**Binding Motifs in the CBP Bromodomain: An Analysis of 20 Crystal Structures of Complexes with Small Molecules**

Jian Zhu, ‡ Jing Dong, ‡ Laurent Batiste, ‡ Andrea Unzue, ‡ Aymeric Dolbois, ‡ Vlad Pascanu, ‡ Pawel Śledź, ‡ Cristina Nevado, ‡,*‡ and Amedeo Calisch ‡,*‡

†Department of Biochemistry, and ‡Department of Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Supporting Information

**ABSTRACT:** We analyze 20 crystal structures of complexes between the CBP bromodomain and small-molecule ligands that belong to eight different chemotypes identified by docking. The binding motif of the moiety that mimics the natural ligand (acetylated side chain of lysine) at the bottom of the binding pocket is conserved. In stark contrast, the rest of the ligands form different interactions with different side chains and backbone polar groups on the outer rim of the binding pocket. Hydrogen bonds are direct or water-bridged. van der Waals contacts are optimized by rotations of hydrophobic side chains and a slight inward displacement of the ZA loop. Rare types of interactions are observed for some of the ligands.

**KEYWORDS:** CBP bromodomain, docking, fragment-based ligand design, structure-based hit optimization, cancer

**RESULTS**

We use the term headgroup (or simply head) for the part of the ligands that mimics the side chain of the acetylated lysine of the endogenous ligand. The remaining part of the ligand is described as linker and tail group. Despite the variability of head groups there are two conserved hydrogen bonds at the bottom of the acetyl-lysine pocket with two side chains that are present in most of the 61 human bromodomains. These are a direct hydrogen bond with the side chain of Asn1168 and a water-bridged polar interaction with the Tyr1125 (Figure 1A). While in compounds 1–19 a single carbonyl oxygen forms both hydrogen bonds, the N3 and N9 atoms of the adenine headgroup of compound 20 are involved in the direct and water-bridged hydrogen bonds (Figure 1B), respectively.

Another conserved interaction in the bottom of the pocket is the optimal packing of a methyl with the side chain of Phe1111. The only two exceptions are the azobenzene derivative 1 which has a chlorine atom next to the phenyl ring of Phe1111 (Figure 1D) and the adenine derivative 20 which has its aromatic C8 atom close to Phe1111.

In stark contrast to the conserved binding motif of the head groups, the tail groups show a large heterogeneity of orientations and interactions (Figures 1A and 2). The intermolecular contacts involve residues of the CBP bromodomain with different flexibility. The side chains of the binding pockets can be grouped into three classes: fully rigid (Pro1106, Pro1110, Phe1111, Val1115, Tyr1125, Phe1115).
Table 1. 2D Structures and Assay Results of CBP Bromodomain Inhibitors

<table>
<thead>
<tr>
<th>Cpd</th>
<th>2D Structure</th>
<th>$K_D$ (pM)</th>
<th>IC$_{50}$ (pM)</th>
<th>LE</th>
<th>PDB ID (Res.)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>ND</td>
<td>ND</td>
<td>-</td>
<td>SEIC (1.50)</td>
<td>This study</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>ND</td>
<td>ND</td>
<td>-</td>
<td>50WK (1.25)</td>
<td>This study</td>
</tr>
<tr>
<td>3</td>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>4</td>
<td>0.37</td>
<td>5EP (1.20)</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>4</td>
<td>0.74</td>
<td>5MQE (1.65)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>85</td>
<td>455</td>
<td>0.40</td>
<td>5MPZ (1.40)</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td><img src="image6.png" alt="Structure 6" /></td>
<td>ND</td>
<td>31% @ 0.5M</td>
<td>-</td>
<td>5MQK (1.53)</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td><img src="image7.png" alt="Structure 7" /></td>
<td>20</td>
<td>0.34</td>
<td>4TS (2.00)</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td><img src="image8.png" alt="Structure 8" /></td>
<td>ND</td>
<td>37% @ 0.5M</td>
<td>-</td>
<td>5MQG (1.33)</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td><img src="image9.png" alt="Structure 9" /></td>
<td>1.4</td>
<td>0.30</td>
<td>5MME (1.35)</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td><img src="image10.png" alt="Structure 10" /></td>
<td>0.970</td>
<td>7.7</td>
<td>0.32</td>
<td>5MMG (1.23)</td>
<td>This study</td>
</tr>
<tr>
<td>11</td>
<td><img src="image11.png" alt="Structure 11" /></td>
<td>0.770</td>
<td>ND</td>
<td>0.35</td>
<td>4TQN (1.70)</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td><img src="image12.png" alt="Structure 12" /></td>
<td>1.4</td>
<td>ND</td>
<td>0.24</td>
<td>5ENG (1.30)</td>
<td>This study</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cpd</th>
<th>2D Structure</th>
<th>$K_D$ (pM)</th>
<th>IC$_{50}$ (pM)</th>
<th>LE</th>
<th>PDB ID (Res.)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td><img src="image13.png" alt="Structure 13" /></td>
<td>0.003</td>
<td>0.015</td>
<td>0.38</td>
<td>5MPK (1.90)</td>
<td>This study</td>
</tr>
<tr>
<td>14</td>
<td><img src="image14.png" alt="Structure 14" /></td>
<td>0.035</td>
<td>0.019</td>
<td>0.31</td>
<td>5NLK (1.80)</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td><img src="image15.png" alt="Structure 15" /></td>
<td>ND</td>
<td>2.2</td>
<td>0.30</td>
<td>6FQO (1.43)</td>
<td>This study</td>
</tr>
<tr>
<td>16</td>
<td><img src="image16.png" alt="Structure 16" /></td>
<td>ND</td>
<td>0.140</td>
<td>0.28</td>
<td>6FQO (1.35)</td>
<td>This study</td>
</tr>
<tr>
<td>17</td>
<td><img src="image17.png" alt="Structure 17" /></td>
<td>ND</td>
<td>0.146</td>
<td>0.27</td>
<td>6FR0 (1.50)</td>
<td>This study</td>
</tr>
<tr>
<td>18</td>
<td><img src="image18.png" alt="Structure 18" /></td>
<td>ND</td>
<td>0.040</td>
<td>0.31</td>
<td>6FRF (2.10)</td>
<td>This study</td>
</tr>
<tr>
<td>19</td>
<td><img src="image19.png" alt="Structure 19" /></td>
<td>0.085</td>
<td>0.059</td>
<td>0.42</td>
<td>5MPN (1.23)</td>
<td>This study</td>
</tr>
<tr>
<td>20</td>
<td><img src="image20.png" alt="Structure 20" /></td>
<td>20</td>
<td>ND</td>
<td>0.29</td>
<td>5H85 (1.70)</td>
<td>This study</td>
</tr>
</tbody>
</table>

Diffraction resolutions are reported in angstroms. The head groups of compounds 1 to 20 can be clustered in eight chemotypes (horizontal lines). The 2D structures are oriented with the mimic of the acetylated lysine on the bottom right. The values of the dissociation constant ($K_D$) were determined in duplicate by the BROMOscan competition binding assay (Figure S5). The IC$_{50}$ values were measured by AlphaScreen (Figure S6).\textsuperscript{14} Ligand efficiency\textsuperscript{15} is calculated as $LE = -(1.37/HA) \times \log K_D$, in units of kcal/mol per heavy atom (HA). The IC$_{50}$ value is used if the $K_D$ value is not available.
Tyr1167, Asn1168, the gatekeeper Val1174, and Phe1177), partially flexible (Leu1120 and Ile1122), and fully flexible (Gln1113, Leu1119, and Arg1173) (Figure 1C). Ligand binding is stabilized by van der Waals interactions with the rigid side chains and the partially flexible Leu1120 and Ile1122.

The side chains of Leu1120 and Ile1122 are oriented according to the steric requirements of the different ligands. Furthermore, the ZA loop (residues Arg1103 to Asp1134) adapts to the size of the ligands and shows a slight displacement toward the center of the binding site with respect to the apo structure (Figure 1A).

Concerning the highly flexible side chains, Gln1113 shows two main orientations; it points toward the binding site or completely outside. The inside-orientation of Gln1113 is stabilized by a hydrogen bond with the carbonyl of Leu1109 as in the apo structure (PDB 3DYW; Figure 2P). Furthermore, there are stacking interactions (e.g., ligands 13, 14, and 18; Figure 2g, 2H and 2L) or hydrogen bonds (e.g., ligands 9, 10, 12, and 14; Figure 2C, 2D, 2F, and 2H) with the tail group. Gln1113 points outward from the binding site in half of the 20 complexes.

The isoxazol-3-one ring in ligands 16 and 17 and the dimethylisoxazole ring in compound 18 are rotated by about 40 degrees with respect to the phenyl. As a consequence, the methyl next to the oxygen in the ring points downward from the binding site and displaces a water molecule which, in other structures (e.g., apo), is involved in a hydrogen bond to the carbonyl of Pro1114. Thus, it could be beneficial to further modify the isoxazole group to achieve polar interactions with the residues lining the ZA channel, e.g., to form a hydrogen bond with the carbonyl group of Pro1114.

The side chain of Leu1119 is involved in van der Waals contacts with three of the 20 ligands, i.e., 13 (closest distance 4.6 Å), 14 (closest distance 4.4 Å), and one conformation of 18 (closest distance 4.1 Å).

The interaction with the very flexible Arg1173 can be ionic (i.e., salt bridge with the benzoate of compounds 3 and 11 or the tetrazole of 13 and 19), hydrogen bond-mediated as in compounds 1, 14, 16, 17, and 18, charge-dipole as with the −CF$_2$− of compound 15 (a similar interaction has been reported for an ATAD2 bromodomain ligand), water-bridged (e.g., compound 9), or absent (compound 20).

The side chains of Pro1123 and Asp1124 assume a different orientation only in the complex with the azobenzene derivative 1, which is due to van der Waals attraction between the exposed benzene ring of the ligand and the Pro1123 side chain (Figure S2). The side chain of Asp1124 is involved in a favorable polar interaction with the hydroxyl of Tyr1167 in all structures except for the complex with compound 1.

Six of the 20 ligands (viz., compounds 3, 11–14, and 20) have an amide linker which is involved in water-bridged interactions with the backbone carbonyl of Gln1113 (3, 12–
The direct comparison of the amide linkers in compounds 3 and 20, which point in opposite directions in the 2D structures, shows that the same water-bridged interaction with the backbone carbonyl of Gln1113 can be formed with a donor -NH- or acceptor -CO- (Figure 2B and O). Ligands 9 and 10 have a -NH2- linker which is involved in water-bridged hydrogen bond interactions with the backbone carbonyl of Gln1113 similar to the amide linker containing ligands.

It is interesting to analyze the influence of the linker in the 12 acetyl-benzene derivatives (compounds 8–19). A comparison of compounds 13 and 19 shows that the difference between direct covalent bond and amide linker does not influence the orientation of the tetrazole tail group which makes in both cases an ionic interaction with the side chain of Arg1173 (Figure 2G and 2N). The amide linker in ligands 11 and 14 shows opposite orientation as compared to that of the ligands 12 and 13, thereby influencing the position of the

Figure 2. Interactions between the CBP bromodomain and small-molecule ligands. The spheres represent four conserved water molecules (red) and other water molecules involved in ligand binding (green). Conserved hydrogen bonds with the side chains of Asn1168 and Tyr1125 are shown (dashed lines in salmon). Other interactions between ligands and CBP bromodomain are emphasized (dashed lines in blue). (M) Ligand 18 shows alternate binding modes in different protein molecules in the asymmetric unit. (P) The conformation of Gln1113 is stabilized by hydrogen bonding interactions in the apo crystal structure of CBP bromodomain (PDB ID 3DWY).
benzene ring (Figure 2 and S3A). Concerning the \(-\text{NH}_2\)-linker, the overlap of compounds 9 and the amide-linked 12 indicates that there is a substantial difference in the orientation of the benzene ring and the methyl ester in the meta position of the tail group (Figure S3B).

The purine ring in compound 20 shows a similar binding mode to the conserved asparagine (Asn1168 in CBP bromodomain) and tyrosine (Tyr1125 in CBP bromodomain) as the 2-amine-9H-purine containing ligand 7d in the bromodomain of BRD9 (Figure S4A and S4B). The mercaptopurine compound 4 binds to the BRPF1 bromodomain in a different way as the carbonyl group on the purine ring forms a bifurcated hydrogen bond with the conserved asparagine and tyrosine (Asn708 and Tyr665 in BRPF1 bromodomain, respectively) (Figure S4C). The orientation of the purine scaffold in these three compounds is influenced by the size of the so-called “gatekeeper” residue (Val1174, Tyr106, and Phe714 in CBP, BRD9, and BRPF1 bromodomain, respectively) (Figure S4D), which is consistent with a previous report.

We do not discuss the structure–activity relationship in detail for the 20 compounds because of their heterogeneity. Concerning the series of 12 derivatives of acetylbenzene (compounds 8 to 19), it emerges that low nM affinity can be reached by engaging Arg1173 via a salt bridge (compounds 13 and 19) or charge-dipole(s) interactions (compounds 14 and 18). The linker does not contribute directly to the binding energy as compounds 13 and 14 have an amide linker while compounds 18 and 19 feature a single covalent bond between the head and tail groups. The linker does contribute to the selectivity as the acetylbenzene derivatives with the amide linker are more selective against the N-terminal domain of BRD4 (BRD4(1)) because their tail group clashes into the Trp81 side chain of BRD4(1).12

## CONCLUSIONS

We have analyzed the binding motifs of 20 ligands of the CBP bromodomain. Nine of these ligands are docking hits9,21,22 while the remaining ligands (compounds 9–19) were synthesized in an optimization campaign.12 Most of these ligands have favorable ligand efficiency (LE > 0.30 kcal/(mol HA)) irrespective of their size (Table 1). The conserved hydrogen bond with the NH3 of Asn1168 is observed for the 20 ligands similarly to the hydrogen bond formed by the natural ligand acetyl-lysine.

Three new observations emerge from the analysis of the structural alignment. First, the interactions with the rim of the binding site are very heterogeneous. The variability in these interactions is congruent with the chemical diversity of the tail groups and flexibility of the ZA loop as observed in crystal structures and molecular dynamics simulations.23–28 Compared to the apo structure, the ZA loop is slightly displaced toward the BC loop to optimize the van der Waals contacts with the ligands.

Second, there are multiple water-bridged interactions, and some rarely observed interaction types. The latter include a charge–dipole interaction between -CF2- on ligand 15 and guanidinium of Arg1173, halogen–aromatic interaction between the chlorine of ligand 1 and phenyl of Phe1111, and aromatic–amide packing of the thiazole of compound 13 and the side chain of Gln1113.

Third, the 12 compounds that share an acetyl-benzene moiety as head can be clustered in three different groups according to the linker that connects the head and tail groups. The binding modes of the compounds with the amide linker are more similar to the poses of the compounds with a direct covalent bond than the derivatives with an amino linker. Thus, the relative orientation of head and tail groups has a stronger influence on the binding mode than their distance.

All crystal structures have been released in the PDB which should facilitate the structure-based optimization of CBP ligands.

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsmedchemlett.8b00286.

Experimental methods, X-ray crystal structure statistics data, electron density maps of bound ligands, characterization data of all synthesized compounds, structural analyses of all reported crystal structures, and dose–response curves for both BROMOscan and AlphaScreen assays (PDF)

Molecular formula strings for ligands 1 to 20 (XLSX)

### Accession Codes

PDB accession codes of the structures of the CBP bromodomain in complex with small molecules are I(S5IC), 2(S7WK), 3(S5EP7), 9(S5MME), 10(S5MMG), 12(S5ENG), 13(S5PK), 15(S5FUQ), 16(S5QO), 17(S6R0), 18(S6RFF), 19(S5PN), and 20(S5H8S). Authors will release the atomic coordinates and experimental data upon article publication.

### AUTHOR INFORMATION

#### Corresponding Authors

* Phone: +41 44 635 55 21. E-mail: caflisch@bioc.uzh.ch.
* Phone: +41 44 635 39 45. E-mail: cristina.nevado@chem.uzh.ch.

#### ORCID

Jian Zhu: 0000-0002-2486-3658
Cristina Nevado: 0000-0002-3297-581X
Amedeo Caflisch: 0000-0002-2317-6792

#### Funding

We acknowledge financial support from the Swiss National Science Foundation (to A.C.), the Sinergia program (to C.N. and A.C.), the Synapsis Foundation-Alzheimer Research Switzerland, and the Heidi Seiler-Stiftung (to A.C., C.N., and P.S.).

#### Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

We thank the staff at X06DA and X06SA beamlines, Swiss Light Source, Paul Scherrer Institute (Villigen, Switzerland) for assistance in data collection.

### ABBREVIATIONS

Alpha, amplified luminescence proximity homogeneous assay; BRD4, bromodomain 4; BRD9, bromodomain 9; BRPF1, Bromodomain and PHD Finger Containing Protein 1; CBP, CREB Binding Protein; CREB, cyclic-AMP response element binding protein.
Inhibitor of Cyclic Adenosine Monophosphate Response Element
GNE-781, A Highly Advanced Potent and Selective Bromodomain
Ly, J.; Maher, J.; Masui, C.; Merchant, M.; Ran, Y.; Taylor, A. M.;
Grogan, J. L.; Hatzivassiliou, G.; Huang, W.; Hunsaker, T. L.;
Crawford, T. D.; Cyr, P.; de Almeida Nagata, D.; Gascoigne, K. E.;
An, L.; Beresini, M. H.; de Leon Boenig, G.; Bronner, S. M.; Chan, E.;
Li, C.; Wu, C.; Li, K.; Hui, X.; Zhou, Y.; Smaill, J. B.; Patterson, A. V.;
(14) Philpott, M.; Yang, J.; Tumber, T.; Fedorov, O.; Uttarkar, S.;
Filippakopoulos, P.; Picaud, S.; Keates, T.; Felletter, I.; Ciulli, A.;
(18) Bamborough, P.; Chung, C. W.; Demont, E. H.; Furze, R. C.;
Bannister, A. J.; Che, K. H.; Di交流合作, C.; Grandi, P.;
Kouzarides, T.; Michon, A. M.; Mitchell, D. J.; Prinjha, R. K.; Rau, C.;
(22) Spiliotopoulos, D.; Wamhoff, E. C.; Lolli, G.; Rademacher, C.;
(23) Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.;