# Supplementary Information

## Molecular determinants of ligand binding modes in the histamine $H_4$ receptor: Linking ligand-based 3D-QSAR models to *in silico* guided receptor mutagenesis studies

*Enade P. Istyastono[‡*], Saskia Nijmeijer[‡*], Herman D. Lim[¶], Andrea C. van de Stolpe[‡], Luc Roumen[‡], Albert J. Kooistra[‡], Henry F. Vischer[‡], Iwan J.P. de Esch[‡], Rob Leurs[‡], Chris de Graaf[‡**]*

[‡]Leiden/Amsterdam Center for Drug Research (LACDR), Division of Medicinal Chemistry, Department of Pharmacochemistry, Faculty of Exact Sciences, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

[¶]Griffin Discoveries BV, Department of Medicinal Chemistry, Room P-246, De Boelelaan 1083, 1081 HV, Amsterdam, The Netherlands

Contents

## Purity data for compounds 2, 5, 6a-l as determined by LCMS.

**Supplementary Table 1:** Purity and retention times of the synthesised compounds determined by analytical HPLC-MS.[a]

| Compound | Retention time (min) | Purity[b] |
|---|---|---|
| **2** | 2.40 | 100.0 |
| **5** | 2.79 | 96.5 |
| **6a** | 2.48 | 98.6 |
| **6b** | 2.64 | 99.0 |
| **6c** | 2.45 | 95.4 |
| **6d** | 3.00 | 98.6 |
| **6e** | 12.54 [c] | 96.8 |
| **6f** | 2.73 | 97.0 |
| **6g** | 2.95 | 100.0 |
| **6h** | 2.57 | 99.7 |
| **6i** | 2.39 | 98.3 |
| **6j** | 2.85 | 97.9 |
| **6k** | 2.57 | 98.4 |
| **6l** | 2.69 | 98.9 |

[a] The conditions can be found in the experimental section of the main article; [b] The purities were calculated as the percentage peak area of the analyzed compound by UV detection.

[c] different method:

LCMS analysis of compound 6[e] was performed with a Shimadzu LC-8A preparative liquid chromatograph pump system with a Shimadzu SPD-10AV UV–vis detector with the MS detection performed with a Shimadzu LCMS-2010 liquid chromatograph–mass spectrometer. using an XBrigde column (C18, 5 µm, 4,6 x 100 mm). An aqueous buffer (pH 8) of 0.04% $NH_4HCO_3$ (solvent A) and a mixture of 90% MeCN and 10% of a 0.4% $NH_4HCO_3$ buffer (pH 8, solvent B) were used. The runs started with 5% B with a linear gradient to 90% B in 10 minutes, then continuing for 10 minutes with 90% B and finally a linear gradient to 5% B in 10 minutes. Total run time 30 minutes.

## HRMS data for compounds 1, 8, 9a-l.

**Supplementary Table 2:** HRMS data for compounds **1**, **8**, **9a-l**.[a]

| Compound | MF | MW Calc.[M+1] | MW Found [M+1] |
|---|---|---|---|
| **2** | $C_{14}H_{17}ClN_4S$ | 309.0940 | 309.0928 |
| **5** | $C_{20}H_{27}ClN_4S$ | 391.1723 | 391.1698 |
| **6a** | $C_{19}H_{26}N_4S$ | 343.1956 | 343.1937 |
| **6b** | $C_{20}H_{28}N_4S$ | 357.2113 | 357.2093 |
| **6c** | $C_{16}H_{28}N_4S$ | 309.2113 | 309.2090 |
| **6d** | $C_{20}H_{26}Cl_2N_4S$ | 425.1335 | 425.1321 |
| **6e** | $C_{18}H_{25}ClN_4S$ | 365.1566 | 365.1547 |
| **6f** | $C_{21}H_{23}ClN_4S$ | 399.1410 | 399.1385 |
| **6g** | $C_{21}H_{22}Cl_2N_4S$ | 433.1020 | 433.0981 |
| **6h** | $C_{18}H_{30}N_4S$ | 335.2269 | 335.2245 |
| **6i** | $C_{16}H_{28}N_4S$ | 309.2113 | 309.2087 |
| **6j** | $C_{18}H_{24}Cl_2N_4S$ | 399.1177 | 399.1152 |
| **6k** | $C_{21}H_{27}N_4S$ | 382.2065 | 382.2048 |
| **6l** | $C_{20}H_{27}N_5O_2S$ | 402.1963 | 402.1944 |

[a] The conditions can be found in the experimental section of the main article.

## Value of the most influential descriptors for equations 1-5.

**Supplementary Table 3:** Value of the most influential descriptors for equations 1 and 2

| Cmpd | diameter | BCUT_SMR_0 | vdw_area | E_strain |
|---|---|---|---|---|
| **2** | 14.000 | -2.256 | 288.752 | 32.970 |
| **5** | 14.000 | -2.513 | 362.434 | 12.681 |
| **6a** | 13.000 | -2.262 | 328.333 | 45.480 |
| **6b** | 13.000 | -2.513 | 344.141 | 37.308 |
| **6c** | 12.000 | -2.347 | 308.107 | 23.685 |
| **6d** | 14.000 | -2.528 | 380.728 | 28.092 |
| **6e** | 14.000 | -2.346 | 358.017 | 28.542 |
| **6f** | 14.000 | -2.256 | 378.243 | 34.826 |
| **6g** | 14.000 | -2.256 | 396.536 | 34.201 |
| **6h** | 12.000 | -2.529 | 312.524 | 43.572 |
| **6i** | 11.000 | -2.390 | 315.008 | 43.662 |
| **6j** | 14.000 | -2.346 | 376.311 | 27.665 |
| **6k** | 15.000 | -2.528 | 368.583 | 28.796 |
| **6l** | 15.000 | -2.528 | 372.991 | 26.652 |

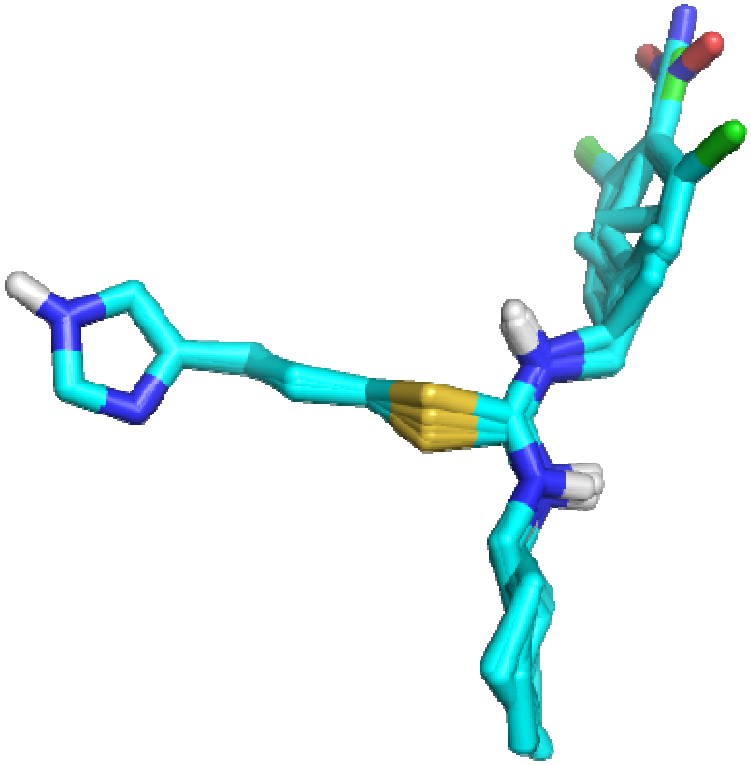
**Supplementary Table 4:** Value of the most influential descriptors for equations 3-5

| Cmpd | DRY.8331 | O.6874 | DRY.8195 | DRY.3376 | DRY.9361 |
|---|---|---|---|---|---|
| **2** | 0.180 | 3.618 | -2.738 | -11.144 | -7.416 |
| **5** | 0.180 | 3.154 | 0.180 | -10.979 | 0.180 |
| **6a** | -4.234 | 2.744 | -1.630 | -6.631 | 0.180 |
| **6b** | 0.180 | 1.909 | -4.011 | -6.664 | -7.573 |
| **6c** | -4.168 | 3.628 | -1.589 | -2.093 | 0.180 |
| **6d** | 0.180 | 3.780 | -4.085 | -11.243 | -7.598 |
| **6e** | -0.316 | 1.304 | -4.251 | -11.037 | -2.614 |
| **6f** | -2.440 | 0.872 | -6.127 | -11.136 | -5.193 |
| **6g** | -2.416 | 2.066 | -6.259 | -11.301 | -5.185 |
| **6h** | 0.180 | -0.371 | -2.449 | -1.928 | -7.433 |
| **6i** | -4.151 | 1.268 | -1.482 | -1.217 | 0.180 |
| **6j** | -0.308 | 3.583 | -4.275 | -11.177 | -2.639 |
| **6k** | 0.180 | 2.145 | -4.019 | -16.566 | -7.573 |
| **6l** | 0.180 | 3.545 | -4.085 | -13.549 | -7.606 |

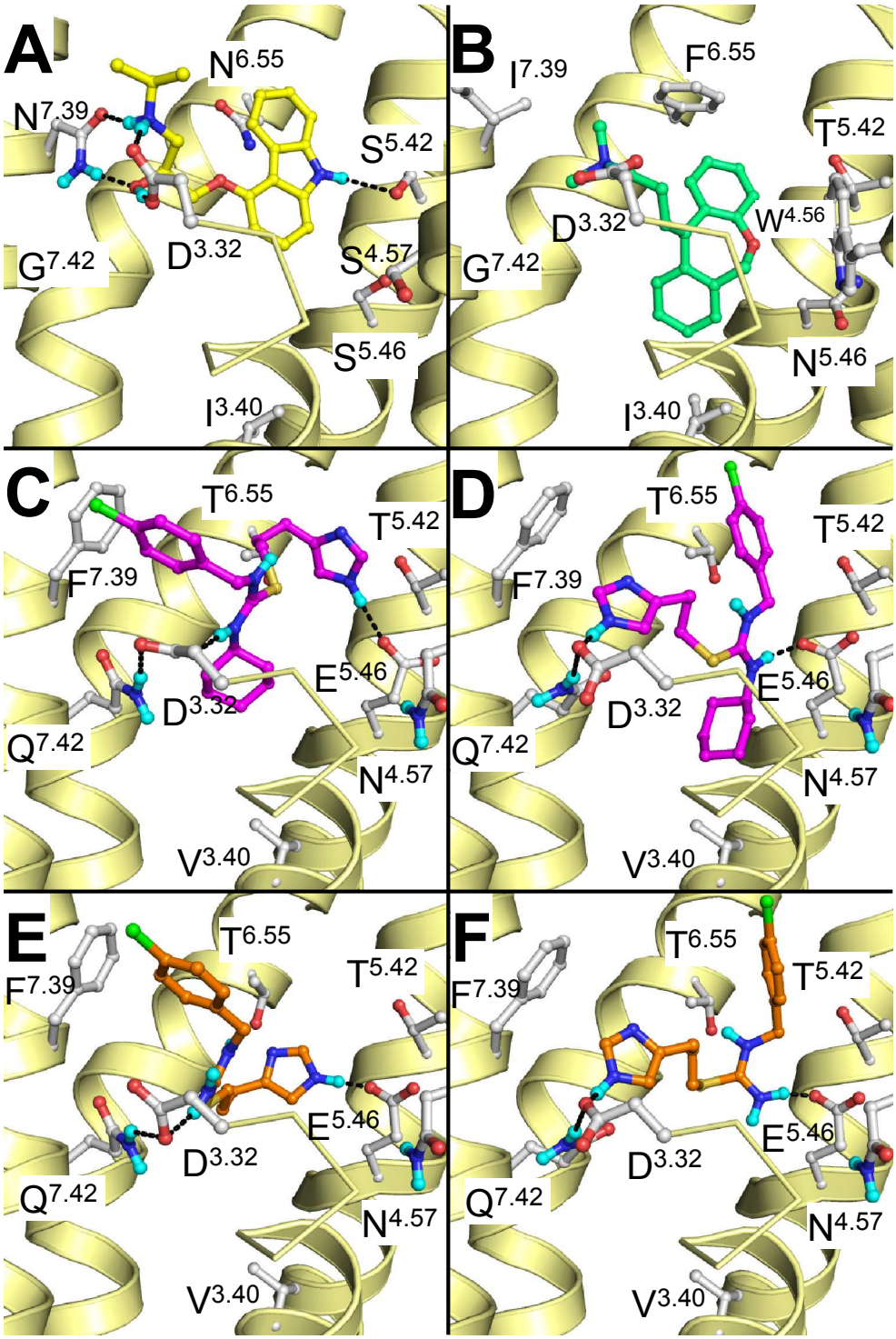
**Supplementary Table 4.** The frequency of the ligand forming hydrogen bond to $D^{3.32}$ in 1 ns MD simulations

| Compound | Binding Mode | |
|---|---|---|
| | **I** | **II** |
| Clobenpropit **(2)** | 98.5%[a] | 97.0%[b] |
| VUF5228 **(6)** | 0%[c] | 98.8%[d] |

Molecular dynamics simulations starting from binding modes depicted in Fig. 5A[a], 5C[b], Fig. 5B,[c] and Fig. 5D[d]



**Supplementary Figure 1**: Alignment of compounds **2**, **5**, and **6a-l** used in 3D-QSAR




**Supplementary Figure 2**: Comparison of: **A**) carazalol (yellow carbon atoms) bound adrenergic beta-2 receptor (ADRB2) crystal structure[1]; **B**) doxepin (green) bound histamine H1 receptor ($H_1R$) crystal structure[2]; **C-D**) VUF5228 (cpd **5**, magenta) bound $H_4R$ homology models; **E-F**) clobenpropit (cpd **2**, orange) bound $H_4R$ homology models.

## References Supplementary Information

1. Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007,** 318, 1258-65.

2. Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S. Structure of the human histamine H1 receptor complex with doxepin. *Nature* **2011,** 475, 65-70.