Open Access
ChemistryOPEN

Supporting Information

Synthetic Activators of Cell Migration Designed by Constructive Machine Learning

Dominique Bruns, Daniel Merk, Karthiga Santhana Kumar, Martin Baumgartner,* and Gisbert Schneider*

Author Contributions

D.B. Data curation:Lead; Formal analysis:Lead; Investigation:Equal; Methodology:Equal; Writing - Original Draft:Equal
D.M. Conceptualization:Equal; Investigation:Equal; Methodology:Lead; Supervision:Equal; Visualization:Equal; Writing - Original Draft:Equal
K.K. Formal analysis:Supporting; Investigation:Equal; Validation:Supporting; Visualization:Supporting; Writing - Original Draft:Supporting
M.B. Methodology:Equal; Resources:Equal; Writing - Review & Editing:Equal
G.S. Conceptualization:Lead; Formal analysis:Equal; Funding acquisition:Lead; Methodology:Equal; Supervision:Lead; Writing - Review & Editing:Equal

Supporting Information

Table of contents

Supplementary figures and tables

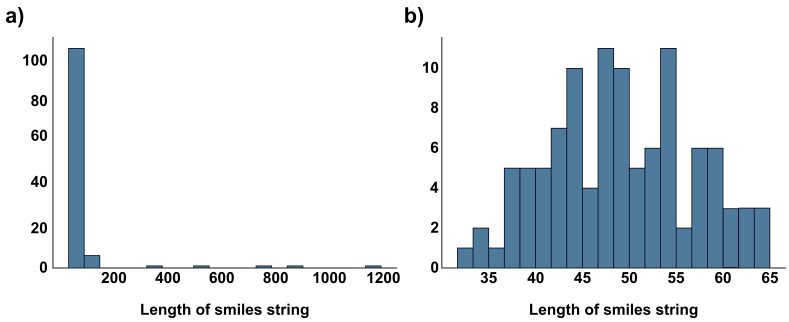

Figure S1: SMILES string length distribution of the full CXCR4 active dataset (a) and the set used for fine-tuning with a cutoff of 74 characters (b).

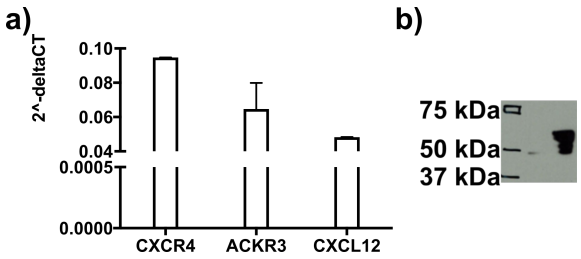

Figure S2: Expression of CXCR4 in DAOY cells. (a) CXCR4, ACKR3 and CXCL12 expression in DAOY cells determined on mRNA level by q-RT-PCR. (b) CXCR4 expression in DAOY cells determined by western blot (AB1846).

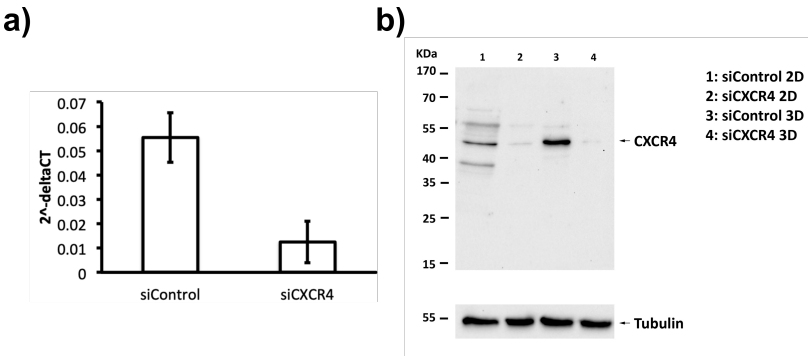

Figure S3: Expression of CXCR4 in DAOY cells after RNAi mediated knockdown. CXCR4 targeting siRNA markedly reduced CXCR4 expression in DAOY cells on mRNA (qRT-PCR, left) and protein level (western blot, right). Non-targeting siRNA had no effect on CXCR4 expression.

Table S1: Composition of the active set used to fine-tune the generative model. Activity is presented as the negative logarithmic IC_{50} value of the compounds against CXCR4 from literature.

SMILES	pIC_{50}	Literature
CC1(C)C[n+]2c(CSC(NC3CCCCC3)=[NH+]C3CCCCC3)csc2N1	9	22
c1sc2[n+](c1CSC(NC1CCCCC1)=[NH+]C1CCCCC1)CCN2	8	22
CC1(C)C[n+]2c(CSC(NC3CCCCCC3)=[NH+]C3CCCCCC3)csc2N1	8.7	22
c1sc2[n+](c1CSC(NC1CCCCCC1)=[NH+]C1CCCCCC1)CCN2	8.5	22
NCCCN(Cc1nc2ccccc2[nH]1)C1CCCc2cccnc21	7.2	23
NCCCCCCN(Cc1nc2ccccc2[nH]1)C1CCCc2cccnc21	7.1	23
NCCCCN(Cc1nc2ccccc2[nH]1)C1CCCc2cccnc21	8	23, 24, 27
CON=Cc1ccc(CN(Cc2nc3ccccc3[nH]2)C2CCCc3cccnc32)c(CN)c1	8.3	25
COc1ccc(CN(Cc2nc3ccccc3[nH]2)C2CCCc3cccnc32)c(CN)c1	8.3	25
NCc1ccc(CN(Cc2nc3ccccc3[nH]2)C2CCCc3cccnc32)c(CN)c1	8	25
NCc1ncccc1CN(Cc1nc2ccccc2[nH]1)C1CCCc2cccnc21	7.7	25
NCc1ccncc1CN(Cc1nc2ccccc2[nH]1)C1CCCc2cccnc21	6.8	25
Cc1ccnc(C(C)N(CCCCN)Cc2ncccc2C)c1	8.1	26
CC(c1ccccn1)N(CCCCN)Cc1ncccc1C(C)(C)C	7.7	26
Cc1cnc(CN(CCCCN)C(C)c2ccccn2)c(C)c1	7.7	26
Cc1cccnc1CN(CCCCN)C1CCCc2cccnc21	7.7	26
Cc1cnc(CN(CCCCN)C2CCCc3cccnc32)c(C)c1	7.7	26
Nc1cccnc1CN(CCCC[NH3+])C1CCCc2cccnc21	7.6	26
NC(=O)NCCCCN(CC1Cc2ccccc2CN1)C1CCCc2cccnc21	8.5	27
NCCCCN(CC1Cc2ccccc2CN1)C1CCCc2cccnc21	8.5	27
CC(=O)NCCCCN(CC1Cc2ccccc2CN1)C1CCCc2cccnc21	8.4	27
NCCCCN(CC1CN(S(=O)(=O)c2ccc(Cl)cc2)CCN1)C1CCCc2cccnc21	8.7	28
NCCCCN(CC1CN(S(=O)(=O)c2ccccc2)CCN1)C1CCCc2cccnc21	8.7	28
Oc1cccc(CN(CCN2CCCCC2)CC2CCCN(C3CCCC3)C2)c1	7	29
COc1ccc(CN(CCN2CCCCC2)CC2CCCN(C3CCCC3)C2)cc1O	6.6	30

Table S2: Activity profiling of computer-designed compounds **1** and **2** on molecular targets involved in cell migration. Data are presented as individual percent (%) activity values of the respective control. Compounds were tested in a single concentration of 50 μM with two technical replicates. n/d, not determined.

Target (assays type)	1	2
CCR1 (radioligand binding)	<0	<0
CCR2 (agonist effect)	<0	11, 31
CCR2 (antagonist effect)	-1, 4	1, 4
CCR4 (agonist effect)	<0	<0
CCR4 (antagonist effect)	-2, 1	9, 10
CCR6 (agonist effect)	<0	<0
CCR6 (antagonist effect)	17, 18	9, 16
CCR7 (agonist effect)	0	<0
CCR7 (antagonist effect)	3, 6	14, 15
CCR8 (agonist effect)	<0	<0
CCR8 (antagonist effect)	9, 12	2, 31
CCR9 (agonist effect)	0, 2	<0
CCR9 (antagonist effect)	-9, 2	8, 10
CCR10 (agonist effect)	0	0
CCR10 (antagonist effect)	14, 21	26, 36
CX3CR1 (agonist effect)	2	<0
CX3CR1 (antagonist effect)	<0	2, 12
CXCR1 (agonist effect)	<0	<0
CXCR1 (antagonist effect)	-1, 6	2, 9
CXCR2 (agonist effect)	0	0
CXCR2 (antagonist effect)	0, 8	4, 9
CXCR3 (agonist effect)	<0	<0
CXCR3 (antagonist effect)	<0	1, 12
CXCR5 (agonist effect)	<0	<0
CXCR5 (antagonist effect)	-6, 14	-4, 2
CXCR6 (agonist effect)	1	<0
CXCR6 (antagonist effect)	<0	<0
D1 (agonist effect)	<0	0
D1 (antagonist effect)	<0	<0
D2L (antagonist radioligand)	0, 1	0, 2
D2S (agonist effect)	17, 20	42, 52
D2S (antagonist effect)	15, 26	n/d
HDAC4 (inhibition)	<0	<0
HDAC6 (inhibition)	<0	<0
HDAC7 (inhibition)	0, 14	-1, 13
HDAC9 (inhibition)	-8, 16	-4
MMP-1 (inhibition)	3, 7	6, 13
sst1 (agonist effect)	<0	26, 29
sst1 (antagonist effect)	0, 2	-10, 16