**Supplementary Information-PC**

**Combining Structure-based and 3D QSAR Pharmacophore Models to Discover Diverse Ligands against EGFR in Oral Cancer**

**Running Heading:**

**Design of EGFR inhibitors for oral cancer therapy**

**Table of Contents**

**Table S1**. The ligands employed for 3D-QSAR modeling with their pIC50 values.

**Table S2**. The primers used for the gene expression studies.

**Table S3**. The energy scores of the e-pharmacophore features.

**Table S4**. Ligand based pharmacophore hypotheses with scores.

**Table S5**. The GI50 of lead compounds derived from virtual screening and % pEGFR in FaDu.

**Table S6**. The predicted ADME properties of hit compounds.

**Figure S1**. The crystal structures used for the screening protocol.

**Figure S2**.The GI50 curves from FaDu cells and Cal27 cells treated for 72 h.

**Figure S3**. Gene expression analyses of **H2** treated FaDu cells and Cal27 cells

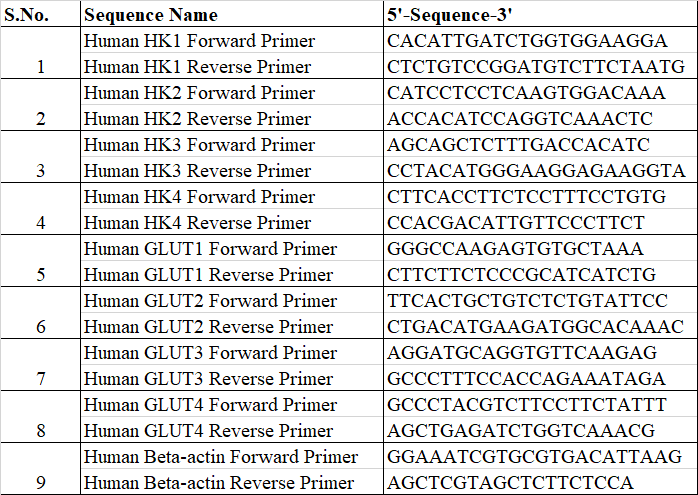
**Figure S4**. 2NBDG uptake assays in FaDu and Cal27 cells

**Figure S5**. Lactate dehydrogenase (LDH) activity and extracellular lactate in FaDu and Cal27 cells.

**Table S1**. The ligands employed for 3D-QSAR modeling with their pIC50 values.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S. No.** | **2D Structure of compounds** | **pIC50 experimental** | **pIC50 predicted** | **Difference** | **Dataset** |
| 1 |  | 7.26 | 7.25 | 0.01 | Training |
| 2 |  | 7.05 | 7.05 | 0 | Training |
| 3 |  | 8.15 | 8.13 | 0.02 | Training |
| 4 |  | 9.52 | 9.47 | 0.05 | Training |
| 5 |  | 7.64 | 7.63 | 0.01 | Training |
| 6 |  | 8.4 | 8.37 | 0.03 | Training |
| 7 |  | 6.64 | 6.65 | 0.01 | Training |
| 8 |  | 7.89 | 7.87 | 0.02 | Training |
| 9 |  | 6.52 | 6.53 | 0.01 | Training |
| 10 |  | 8.3 | 8.27 | 0.03 | Training |
| 11 |  | 8.17 | 8.15 | 0.02 | Training |
| 12 |  | 7.99 | 7.97 | 0.02 | Test |
| 13 |  | 7.82 | 7.80 | 0.02 | Training |
| 14 |  | 8.83 | 8.79 | 0.04 | Training |
| 15 |  | 8.52 | 8.49 | 0.03 | Test |
| 16 |  | 8.7 | 8.67 | 0.03 | Training |
| 17 |  | 8.22 | 8.20 | 0.02 | Training |
| 18 |  | 8.08 | 8.06 | 0.02 | Training |
| 19 |  | 7.12 | 7.12 | 0 | Training |
| 20 |  | 7.54 | 7.53 | 0.01 | Training |
| 21 |  | 8.13 | 8.11 | 0.02 | Test |
| 22 |  | 8.3 | 8.27 | 0.03 | Training |
| 23 |  | 9.52 | 9.47 | 0.05 | Training |
| 24 |  | 8.02 | 8 | 0.02 | Training |
| 25 |  | 7.7 | 7.69 | 0.01 | Training |
| 26 |  | 9.3 | 9.25 | 0.05 | Training |
| 27 |  | 9.12 | 9.08 | 0.04 | Training |
| 28 |  | 8.11 | 8.09 | 0.02 | Training |
| 29 |  | 9.3 | 9.25 | 0.05 | Training |
| 30 |  | 8.7 | 8.67 | 0.03 | Training |
| 31 |  | 6.46 | 6.47 | 0.01 | Test |
| 32 |  | 9.15 | 9.11 | 0.04 | Test |
| 33 |  | 8.49 | 8.46 | 0.03 | Test |
| 34 |  | 7.89 | 7.87 | 0.02 | Training |
| 35 |  | 6.72 | 6.73 | 0.01 | Training |
| 36 |  | 8.3 | 8.27 | 0.03 | Training |
| 37 |  | 7.72 | 7.71 | 0.01 | Test |
| 38 |  | 11.1 | 11.02 | 0.08 | Test |
| 39 |  | 7.14 | 7.14 | 0.00 | Training |
| 40 |  | 7.49 | 7.48 | 0.01 | Training |
| 41 |  | 7.05 | 7.05 | 0.00 | Test |
| 42 |  | 7.26 | 7.25 | 0.01 | Test |
| 43 |  | 7 | 7 | 0.00 | Test |
| 44 |  | 8.82 | 8.78 | 0.04 | Test |
| 45 |  | 9.92 | 9.86 | 0.06 | Training |
| 46 |  | 8.1 | 8.08 | 0.02 | Training |
| 47 |  | 8.62 | 8.59 | 0.03 | Training |
| 48 |  | 8 | 7.98 | 0.02 | Training |
| 49 |  | 7.7 | 7.69 | 0.01 | Training |
| 50 |  | 7.57 | 7.56 | 0.01 | Training |
| 51 |  | 9 | 8.96 | 0.04 | Training |
| 52 |  | 7.22 | 7.22 | 0.00 | Training |
| 53 |  | 5.7 | 5.73 | 0.03 | Test |

**Table S2**. The primers used for the gene expression studies.



**Table S3**. The energy scores of the e-pharmacophore features.

|  |  |  |
| --- | --- | --- |
| **PDB** | **Subset Pharmacophore** | **Energy Score (kcal/ mol)** |
| 4I22 | A1 | -2.20 |
|  | R15 | -1.27 |
|  | H8 | -0.12 |
|  | H7 | -0.05 |
|  | R13 | -1.02 |
|  | R14 | -0.79 |
| 3W2S | A1 | -2.13 |
|  | D7 | -0.8 |
|  | A3 | -0.7 |
|  | D10 | -0.63 |
|  | H12 | -0.38 |
|  | A5 | -0.18 |
|  | R16 | -1.2 |
|  | R15 | -1.18 |
|  | R17 | -0.96 |
|  | R14 | -0.88 |
| 3W33 | A1 | -2.13 |
|  | D7 | -1.98 |
|  | R14 | -1.37 |
|  | R12 | -1.15 |
|  | D6 | -0.31 |
|  | A3 | -0.2 |
|  | R13 | -1.11 |
|  | R15 | -0.84 |
| 5X2C | A1 | -2.17 |
|  | R12 | -1.46 |
|  | D4 | -0.6 |
|  | R10 | -0.82 |
|  | R11 | -0.74 |

**Table S4**. Ligand based pharmacophore hypotheses with scores.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Hypothesis** | **Survival score** | **Survival inactive score** | **Vector score** | **Volume score** | **Site score** |
| AAHRR.322 | 2.431 | 0.835 | 0.680 | 0.357 | 0.39 |
| AAHRR.189 | 2.725 | 1.327 | 0.798 | 0.440 | 0.49 |
| ADHRR.386 | 2.801 | 1.493 | 0.812 | 0.440 | 0.55 |
| ADHRR.662 | 2.918 | 1.261 | 0.844 | 0.490 | 0.58 |
| AADHR.401 | 2.911 | 1.784 | 0.845 | 0.477 | 0.59 |

**Table S5.**The GI50 of lead compounds derived from virtual screening and % inhibition of pEGFR (Y1068) in FaDu (ELISA)a

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Compd** | **Docking scoreb** | **GI50 for FaDu (µM)** | **GI50 for A549 (µM)** | **GI50 for Cal27 (µM)** | **% inhibition pEGFR (ELISA)c** |
| **H1** | -7.926 | 87.24 ± 0.69 | 65.55 ± 1.19 | 86.25 ± 2.36 | 28.69 ± 11.71 |
| **H2** | -7.703 | 63.06 ± 1.09 | >100 | 20.47 ± 2.20 | 71.72 ± 10.50 |
| **H3** | -7.490 | 82.67 ± 0.89 | 86.01 ± 0.95 | 43.79 ± 1.66 | 45.34 ± 13.63 |
| **H4** | -7.212 | 52.96 ± 0.45 | 27.33 ± 1.05 | 22.00 ± 2.41 | 67.56 ± 11.25 |
| **H5** | -10.008 | 37.35 ± 0.39 | 83.08 ± 1.25 | 28.60 ± 3.15 | 42.16 ± 13.05 |
| **H6** | -7.567 | 91.02 ± 0.55 | 78.22 ± 1.29 | 2.7 ± 3.31 | 29.15 ± 13.03 |
| **H7** | -9.687 | >100 | 5.122 ± 0.96 | >100 | 33.92 ± 11.62 |
| **H8** | -7.289 | >100 | 7.799 ± 1.42 | >100 | 26.56 ± 12.24 |
| **H9** | -6.953 | 45.77 ± 0.33 | 51.07 ± 1.09 | 87.72 ± 3.68 | 56.18 ± 13.09 |
| **H10** | -9.366 | 65.92 ± 0.79 | 41.5 ± 0.93 | 28.29 ± 2.25 | 41.18 ± 14.18 |
| Erlotinibd | - | >100 | 34.83 ± 0.81 | >100 | 58.36 ± 11.28 |
| Doxorubicine | - | <3 | 6.74 ± 0.59 | 9.24 ± 2.44 | 87.09 ± 11.70 |

aAll the compounds tested for GI50 were initially studied at 100 µM and leads showing more 60% inhibition were further diluted to lower concentration 3 µM to calculate GI50 values with each experiment performed in triplicates and the data are shown in mean ± SEM. b The best docking scores obtained when docked to four crystal structures cThe pEGFR (Y1068) inhibition (ELISA) was tested at 30 µM in FaDu cells and data presented with respect to DMSO control. dErlotinib was used as a standard EGFR inhibitor and eDoxorubicin was used as anticancer drug standard.

**Table S6**. The predicted ADME properties of hit compounds.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Compound** | **MW** | **QP logP o/w** | **QP logS** | **QP log BB** | **HERG logIC50** | **CaCO2 permeability (nm/sec)** | **MDCK permeability (nm/sec)** | **QP log Kp** | **Lipinski rule of 5 violations** |
| **H2** | 314.36 | 2.851 | -4.3 | -0.645 | -5.922 | 1063 | 905 | -1.826 | 0 |
| **H4** | 285.34 | 3.293 | -3.706 | -0.144 | -4.918 | 3879 | 2141 | -1.249 | 0 |
| **H9** | 366.48 | 3.613 | -4.13 | -0.687 | -5.558 | 1270 | 655 | -1.188 | 0 |
| **Erlotinib** | 393.44 | 4.13 | -4.571 | -0.461 | -6.236 | 4637 | 2597 | -0.015 | 0 |

D:\oral cancer\target 1 EGFR paper\Images for future medicinal chemistry\Figure S1 in 300 (2).tiff

**Figure S1**: **A**. The crystal structures with the ligand binding pocket used for the screening protocol. **B**. The 2D structure of respective crystal ligands with their reported IC50 values.

D:\oral cancer\target 1 EGFR paper\Images for future medicinal chemistry\Figure S2 modified300.tiff

**Figure S2**. The GI50 curves from A. FaDu cells and B. Cal27 cells treated for 72 h in triplicates fitted in logistic dose response curves. Hill slopes of cytotoxicity assay have been shown post treatment with H2 (GI50 for FaDu = 63.06 ± 1.09 µM, GI50 for Cal27 =20.47 ± 2.20 µM), H4 (GI50 for FaDu = 52.96 ± 0.45 µM, GI50 for Cal27 = 22.00 ± 2.41 µM), H9 (GI50 for FaDu = 45.77 ± 0.33 µM, GI50 for Cal27 = 87.72 ± 3.68 µM) and erlotinib (GI50 for both cell lines >100 µM). % Cell viability compared to DMSO control in C. FaDu and D. Cal27 cells. 2-way ANOVA followed by Dunnett’s multiple comparisons test #p<0.0001, @p<0.001.

D:\oral cancer\target 1 EGFR paper\Images for future medicinal chemistry\Figure S3 in 300 modified again.tiff

**Figure S3**: **A**. Gene expression analyses of **H2** treated FaDu cells. **B**. Gene expression analyses of **H2** treated Cal27 cells. The representative genes of glycolysis have been shown here. The treatment period was 72 h. Data are mean ± SEM of 3 individual experiments. Student’s t-test \*\**p*<0.01, \*\*\**p*<0.001 and \*\*\*\**p*<0.0001.

D:\oral cancer\target 1 EGFR paper\Images for future medicinal chemistry\Figure S4 in 300 modified again.tiff

**Figure S4**: **A**. FaDu cells were treated with **H2** and then incubated with 100 µM of 2-NBDG for 1 hour. Cells were washed and fluorescence was analysed by microscopy **B**. Relative fluorescence intensity of 2-NBDG was plotted for treated FaDu cells. **C**. 2-NBDG fluorescence analysis by microscopy in Cal27 cells. **D**. Relative fluorescence intensity of 2-NBDG plotted for Cal27 cells. The treatment period was 72 h. Data are mean ± SEM of 3 individual experiments. Student’s t-test \*\**p*<0.01.

D:\oral cancer\target 1 EGFR paper\Images for future medicinal chemistry\Figure S5 in 300 modified again.tiff

**Figure S5**: **A**. Lactate dehydrogenase (LDH) activity in FaDu cells treated with **H2** tested at 10 and 30 µM concentrations. **B**. Area under curve (AUC) for extracellular lactate in culture medium of FaDu cells treated with **H2**. **C**. LDH activity in Cal27 cells treated with **H2**. **D**. AUC for extracellular lactate in culture medium of Cal27 cells treated with **H2**. The treatment period was 72 h. Data are mean ± SEM of 3 individual experiments. Student’s t-test \*\*\*p<0.001, \*\*\*\*p<0.0001.