# Supporting Information

# Discovery of Novel CCR5 Ligands as Anti-Colorectal Cancer Agents by Sequential Virtual Screening

Mariam A. El-Zohairy[†]; Darius P. Zlotos[†]; Martin R. Berger[‡]; Hassan H. Adwan[⊥*]; Yasmine M. Mandour[†§*]

[†] Pharmaceutical Chemistry Department, Faculty of Pharmacy and Biotechnology, The German University in Cairo, New Cairo City, 11835 Cairo, Egypt

[‡] Toxicology and Chemotherapy Unit, German Cancer Research Centre (DKFZ), DE 69120 Heidelberg, Germany

[⊥] Pharmacology and Toxicology Department, Faculty of Pharmacy and Biotechnology, The German University in Cairo, New Cairo City, 11835 Cairo, Egypt

[§] School of Life and Medical Sciences, University of Hertfordshire hosted by Global Academic Foundation, New Administrative Capital, 11578 Cairo, Egypt

**Keywords:**

CCR5; Colorectal cancer; Virtual Screening; Pharmacophore; Docking.

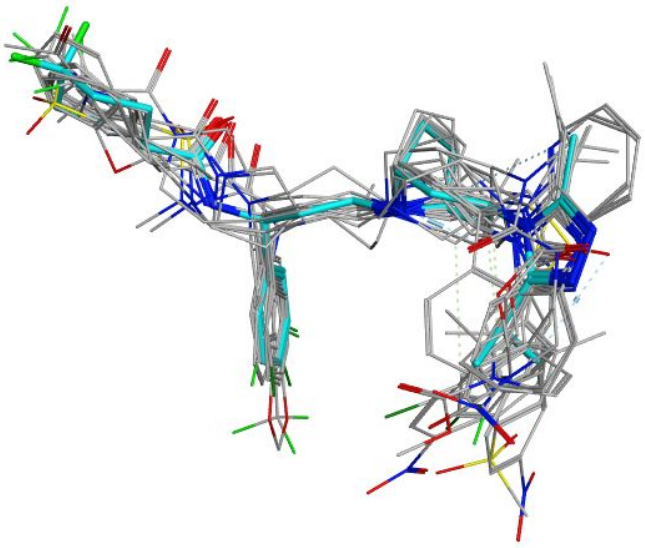


**Figure S1.** Alignment of the 16 training set compounds (grey, lines) on the co-crystallized structure of MVC (cyano, sticks).

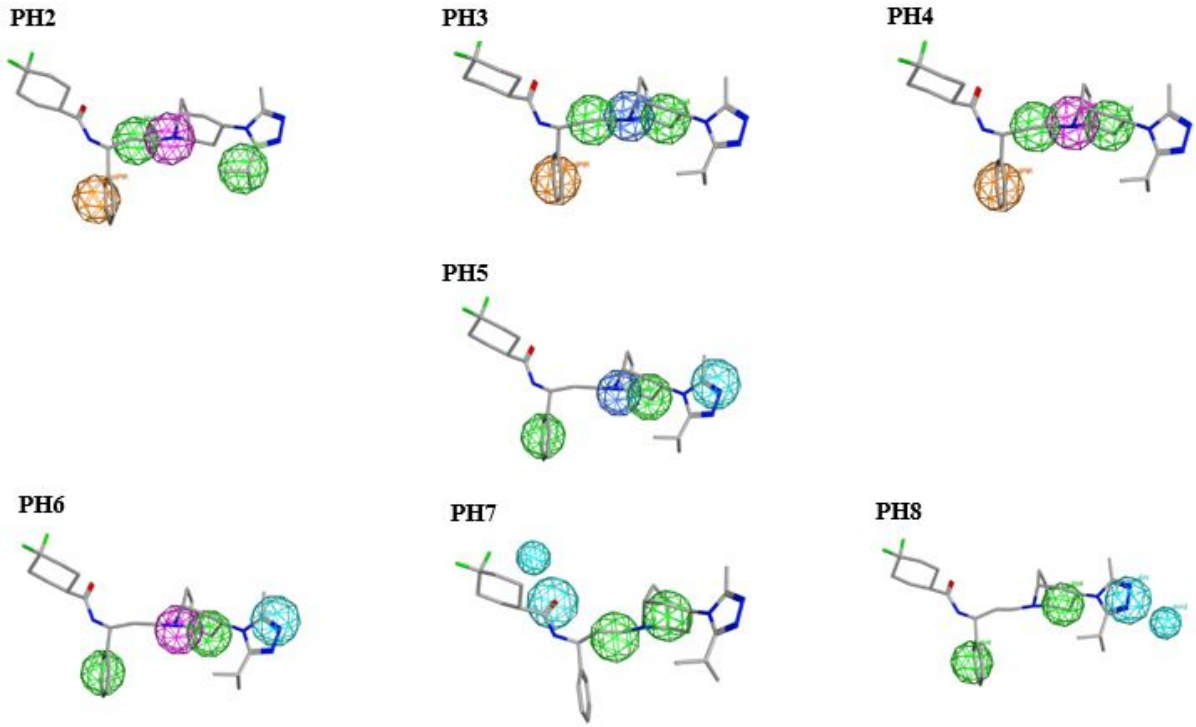



**Figure S2.** Ligand-based pharmacophore model (PH2-PH8) for CCR5 inhibitors overlaid on the crystal coordinates of MVC. 3D spatial arrangement between the chemical features of the pharmacophore model represented by blue (cationic center), orange (aromatic center), purple (Hydrogen bond donor), and green (hydrophobic center) spheres.

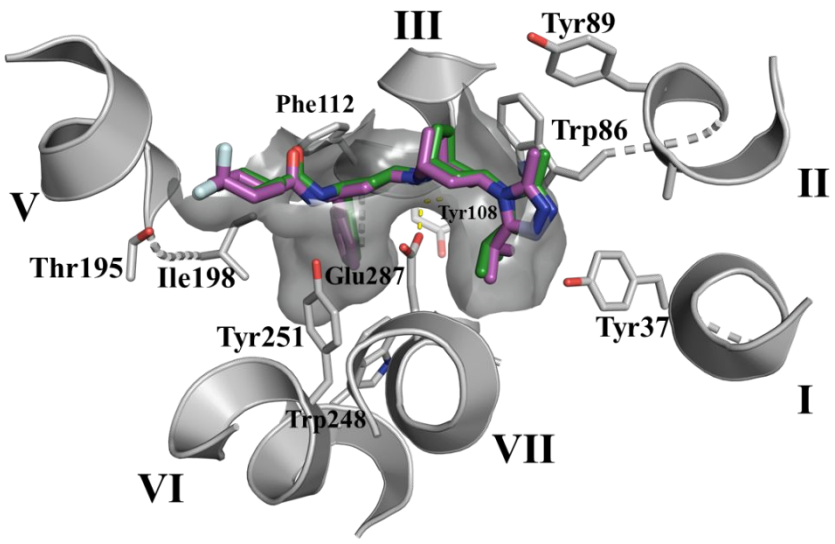



**Figure S3.** Graphical representation of MVC binding pocket of CCR5 (PDB code: 4MBS). Overlay of the docked pose (green) and PDB coordinates (magenta) of MVC showing residues from TMS1 and TMS2 (grey sticks). Only side chain atoms are shown for clarity.

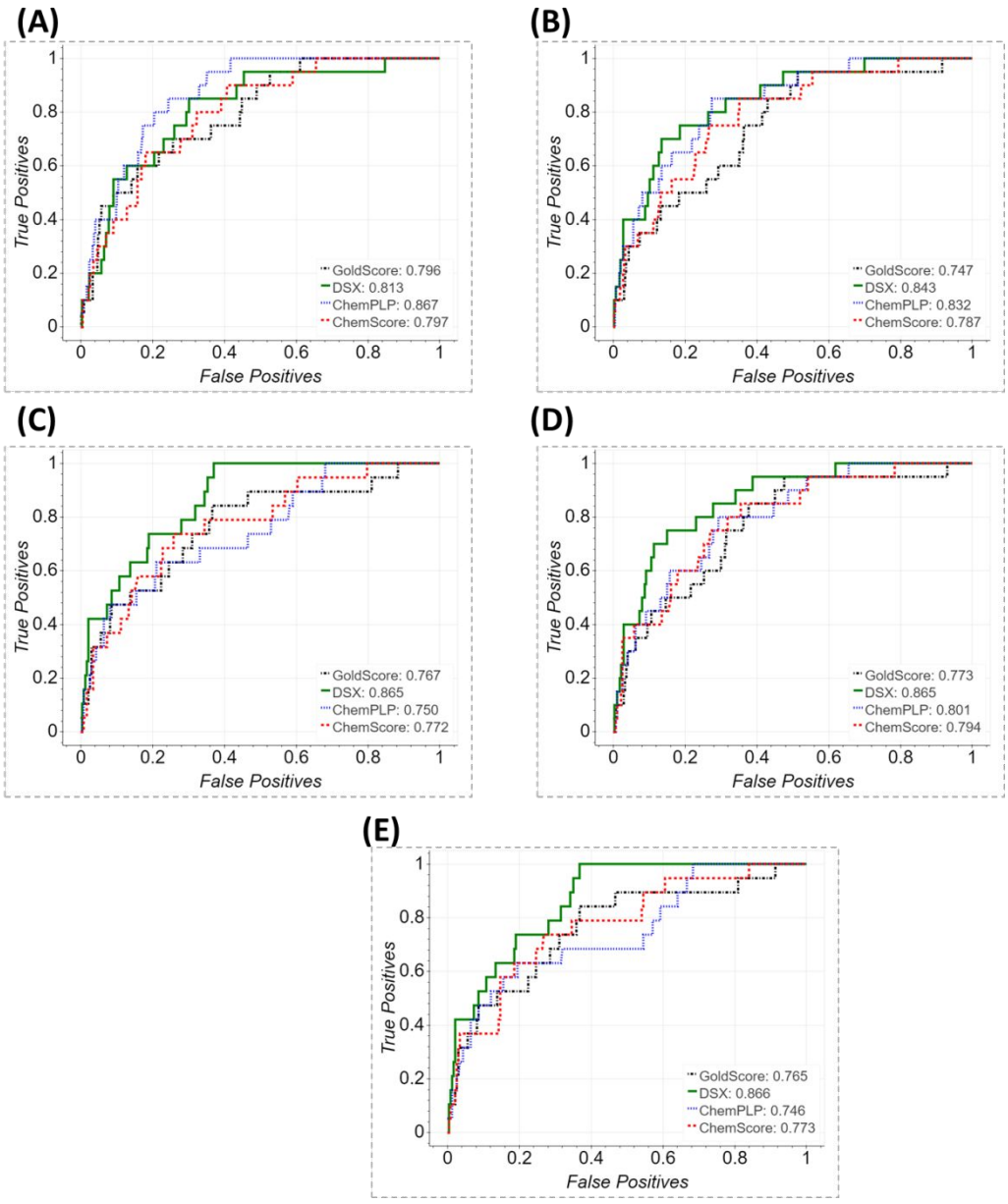



**Figure S4.** ROC curves obtained for the docked poses of validation database filtered by **(A)** none, **(B)** PLIF-M1, **(C)** PLIF-M2, **(D)** PLIF-M3, **(E)** PLIF-M4, and plotted for various scoring functions.

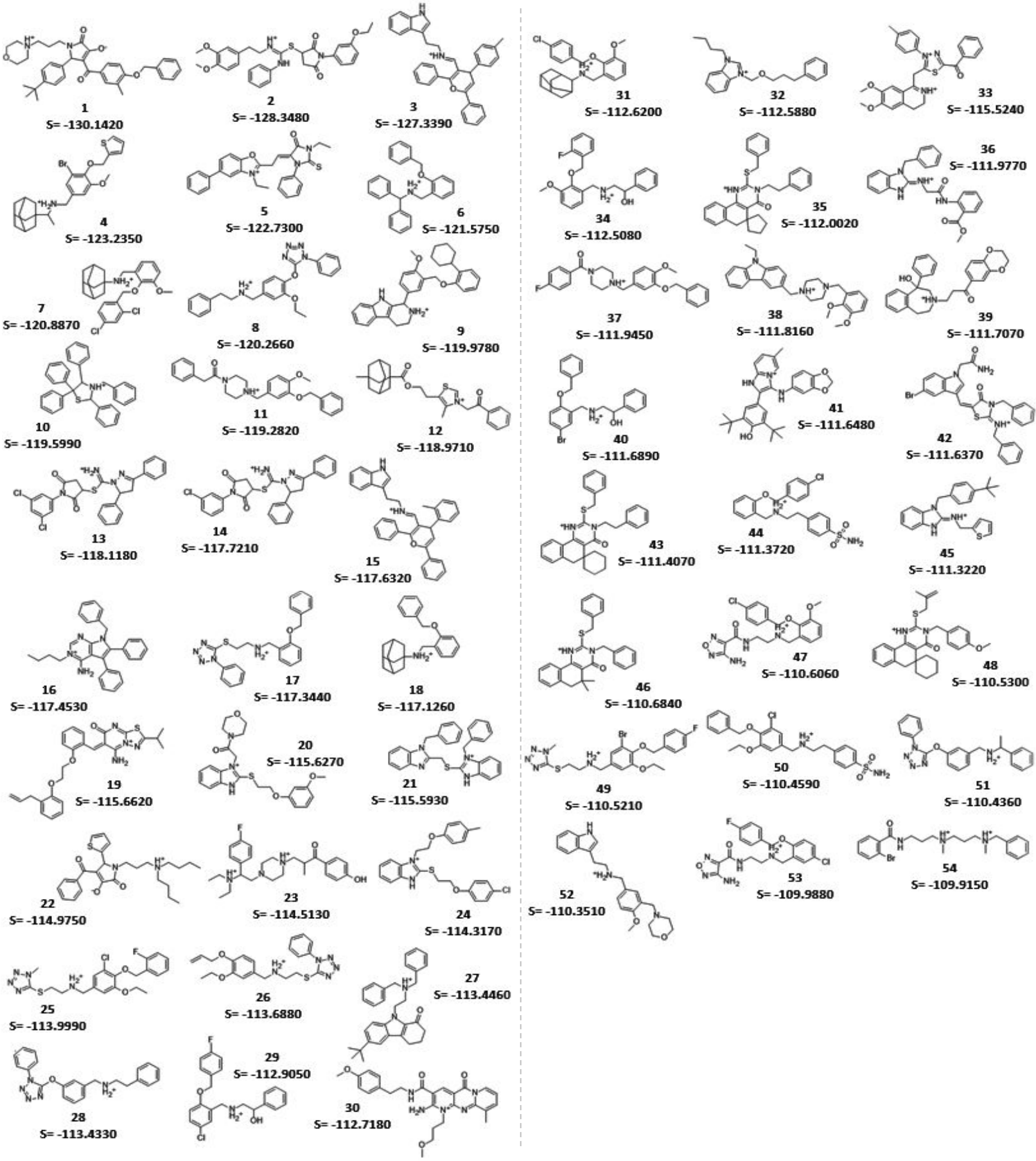
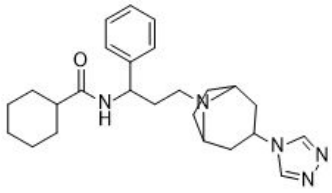
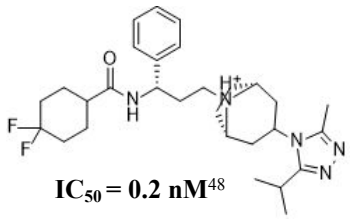
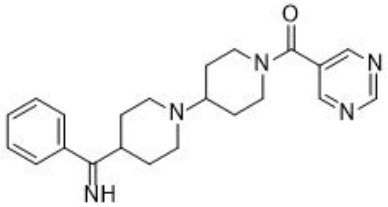
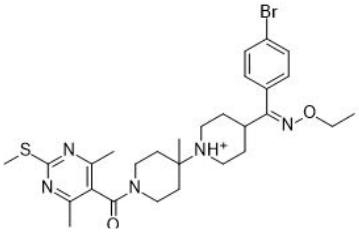
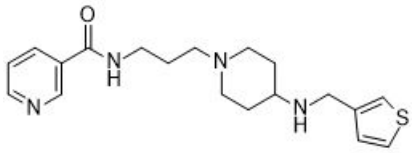
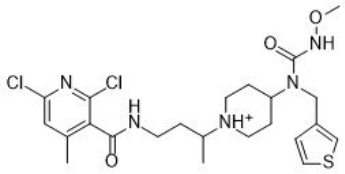
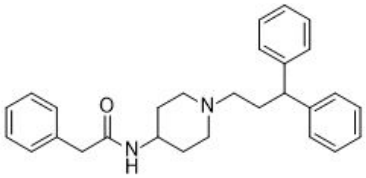
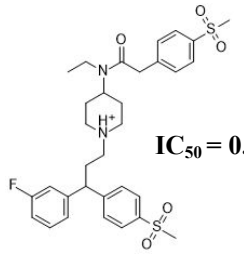
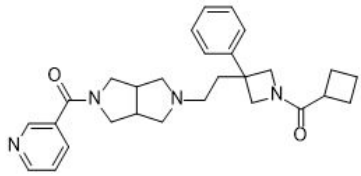
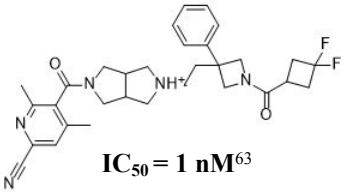
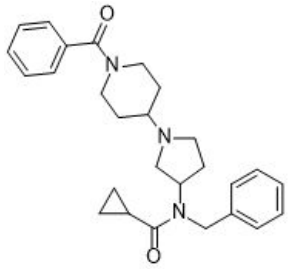
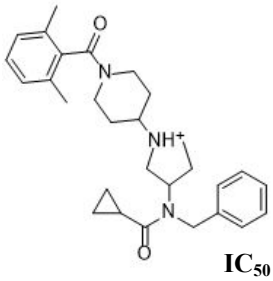
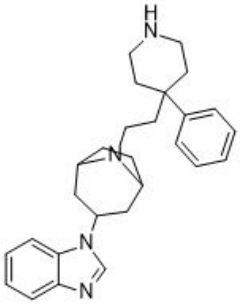
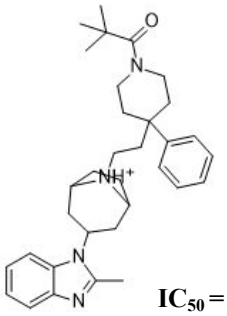
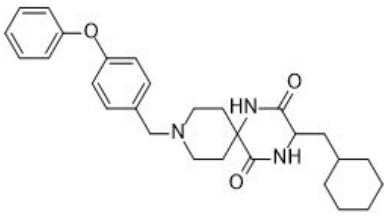
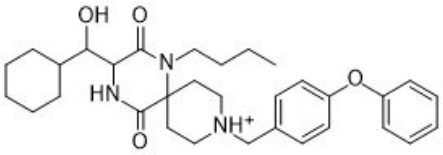
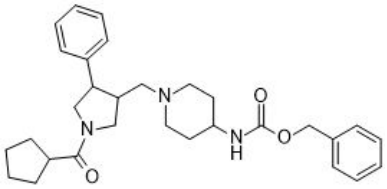
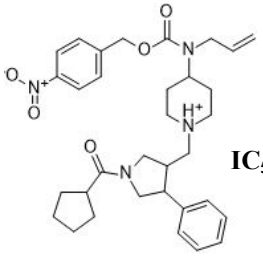
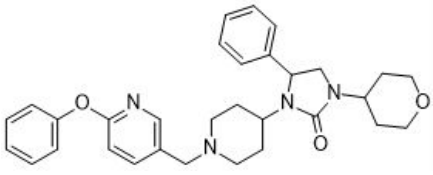
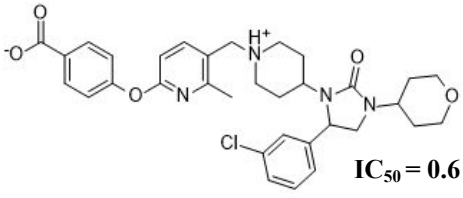
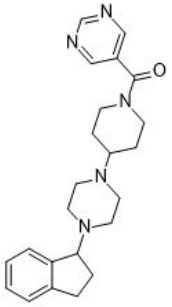
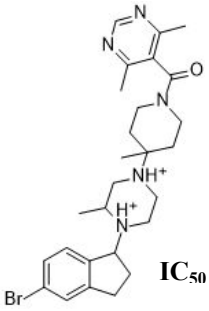
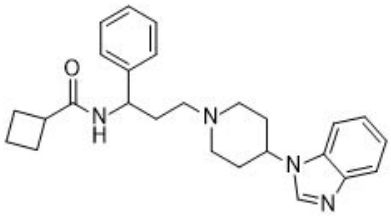
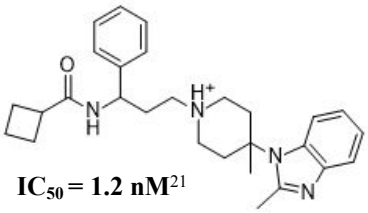
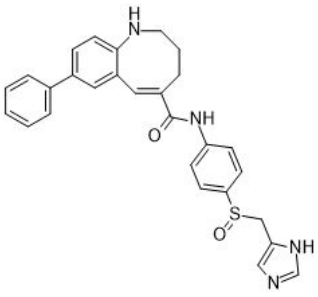
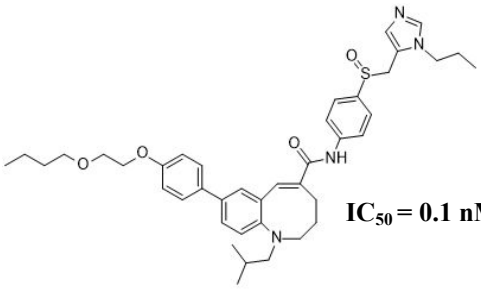
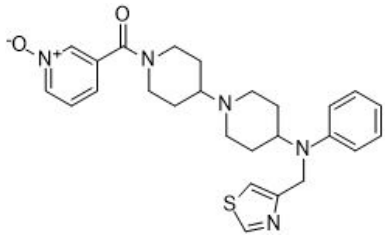
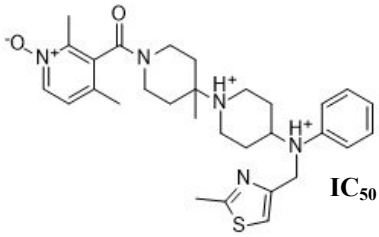
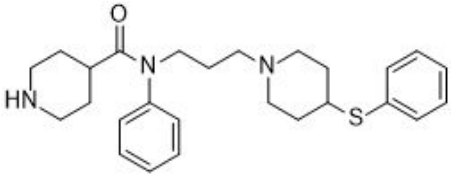
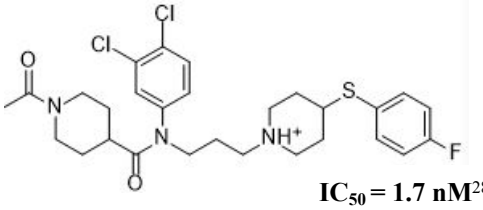


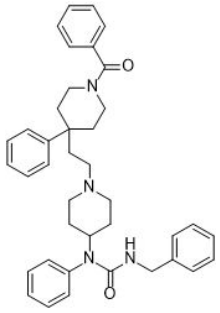
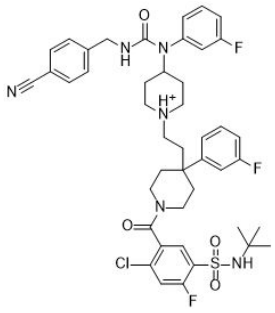
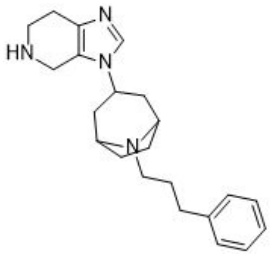
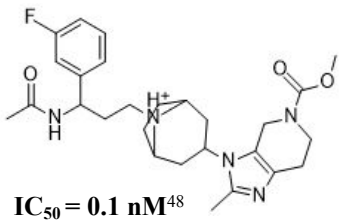
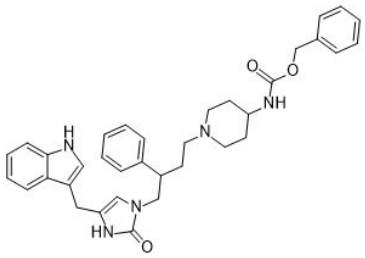
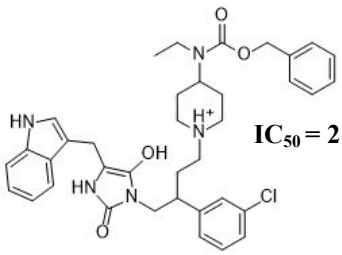
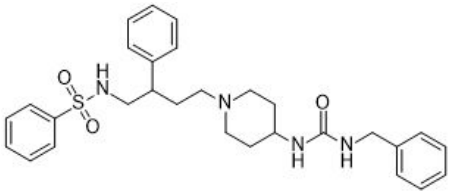
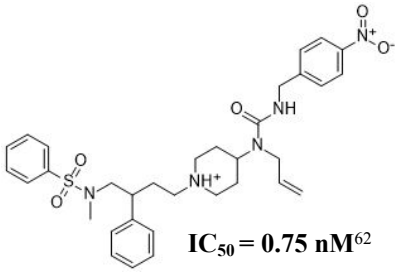
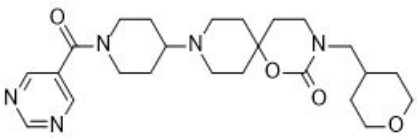
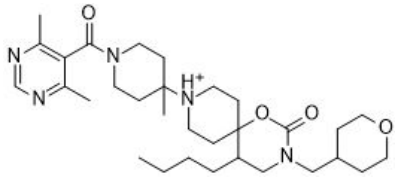

**Figure S5.** Chemical structure of the top 1% scoring compounds obtained by VS of Specs database. S represents the DSX scores.

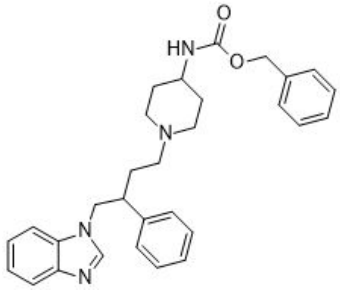
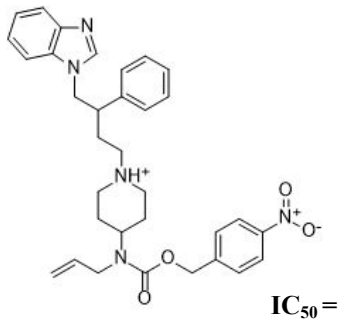
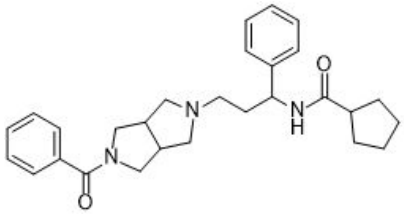
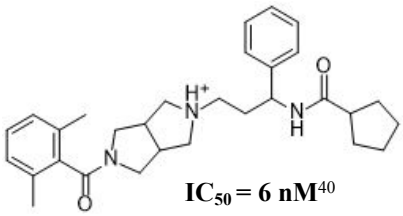
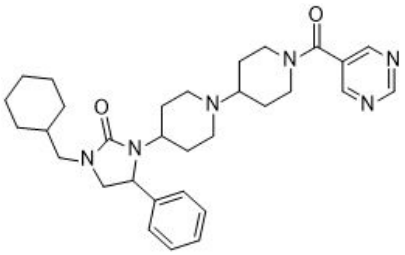
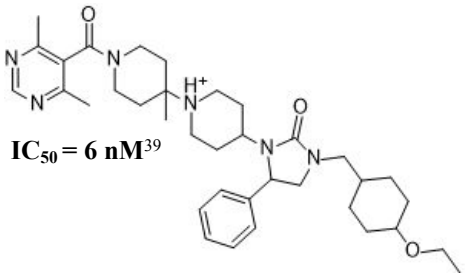
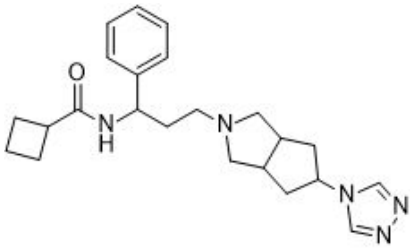
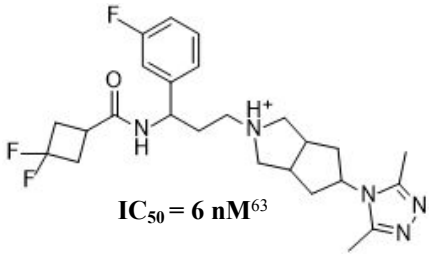
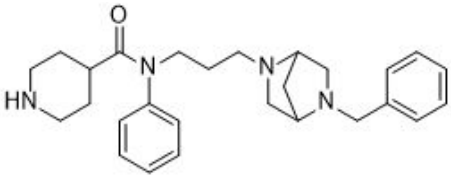
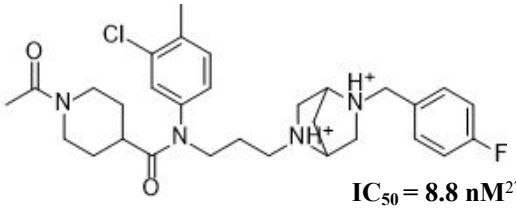
**Table S1.** Structural details of the 39 training set compounds for each chemical scaffold cluster

| Clusters Number | Scaffolds | Representative Structure |
|---|---|---|
| 1 |  |  $IC_{50}$ = 0.2 nM[48] |
| 2 |  |  Ki = 1.3 nM[36] |
| 3 |  |  $IC_{50}$ = 0.08 nM[44] |
| 4 |  |  $IC_{50}$ = 0.22 nM[60] |
| 5 |  |  $IC_{50}$ = 1 nM[63] |

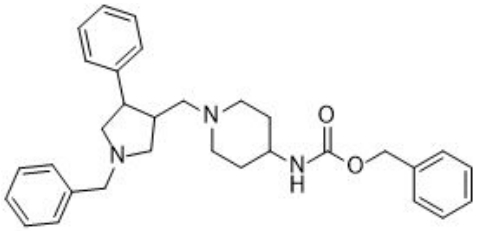
| | | | |
|---|---|---|---|
| **6** |  |  | **$IC_{50}$ = 1 nM**[38] |
| **7** |  |  | **$IC_{50}$ = 0.35 nM**[29] |
| **8** |  |  | **$IC_{50}$ = 1.1 nM**[35] |
| **9** |  |  | **$IC_{50}$ = 0.2 nM**[61] |
| **10** |  |  | **$IC_{50}$ = 0.6 nM**[46] |

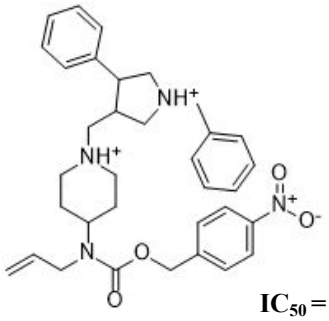
| | | | |
|---|---|---|---|
| **11** |  |  | **$IC_{50}$ = 1.1 nM**[52] |
| **12** |  |  | **$IC_{50}$ = 1.2 nM**[21] |
| **13** |  |  | **$IC_{50}$ = 0.1 nM**[41] |
| **14** |  |  | **$IC_{50}$ = 1.7 nM**[50] |
| **15** |  |  | **$IC_{50}$ = 1.7 nM**[28] |

| | | | |
|---|---|---|---|
| **16** |  |  | **$IC_{50}$ = 0.4 nM**[55] |
| **17** |  |  | **$IC_{50}$ = 0.1 nM**[48] |
| **18** |  |  | **$IC_{50}$ = 2 nM**[56] |
| **19** |  |  | **$IC_{50}$ = 0.75 nM**[62] |
| **20** |  |  | **$IC_{50}$ = 4 nM**[57] |

| | | | |
|---|---|---|---|
| **21** |  |  | **$IC_{50}$ = 4 nM**[58] |
| **22** |  |  | **$IC_{50}$ = 6 nM**[40] |
| **23** |  |  | **$IC_{50}$ = 6 nM**[39] |
| **24** |  |  | **$IC_{50}$ = 6 nM**[63] |
| **25** |  |  | **$IC_{50}$ = 8.8 nM**[27] |

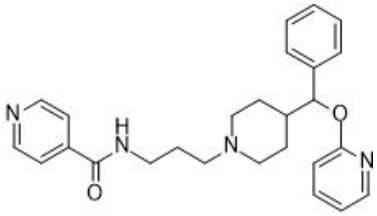
**26**

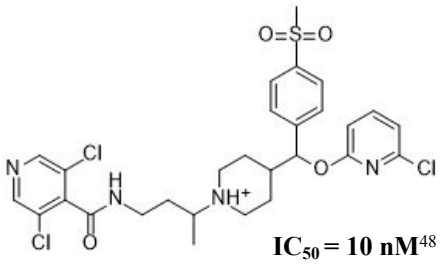




$IC_{50}$ = 0.1 nM[61]

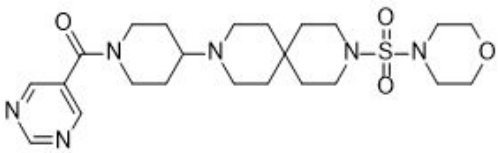
**27**

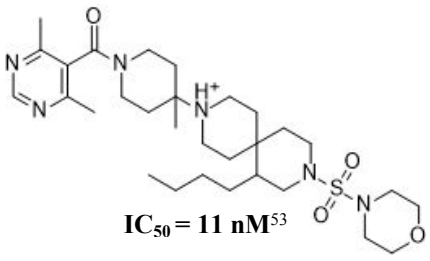




$IC_{50}$ = 10 nM[48]

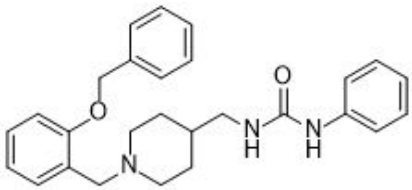
**28**

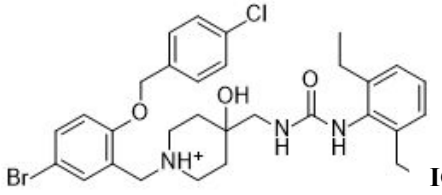




$IC_{50}$ = 11 nM[53]

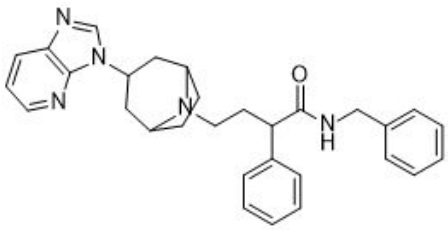
**29**

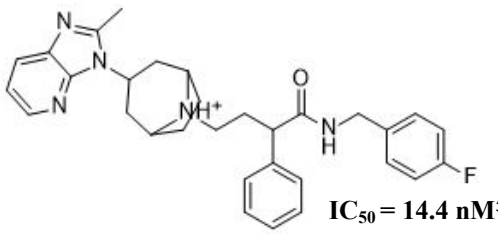




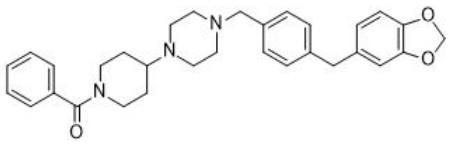
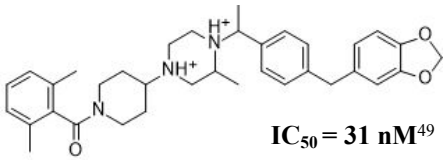
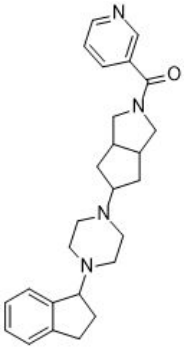
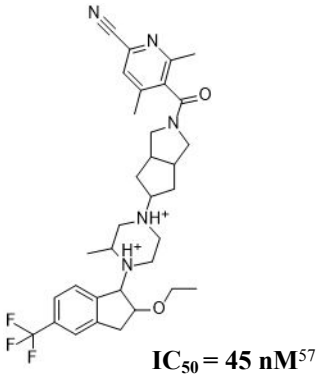
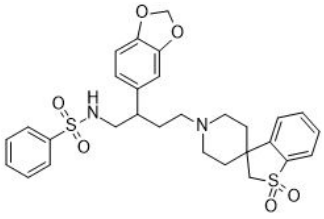
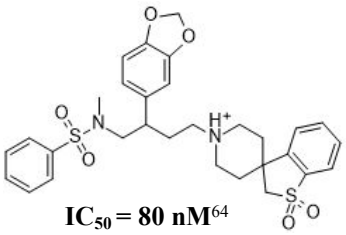
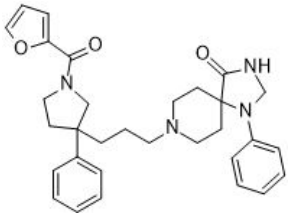
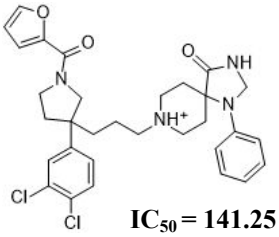
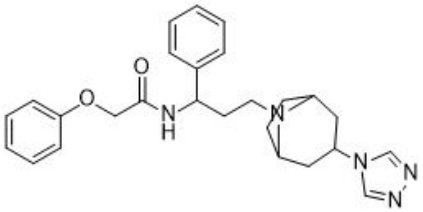
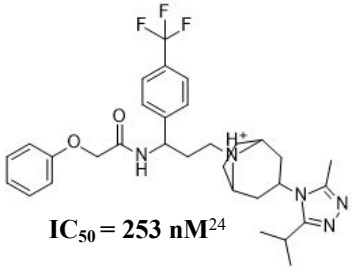
$IC_{50}$ = 11 nM[34]

**30**





$IC_{50}$ = 14.4 nM[54]

| | | |
|---|---|---|
| **31** |  |  **$IC_{50}$ = 31 nM**[49] |
| **32** |  |  **$IC_{50}$ = 45 nM**[57] |
| **33** |  |  **$IC_{50}$ = 80 nM**[64] |
| **34** |  |  **$IC_{50}$ = 141.25 nM**[30] |
| **35** |  |  **$IC_{50}$ = 253 nM**[24] |

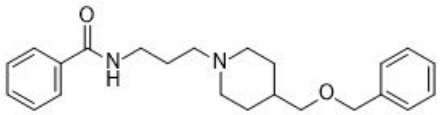
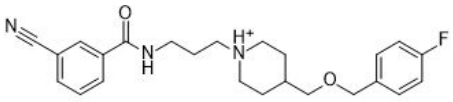
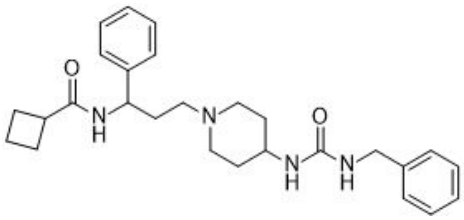
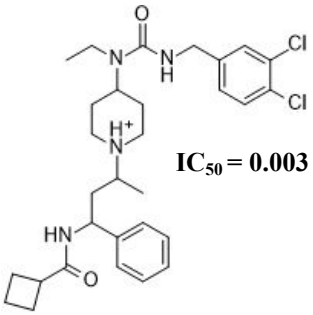
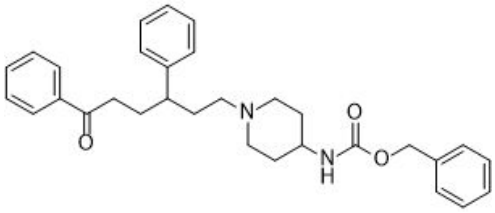
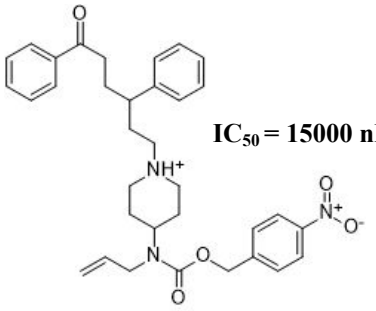
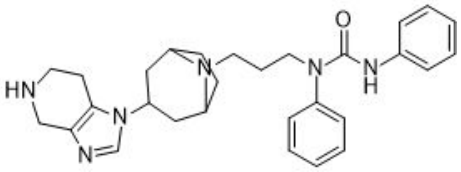
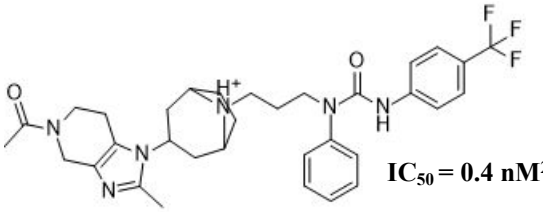
| | | |
|---|---|---|
| **36** |  |  **$IC_{50}$ = 1000 nM**[59] |
| **37** |  |  **$IC_{50}$ = 0.003 nM**[22] |
| **38** |  |  **$IC_{50}$ = 15000 nM**[43] |
| **39** |  |  **$IC_{50}$ = 0.4 nM**[23] |

**Table S2.** Pose-retrieval docking data of MVC

| Scoring | No. of clusters | Best-Ranked Pose | | Largest cluster best-ranked pose | |
|---|---|---|---|---|---|
| | | Score | RMSD* (Å) | Score | RMSD* (Å) |
| **ASP** | 34 | 78.63 | 0.68 | 77.81 | 0.82 |
| **GoldScore** | **22** | **101.49** | **0.55** | **101.49** | **0.55** |
| **ChemPLP** | 38 | 147.77 | 0.46 | 147.77 | 0.46 |
| **ChemScore** | 29 | 69.10 | 0.58 | 69.01 | 0.58 |

*RMSD = root mean square deviation between docked pose and the co-crystallized coordinates of MVC