Supporting Information

for Adv. Sci., DOI: 10.1002/advs.202102435

Discovery of Novel GR Ligands towards the Druggable GR Antagonist Conformations Identified by MD Simulations and Markov State Model Analysis

Xueping Hu†, Jinping Pang†, Jintu Zhang†, Chao Shen, Xin Chai, Ercheng Wang, Haiyi Chen, Xuwen Wang, Mojie Duan, Weitao Fu, Lei Xu, Yu Kang, Dan Li*, Hongguang Xia*, Tingjun Hou*
Discovery of Novel GR Ligands towards the Druggable GR Antagonist Conformations Identified by MD Simulations and Markov State Model Analysis

Xueping Hu\textsuperscript{1,2,†}, Jinping Pang\textsuperscript{1,†}, Jintu Zhang\textsuperscript{1,†}, Chao Shen\textsuperscript{1}, Xin Chai\textsuperscript{1}, Ercheng Wang\textsuperscript{1}, Haiyi Chen\textsuperscript{1}, Xuwen Wang\textsuperscript{1}, Mojie Duan\textsuperscript{3}, Weitao Fu\textsuperscript{1}, Lei Xu\textsuperscript{4}, Yu Kang\textsuperscript{1}, Dan Li\textsuperscript{1,*}, Hongguang Xia\textsuperscript{5,*}, Tingjun Hou\textsuperscript{1,2,*}

\textsuperscript{1} Innovation Institute for Artificial Intelligence in Medicine of Zhejiang University, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China
\textsuperscript{2} State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310058, Zhejiang, China
\textsuperscript{3} Key Laboratory of magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhu, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, Hubei, China
\textsuperscript{4} Institute of Bioinformatics and Medical Engineering, School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, 213001, China
\textsuperscript{5} Department of Biochemistry & Research Center of Clinical Pharmacy of The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, China

\textsuperscript{†}These authors contribute equally to this work.

\textsuperscript{*} Corresponding authors

Dan Li
Email: lidancps@zju.edu.cn

Hongguang Xia
Email: hongguangxia@zju.edu.cn

Tingjun Hou
E-mail: tingjunhou@zju.edu.cn
Table S1. The conformational ensemble of AF2 induced by different ligands in the 10 μs MD simulations based on the C\textsubscript{\alpha} atoms RMSD of residues 567-579, 590-597 and 754-764.

| Ligand  | States  | Populations (%), RMSD<2.5Å | MIN\textsubscript{RMSD} (Å) |
|---------|---------|-----------------------------|-----------------------------|
| apo     | 3H52_A  | ~ 0.0                        | 2.05                        |
|         | 3H52_B  | 2.0                          | 1.68                        |
|         | 3H52_C  | 2.0                          | 1.15                        |
|         | 1M2Z    | 0.0                          | 2.53                        |
| DEX     | 3H52_A  | 2.0 ↑                        | 1.74                        |
|         | 3H52_B  | 0.5 ↓                        | 1.95                        |
|         | 3H52_C  | 3.0 ↑                        | 1.42                        |
|         | 1M2Z    | ~ 0.0 ↑                      | 2.20                        |
| AZD9567 | 3H52_A  | 1.5 ↑                        | 1.77                        |
|         | 3H52_B  | 1.5 ↓                        | 1.90                        |
|         | 3H52_C  | 1.0 ↓                        | 1.47                        |
|         | 1M2Z    | ~ 0.0 ↑                      | 2.46                        |
| RU486   | 3H52_A  | ~ 0.0 ~                      | 1.95                        |
|         | 3H52_B  | ~ 0.0 ↓                      | 2.18                        |
|         | 3H52_C  | 6.0 ↑                        | 1.27                        |
|         | 1M2Z    | ~ 0.0 ↑                      | 2.16                        |

Note: active antagonist state, 3H52_A; partial active antagonist/agonist state, 3H52_B; passive antagonist state, 3H52_C; agonist state, 1M2Z.
Table S2. The chemical structures of the 88 compounds identified by structure-based virtual screening.

| No. | SMILES                                                                 | ChemDiv ID | MW  |
|-----|------------------------------------------------------------------------|------------|-----|
| HP-1| c1ccccc1CC(CC)NC(=O)CSCc2c(-n3cccc3)n(ne2)-c4ccccc4                     | E008-0202  | 444.6 |
| HP-2| Cc1cc(c(c1)OC)NC(=O)C2CCN(CC2)S(=O)(=O)c(c3)ccc(c34)SCc(c=O)N4        | E959-1620  | 475.59 |
| HP-3| c1cc(S(=O)(=O)N)cc1CCNC(C2)=OCC(=O)N2c3ccc(c3)OCC                      | 5593-1275  | 417.49 |
| HP-4| CC(C1)CCc(c12)sc(c2C(=O)N)NC(=O)c3ccc(c3)S(=O)(=O)N(C)c4cccc4        | 2513-0642  | 483.61 |
| HP-5| CC(=O)Nc1ccc(cc1)NS(=O)(=O)c(c2C)c(c2)-c3cnc(o3)C4CC4                 | G408-2569  | 411.48 |
| HP-6| Cc1cc(cc1)N2CCN(CC2)C(=O)c(c3)nc(c34)5c5c(nC)c4=O)cccc5              | C593-0520  | 428.54 |
| HP-7| CC(=O)c1cc(c(c1)NS(=O)(=O)c(c2C)c(c2)-c(nn3)cc3-c(cc4)cccc4          | M348-0226  | 461.55 |
| HP-8| CS(=O)(=O)N(C(=O)c1c(c12)cc(c2C(=O)N)c3(c34)cccc3-c4cc(OC)cccc4     | Y041-3684  | 439.49 |
| HP-9| Cc1cc(c(c1)Sc(c1)=O)(=O)N2C(CC2=O)C(=O)N3CCCN(CC3)c4cc(Cl)cccc4      | G631-1649  | 476   |
| HP-10| CC1CCN(CC1)C(=O)C2CCN(CC2)S(=O)(=O)c(c(n3)cc3-C=C/C4(c4)(F)cccc4   | G637-0921  | 475.59 |
| HP-11| CC(C)C1ccc(cc1)Sc(c1)=O)c(c2C)c(N3CCCNCC3)cc(c24)n(C)c(c1)n4CC        | G266-0187  | 484.62 |
| HP-12| c1ccccc1CN(c(c1)2)c(c(c23)n(C)c=c4cc(CC1)c4cc(cc4)cccc4            | C301-5938  | 436.54 |
| HP-13| FC(F)c1cc(c(c1)N2CCN(CC2)S(=O)(=O)c(c(c(C)c3)cc(c34)4[nH]c(c1)nH]4 | G801-0347  | 440.45 |
| HP-14| c1ccccc1CC(=O)c(c(c3)cc(c23)sc(c6)=O)cccc4                              | 6807-1451  | 485.54 |
| HP-15| c1cccc(c12)ccc(c2)NC(=O)CN(c3cccc(C)c3)S(=O)(=O)c4cccc4              | 4577-1565  | 444.56 |
| HP-16| Fe1cc(Cl)cc(c1)NC(=O)CN(C)S(=O)(=O)c4cc4                                | G855-4250  | 453.92 |
| HP-17| Cc1cc(c(c1)N(C)=O)C(C)n(n2)cc2-c(c3)cc(c34)cccc3(c=O)cccc4           | D315-1286  | 494.62 |
| HP-18| CC(C)CNC(=O)C1CCCN1CC(=O)e2cc(c(c2)cc(C)=O)c(c3)cc(C)c3            | L027-0277  | 457.6 |
| HP-19| o1ccccc1C(=O)N(CCC2)c(c23)cc(c3)NS(=O)(=O)c(c4)cc(c4)OC                | G503-0123  | 440.52 |
| HP-20| c1ccccc(Cl)c1CN(C)=O)C(C(C)C)SC(=O)(=O)c2cc(c23) CNC(C(=O)CC          | C464-0916  | 492.04 |
| HP-21| c1cccc(c1)O)C(C(=O)OCC)NS(=O)(=O)c2cc3C(C)=O)c(c4)cccc4              | 3063-0200  | 451.5 |
| HP-22| Cc1cc(c(c1)=O)(=O)NC(C)c2nc(2o2)-c3ccc(c3)N4CCCC4                  | F373-0024  | 412.51 |
| HP-23| o1ccccc1CN(C)=O)C2(CC2)c3ccc(c3)(4S(=O)(=O)c(c4)cccc4                | L164-0284  | 570.55 |
| HP-24| c1ccccc1(C)CNC(=O)C(C)S(=O)(=O)c(c2)ccc(c23)N(C(=O)CC                 | E746-0740  | 428.55 |
| HP-25| c1cc(F)c(c(1OC)S(=O)(=O)Nc(c2c)cce2-c3ccc(n3)N4CCCC4                | G620-0764  | 442.52 |
| HP-84  | c1ccccc(C)c1C(=O)NC(=O)N2CCCN(2)C3CCCN(CC3)C(=O)CCc4ccccc4 | V015-6088 | 461.61 |
| HP-85  | c1ccccc1C(C)CNS(=O)(=O)c(cc2)ccc2CCN(C3=O)C(=O)c(c34)nccc4 | E734-2528 | 449.53 |
| HP-86  | c1ccccc1C(O)(c2ccccc2)C3CCCN(CC3)C(=S)Ne4ccc(cc4)S(=O)(=O)N | 6208-0881 | 481.64 |
| HP-87  | COC(=O)c1c(cccc1)NC(=O)CSe(c2S(=O)(=O)c3ccccc3)[nH]c(n2)-c4ccccc4 | D406-0335 | 507.59 |
| HP-88  | COC(=O)c1c(cccc1)NC(=O)CN(S(=O)(=O)C)c(c(Cl)cc2)cc2C(F)(F)F | 4577-1365 | 464.85 |
Figure S1. The evolution of the RMSD of Ile581-Lys777, and the distance between H12 (Ala754-Tyr764) and Leu589 over time for (a) apo-LBD, (b) dex-LBD, (c) azd-LBD, and (d) ru486-LBD.
Figure S2. The five metastable macrostates calculated by PCCA++ for apo-LBD (a), dex-LBD (b), azd-LBD (c), and ru486-LBD (d).
Figure S3. The population distributions of the four experimental AF2 conformations (3H52_A, 3H52_B, 3H52_C, and 1M2Z.) in the four systems (apo-LBD, dex-LBD, azd-LBD, and ru486-LBD). The RMSD of the AF2 C$_\alpha$ atoms (including the residues 567-579, 590-597 and 754-764) was used as the criterion to compare the similarity between the conformation from the 10 μs MD simulation trajectories and the experimental structure.
The implied timescales versus lag time for the (a) apo-LBD, (b) dex-LBD, (c) azd-LBD, and (d) ru486-LBD systems. The implied timescales of a dynamic system will tend to be constant with the increasing of the lag time if the system satisfies the Markov State Model. When estimating the Markov State Model, the smallest lag time was used to obtain the model with highest time resolution. Thus, we chose 20 steps (0.2 ns) as the lag time.

**Figure S4.**
Figure S5. Generalized Chapman-Kolmogorov tests for the eight microstates of the apo-LBD system. We compare the estimated transition probabilities calculated from the MD data (circles) and the predictions of the MSMs with different lag times. For the states 1 to 8, a nearly perfect agreement were observed for all the lag times, indicating the high Markovianity of the microstates.
**Figure S6.** Generalized Chapman-Kolmogorov tests for the eight microstates of the dex-LBD system.
Figure S7. Generalized Chapman-Kolmogorov tests for the eight microstates of the azd-LBD system.
Figure S8. Generalized Chapman-Kolmogorov tests for the eight microstates of the ru486-LBD system.
Figure S9. The structures of the main pathways in (a) apo-GR(a), (b) dex-LBD(b), (c) azd-LBD and (d) ru486-LBD.
Figure S10. The passive antagonist states (in cyan) in the azd-LBD system. AZD9567 in the passive antagonist states was shown in green. The RMSD was calculated using the CA atoms of AF2. The structure of PDB 6EL9 was shown in gray.

Figure S11. Inhibition percentage of cell viability in HeLa cell lines treated with 25 μM tested compounds for 48 h.
Figure S12. Generalized Chapman-Kolmogorov tests for the eight microstates of the HP19-LBD system.