CH), 111.75 (CH), 110.79 (C), 102.80 (CH), 93.52 (CH), 58.68 (OCH₃), 50.49 (CH₂),
8 34.48 (CH₂). HRMS (EI) m/z calc for [C₂₂H₁₉ClFN₃O] 395.1201; found: 395.1190.
9
10
11
12
13
14
15

16
17 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-2-(pyridin-4-**
18 **yl)ethanaminium chloride (14.HCl)**. Yield: 73%. ^1H NMR (300 MHz, DMSO- d_6) $\delta =$
19 13.36 (s, 1H), 11.96 (s, 1H), 11.42 (s, 1H), 8.59–8.46 (m, 2H), 8.11 (s, 1H), 7.64 (d, $J =$
20 7.9 Hz, 1H), 7.45 (d, $J = 8.2$ Hz, 1H), 7.38–7.32 (m, 3H), 7.22 (t, $J = 7.6$ Hz, 1H), 7.08
21 (t, $J = 7.5$ Hz, 1H), 6.70 (s, 1H), 4.01–3.89 (m, 5H), 3.03 (t, $J = 7.0$ Hz, 2H). ^{13}C NMR
22 (75 MHz, DMSO- d_6) $\delta = 162.70$ (C), 149.63 (2CH), 146.75 (C), 145.16 (CH), 139.33
23 (C), 137.63 (C), 128.03 (C), 127.76 (C), 124.43 (2CH), 123.75 (CH), 121.09 (CH),
24 120.35 (CH), 111.74 (CH), 110.85 (C), 102.81 (CH), 93.54 (CH), 58.66 (OCH₃), 49.95
25 (CH₂), 34.44 (CH₂). HRMS (EI) m/z calc for [C₂₁H₂₀N₄O] 344.1637; found: 344.1644.
26
27
28
29
30
31
32
33
34
35
36
37

38 **N-((5-(1H-indol-2-yl)-3-methoxy-1H-pyrrol-2-yl)methylene)-2-(4-hydroxy-3,5-**
39 **dimethylphenyl)ethanaminium chloride (15.HCl)**. Yield: 41%. ^1H NMR (300 MHz,
40 DMSO- d_6) $\delta = 13.21$ (s, 1H), 11.90 (s, 1H), 11.22 (d, $J = 14.9$ Hz, 1H), 8.13–8.03 (m,
41 2H), 7.65 (d, $J = 7.9$ Hz, 1H), 7.45 (d, $J = 8.2$ Hz, 1H), 7.32 (s, 1H), 7.22 (t, $J = 7.6$ Hz,
42 1H), 7.08 (t, $J = 7.6$ Hz, 1H), 6.84 (s, 2H), 6.69 (d, $J = 2.1$ Hz, 1H), 3.96 (s, 3H, OCH₃),
43 3.80 (app q, 2H), 2.81 (t, $J = 7.0$ Hz, 2H), 2.11 (s, 6H). ^{13}C NMR (75 MHz, DMSO- d_6)
44 $\delta = 162.55$ (C), 151.79 (C), 145.17 (CH), 139.06 (C), 137.61 (C), 128.71 (2CH), 128.08
45 (C), 127.77 (C), 127.75 (C), 124.18 (2C), 123.71 (CH), 121.07 (CH), 120.34 (CH),
46 111.72 (CH), 110.71 (C), 102.65 (CH), 93.46 (CH), 58.64 (OCH₃), 51.42 (CH₂), 34.75
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(CH₂), 16.61 (2CH₃). HRMS (EI) m/z calc for [C₂₄H₂₅N₃O₂] 387.1947; found: 387.1931.

5. Additional Information

Supporting Information

Binding sites found with SiteMap and docking scores for compounds **1-11** in all of them. Selected 19 molecules based on their docking scores after induced fit simulations with PELE. Compounds characterization data. Enzymatic assay dose response curves for all compounds. LogP and LLE values for all compounds.

Author Information

Corresponding authors: Vanessa Soto-Cerrato vsoto@ub.edu; Victor Guallar, victor.guallar@bsc.es

[†]Martin Kotev, Pilar Manuel-Manresa and Elsa Hernando, contributed equally to this manuscript.

Acknowledgment

This work was supported by grants from the Spanish Government (CTQ2013-48287; FIS PI13/00089), the Consejería de Educación de la Junta de Castilla y León (Projects BU340U13 and BU092U16) and from La Marató de TV3 Foundation (20132730). E.H. thanks the Junta de Castilla y León (Consejería de Educación) and Fondo Social Europeo for a PIRTU contract. OIDD screening data supplied courtesy of Eli Lilly and Company—used with Lilly's permission. To learn more about the Lilly Open Innovation Drug Discovery program, please visit the program website at <https://openinnovation.lilly.com>. Anti-trimethyl-Histone H3 (Lys27) antibody was a

1
2
3 kind gift from Dr. Marian Martínez-Balbás (IBMB-CSIC), who helped us with technical
4
5 assistance. We thank Dr. Benjamín Torrejón from CCiTUB (Centres Científics i
6
7 Tecnològics Universitat de Barcelona, Barcelona, Spain) for technical assistance.
8
9

10 **Abbreviations Used:**

11
12 EZH2, enhancer of zeste homolog 2; PRC2, polycomb repressive complex 2; SAM, S-
13
14 adenosyl-L-methionine; SAH, S-adenosyl homocysteine; PELE, protein energy
15
16 landscape exploration; SUZ12, supressor of zeste 12; EED, embryonic ectoderm
17
18 development; SET, su(var)3-9, enhancer-of-zeste and trithorax; AEBP2, adipocyte
19
20 enhancer-binding protein; RbAp48, retinoblastoma binding protein 4.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 **References**

- 37
38
39 1. Vire, E.; Brenner, C.; Deplus, R.; Blanchon, L.; Fraga, M.; Didelot, C.; Morey, L.; Van
40
41 Eynde, A.; Bernard, D.; Vanderwinden, J.-M.; Bollen, M.; Esteller, M.; Di Croce, L.; de Launoit,
42
43 Y.; Fuks, F. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **2006**,
44
45 439, 871-874.
46
47
48 2. Varambally, S.; Dhanasekaran, S. M.; Zhou, M.; Barrette, T. R.; Kumar-Sinha, C.; Sanda,
49
50 M. G.; Ghosh, D.; Pienta, K. J.; Sewalt, R. G. A. B.; Otte, A. P.; Rubin, M. A.; Chinnaiyan, A. M.
51
52 The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **2002**,
53
54 419, 624-629.
55
56
57 3. Kim, K. H.; Roberts, C. W. M. Targeting EZH2 in cancer. *Nat. Med.* **2016**, 22, 128-134.
58
59
60

- 1
2
3 4. Takawa, M.; Masuda, K.; Kunizaki, M.; Daigo, Y.; Takagi, K.; Iwai, Y.; Cho, H. S.;
4
5 Toyokawa, G.; Yamane, Y.; Maejima, K.; Field, H. I.; Kobayashi, T.; Akasu, T.; Sugiyama, M.;
6
7 Tsuchiya, E.; Atomi, Y.; Ponder, B. A.; Nakamura, Y.; Hamamoto, R. Validation of the histone
8
9 methyltransferase EZH2 as a therapeutic target for various types of human cancer and as a
10
11 prognostic marker. *Cancer Sci.* **2011**, 102, 1298-305.
12
13
14 5. Jiang, T.; Wang, Y.; Zhou, F.; Gao, G.; Ren, S.; Zhou, C. Prognostic value of high EZH2
15
16 expression in patients with different types of cancer: a systematic review with meta-analysis.
17
18 *Oncotarget* **2016**, 7, 4584-97.
19
20
21 6. Yu, J.; Yu, J.; Rhodes, D. R.; Tomlins, S. A.; Cao, X.; Chen, G.; Mehra, R.; Wang, X.;
22
23 Ghosh, D.; Shah, R. B.; Varambally, S.; Pienta, K. J.; Chinnaiyan, A. M. A polycomb repression
24
25 signature in metastatic prostate cancer predicts cancer outcome. *Cancer Res.* **2007**, 67, 10657-
26
27 63.
28
29 7. McCabe, M. T.; Ott, H. M.; Ganji, G.; Korenchuk, S.; Thompson, C.; Van Aller, G. S.; Liu,
30
31 Y.; Graves, A. P.; Iii, A. D. P.; Diaz, E.; LaFrance, L. V.; Mellinger, M.; Duquenne, C.; Tian, X.;
32
33 Kruger, R. G.; McHugh, C. F.; Brandt, M.; Miller, W. H.; Dhanak, D.; Verma, S. K.; Tummino, P. J.;
34
35 Creasy, C. L. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating
36
37 mutations. *Nature* **2012**, 492, 108-112.
38
39
40 8. Knutson, S. K.; Kawano, S.; Minoshima, Y.; Warholic, N. M.; Huang, K.-C.; Xiao, Y.;
41
42 Kadowaki, T.; Uesugi, M.; Kuznetsov, G.; Kumar, N. Selective inhibition of EZH2 by EPZ-6438
43
44 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma. *Mol. Cancer Ther.*
45
46 **2014**, 13, 842-854.
47
48
49 9. Brooun, A.; Gajiwala, K. S.; Deng, Y. L.; Liu, W.; Bolanos, B.; Bingham, P.; He, Y. A.;
50
51 Diehl, W.; Grable, N.; Kung, P. P.; Sutton, S.; Maegley, K. A.; Yu, X.; Stewart, A. E. Polycomb
52
53 repressive complex 2 structure with inhibitor reveals a mechanism of activation and drug
54
55 resistance. *Nat. Commun.* **2016**, 7, 11384.
56
57
58
59
60

- 1
2
3 10. Gale, P. A.; Pérez-Tomás, R.; Quesada, R. Anion Transporters and Biological Systems.
4
5 *Acc. Chem. Res.* **2013**, *46*, 2801-2813.
6
7 11. Hernando, E.; Soto-Cerrato, V.; Cortes-Arroyo, S.; Perez-Tomas, R.; Quesada, R.
8
9 Transmembrane anion transport and cytotoxicity of synthetic tambjamine analogs. *Org.*
10
11 *Biomol. Chem.* **2014**, *12*, 1771-8.
12
13 12. Soto-Cerrato, V.; Manuel-Manresa, P.; Hernando, E.; Calabuig-Fariñas, S.; Martínez-
14
15 Romero, A.; Fernández-Dueñas, V.; Sahlholm, K.; Knöpfel, T.; García-Valverde, M.; Rodilla, A.
16
17 M.; Jantus-Lewintre, E.; Farràs, R.; Ciruela, F.; Pérez-Tomás, R.; Quesada, R. Facilitated Anion
18
19 Transport Induces Hyperpolarization of the Cell Membrane That Triggers Differentiation and
20
21 Cell Death in Cancer Stem Cells. *J. Am. Chem. Soc.* **2015**, *137*, 15892-15898.
22
23 13. Rodilla, A. M.; Korrodi-Gregorio, L.; Hernando, E.; Manuel-Manresa, P.; Quesada, R.;
24
25 Perez-Tomas, R.; Soto-Cerrato, V. Synthetic tambjamine analogues induce mitochondrial
26
27 swelling and lysosomal dysfunction leading to autophagy blockade and necrotic cell death in
28
29 lung cancer. *Biochem. Pharmacol.* **2017**, *126*, 23-33.
30
31 14. Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and
32
33 ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J.*
34
35 *Comput. Aid. Mol. Des.* **2013**, *27*, 221-234.
36
37 15. Anandkrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and
38
39 the preparation of biomolecular structures for atomistic molecular modeling and simulations.
40
41 *Nucleic Acids Res.* **2012**, *40*, W537-W541.
42
43 16. Jaguar 4.1. Schrödinger, Inc., Portland, Oregon, 2000.
44
45 17. *SiteMap*, version 2.5; Schrödinger, LLC: New York, NY, 2011.
46
47 18. *Glide*, version 5.6; Schrödinger: New York, NY: LLC, 2010.
48
49 19. Borrelli, K. W.; Vitalis, A.; Alcantara, R.; Guallar, V. PELE: Protein Energy Landscape
50
51 Exploration. A Novel Monte Carlo Based Technique. *J. Chem. Theory Comput.* **2005**, *1*, 1304-
52
53 1311.
54
55
56
57
58
59
60

- 1
2
3 20. Carlson, H. A.; Smith, R. D.; Damm-Ganamet, K. L.; Stuckey, J. A.; Ahmed, A.; Convery,
4 M. A.; Somers, D. O.; Kranz, M.; Elkins, P. A.; Cui, G.; Peishoff, C. E.; Lambert, M. H.; Dunbar, J.
5 B. CSAR 2014: A Benchmark Exercise Using Unpublished Data from Pharma. *J. Chem. Inf.*
6 *Model.* **2016**.
- 7
8
9
10
11 21. Borrelli, K. W.; Cossins, B.; Guallar, V. Exploring hierarchical refinement techniques for
12 induced fit docking with protein and ligand flexibility. *J. Comput. Chem.* **2010**, 31, 1224-1235.
- 13
14
15 22. Bahar, I.; Atilgan, A. R.; Erman, B. Direct evaluation of thermal fluctuations in proteins
16 using a single-parameter harmonic potential. *Folding & Design* **1997**, 2, 173-181.
- 17
18
19
20 23. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and
21 Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate
22 Quantum Chemical Calculations on Peptides†. *J. Phys. Chem. B* **2001**, 105, 6474-6487.
- 23
24
25 24. Zhang, L. Y.; Gallicchio, E.; Friesner, R. A.; Levy, R. M. Solvent Models for Protein-Ligand
26 Binding: Comparison of Implicit Solvent Poisson and Surface Generalized Born Models With
27 Explicit Solvent Simulations. *J. Comput. Chem.* **2001**, 22, 591-607.
- 28
29
30
31
32 25. Espona-Fiedler, M.; Soto-Cerrato, V.; Hosseini, A.; Lizcano, J. M.; Guallar, V.; Quesada,
33 R.; Gao, T.; Perez-Tomas, R. Identification of dual mTORC1 and mTORC2 inhibitors in
34 melanoma cells: prodigiosin vs. obatoclox. *Biochem. Pharmacol.* **2012**, 83, 489-96.
- 35
36
37
38 26. Hosseini, A.; Espona-Fiedler, M.; Soto-Cerrato, V.; Quesada, R.; Pérez-Tomás, R.;
39 Guallar, V. Molecular interactions of prodiginines with the BH3 domain of anti-apoptotic Bcl-2
40 family members. *PLoS ONE* **2013**, 8, e57562.
- 41
42
43
44 27. Wu, H.; Zeng, H.; Dong, A.; Li, F.; He, H.; Senisterra, G.; Seitova, A.; Duan, S.; Brown, P.
45 J.; Vedadi, M.; Arrowsmith, C. H.; Schapira, M. Structure of the Catalytic Domain of EZH2
46 Reveals Conformational Plasticity in Cofactor and Substrate Binding Sites and Explains
47 Oncogenic Mutations. *PLoS ONE* **2013**, 8, e83737.
- 48
49
50
51
52
53
54
55 28. Italiano, A. Role of the EZH2 histone methyltransferase as a therapeutic target in
56 cancer. *Pharmacol. Ther.* **2016**, 165, 26-31.
- 57
58
59
60

1
2
3 29. Dairi, K.; Tripathy, S.; Attardo, G.; Lavallee, J.-F. Two-step synthesis of the bipyrrrole
4 precursor of prodigiosins. *Tetrahedron Lett.* **2006**, 47, 2605-2606.
5
6

7
8 30. Pinkerton, D. M.; Banwell, M. G.; Willis, A. C. Total Syntheses of Tambjamines C, E, F, G,
9 H, I and J, BE-18591, and a Related Alkaloid from the Marine Bacterium *Pseudoalteromonas*
10 tunicate. *Org. Lett.* **2007**, 9, 5127-5130.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60