

# Probing the Molecular Interactions Between CXC Chemokine Receptor 4 (CXCR4) and an Arginine-Based Tripeptidomimetic Antagonist (KRH-1636)

*Zack G. Zachariassen,<sup>†</sup> Stefanie Karlshøj,<sup>§</sup> Bengt Erik Haug,<sup>‡</sup> Mette M. Rosenkilde,<sup>§</sup> and Jon Våbenø<sup>\*,†</sup>*

<sup>†</sup>Department of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Breivika, NO-9037 Tromsø, Norway. <sup>§</sup>Laboratory for Molecular Pharmacology, Department of Neuroscience and Pharmacology, Faculty of Health and Medical Sciences, The Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen, Denmark. <sup>‡</sup>Department of Chemistry and Centre for Pharmacy, University of Bergen, Allégaten 41, NO-5007 Bergen, Norway.

**Abstract:** We here report an experimentally verified binding mode for the known tripeptidomimetic CXCR4 antagonist KRH-1636 (**1**). A limited SAR study was first conducted based on the three functionalities of **1**, followed by site-directed mutagenesis studies. The receptor mapping showed that both the potency and affinity of **1** were dependent on the transmembrane residues His<sup>113</sup>, Asp<sup>171</sup>, Asp<sup>262</sup>, and His<sup>281</sup>, and also suggested the involvement of Tyr<sup>45</sup> and Gln<sup>200</sup> (potency) and Tyr<sup>116</sup> and Glu<sup>288</sup> (affinity). Molecular docking of **1** to an X-ray structure of CXCR4 showed that the L-Arg guanidino group of **1** forms polar interactions with His<sup>113</sup> and Asp<sup>171</sup>, the (pyridin-2-ylmethyl)amino moiety is anchored by Asp<sup>262</sup> and His<sup>281</sup>, while the naphthalene ring is tightly packed in a hydrophobic subpocket formed by the aromatic side chains of Trp<sup>94</sup>, Tyr<sup>45</sup>, and Tyr<sup>116</sup>. The detailed picture of ligand-receptor interactions provided here will assist in structure-based design and further development of small-molecule peptidomimetic CXCR4 antagonists.

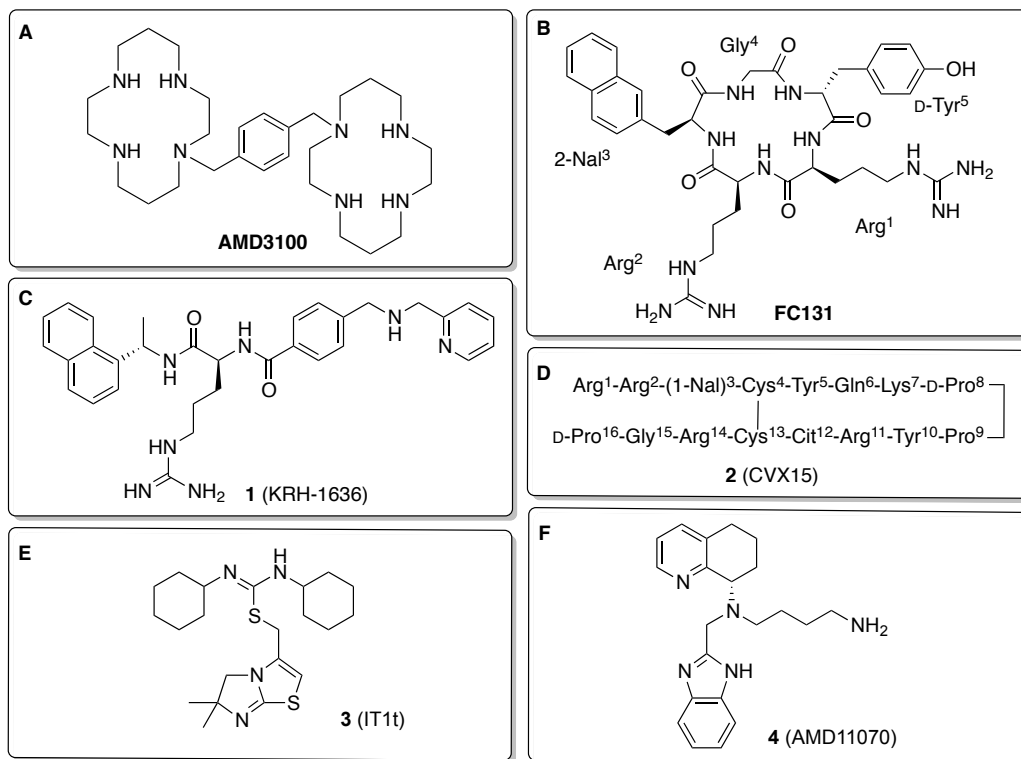
## INTRODUCTION

Chemokines are small (8-14 kDa) proteins from the cytokine superfamily that exert their biological effects through interaction with GPCRs.<sup>1</sup> The involvement and physiological role of the chemokine ligand-receptor system extends from control of chemotaxis (directional migration of leukocytes)<sup>2</sup> and hematopoiesis<sup>3</sup> to embryonic development, angiogenesis and cancer metastasis.<sup>4,5</sup> The human CXC chemokine receptor 4 (CXCR4) is activated by CXC chemokine ligand 12 (CXCL12), which is a 68-residue polypeptide.<sup>6,7</sup> The CXCL12: CXCR4 interaction is involved in development of the hematopoietic, nervous, and gastrointestinal systems<sup>8,9</sup> and plays an essential role in immune defense and inflammation.<sup>10,11</sup> In addition to these physiological roles, CXCR4 is the co-receptor for cellular entry of T-tropic HIV strains;<sup>12,13</sup> thus, inhibition of

CXCR4 has emerged as a promising strategy for anti-HIV therapy. CXCR4 has also been shown to play an important role in angiogenesis, metastasis and stem-cell mobilization.<sup>2,4</sup> Consequently, development of CXCR4 antagonists for therapeutic applications remains an important goal.

Significant advances in the development of novel CXCR4 antagonists have been made and several different classes of CXCR4 antagonists, both peptides and non-peptides, have been reported.<sup>14-16</sup> So far, the hematopoietic stem cell mobilizing agent plerixafor (AMD3100, Figure 1A)<sup>17</sup> is the only CXCR4 antagonist that has reached the market.<sup>18</sup> Due to its high positive charge at physiological pH, AMD3100 exhibits poor oral bioavailability, and is therefore administered by subcutaneous injection.

Extensive studies by Fujii and co-workers on polyphemusin II-derived peptide CXCR4 antagonists culminated in the discovery of the potent cyclopentapeptide antagonist FC131 (Figure 1B).<sup>19</sup> Successive SAR studies of FC131 have since shown that the Arg<sup>1</sup>-Arg<sup>2</sup>-2-Nal<sup>3</sup> tripeptide fragment of the cyclopentapeptide serves as an essential recognition motif for peptide-based CXCR4 antagonists,<sup>20-24</sup> providing impetus for development of tripeptidomimetic ligands that block CXCR4. Indeed, a series of tripeptide-like CXCR4 antagonists containing a central arginine residue has been reported by Kureha Chemical Industries<sup>25-27</sup> with the potent tripeptidomimetic **1** (KRH-1636, Figure 1C)<sup>28,29</sup> as a prototype compound. These small-molecule peptidomimetics represent a particularly interesting class of CXCR4 antagonists due to their potentially improved pharmacokinetic properties compared to their peptide counterparts.<sup>30, 31</sup> Based on structural comparison of low-energy conformations of **1** with a 3D-pharmacophore model for the cyclopentapeptide CXCR4 antagonists, **1** has been suggested to mimic the Arg<sup>1</sup>-Arg<sup>2</sup>-2-Nal<sup>3</sup> fragment of FC131.<sup>32</sup> However, it has not been shown experimentally that the two ligands bind to CXCR4 in the same way.



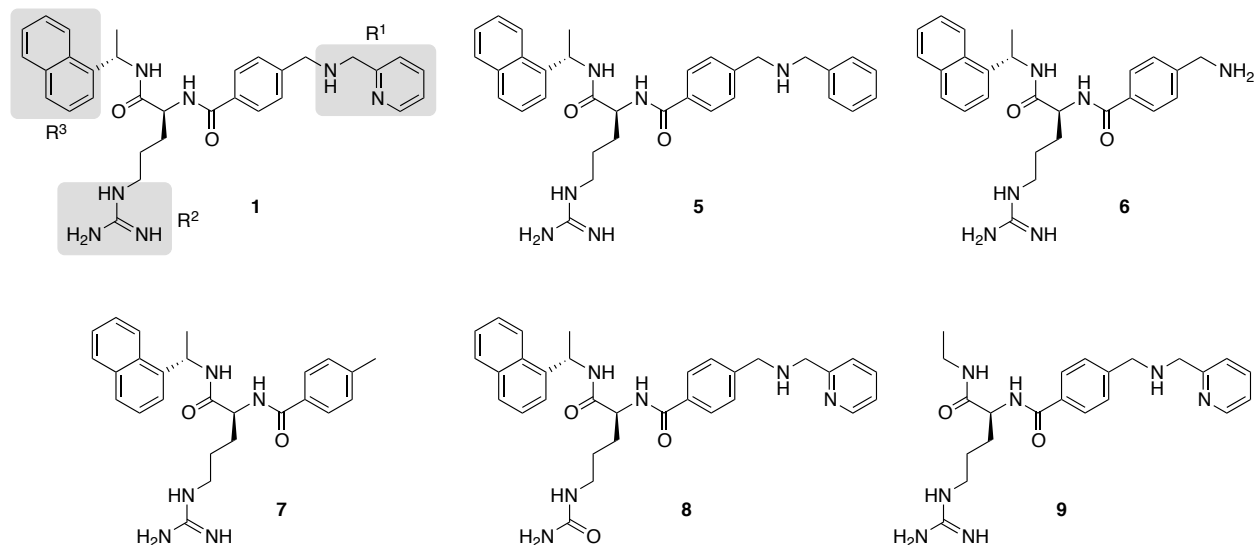
**Figure 1.** Structures of selected CXCR4 antagonists: (A) bicyclam AMD3100, (B) cyclopentapeptide FC131, (C) prototype arginine-based tripeptidomimetic **1**, (D) 16-mer peptide **2**, (E) isothioureia derivative **3**, and (F) tetrahydroquinoline derivative **4**.

The first X-ray structures of CXCR4 were reported in 2010, and are co-crystallized complexes containing the 16-mer peptide antagonist **2** (CVX15, Figure 1D) and the small-molecule antagonist **3** (IT1t, Figure 1E).<sup>33</sup> Very recently, an X-ray structure of a covalent complex between CXCR4 and the viral chemokine vMIP-II was also released.<sup>34</sup> The CXCR4:**2** complex provided the first detailed insight into binding of a peptide antagonist to CXCR4, and based on this structure we have recently reported an experimentally verified binding mode for the cyclopentapeptide FC131 (Figure 1B).<sup>35</sup> Prior to the publication of the X-ray structures of CXCR4, extensive receptor mapping studies were performed for the AMD-compound series,

which resulted in the identification of plausible binding modes for the bicyclam AMD3100,<sup>36-38</sup> its monocyclam analogs,<sup>39, 40</sup> and the non-cyclam **4** (AMD11070, Figure 1F).<sup>40</sup> However, while the molecular pharmacology of the prototype small-molecule peptide and non-peptide CXCR4 antagonists has been extensively characterized, experimental binding mode studies for the more drug-like tripeptidomimetic CXCR4 antagonists have not been reported. Thus, in order to facilitate rational design of novel peptidomimetic CXCR4 antagonists, we here report the binding mode of the known tripeptidomimetic CXCR4 antagonist **1** using a combination of SAR studies, receptor mapping by site-directed mutagenesis, and molecular docking.

## RESULTS AND DISCUSSION

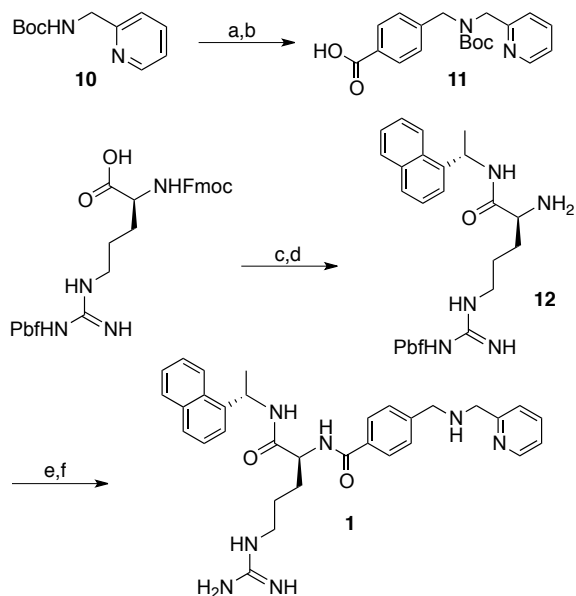
**SAR Study. Design.** The lead compound **1** was developed from a hit compound identified through screening of a chemical library,<sup>28, 29</sup> and although recent patent literature discloses compounds related to **1**,<sup>25-27</sup> no SAR data has been reported for this compound class. In order to confirm the importance of the functionalities in **1** that could be assumed to represent pharmacophoric groups, we first carried out a limited SAR study focusing on the two positively charged groups (R<sup>1</sup> and R<sup>2</sup>) and the aromatic group (R<sup>3</sup>) (Figure 2). In order to probe the R<sup>1</sup> position, the pyridine ring was replaced with a phenyl ring (**5**), the pyridin-2-ylmethyl substituent was removed to give the primary amine (**6**), and the charged (pyridin-2-ylmethyl)amino moiety was removed altogether (**7**). The importance of the charged guanidino group (R<sup>2</sup>) was investigated by replacing it with a urea group (**8**), while the aromatic position (R<sup>3</sup>) was probed by removal of the naphthalene unit (**9**).



**Figure 2.** Structures of compounds **1** and **5–9**; the modified groups ( $R^1$ ,  $R^2$ , and  $R^3$ ) in the lead compound **1** are highlighted.

*Chemistry.* Compounds **1** and **5–9** (Figure 2) were all prepared by adopting a procedure previously reported for the synthesis of **1** (Scheme 1).<sup>29</sup> For compound **1**, the right-hand side was prepared by N-alkylation of Boc-protected pyridin-2-ylmethanamine (**10**) with methyl 4-(bromomethyl)benzoate followed by hydrolysis of the methyl ester to give carboxylic acid **11**. The left-hand side was assembled by coupling of Fmoc-L-Arg(Pbf)-OH with (*S*)-1-(naphthalen-1-yl)ethanamine using PyBOP as the coupling reagent, followed by Fmoc-deprotection with diethylamine in DMF to afford amine **12**. Coupling of the two halves (**11** and **12**) was facilitated by EDCI/HOBt, and subsequent removal of the Boc and Pbf protecting groups by treatment with a TFA cocktail gave **1**. For the preparation of **6** and **7**, the left hand side (**12**) was coupled with *N*-Boc 4-(aminomethyl)benzoic acid and 4-methylbenzoic acid, respectively. Analog **5** was prepared by N-alkylation of **6** with benzyl bromide using  $K_2CO_3$  as base and DMF as solvent. The  $R^2$  analog **8** was prepared using Fmoc-protected citrulline as the starting material while the  $R^3$  analog **9** was prepared by use of ethanamine in the first coupling step (Scheme 1).

**Scheme 1. Synthesis of compound 1.**



*Reagents and conditions:* (a) i) NaH, DMF, 0 °C; ii) Methyl 4-(bromomethyl)benzoate, DMF, 0 °C to rt; (b) 1M NaOH/THF/MeOH (1:1:1); (c) (*S*)-1-(naphthalen-1-yl)ethanamine, PyBOP, DIPEA, CH<sub>2</sub>Cl<sub>2</sub> 0 °C to rt.; (d) Et<sub>2</sub>NH, DMF; (e) **11**, EDCI, HOBt, DMF; (f) TFA/TIS/H<sub>2</sub>O (95:2.5:2.5).

*Biology and SAR.* The antagonistic potency of compounds **1** and **5–9** (Table 1) was assessed in a functional assay measuring the inhibition of CXCL12-induced activation of human CXCR4, transiently expressed in COS-7 cells. The known lead compound **1** was originally reported by Ichiyama *et al.* to efficiently inhibit binding of [<sup>125</sup>I]CXCL12 (IC<sub>50</sub> = 0.013 μM),<sup>28</sup> and our own functional data validates **1** as a potent CXCR4 antagonist (EC<sub>50</sub> = 0.50 μM, Table 1) with respect to the endogenous ligand CXCL12. The phenyl-analog **5** (EC<sub>50</sub> = 8.5 μM) displayed 17-fold lower potency than **1**, demonstrating an important role of the pyridine nitrogen in **1**. The potency of the primary amine **6** (EC<sub>50</sub> = 12 μM) was 24-fold lower than for **1**, representing the overall contribution from the pyridin-2-ylmethyl substituent to the antagonistic potency, while the loss

of potency for compound **7** ( $EC_{50} > 100 \mu\text{M}$ ) confirms the positively charged amino-group in R<sup>1</sup> as an essential pharmacophoric element. Compound **8**, containing a neutral urea group that partly retains the H-bonding properties of the original guanidino group, failed to show antagonistic potency ( $EC_{50} > 100 \mu\text{M}$ ), demonstrating the importance of the positive charge in R<sup>2</sup> position. Similarly, the loss of potency for the R<sup>3</sup> analog **9** ( $EC_{50} > 100 \mu\text{M}$ ) confirms an important role for the naphthyl group.

**Table 1.** Antagonistic potency of **1** and **5–9** on human CXCR4.

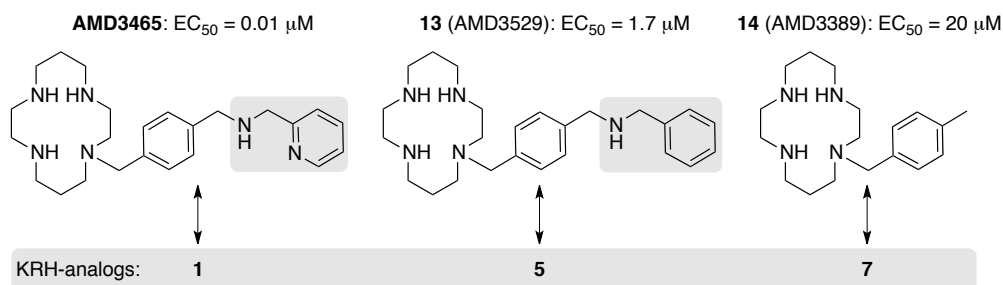
compd	log $EC_{50} \pm \text{SEM}^a$	$EC_{50}$ ( $\mu\text{M}$ )
<b>1</b> <sup>b</sup>	$-6.30 \pm 0.07$	0.50
<b>5</b>	$-5.12 \pm 0.08$	8.5
<b>6</b>	$-4.93 \pm 0.08$	12
<b>7</b>	$> -4$	$> 100$
<b>8</b>	$> -4$	$> 100$
<b>9</b>	$> -4$	$> 100$

<sup>a</sup> $EC_{50}$  values represent the mean of at least three independent experiments performed in duplicates. <sup>b</sup>Known compound.

Earlier attempts to improve the pharmacokinetic profile of the bicyclam CXCR4 antagonists, i.e. AMD3100 (Figure 1A), led to the discovery of the monocyclam AMD3465 (Figure 3),<sup>41</sup> which was shown to be 22-fold more potent than its precursor AMD3100<sup>39</sup> and contains the same (pyridin-2-ylmethyl)amino moiety as **1**. The importance of this group for the antagonistic activity of AMD3465 has been investigated previously using the analogs **13** (AMD3529) and **14** (AMD3389) (Figure 3),<sup>39</sup> which are structural counterparts of the KRH-analogs **5** and **7** in the



present study (Figure 2). In line with the present findings for **5** and **7** (Table 1), **13** and **14** were shown to display significantly lower potencies for CXCR4 than AMD3465 as determined by a functional assay inhibiting CXCL12-induced receptor activation.<sup>39</sup>



**Figure 3.** Structures and reported potencies of the monocyclam AMD3465 and the two analogs **13** and **14**,<sup>39</sup> also showing the structural analogy of these monocyclams with the R<sup>1</sup> substituent in compounds **1**, **5**, and **7** in the present study.

Regarding the central Arg residue in **1** (R<sup>2</sup>), it has been suggested that this position could correlate to Arg<sup>2</sup> in the cyclopentapeptide antagonist FC131 (Figure 1).<sup>32</sup> SAR studies of the cyclopentapeptides have shown that the Arg<sup>2</sup> position is very sensitive to modifications,<sup>20, 23</sup> and that replacement of the guanidino group with a urea group (Cit<sup>2</sup>) leads to complete loss of antagonistic potency,<sup>23</sup> which is in agreement with the present data for the KRH-analog **8**.

Collectively, the findings from this initial SAR study confirmed that the secondary amine, the guanidino group, and the aromatic naphthyl group are essential for the antagonistic potency of **1**, while the pyridine ring in R<sup>1</sup> position contributes to increased potency.

**Identification of Binding Site for 1: Receptor Mutagenesis.** Next, we carried out mutagenesis studies to map key binding interactions between **1** and CXCR4 using a library of 25 CXCR4-mutants (Table 2). Most of the mutations were located in the transmembrane helices (TMs), but some were also introduced in extracellular loop 2 (ECL2) that connects TM4 and

TM5. In most cases, the mutations included replacement with Ala; however, certain acidic Asp residues were alternatively substituted with the structurally similar and uncharged Asn residue. Selective substitutions with Trp were also done in order to introduce more bulk. We have recently utilized the same receptor library to map the binding of the cyclopentapeptide CXCR4 antagonist FC131, where we reported the surface expression (using ELISA) and functional response to CXCL12 of WT-CXCR4 and all mutant receptors<sup>35</sup> (the same data is provided in Table 2). While there was some variation in surface expression and ability to be activated by CXCL12, the receptors in the mutant library were considered suitable for mapping of the binding site, as previously discussed.<sup>35</sup>

*Inhibition of CXCL12-induced Activation of CXCR4.* We first assessed the antagonistic potency of **1** on the entire CXCR4 mutant library (Table 2). The H113A, D262N, and H281A mutations were found to have the largest impact (>25-fold reduction) on the antagonistic potency of **1** relative to WT-CXCR4 (the position of residues based on the generic numbering system proposed by Baldwin<sup>42</sup> and modified by Schwartz<sup>43</sup> followed by the Ballesteros/Weinstein numbering system<sup>44</sup> is given in Tables 2 and 3). The potency of **1** was also reduced (by 5-25 fold) by mutations of Tyr<sup>45</sup> (Y45A), Asp<sup>171</sup> (D171N), Gln<sup>200</sup> (Q200A and Q200W), Trp<sup>252</sup> (W252A), Ile<sup>259</sup> (I259A), and Ile<sup>284</sup> (I284A). In contrast, the W94A, D97A, T117A, and F189A mutants led to *increased* potency (>2-fold).

As the mutagenesis data for **1** alone did not provide information about which parts of the ligand that interact with the different parts of the receptor, the R<sup>1</sup> analogs **5** and **6** (Figure 2), which showed decent activity on WT-CXCR4 (Table 1), were also included in the mutagenesis study to further probe the molecular interactions of the R<sup>1</sup>-side chain with CXCR4. The collective functional data analysis for **1**, **5**, and **6** (Table 2) showed that the mutations in TM1

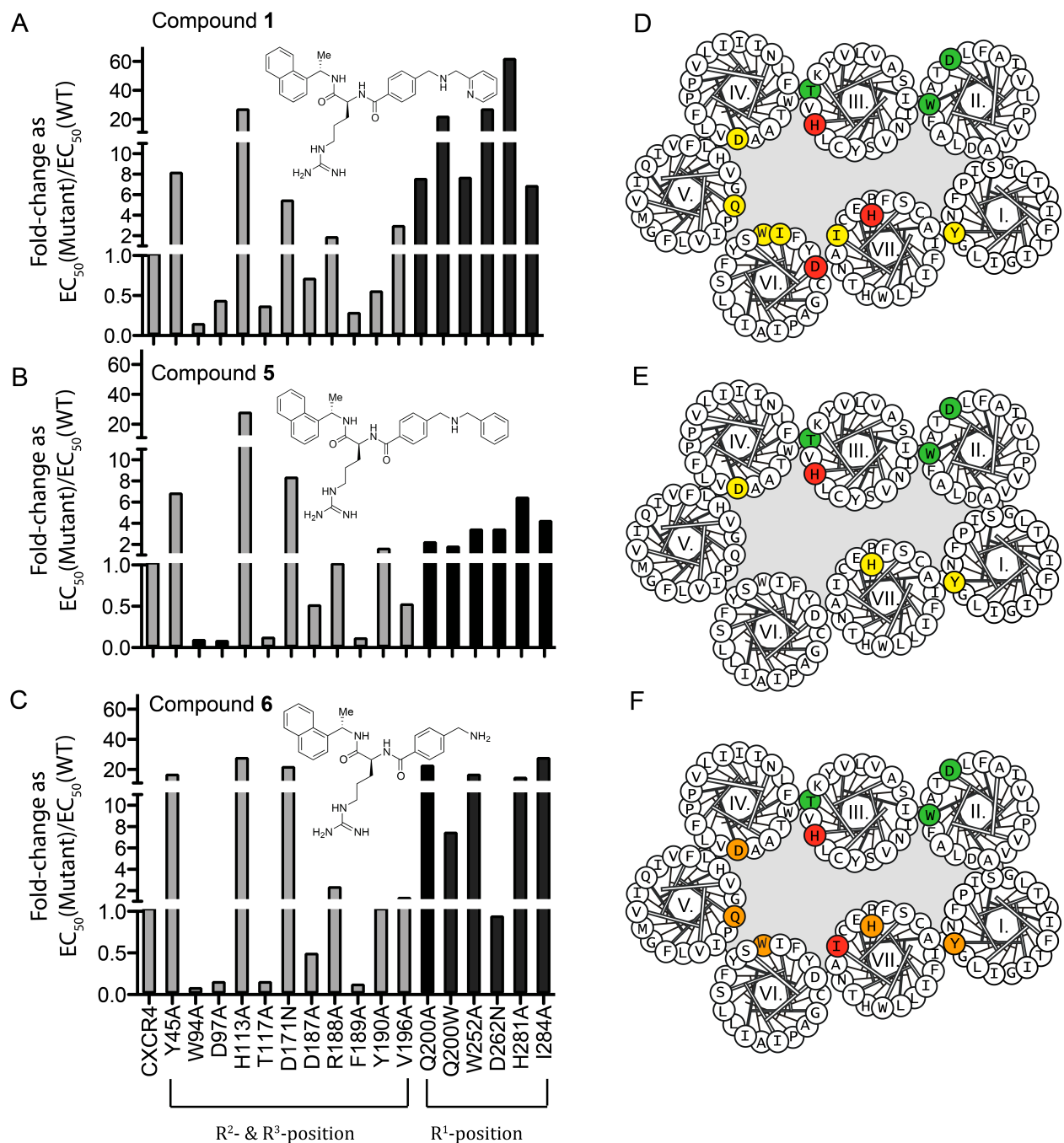
(Y45A), TM2 (W94A, D97A), TM3 (H113A, T117A), TM4 (D171N), and ECL2 affected the three ligands in a similar manner. However, several mutations in TMs 5-7 affected the three ligands in a distinctive manner. Specifically, the H281A mutation in TM7 was the mutation with largest effect on the potency of **1** (61-fold reduction), but affected analog **5** to a much lesser extent (6-fold). Thus, the overall trend identified in the phosphatidylinositol (PI) turnover experiments for **1**, **5**, and **6** (Table 2) suggests that the structurally modified R<sup>1</sup> group contacts TMs 5-7; this trend is further illustrated in Figure 4.

The absolute majority of the mutations had no or little effect on the potency of CXCL12 (Table 2), demonstrating that most of the mutants were functional. However, the W94A, D97A, and D187A mutations reduced the agonistic potency of CXCL12 by 8-14 fold (Table 2). Furthermore, the Y116A and E288A mutants were not activated by CXCL12 in spite of expression levels of 69 and 77%, respectively (Table 2); consequently, no data on the antagonistic action of **1**, **5**, and **6** on these mutants are available. In line with the findings in the present study, the D97A, D187A, and E288A mutations have previously been found to affect both the binding and signaling of CXCL12.<sup>37, 45-49</sup> Moreover, Trp<sup>94</sup> and Tyr<sup>116</sup>, together with the three aforementioned residues, have been implicated as part of “site two” in the “two-site model” for binding of CXCL12 to CXCR4.<sup>33, 45, 50, 51</sup> Additionally, the direct interaction of Tyr<sup>116</sup> with agonists has been suggested in activation of GPCRs.<sup>34, 52</sup> Thus, for these residues there are certain limitations in using CXCL12 as a probe.

**Table 2.** Potency of the lead compound **1** and the R<sup>1</sup>-analogs **5** and **6** on WT-CXCR4 and 25 mutant receptors, measured as inhibition of CXCL12-induced activation in COS-7 cells transiently co-transfected with CXCR4 receptor mutants and chimeric G protein G<sub>qi4myr</sub> (functional assay).

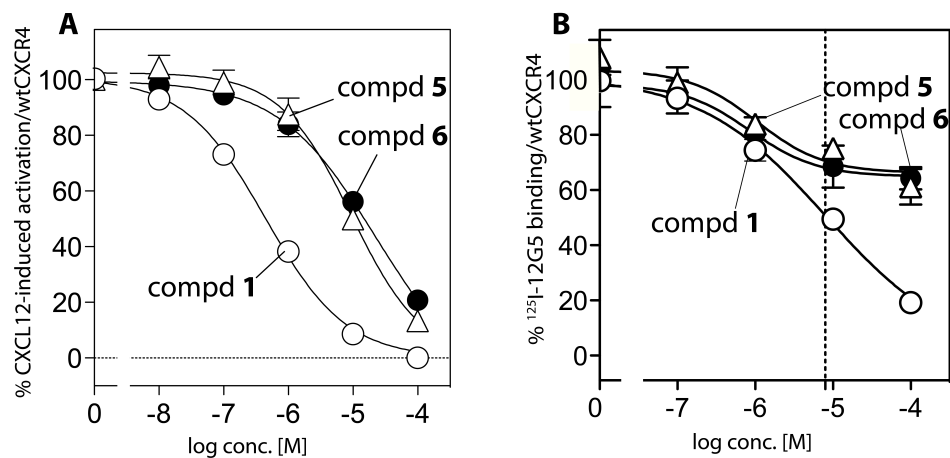
Helix	Receptor			Surface expression <sup>a</sup>		CXCL12 <sup>a</sup>		compd 1				compd 5				compd 6				
	Position <sup>b</sup>	Mutation	% ± SEM (n)	EC <sub>50</sub> ± SEM (log)	EC <sub>50</sub> (nM)	F <sub>mut</sub> <sup>c</sup>	Response at 0.1 μM (% of WT ± SEM)	(n) <sup>d</sup>	EC <sub>50</sub> ± SEM (log)	EC <sub>50</sub> (μM)	F <sub>mut</sub>	(n)	EC <sub>50</sub> ± SEM (log)	EC <sub>50</sub> (μM)	F <sub>mut</sub>	(n)	EC <sub>50</sub> ± SEM (log)	EC <sub>50</sub> (μM)	F <sub>mut</sub>	(n)
WT	WT	WT	100 ± 0.0 (5)	-8.8 ± 0.04	1.5	1.0		(69)	-6.3 ± 0.07	0.50	1.0	(56)	-5.1 ± 0.08	8.5	1.0	(16)	-4.9 ± 0.08	12	1.0	(16)
TM-1	I:07/I:39	Y45A	90 ± 7.0 (3)	-8.5 ± 0.20	3.4	2.3**	40 ± 5.5	(10)	-5.4 ± 0.21	4.0	8.0***	(6)	-4.3 ± 0.22	56	6.6***	(5)	-3.7 ± 0.15	181	15***	(4)
TM-2	II:20/2:60	W94A <sup>e</sup>	38 ± 6.4 (3)	-7.7 ± 0.10	18	13***	31 ± 11	(17)	-7.2 ± 0.18	0.06	0.12***	(16)	-6.3 ± 0.46	0.55	0.07***	(3)	-6.2 ± 0.46	0.64	0.05***	(3)
	II:23/2:63	D97A <sup>e</sup>	99 ± 6.3 (4)	-7.7 ± 0.07	20	14***	34 ± 3.5	(10)	-6.7 ± 0.12	0.21	0.41	(4)	-6.3 ± 0.21	0.54	0.06***	(5)	-5.9 ± 0.09	1.4	0.12***	(5)
TM-3	III:05/3:29	H113A <sup>e</sup>	80 ± 9.0 (3)	-9.1 ± 0.08	0.84	0.58**	83 ± 8.1	(21)	-4.9 ± 0.14	13	26***	(13)	> -4	> 100	> 12	(5)	> -4	> 100	5.8	(4)
	III:08/3:32	Y116A <sup>e</sup>	69 ± 12 (3)	> -7			-0.8 ± 5.9	(5)	not determined				not determined				not determined			
	III:09/3:33	T117A	74 ± 5.5 (3)	-8.8 ± 0.09	1.7	1.2	57 ± 13	(4)	-6.8 ± 0.23	0.17	0.34	(2)	-6.1 ± 0.19	0.87	0.10***	(2)	-5.9 ± 0.73	1.4	0.12***	(2)
TM-4	IV:20/4:60	D171N <sup>e</sup>	67 ± 11 (3)	-8.5 ± 0.08	3.2	2.2***	38 ± 4.7	(25)	-5.6 ± 0.13	2.7	5.4***	(15)	-4.2 ± 0.54	68	8.1***	(3)	-3.6 ± 0.54	231	20***	(3)
	Cys-3	R183A	17 ± 1.8 (3)	-10 ± 0.20	0.1	0.07***	24 ± 1.4	(3)	-6.0 ± 0.18	1.0	2.0	(3)	not determined				not determined			
	Cys+1	D187A <sup>e</sup>	49 ± 8.7 (4)	-7.9 ± 0.06	13	8.0***	41 ± 5.2	(3)	-6.5 ± 0.05	0.35	0.69	(3)	-5.4 ± 0.02	4.1	0.49	(3)	-5.3 ± 0.05	5.4	0.46	(3)
	Cys+2	R198A	174 ± 21 (3)	-9.3 ± 0.08	0.53	0.36***	44 ± 3.5	(4)	-6.1 ± 0.53	0.96	1.7	(2)	-5.1 ± 0.08	8.4	0.99	(2)	-4.6 ± 0.36	25	2.1	(2)
ECL-2	Cys+3	F189A	97 ± 8.7 (3)	-8.7 ± 0.11	1.8	1.2	73 ± 9.5	(9)	-6.9 ± 0.10	0.13	0.26**	(7)	-6.1 ± 0.17	0.74	0.09***	(2)	-6.0 ± 0.62	1.0	0.09***	(2)
	Cys+4	Y190A	105 ± 19 (3)	-8.9 ± 0.18	1.2	0.82	69 ± 7.0	(8)	-6.6 ± 0.43	0.27	0.53	(6)	-4.9 ± 0.11	12	1.4	(2)	-4.9 ± 0.56	12	1.0	(2)
	V:01/5:35	V196A	101 ± 13 (3)	-8.9 ± 0.17	1.4	0.96	67 ± 15	(8)	-5.8 ± 0.13	1.4	2.8	(5)	-5.4 ± 0.27	4.2	0.50	(3)	-4.9 ± 0.27	13	1.1	(3)
	V:04/5:38	F199A	79 ± 7.2 (3)	-8.8 ± 0.06	1.4	0.98	76 ± 16	(4)	-6.2 ± 0.03	0.69	1.4	(2)	not determined				not determined			
TM-5	V:05/5:39	Q200A	4.7 ± 2.8 (3)	-8.9 ± 0.07	1.4	0.95	71 ± 5.1	(10)	-5.4 ± 0.18	3.7	7.5***	(6)	-4.8 ± 0.30	17	2.0	(3)	-3.6 ± 0.26	250	21***	(3)
	V:05/5:39	Q200W	75 ± 5.2 (3)	-8.7 ± 0.08	1.8	1.2	34 ± 3.4	(11)	-5.0 ± 0.21	10	21***	(7)	-4.9 ± 0.16	14	1.6	(4)	-4.1 ± 0.12	85	7.2***	(4)
	V:08/5:42	H203A	111 ± 4.6 (3)	-8.9 ± 0.17	1.4	1.0	108 ± 24	(5)	-6.1 ± 0.33	0.71	1.4	(3)	not determined				not determined			
	VI:13/6:48	W252A	51 ± 4.3 (3)	-9.1 ± 0.06	0.78	0.53**	75 ± 6.2	(11)	-5.4 ± 0.26	3.8	7.5***	(9)	-4.6 ± 0.41	27	3.2	(4)	-3.8 ± 0.24	173	15***	(4)
TM-6	VI:16/6:51	Y255A	47 ± 3.5 (3)	-8.9 ± 0.11	1.1	0.77	32 ± 9.7	(7)	-5.8 ± 0.26	1.7	3.3	(3)	not determined				not determined			
	VI:20/6:55	I259A	3.0 ± 0.3 (3)	-8.7 ± 0.09	2.1	1.4	59 ± 4.6	(7)	-5.6 ± 0.28	2.6	5.1*	(3)	not determined				not determined			
	VI:20/6:55	I259W	34 ± 4.3 (3)	-8.9 ± 0.06	1.3	0.91	28 ± 4.2	(6)	-5.8 ± 0.22	1.5	2.9	(3)	not determined				not determined			
	VI:23/6:58	D262N <sup>e</sup>	54 ± 4.1 (3)	-8.2 ± 0.04	5.8	4.0***	63 ± 7.0	(23)	-4.9 ± 0.12	13	26***	(14)	-4.6 ± 0.23	27	3.2*	(3)	-5.0 ± 0.65	11	0.91	(3)
TM-7	VII:02/7:32	H281A <sup>e</sup>	169 ± 29 (3)	-8.7 ± 0.13	1.8	1.2	33 ± 9.5	(18)	-4.5 ± 0.28	31	81***	(7)	-4.3 ± 0.17	53	6.2***	(3)	-3.8 ± 0.17	155	13***	(4)
	VII:02/7:35	I284A	56 ± 9.7 (4)	-8.6 ± 0.05	2.3	1.6*	38 ± 4.3	(13)	-5.5 ± 0.22	3.4	6.7***	(7)	-4.5 ± 0.28	34	4.0***	(5)	-3.5 ± 0.61	306	26***	(3)
	VII:06/7:39	E288A <sup>e</sup>	77 ± 16 (5)	> -7			11 ± 4.0	(10)	not determined				not determined				not determined			

<sup>a</sup>Data for CXCL12 activation are from ref. 35. <sup>b</sup>The position of each residue is given based on the generic numbering system proposed by Baldwin and modified by Schwartz, followed by the Ballesteros/Weinstein numbering system. <sup>c</sup>F<sub>mut</sub> is the ratio of the mutant and WT-CXCR4 potencies. Red: F<sub>mut</sub> >25; orange: F<sub>mut</sub> from 10-25; yellow: F<sub>mut</sub> from 5-10; green: F<sub>mut</sub> <0.5. EC<sub>50</sub> values >100 μM are underlined and also colored red. P < 0.001\*\*\*, P < 0.01\*\*, P < 0.05\*. <sup>d</sup>The number of independent experiments is shown in parentheses (n). <sup>e</sup>These mutant receptors were also tested in binding assay (Table 3).



**Figure 4.** (A-C) Bar-graph representations of fold-reduction in potency ( $F_{mut}$ ) for compounds **1**, **5**, and **6** for each mutant with respect to WT-CXCR4 (functional data, Table 2). Grey bars: mutants in TMs 1-4 and ECL2; black bars: mutants in TMs 5-7. (D-F) Residues identified to be important for the potency of compounds **1**, **5**, and **6** (functional data, Table 2), shown in a helical

wheel diagram of CXCR4. Red background color: >25 fold decrease; orange: 10-25 fold decrease; yellow: 5-10 fold decrease.



**Figure 5.** (A) Ability of the lead compound **1** and analogs **5** and **6** to inhibit CXCL12 induced activation of WT-CXCR4, and (B) to displace <sup>125</sup>I-12G5 from WT-CXCR4.

*Displacement of Radiolabeled Monoclonal Antibody (<sup>125</sup>I-12G5).* To address the issue of reduced activity of CXCL12 in certain mutants (W94A, D97A, Y116A, D187A, and E288A) a competitive binding assay using the radiolabeled monoclonal antibody 12G5 (<sup>125</sup>I-12G5) was employed to probe ligand affinity (Table 3). A good correlation between anti-HIV potency and binding affinity measured as displacement of <sup>125</sup>I-12G5 has previously been demonstrated for the bicyclam antagonists,<sup>36, 38, 53</sup> which represents an additional advantage of using 12G5 as radioligand. In addition to the five mutants mentioned above, the binding experiments were extended to also include four of the mutants that were identified as important for the potency of **1** in the functional assay (H113A, D171N, D262N, and H281A). All nine mutants had similar affinities for 12G5 as the WT receptor (Table 3), meaning that the mutations did not significantly affect receptor folding.

Compared to the strong antagonistic potency ( $EC_{50} = 0.50 \mu\text{M}$ ) measured in the functional assay (Figure 5A), **1** displayed lower affinity to WT-CXCR4 ( $IC_{50} = 8.0 \mu\text{M}$ ) measured against  $^{125}\text{I}$ -12G5 (Figure 5B), suggesting that this compound binds allosterically with respect to 12G5. This finding clearly illustrates the differences between displacing  $^{125}\text{I}$ -12G5 (Table 3) and inhibiting the action of CXCL12 (Tables 1 and 2). The two  $R^1$  analogs **5** and **6** did not completely displace 12G5 (Figure 5B) – which also indicates an allosteric binding mode – and were therefore excluded from further binding experiments. The affinity of **1** determined against  $^{125}\text{I}$ -12G5 was strongly (8- to >12-fold) negatively affected by the TM3 mutants H113A, the TM4 mutant D171N, the TM5 mutant D262N, and the TM7 mutants H281A and E288A (Table 3). Moreover, **1** could not displace 12G5 from the Y116A-mutant. Overall, the binding data show that His<sup>113</sup>, Tyr<sup>116</sup>, and Asp<sup>171</sup> at the interface between TM3 and TM4, and Asp<sup>262</sup>, His<sup>281</sup>, and Glu<sup>288</sup> at the interface between TM6 and TM7 in the main binding pocket are important for the displacement of 12G5 by **1**.

To summarize, the functional studies (Tables 1 and 2) and the binding studies (Table 3) show that Tyr<sup>45</sup> (TM1), His<sup>113</sup> and Tyr<sup>116</sup> (TM3), Asp<sup>171</sup> (TM4), Gln<sup>200</sup> (TM5), Asp<sup>262</sup> (TM6), and His<sup>281</sup> and Glu<sup>288</sup> (TM7) are important for the function and/or binding of **1**.

**Table 3.** Affinity of **1** for WT-CXCR4 and nine mutant receptors measured as displacement of <sup>125</sup>I-labeled antibody 12G5 in transiently transfected COS-7 cells (competition binding assay).

Receptor			12G5 <sup>a</sup>						compd 1				
Helix	Position <sup>o</sup>	Mutation	IC <sub>50</sub> ± SEM (log)	IC <sub>50</sub> (nM)	F <sub>mut</sub> <sup>b</sup>	Bmax ± SEM (fmol/10 <sup>5</sup> cells)	(n) <sup>c</sup>	P <sup>d</sup> B <sub>max</sub>	IC <sub>50</sub> ± SEM (log)	IC <sub>50</sub> (μM)	F <sub>mut</sub>	P	(n)
WT	WT	WT	-8.3 ± 0.13	4.7	1.0	0.096 ± 0.018	(12)		-5.1 ± 0.116	8.0	1.0		(12)
TM-2	II:20/2.60	W94A	-8.7 ± 0.15	1.9	0.40	0.053 ± 0.022	(8)		-5.8 ± 0.217	1.7	0.21	**	(9)
	II:23/2.63	D97A	-8.0 ± 0.12	9.4	2.0	0.086 ± 0.013	(3)		-5.1 ± 0.129	8.5	1.1		(3)
TM-3	III:05/3.29	H113A	-8.6 ± 0.15	2.7	0.6	0.037 ± 0.006	(7)	*	> -4	> 100	>12		(7)
	III:08/3.32	Y116A	-8.1 ± 0.08	8.7	1.9	0.036 ± 0.016	(5)		No displacement				(3)
TM-4	IV:20/4.60	D171N	-8.7 ± 0.18	2.2	0.47	0.034 ± 0.017	(7)	*	-4.2 ± 0.112	67	8.3	***	(8)
ECL-2	Cys+1	D187A	-7.8 ± 0.04	16	3.4	0.161 ± 0.012	(3)		-4.6 ± 0.097	26	3.2		(3)
TM-6	VI:23/6.58	D262N	-8.4 ± 0.15	3.7	0.8	0.101 ± 0.020	(8)		> -4	> 100	>12		(8)
TM-7	VII:-02/7.32	H281A	-8.6 ± 0.15	2.4	0.52	0.028 ± 0.008	(7)	*	-4.2 ± 0.13	63	7.8	***	(8)
	VII:06/7.39	E288A	-8.7 ± 0.16	2.1	0.45	0.035 ± 0.011	(7)	*	> -4	> 100	>12		(8)

<sup>a</sup>Data for 12G5 binding are from ref. 35. <sup>b</sup>F<sub>mut</sub> is the ratio of the mutant and WT-CXCR4 affinities. <sup>c</sup>The number of independent experiments is shown in parentheses (n). <sup>d</sup>P<0.001\*\*\*, P<0.01\*\*, P<0.05\*. <sup>e</sup>The position of each residue is given based on the generic numbering system proposed by Baldwin and modified by Schwartz, followed by the Ballesteros/Weinstein numbering system.

**Molecular Docking and Derived Binding Model.** *Binding Model: Key Interactions.* In order to rationalize the experimental data, **1** was docked to the X-ray structure of CXCR4 (PDB code 3OE0)<sup>33</sup> using Schrödinger's induced-fit docking protocol,<sup>54</sup> which models the conformational changes induced by ligand binding. In order to avoid generation of irrelevant poses, a H-bond constraint was set on His<sup>113</sup> (N<sup>δ1</sup>), which was shown to be highly important for both the affinity and potency of **1**. Visual inspection of the 15 generated ligand-receptor complexes resulted in the identification of a pose that was in agreement with the experimental data. In our proposed binding model (Figure 6), the (pyridin-2-ylmethyl)amino moiety (R<sup>1</sup> side chain) of **1** is oriented towards the TM6 and TM7 region, and the R<sup>1</sup> side chain adopts a bent conformation around the



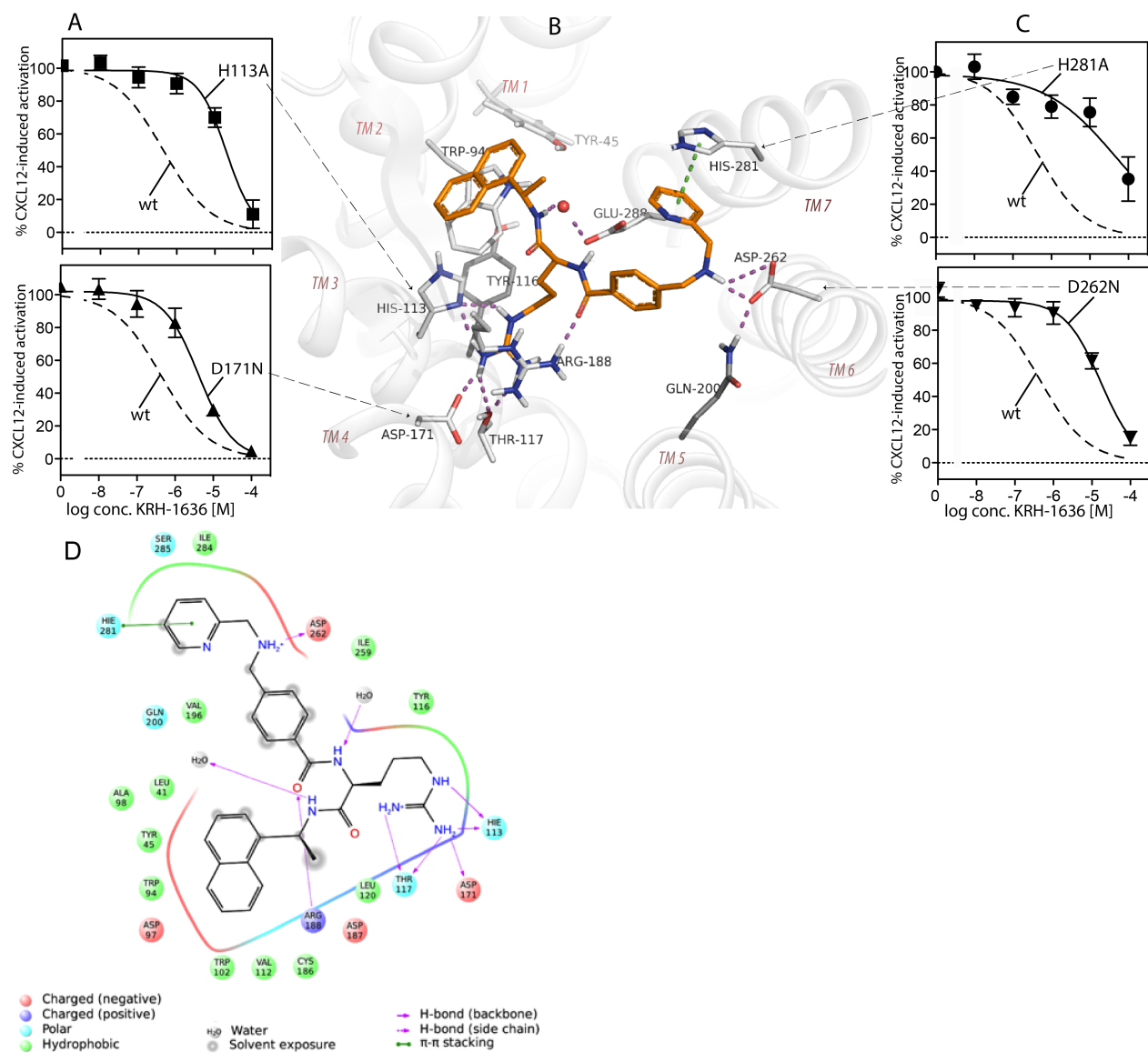
secondary amino group, where the pyridine and phenyl rings lie in almost parallel planes. His<sup>281</sup> in TM7 forms an aromatic  $\pi$ - $\pi$  stacking interaction with the pyridine ring, and Asp<sup>262</sup> in TM6 is involved in a bidentate interaction with the charged secondary amine in the R<sup>1</sup> side chain of **1**. The guanidino group of Arg (R<sup>2</sup>-position) forms polar interactions with His<sup>113</sup> and Thr<sup>117</sup> (TM3) and Asp<sup>171</sup> (TM4). The aromatic naphthyl ring (R<sup>3</sup>-position) is embedded in a hydrophobic pocket formed by the aromatic side-chains of Tyr<sup>45</sup> and Trp<sup>94</sup>, and further stabilized by the Tyr<sup>116</sup> side-chain located directly below Trp<sup>94</sup>. Interestingly, previous studies indicate that the three aromatic residues Tyr<sup>45</sup>, Trp<sup>94</sup> and Tyr<sup>116</sup> of CXCR4, which are also found in the CCR5 chemokine receptor as Tyr<sup>37</sup>, Trp<sup>86</sup>, and Tyr<sup>108</sup>, respectively, might form a hydrophobic pocket that is involved in the binding of small-molecule inhibitors.<sup>55, 56</sup> Thus, mutations of the aforementioned three residues could destabilize the hydrophobic subpocket that interacts with the naphthyl ring of **1**. Furthermore, the observation that Tyr<sup>116</sup> is involved in a direct interaction with Glu<sup>288</sup> (see Supporting Information, Figure S1) implies that the Y116A mutation can affect the overall geometry of the main binding crevice, specifically the orientations of TM3 and TM7. Thus, the Y116A mutation may indirectly affect the potential interactions between the ligand and His<sup>113</sup>, Thr<sup>117</sup>, His<sup>281</sup>, and Glu<sup>288</sup>.

*Binding Model: Indirect Effects.* The docking pose for **1** also reveals the involvement of Glu<sup>288</sup>, which is a key residue as shown in the binding assay (Table 3). A Glu residue at this position (VII:06/7.39) is highly conserved among chemokine receptors and acts as an anchor point for several small molecule ligands.<sup>57</sup> Glu<sup>288</sup> in CXCR4 has been consistently implicated in binding modes of several different classes of CXCR4 antagonists, *i.e.* AMD3100, AMD3465, **4**, FC131, and also the CXCR4 co-crystallized ligands **2** and **3**,<sup>33, 35, 37, 39, 40</sup> either through direct or water-

mediated interactions with the ligand. Similarly, our binding model shows a water-mediated interaction between Glu<sup>288</sup> and the ligand backbone (Figure 6).

Substitution of Gln<sup>200</sup> with Ala and Trp (Q200A and Q200W) caused a 7- and 21-fold reduction in potency of **1** with respect to WT-CXCR4; the involvement of this residue was also seen in previous binding mode studies of the monocyclam antagonist AMD3465.<sup>39</sup> In our binding model for **1** (Figure 6), an intra-receptor H-bond between Gln<sup>200</sup> and Asp<sup>262</sup> was identified, suggesting that substitution of Gln<sup>200</sup> can affect the interaction between Asp<sup>262</sup> and the ligand.

The model further shows an interaction between the guanidino group of Arg<sup>188</sup> and the ligand backbone (Figure 6), which was also previously identified in the binding modes of the peptide antagonists FC131 and **2**.<sup>23, 33, 58</sup> However, the R188A mutation did not have any effect on the antagonistic potency of **1** or the two analogs (Table 2); interestingly, the R188A-mutant was also reported to have no effect on the potency of FC131.<sup>35</sup> A possible explanation is that both **1** and FC131 are doubly positively charged, which results in electrostatic repulsion by Arg<sup>188</sup>. Substitution of Arg<sup>188</sup> with Ala (R188A) therefore results in two opposing effects: an unfavorable removal of the H-bond interaction with the ligand, and a favorable removal of the electrostatic repulsion. Thus, it can be hypothesized that the net result is no overall effect of the R188A mutation on **1** and FC131.



**Figure 6.** (A) Effects of the H113A and D171N mutations on the inhibition of CXCL12-induced PI turnover by **1**. (B) Proposed binding model for **1**, showing the key interacting residues in CXCR4; the ligand is shown in orange sticks, the receptor in grey-white sticks and ribbons, and the water molecule as a red sphere; H-bonds are shown in magenta, and  $\pi$ - $\pi$  interactions in green. (C) Effects of the H281A and D262N mutations on the inhibition of CXCL12-induced PI turnover by **1**. (D) 2D-representation of the docking pose for **1**.

While the effect of most mutations (Tables 2 and 3) can be explained by the binding model, it is not possible to rationalize every single one. For example, the W94A and D97A mutations both led to increased potencies for all three analogs (Table 2). Earlier studies have suggested the involvement of these residues in the second binding step (“site two”) for the endogenous ligand CXCL12,<sup>33, 45</sup> and the present mutagenesis data showed a reduced ability (13-14 fold) of CXCL12 to induce signaling for the W94A and D97A mutants. Moreover, the X-ray structure of CXCR4 suggests that Asp<sup>97</sup> constitutes a gate-entry position that controls ligand access to the major binding pocket;<sup>33, 35</sup> thus, a pre-interaction with Asp<sup>97</sup> is implicated in the passage of the ligands to the TM receptor region. Taken together, the existing data supported by our mutagenesis suggest that the TM2 residues (Trp<sup>94</sup> and Asp<sup>97</sup>) have a less distinctive involvement in the binding of these ligands to the receptor. First, mutation of Trp<sup>94</sup> and Asp<sup>97</sup> will weaken the binding of CXCL12, thereby increasing the antagonistic potency of the three analogs as reflected in Table 2. Second, the potential pre-interaction with Asp<sup>97</sup> is removed in the D97A mutant, which allows easier access of the ligands to the TM binding pocket. Third, the involvement of the Trp<sup>94</sup> in the hydrophobic pocket surrounding the naphthalene ring (R<sup>3</sup>-position) in **1** (Figure 6), suggests that removal of the bulky side chain of Trp (W94A-mutant) allows a more energetically favorable position of the naphthyl group in the binding pocket.

Conversely, mutation of hydrophobic residues in TM6 and TM7 reduced the potency of **1** (W252A, I259A and I284A) and **6** (W252A and I284A). However, residues Trp<sup>252</sup>, Ile<sup>259</sup>, and Ile<sup>284</sup> are not directly involved in ligand binding, meaning that the observed effects of the mutations most likely are secondary, *i.e.* implemented through local or global changes in TM6 and TM7.

Despite the lower WT potency of **6** (Table 2), the effects of mutants Y45A and D171N were more pronounced for this analog. Our SAR data showed that gradual removal of the R<sup>1</sup> group led to reduced potency, while the receptor mapping (Table 2) suggested that the R<sup>1</sup> group contacts TMs 5-7. It could be reasoned that the weakened interaction of **6** with TMs 5-7 (R<sup>1</sup> group interaction) makes the remaining interactions more important, which would explain the stronger dependence on Tyr<sup>45</sup> (R<sup>3</sup> group interaction) and Asp<sup>171</sup> (R<sup>2</sup> group interaction) for **6** (Table 2).

**Comparison with Experimentally Determined Binding Modes for Other Small-molecule CXCR4 Antagonists. Isothiourea Derivatives.** Gebhard *et al.* have reported a series of small-molecule isothiourea derivatives that selectively block CXCL12 and HIV gp120 binding to CXCR4.<sup>59, 60</sup> The X-ray structures of CXCR4 in complex with the most potent compound **3** (Figure 1E)<sup>33</sup> later showed that **3** exhibits a unique binding mode by occupying the so-called “minor binding pocket” between TMs 1, 2, 3, and 7,<sup>61</sup> *i.e.* without interacting with TMs 4-6. Consequently, the binding mode of **3** differs from other reported binding modes of CXCR4 antagonists (see below), and is distinct from the present binding mode for **1**.

*Peptide Antagonists.* Prior to the publication of the X-ray structure of CXCR4, computational binding mode studies had been performed for the cyclopentapeptide CXCR4 antagonist FC131<sup>62, 63</sup> and **1**<sup>63</sup> based on molecular docking to homology models of CXCR4 that were built on the crystal structure of bovine rhodopsin. These studies were purely theoretical, *i.e.* not based on any experimentally verified ligand-receptor interactions. Moreover, when the X-ray structure of CXCR4<sup>33</sup> was revealed in 2010 it showed that bovine rhodopsin was a poor template for CXCR4,<sup>64</sup> questioning the validity of the previously proposed binding models for the small molecule peptide-based antagonists.

However, we have recently proposed an experimentally verified binding mode for FC131 based on ligand modifications, receptor mutagenesis, and molecular docking to the X-ray structure of CXCR4.<sup>35</sup> The binding model for FC131 was generally consistent with the binding mode for the Arg<sup>1</sup>-Arg<sup>2</sup>-1-Nal<sup>3</sup> fragment of the 16-mer peptide antagonist **2** in complex with CXCR4 (PDB code 3OE0),<sup>33</sup> *i.e.* the Arg<sup>1</sup> side-chain in FC131 projected towards ECL2 of the receptor to interact with Asp<sup>187</sup>, the Arg<sup>2</sup> guanidino group interacted with His<sup>113</sup>, Thr<sup>117</sup>, and Asp<sup>171</sup> in TMs 3-4, while the aromatic 2-Nal<sup>3</sup> side-chain was accommodated in a hydrophobic pocket around TM5.

Thus, the Arg<sup>2</sup> guanidino group in FC131 and the guanidino group (R<sup>2</sup> position) in **1** both interact with His<sup>113</sup>, Thr<sup>117</sup>, and Asp<sup>171</sup> (Figure 6); structural comparison of the reported binding mode for FC131<sup>35</sup> and the present binding mode for **1** (Supporting Information, Figure S2) also shows that the L-Arg<sup>2</sup> side chain of FC131 overlaps relatively well with the central arginine residue of **1**. However, the orientation of the naphthyl groups in FC131 and **1** are different; the 2-naphthyl group in FC131 is oriented towards TM5, while the 1-naphthyl group in **1** is facing TM2.

*AMD-compound Series.* Based on site-directed mutagenesis studies of CXCR4 and binding site transfer to CXCR3, the critical CXCR4 receptor residues for binding of the bicyclam AMD3100 (Figure 1) have been shown to be Asp<sup>171</sup>, Asp<sup>262</sup>, and Glu<sup>288</sup>.<sup>36,37</sup> Based on the fact that AMD3100 and **1** inhibit binding of the same set of monoclonal antibodies, directed against the extracellular loops of CXCR4, it has been suggested that they have similar binding sites and molecular mode of action.<sup>28,65</sup> As Asp<sup>171</sup>, Asp<sup>262</sup>, and indirectly Glu<sup>288</sup> are also involved in binding of **1** (Figure 6), the present study confirms that these two compounds bind to the same receptor region.

In addition to the key interactions seen for AMD3100, the monocyclam analog AMD3465 (Figure 3) has been shown to interact with His<sup>113</sup> in TM3 and His<sup>281</sup> in TM7,<sup>39,40</sup> which were also identified as important for **1** in the present study. The reduced potency observed for **1** in the D262N and H281A mutants are of particular interest as the same mutations were shown to have selective importance for the receptor interactions of AMD3465,<sup>39</sup> which contains the same (pyridin-2-ylmethyl)amino moiety as **1**. Thus, in addition to the key interactions with His<sup>113</sup> and Asp<sup>171</sup> (R<sup>2</sup> guanidino group) and Asp<sup>262</sup> (R<sup>1</sup> amino group), the present study suggests that the pyridine ring in the R<sup>1</sup> side chain of **1** picks up an additional interaction with His<sup>281</sup> in TM7, in the same way as AMD3465.

The interaction pattern for the non-cyclam **4** (Figure 1F) is more unclear, but it has been suggested that it binds in a similar fashion as AMD3100 and AMD3465; however, Asp<sup>97</sup> and Asp<sup>171</sup> appeared to have a unique importance for the binding of this analog.<sup>40,66</sup>

## CONCLUSIONS

The present study describes the binding of the prototype tripeptidomimetic antagonist **1** to CXCR4 based on SAR studies, receptor mapping, and molecular docking. This ternary approach resulted in the identification of a binding model for **1** in which the secondary amino group is anchored by interaction with Asp<sup>262</sup> in TM6, with the distal pyridine ring further involved in a  $\pi$ - $\pi$  stacking interaction with His<sup>281</sup> in TM7; the Arg-guanidino group engages in polar interactions mainly with His<sup>113</sup> in TM3 and Asp<sup>171</sup> in TM4, while the naphthyl group is tightly packed in a hydrophobic subpocket formed by residues Trp<sup>94</sup>, Tyr<sup>45</sup> and Tyr<sup>116</sup>. Our findings contradict earlier suggestions that **1** mimics the binding of the Arg<sup>1</sup>-Arg<sup>1</sup>-2-Nal<sup>3</sup> fragment in the cyclopentapeptide antagonist FC131. Instead, we propose a “hybrid” binding mode for **1** where the guanidino group

of the central arginine residue (R<sup>2</sup> side chain) overlaps with the binding mode of the Arg<sup>2</sup> residue in the peptide antagonist FC131 and **2**, while the (pyridin-2-ylmethyl)amino moiety (R<sup>1</sup> side chain) overlaps with the binding mode for the same moiety in the monocyclam AMD3465. This experimentally verified binding model for **1** provides important insight for future structure-based design of small-molecule CXCR4 antagonists.

## EXPERIMENTAL SECTION

**Chemistry. General.** All reagents and solvents were purchased from Sigma-Aldrich and used as received. <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a 400 MHz Varian spectrometer. Chemical shifts are expressed in ppm relative methanol (<sup>1</sup>H 3.31 ppm, <sup>13</sup>C 49.0 ppm). Coupling constants are given in Hertz (Hz). Mass spectra were recorded on a Micromass Quattro LC (Micromass, Manchester, UK) and high-resolution mass spectra were recorded on an LTQ Orbitrap XL (Thermo Scientific, Bremen, Germany) using positive-mode ESI. Preparative RP-HPLC was performed on a Waters 600 Semi Prep System equipped with an XBridge™ C<sub>18</sub> column (250 mm × 19 mm, 10 μm particle size) using 15 mL/min flow rate or a Waters 2695 system equipped with an XBridge™ C<sub>18</sub> column (250 mm × 10 mm, 5 μm particle size) using 10 mL/min flow rate with mixtures of acetonitrile and water (both containing 0.1% TFA) as the eluent in both cases Analytical RP-HPLC was performed on a Waters 2695 system equipped with an XBridge™ C<sub>18</sub> column (250 mm × 4.6 mm, 5 μm particle size,) using 1 mL/min flow rate with detection at 214 nm and 254 nm facilitated by a PDA detector (210 – 310 nm). In some cases, analyses were performed on a Waters ACQUITY UHPLC H-Class equipped with a Waters ACQUITY UHPLC BEH C<sub>18</sub> column (1 × 150 mm, 1.75 μm particle size) with a 0.120



mL/min flow rate. All compounds undergoing biological evaluation were found to be of >95% purity as judged by analytical RP-HPLC with PDA detection (210 – 310 nm).

***N*-((*S*)-5-guanidino-1-(((*S*)-1-(naphthalen-1-yl)ethyl)amino)-1-oxopentan-2-yl)-4-(((pyridin-2-ylmethyl)amino)methyl)benzamide (1).** The title compound was prepared following an earlier reported procedure.<sup>29</sup> Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 8.12 (d, *J* = 8.3, 1H), 7.96 (d, *J* = 8.0, 2H), 7.91 – 7.85 (m, 2H), 7.81 (d, *J* = 8.2, 1H), 7.64 (d, *J* = 8.0, 2H), 7.60 – 7.37 (m, 7H), 5.94 – 5.82 (m, 1H), 4.62 (app t, *J* = 8.0, 1H), 4.41 (s, 2H), 4.39 (s, 2H), 3.19 – 3.05 (m, 2H), 1.94 – 1.74 (m, 2H), 1.66 (d, *J* = 6.9, 3H), 1.64 – 1.50 (m, 2H); <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 172.7, 169.2, 158.5, 152.3, 150.6, 140.1, 138.8, 136.3, 136.1, 135.4, 132.2, 131.2, 129.9, 129.3, 129.0, 127.2, 126.7, 126.4, 125.1, 124.1, 124.1, 123.7, 54.9, 51.5, 51.3, 46.2, 41.9, 30.5, 26.4, 21.2; HRMS (ESI): *m/z* [M + H]<sup>+</sup> calcd for C<sub>32</sub>H<sub>38</sub>N<sub>7</sub>O<sub>2</sub>: 552.3081, found 552.3081. The purity of the title compound was found to be >99% using analytical RP-HPLC (diode array detection, 210 nm – 310 nm).

**4-(aminomethyl)-*N*-((*S*)-5-guanidino-1-(((*S*)-1-(naphthalen-1-yl)ethyl)amino)-1-oxopentan-2-yl)benzamide (6).** The title compound was prepared adopting a literature procedure<sup>29</sup> where *N*-Boc 4-(methylamino)benzoic acid was used instead of **11** (Scheme 1). Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.89 (d, *J* = 8.3, 1H), 7.72 (d, *J* = 8.3, 2H), 7.66 (app t, *J* = 8.2, 1H), 7.58 (d, *J* = 8.2, 1H), 7.41 – 7.23 (m, 6H), 5.69 – 5.60 (m, 1H), 4.41 (app dd, *J* = 8.3, 6.0, 1H), 3.98 (s, 2H), 3.00 – 2.82 (m, 2H), 1.76 – 1.55 (m, 2H), 1.52 – 1.45 (m, 1H), 1.44 (d, *J* = 6.9, 3H), 1.40 – 1.30 (m, 1H); <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 172.7, 169.1, 158.3, 139.9, 138.0, 135.6, 135.2, 132.0, 129.8, 129.7, 129.1, 128.9, 127.1, 126.6, 126.2, 123.9, 123.5,

54.8, 46.0, 43.6, 41.7, 30.3, 26.3, 21.1; HRMS (ESI) :  $m/z$   $[M + H]^+$  calcd for  $C_{26}H_{33}N_6O_2$ : 461.2660, found 461.2664. The purity of the title compound was found to be >99% using analytical RP-HPLC (diode array detection, 210 nm – 310 nm).

**4-((benzylamino)methyl)-*N*-((*S*)-5-guanidino-1-(((*S*)-1-(naphthalen-1-yl)ethyl)amino)-1-oxopentan-2-yl)benzamide (5).** A solution of **6** (crude product following TFA mediated deprotection, 0.09 mmol) in DMF (2.0 mL) was cooled to 0 °C and  $K_2CO_3$  (24 mg, 0.18 mmol) was added. To this suspension, benzyl bromide (10  $\mu$ L, 0.09 mmol) was added slowly, and the reaction mixture was allowed to stir at rt for 24 h. Water (1 mL) was added and the mixture was extracted with diethyl ether (10 mL), washed with water (2 x 5 mL) and a saturated solution of NaCl (10 mL). The organic layer was dried over anhydrous  $MgSO_4$ , filtered and the solvents were removed under reduced pressure. The crude product was purified using RP-HPLC followed by lyophilization to give the title compound as a fluffy white material.  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta$  8.12 (d,  $J = 8.2$ , 1H), 7.95 (d,  $J = 8.2$ , 2H), 7.90 (d,  $J = 7.8$ , 1H), 7.81 (d,  $J = 8.2$ , 1H), 7.60 (t,  $J = 7.6$ , 3H), 7.57 – 7.45 (m, 8H), 5.95 – 5.82 (m, 1H), 4.62 (app dd,  $J = 7.9$ , 6.3, 1H), 4.33 (s, 2H), 4.28 (s, 2H), 3.21 – 3.06 (m, 2H), 1.97 – 1.73 (m, 2H), 1.66 (d,  $J = 6.9$ , 3H), 1.64 – 1.48 (m, 2H);  $^{13}C$  NMR (100 MHz,  $CD_3OD$ ) 173.0, 169.5, 158.7, 140.3, 136.5, 136.3, 135.7, 132.5, 132.4, 131.4, 131.3, 131.0, 130.6, 130.1, 129.6, 129.3, 127.5, 127.0, 126.6, 124.3, 123.9, 55.2, 52.5, 51.7, 46.5, 42.1, 30.7; HRMS (ESI):  $m/z$   $[M + H]^+$  calcd for  $C_{33}H_{39}N_6O_2$ : 551.3129, found 551.3131. The purity of the title compound was found to be >99% using analytical RP-HPLC (diode array detection, 210 nm – 310 nm).

***N*-((*S*)-5-guanidino-1-(((*S*)-1-(naphthalen-1-yl)ethyl)amino)-1-oxopentan-2-yl)-4-methylbenzamide (7).** The title compound was prepared adopting a literature procedure<sup>29</sup> where 4-methylbenzoic acid was used instead of **11** (Scheme 1). Purification by RP-HPLC afforded

after lyophilization the title compound as a white fluffy material.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.12 (d,  $J = 8.3$ , 1H), 7.90 (d,  $J = 7.7$ , 1H), 7.81 (d,  $J = 8.2$ , 1H), 7.76 (d,  $J = 8.2$ , 2H), 7.60 – 7.45 (m, 4H), 7.30 (d,  $J = 8.2$ , 2H), 5.88 (m, 1H), 4.70 – 4.59 (m, 1H), 3.20 – 3.04 (m, 2H), 2.41 (s, 3H), 1.93 – 1.72 (m, 2H), 1.66 (d,  $J = 6.9$ , 3H), 1.64 – 1.48 (m, 2H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ) 171.5, 168.9, 157.0, 143.1, 138.7, 133.8, 130.5, 129.8, 129.4, 128.9, 128.0, 127.2, 126.3, 125.8, 125.4, 122.6, 122.4, 53.2, 45.5, 40.4, 29.6, 24.5, 21.4, 21.1; HRMS (ESI):  $m/z$   $[\text{M} + \text{H}]^+$  calcd for  $\text{C}_{26}\text{H}_{32}\text{N}_5\text{O}_2$ : 446.2551, found 446.2550. The purity of the title compound was found to be >99% using analytical RP-HPLC (diode array detection, 210 nm – 310 nm).

**N-((S)-1-(((S)-1-(naphthalen-1-yl)ethyl)amino)-1-oxo-5-ureidopentan-2-yl)-4-(((pyridin-2-ylmethyl)amino)methyl)benzamide (8)**. The title compound was prepared adopting a literature procedure<sup>29</sup> where Fmoc-Cit-OH was used instead of Fmoc-Arg(Pbf)-OH. Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.38 (d,  $J = 4.3$ , 1H), 7.81 (d,  $J = 8.3$ , 1H), 7.67 (d,  $J = 8.3$ , 2H), 7.63 – 7.57 (m, 2H), 7.51 (d,  $J = 8.2$ , 1H), 7.32 (m, 3H), 7.27 – 7.13 (m, 5H), 5.56 (app q,  $J = 6.8$ , 1H), 4.37 (app dd,  $J = 8.7$ , 5.6, 1H), 4.12 (s, 2H), 4.09 (s, 2H), 2.89 – 2.73 (m, 2H), 1.65 – 1.43 (m, 2H), 1.36 (d,  $J = 6.9$ , 3H), 1.34 – 1.16 (m, 2H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  173.3, 169.3, 162.3, 152.1, 150.4, 140.2, 139.1, 136.3, 135.9, 135.4, 132.1, 131.2, 129.9, 129.4, 129.0, 127.3, 126.7, 126.4, 125.2, 124.4, 124.1, 123.6, 55.2, 51.5, 51.2, 46.3, 40.4, 30.5, 27.7, 21.4; HRMS (ESI):  $m/z$   $[\text{M} + \text{H}]^+$  calcd for  $\text{C}_{32}\text{H}_{37}\text{N}_6\text{O}_3$ : 553.2922, found 553.2921. The purity of the title compound was found to be >99% using analytical RP-HPLC (diode array detection, 210 nm – 310 nm).

**(S)-N-(1-(ethylamino)-5-guanidino-1-oxopentan-2-yl)-4-(((pyridin-2-ylmethyl)amino)methyl)benzamide (9)**. The title compound was prepared adopting a literature procedure<sup>29</sup> where ethanamine was used instead of (S)-1-(naphthalene-1-yl)ethanamine.

Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 8.66 (d, *J* = 4.8, 1H), 7.97 (d, *J* = 8.2, 2H), 7.87 (td, *J* = 7.7, 1.7, 1H), 7.64 (d, *J* = 8.2, 2H), 7.46 – 7.39 (m, 2H), 4.55 (dd, *J* = 8.6, 5.9 Hz, 1H), 4.40 (s, 2H), 4.39 (s, 2H), 3.29 – 3.19 (m, 4H), 1.97 – 1.61 (m, 4H), 1.14 (t, *J* = 7.3, 3H); <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 173.6, 169.3, 158.8, 152.3, 150.7, 138.8, 136.3, 136.1, 131.2, 129.4, 125.1, 124.2, 55.0, 51.5, 51.3, 42.0, 35.4, 30.4, 26.5, 14.8; HRMS (ESI): *m/z* [M + H]<sup>+</sup> calcd for C<sub>22</sub>H<sub>32</sub>N<sub>7</sub>O<sub>2</sub>: 426.2612, found 426.2611. The purity of the title compound was found to be >99% using analytical RP-HPLC (diode array detection, 210 nm – 310 nm).

**Biology. Materials.** The human CXCR4 receptor cDNA was kindly provided by Tim Wells (GSK, UK). The human chemokine CXCL12 was purchased from Peprotech (NJ, USA). The monoclonal antibody 12G5 was kindly provided by Jim Hoxie (University of Pennsylvania, Philadelphia, PA, USA), and the promiscuous chimeric G protein Gα<sub>Δ6qi4myr</sub> (abbreviated Gα<sub>qi4myr</sub>) was kindly provided by Evi Kostenis (University of Bonn, Germany)<sup>67</sup>. Myo[<sup>3</sup>H]inositol (PT6-271), scintillation proximity assay (SPA) beads, and Bolton-Hunter reagent for 12G5 iodination (in-house procedure)<sup>37</sup> were purchased from Perkin Elmer (MA, USA). Primer-cDNAs for mutagenesis were purchased from TAG Copenhagen (Copenhagen, Denmark), and AG1-X8 anion exchange resin from Bio-Rad (CA, USA). Stock solutions of the ligand antagonists were prepared in different mixtures of DMSO:water, and dilutions were made in water. For dilutions of CXCL12, a buffer with 1 mM acetic acid + 0.1% BSA was used.

**Site-directed Mutagenesis.** Point mutations were introduced in the receptor by the polymerase chain reaction overlap extension technique<sup>68</sup> or the QuikChange<sup>TM</sup> technique (Agilent Technologies, CA, USA), using WT-CXCR4 as a template. All reactions were carried out using *Pfu* polymerase (Stratagene, CA, USA) under conditions recommended by the manufacturer. The

mutations were cloned into the eukaryotic expression vector pcDNA3.1+ (Invitrogen, UK), and were verified by restriction endonuclease digestion and DNA sequencing (Eurofins MWG Operon, Germany).

**Transfections and Tissue Culture.** COS-7 cells were grown at 10% CO<sub>2</sub> and 37 °C in Dulbecco's modified Eagle's medium with glutamax (Invitrogen, U.K.) adjusted with 10% fetal bovine serum, 180 units/mL penicillin, and 45 µg/mL penicillin/streptomycin (PenStrep). Transfection of the COS-7 cells was performed by the calcium phosphate precipitation method.<sup>69</sup>

**Phosphatidylinositol (PI) Turnover Assays.** Transfection of COS-7 cells was carried out according to the abovementioned procedure. Shortly,  $6 \times 10^6$  cells were transfected with 20 µg of receptor cDNA along with 30 µg of the promiscuous chimeric G protein, G $\alpha_{qi4myr}$ , which turns the G $\alpha_i$ -coupled signal, the most common pathway for endogenous chemokine receptors, into the G $\alpha_q$  pathway (phospholipase C activation measured as IP<sub>3</sub> turnover). Two alternative methods, which were found to give the same result, were used:

*Anion-exchange Separation of Phosphoinositides.* In the classic IP<sub>3</sub>-assay which includes an anion-exchange chromatography step, one day after transfection COS-7 cells ( $1.5 \times 10^5$  cells/well) were incubated for 24 h with 2 µCi of myo[<sup>3</sup>H]inositol in 0.3 mL of growth medium per well in a 24-well plate. Next day, the cells were washed twice in PBS, and were incubated (90 min) in 0.2 mL of Hank's balanced salt solution (Invitrogen, U.K.) supplemented with 10 mM LiCl at 37 °C in the presence of various concentrations of the tested compounds. Cells were then extracted by addition of 1 mL of 10 mM formic acid to each well followed by incubation on ice for 30–60 min. The generated [<sup>3</sup>H]inositol phosphates were purified on an AG 1-X8 anion exchange resin. Following the addition of multipurpose liquid scintillation cocktail (Gold Star, Triskem-International, France),  $\gamma$ -radiation was counted in a Beckman Coulter counter LS6500.

*Scintillation Proximity Assay.* The day after transfection, COS-7 cells (35 000 cells/well) were incubated for 24 h with 0.5  $\mu\text{Ci/mL}$  myo[ $^3\text{H}$ ]inositol in 100  $\mu\text{L}$  of growth medium per well in a 96-well plate. The following day, cells were treated as mentioned in the classic assay with volumes adjusted as follows: 100  $\mu\text{L}$  of reaction solution with LiCl, and 50  $\mu\text{L}$  of ice cold formic acid. Next, the generated [ $^3\text{H}$ ]inositol phosphates were quantified by radioactivity measurement; in brief, 20  $\mu\text{L}$  of the formic acid cell extracts were transferred to a white 96-well plate, and 80  $\mu\text{L}$  of 1:8 diluted YSi poly-D-lysine coated beads (SPA-beads, PerkinElmer, MA, USA) was added. Plates were sealed, agitated at maximum speed for at least 30 min, and centrifuged (5 min at 1500 rpm) before  $\gamma$ -radiation was counted in a Packard Top Count NXT counter. In both assays, determinations were made in duplicate.

**Binding Experiments.**  $6 \times 10^6$  COS-7 cells were transfected with 40  $\mu\text{g}$  of receptor cDNA and transferred to culture plates. The number of cells seeded per well was determined by the apparent expression efficiency of the receptors and was aimed at obtaining 5–10% specific binding of the added radioactive ligand. Two days after transfection, cells were assayed by competition binding for 3 h at 4  $^\circ\text{C}$  using 10–15 pM  $^{125}\text{I}$ -12G5 plus unlabeled ligand in 0.2 ml of 50 mM HEPES buffer, pH 7.4, supplemented with 1 mM  $\text{CaCl}_2$ , 5 mM  $\text{MgCl}_2$ , and 0.5% (w/v) bovine serum albumin. Following incubation, cells were washed quickly two times in 4  $^\circ\text{C}$  binding buffer supplemented with 0.5 M NaCl (and on ice). Cells were lysed by addition of 0.5 ml carbamide solution (18% acetic acid, 8M urea, 2% v/v P-40) and radioactivity was counted in Beckman Coulter counter LS6500. Nonspecific binding was determined in the presence of 0.1  $\mu\text{M}$  unlabeled 12G5. Determinations were made in duplicates.

**Cell Surface Expression Measurement (ELISA).** COS-7 cells were transfected with FLAG-tagged receptor DNA in 96-well plates ( $6 \times 10^3$  cells/well). Following the transfection (48

hours), cells were washed 3 times in Tris-buffered saline (TBS), fixed in 3.7% formaldehyde for 15 min at rt, washed and incubated in blocking solution (TBS with 2% BSA) for 30 min. Cells were incubated for 2 hours with anti-FLAG (M1) antibody (Sigma-Aldrich) at 2  $\mu\text{g}/\text{mL}$  in TBS with 1mM  $\text{CaCl}_2$  and 1% BSA. After three washes with TBS/ $\text{CaCl}_2$ /BSA the cells were incubated with goat anti-mouse horseradish peroxidase (HRP)-conjugated antibody at 0.8  $\mu\text{g}/\text{mL}$  (Thermo Fischer Scientific, MA, USA) for 1 hour. After 3 washes, the immunoreactivity was revealed by the addition of TMB Plus substrate (Kem-En-Tec, Taastrup, Denmark) and the reaction was stopped with 0.2 M  $\text{H}_2\text{SO}_4$ . Absorbance was measured at 450 nm on a Wallac platereader (Perkin Elmer).

**Statistical Calculations.** The apparent  $\text{EC}_{50}$  and  $\text{IC}_{50}$  values were determined by nonlinear regression using the GraphPad Prism 4 software (GraphPad Software, San Diego, CA). *P*-values were calculated using unpaired two-tailed *t*-test with 95% confidence intervals. The data analysis was performed as recently described.<sup>35</sup>

**Computational Modeling Procedure.** All molecular modeling calculations were performed using the Schrödinger software suite 2012.<sup>70</sup> Default settings were used, if not specified otherwise. The crystal structure of CXCR4 in complex with **2** (PDB-code: 3OE0)<sup>33</sup> was first prepared and optimized with the Protein Preparation Wizard workflow<sup>71</sup> using the same procedure as previously described.<sup>23</sup> The automatically assigned tautomeric and protonation states (pH 7.4) for ionizable residues were kept, except for Asp<sup>262</sup>, which was manually assigned a negative charge (predicted to be neutral). The ligand (**1**) was docked to the prepared receptor structure using the induced-fit workflow<sup>54</sup> with a docking box of (26 Å)<sup>3</sup> centered on Asp<sup>187</sup>. The ligand was built with two positive charges (the guanidino group and the amino group), and the “penalize nonplanar conformation” option was chosen for amide bonds.

In the first step (Glide docking: flexible ligand, rigid receptor), 50 poses were generated using a softened van der Waals potential (scaling 0.5) for both receptor and ligand. In the second step (Prime refinement: rigid ligand, flexible receptor), receptor residues within a radius of 5.0 Å from the ligand were optimized for the 50 initial poses. In the final step (Glide redocking: flexible ligand, rigid receptor), the ligand was redocked into the optimized binding site using the extra precision (XP) scoring function, and the top 20 poses within an energy window of 30 kcal/mol were kept. In the absence of constraints, a number of irrelevant poses were generated, and a H-bond constraint (applied in both docking steps) was therefore assigned to the N<sup>δ1</sup> atom of His<sup>113</sup>, which resulted in the identification of the binding mode for **1** that is shown in Figure 6.

## ASSOCIATED CONTENT

**Supporting Information.** Characterization data (<sup>1</sup>H and <sup>13</sup>C NMR spectra, HPLC/UPLC chromatograms) for **1** and **5–9** and additional 3D figures from computational modeling. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: +47 77 62 09 09. E-mail: [jon.vabeno@uit.no](mailto:jon.vabeno@uit.no).

### Notes

The authors declare no competing financial interest.



## ACKNOWLEDGMENTS

Financial support for this project was obtained from UiT The Arctic University of Norway (to Z.G.Z. and J.V.), the University of Bergen (to B. E. H.), and from the Danish Council for Independent Research/Medical Sciences, the Novo Nordisk Foundation, the Lundbeck Foundation, and the European Community Sixth Framework programme (INNOCHEM: Grant LSHBCT-2005-518167) (to S.K. and M.M.R). We also thank Inger Smith Simonsen, Randi Thøgersen and Maibritt Sigvardt Baggesen for outstanding technical assistance.

## ABBREVIATIONS

2-Nal, L-3-(2-naphthyl)alanine; CXCL12, CXC chemokine ligand 12; CXCR4, CXC chemokine receptor 4; DIPEA, *N,N*-diisopropylethylamine; ECL, extracellular loop; EDCI, *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide; HOBt, 1-hydroxybenzotriazole; Pbf, 2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl; PI, phosphatidylinositol; PyBOP, (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate; SPA, scintillation proximity assay; TIS, triisopropylsilane; TM, transmembrane helix.

## REFERENCES

- (1) Murphy, P. M.; Baggiolini, M.; Charo, I. F.; Hébert, C. A.; Horuk, R.; Matsushima, K.; Miller, L. H.; Oppenheim, J. J.; Power, C. A. International union of pharmacology. XXII. Nomenclature for chemokine receptors. *Pharmacol. Rev.* **2000**, *52*, 145-176.
- (2) Laing, K. J.; Secombes, C. J. Chemokines. *Dev. Comp. Immunol.* **2004**, *28*, 443-460.
- (3) Sugiyama, T.; Kohara, H.; Noda, M.; Nagasawa, T. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity* **2006**, *25*, 977-988.
- (4) Ward, S. G.; Westwick, J. Chemokines: understanding their role in T-lymphocyte biology. *Biochem. J.* **1998**, *333*, 457-470.
- (5) Horuk, R. Chemokine receptors. *Cytokine Growth Factor Rev.* **2001**, *12*, 313-335.
- (6) Bleul, C. C.; Farzan, M.; Choe, H.; Parolin, C.; Clark-Lewis, I.; Sodroski, J.; Springer, T. A. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* **1996**, *382*, 829-833.
- (7) Oberlin, E.; Amara, A.; Bachelier, F.; Bessia, C.; Virelizier, J.; Arenzana-Seisdedos, F.; Schwartz, O.; Heard, J.; Clark-Lewis, I.; Legler, D. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* **1996**, *382*, 833-835.
- (8) Ma, Q.; Jones, D.; Borghesani, P. R.; Segal, R. A.; Nagasawa, T.; Kishimoto, T.; Bronson, R. T.; Springer, T. A. Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar

neuron migration in CXCR4- and SDF-1-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 9448-9453.

(9) Tachibana, K.; Hirota, S.; Iizasa, H.; Yoshida, H.; Kawabata, K.; Kataoka, Y.; Kitamura, Y.; Matsushima, K.; Yoshida, N.; Nishikawa, S.-i. The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract. *Nature* **1998**, *393*, 591-594.

(10) Busillo, J. M.; Benovic, J. L. Regulation of CXCR4 signaling. *Biochim. Biophys. Acta* **2007**, *1768*, 952-963.

(11) Nagasawa, T.; Kikutani, H.; Kishimoto, T. Molecular cloning and structure of a pre-B-cell growth-stimulating factor. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 2305-2309.

(12) Feng, Y.; Broder, C. C.; Kennedy, P. E.; Berger, E. A. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **1996**, *272*, 872-877.

(13) Berson, J. F.; Long, D.; Doranz, B. J.; Rucker, J.; Jirik, F. R.; Doms, R. W. A seven-transmembrane domain receptor involved in fusion and entry of T-cell-tropic human immunodeficiency virus type 1 strains. *J. Virol.* **1996**, *70*, 6288-6295.

(14) Choi, W. T.; Duggineni, S.; Xu, Y.; Huang, Z.; An, J. Drug discovery research targeting the CXC chemokine receptor 4 (CXCR4). *J. Med. Chem.* **2012**, *55*, 977-994.

(15) Scholten, D.; Canals, M.; Maussang, D.; Roumen, L.; Smit, M.; Wijtmans, M.; de Graaf, C.; Vischer, H.; Leurs, R. Pharmacological modulation of chemokine receptor function. *Br. J. Pharmacol.* **2012**, *165*, 1617-1643.

(16) Singh, I. P.; Chauthe, S. K. Small molecule HIV entry inhibitors: Part I. Chemokine receptor antagonists: 2004-2010. *Expert Opin. Ther. Pat.* **2011**, *21*, 227-269.

(17) De Clercq, E.; Yamamoto, N.; Pauwels, R.; Baba, M.; Schols, D.; Nakashima, H.; Balzarini, J.; Debyser, Z.; Murrer, B. A.; Schwartz, D.; Thornton, D.; Bridger, G.; Fricker, S.; Henson, G.; Abrams, M.; Picker, D. Potent and selective inhibition of human immunodeficiency virus (HIV)-1 and HIV-2 replication by a class of bicyclams interacting with a viral uncoating event. *Proc. Natl. Acad. Sci.* **1992**, *89*, 5286-5290.

(18) De Clercq, E. Recent advances on the use of the CXCR4 antagonist plerixafor (AMD3100, Mozobil™) and potential of other CXCR4 antagonists as stem cell mobilizers. *Pharmacol. Ther.* **2010**, *128*, 509-518.

(19) Fujii, N.; Oishi, S.; Hiramatsu, K.; Araki, T.; Ueda, S.; Tamamura, H.; Otaka, A.; Kusano, S.; Terakubo, S.; Nakashima, H.; Broach, J. A.; Trent, J. O.; Wang, Z. X.; Peiper, S. C. Molecular-size reduction of a potent CXCR4-chemokine antagonist using orthogonal combination of conformation- and sequence-based libraries. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 3251-3253.

(20) Tamamura, H.; Esaka, A.; Ogawa, T.; Araki, T.; Ueda, S.; Wang, Z.; Trent, J. O.; Tsutsumi, H.; Masuno, H.; Nakashima, H. Structure–activity relationship studies on CXCR4 antagonists having cyclic pentapeptide scaffolds. *Org. Biomol. Chem.* **2005**, *3*, 4392-4394.

(21) Tamamura, H.; Araki, T.; Ueda, S.; Wang, Z.; Oishi, S.; Esaka, A.; Trent, J. O.; Nakashima, H.; Yamamoto, N.; Peiper, S. C.; Otaka, A.; Fujii, N. Identification of novel low molecular weight CXCR4 antagonists by structural tuning of cyclic tetrapeptide scaffolds. *J. Med. Chem.* **2005**, *48*, 3280-3289.

(22) Ueda, S.; Oishi, S.; Wang, Z. X.; Araki, T.; Tamamura, H.; Cluzeau, J.; Ohno, H.; Kusano, S.; Nakashima, H.; Trent, J. O.; Peiper, S. C.; Fujii, N. Structure-activity relationships of cyclic peptide-based chemokine receptor CXCR4 antagonists: disclosing the importance of side-chain and backbone functionalities. *J. Med. Chem.* **2007**, *50*, 192-198.

(23) Mungalpara, J.; Thiele, S.; Eriksen, Ø.; Eksteen, J.; Rosenkilde, M. M.; Våbenø, J. Rational design of conformationally constrained cyclopentapeptide antagonists for CXC chemokine receptor 4 (CXCR4). *J. Med. Chem.* **2012**, *55*, 10287-10291.

(24) Mungalpara, J.; Zachariassen, Z. G.; Thiele, S.; Rosenkilde, M. M.; Våbenø, J. Structure-activity relationship studies of the aromatic positions in cyclopentapeptide CXCR4 antagonists. *Org. Biomol. Chem.* **2013**, *11*, 8202-8208.

(25) Yamazaki, T.; Saitou, A.; Ono, M.; Yokoyama, S.; Bannai, K.; Hirose, K.; Yanaka, M. Preparation of amino acid amide derivatives as antagonists of chemokine CXCR4 receptor. WO2003029218A1, 2003.

(26) Yamazaki, T.; Kikumoto, S.; Ono, M.; Saitou, A.; Takahashi, H.; Kumakura, S.; Hirose, K.; Yanaka, M.; Takemura, Y.; Suzuki, S.; Matsui, R. Preparation of bis(heterocyclmethyl)amine compounds as chemokine receptor CXCR4 antagonists. WO2004024697A1, 2004.

(27) Murakami, T.; Kumakura, S.; Yamazaki, T.; Tanaka, R.; Hamatake, M.; Okuma, K.; Huang, W.; Toma, J.; Komano, J.; Yanaka, M. The novel CXCR4 antagonist KRH-3955 is an orally bioavailable and extremely potent inhibitor of human immunodeficiency virus type 1 infection: comparative studies with AMD3100. *Antimicrob. Agents Chemother.* **2009**, *53*, 2940-2948.

(28) Ichiyama, K.; Yokoyama-Kumakura, S.; Tanaka, Y.; Tanaka, R.; Hirose, K.; Bannai, K.; Edamatsu, T.; Yanaka, M.; Niitani, Y.; Miyano-Kurosaki, N.; Takaku, H.; Koyanagi, Y.; Yamamoto, N. A duodenally absorbable CXC chemokine receptor 4 antagonist, KRH-1636, exhibits a potent and selective anti-HIV-1 activity. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 4185-4190.

(29) Yanaka, M.; Yamazaki, T.; Bannai, K.; Hirose, K. CXCR4-antagonistic drugs composed of nitrogen-containing compound. US 0157818, August 12, 2004.

(30) Grande, F.; Garofalo, A.; Neamati, N. Small molecules anti-HIV therapeutics targeting CXCR4. *Curr. Pharm. Des.* **2008**, *14*, 385-404.

(31) Debnath, B.; Xu, S.; Grande, F.; Garofalo, A.; Neamati, N. Small molecule inhibitors of CXCR4. *Theranostics* **2013**, *3*, 47-75.

(32) Våbenø, J.; Nikiforovich, G. V.; Marshall, G. R. A minimalistic 3D pharmacophore model for cyclopentapeptide CXCR4 antagonists. *Biopolymers* **2006**, *84*, 459-471.

(33) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010**, *330*, 1066-1071.

(34) Qin, L.; Kufareva, I.; Holden, L. G.; Wang, C.; Zheng, Y.; Zhao, C.; Fenalti, G.; Wu, H.; Han, G. W.; Cherezov, V.; Abagyan, R.; Stevens, R. C.; Handel, T. M. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science* **2015**, *347*, 1117-1122.

(35) Thiele, S.; Mungalpara, J.; Steen, A.; Rosenkilde, M.; Våbenø, J. Determination of the binding mode for the cyclopentapeptide CXCR4 antagonist FC131 using a dual approach of ligand modifications and receptor mutagenesis. *Br. J. Pharmacol.* **2014**, *171*, 5313-5329.

(36) Gerlach, L. O.; Skerlj, R. T.; Bridger, G. J.; Schwartz, T. W. Molecular interactions of cyclam and bicyclam non-peptide antagonists with the CXCR4 chemokine receptor. *J. Biol. Chem.* **2001**, *276*, 14153-14160.

(37) Rosenkilde, M. M.; Gerlach, L. O.; Jakobsen, J. S.; Skerlj, R. T.; Bridger, G. J.; Schwartz, T. W. Molecular mechanism of AMD3100 antagonism in the CXCR4 receptor. *J. Biol. Chem.* **2004**, *279*, 3033-3041.

(38) Hatse, S.; Princen, K.; Vermeire, K.; Gerlach, L. O.; Rosenkilde, M. M.; Schwartz, T. W.; Bridger, G.; De Clercq, E.; Schols, D. Mutations at the CXCR4 interaction sites for AMD3100 influence anti-CXCR4 antibody binding and HIV-1 entry. *FEBS Lett.* **2003**, *546*, 300-306.

(39) Rosenkilde, M. M.; Gerlach, L. O.; Hatse, S.; Skerlj, R. T.; Schols, D.; Bridger, G. J.; Schwartz, T. W. Molecular mechanism of action of monocyclam versus bicyclam non-peptide antagonists in the CXCR4 chemokine receptor. *J. Biol. Chem.* **2007**, *282*, 27354-27365.

(40) Wong, R. S. Y.; Bodart, V.; Metz, M.; Labrecque, J.; Bridger, G.; Fricker, S. P. Comparison of the potential multiple binding modes of bicyclam, monocyclam, and noncyclam small-molecule CXC chemokine receptor 4 inhibitors. *Mol. Pharmacol.* **2008**, *74*, 1485-1495.

(41) Hatse, S.; Princen, K.; Clercq, E. D.; Rosenkilde, M. M.; Schwartz, T. W.; Hernandez-Abad, P. E.; Skerlj, R. T.; Bridger, G. J.; Schols, D. AMD3465, a monomacrocyclic CXCR4 antagonist and potent HIV entry inhibitor. *Biochem. Pharmacol.* **2005**, *70*, 752-761.

(42) Baldwin, E. T.; Weber, I. T.; St Charles, R.; Xuan, J. C.; Appella, E.; Yamada, M.; Matsushima, K.; Edwards, B. F.; Clore, G. M.; Gronenborn, A. M.; Wlodawer, A. Crystal structure of interleukin 8: symbiosis of NMR and crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 502-506.

(43) Schwartz, T. W. Locating ligand-binding sites in 7TM receptors by protein engineering. *Curr. Opin. Biotechnol.* **1994**, *5*, 434-444.

(44) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366-428.

(45) Brelot, A.; Heveker, N.; Montes, M.; Alizon, M. Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities. *J. Biol. Chem.* **2000**, *275*, 23736-23744.

(46) Crump, M. P.; Gong, J. H.; Loetscher, P.; Rajarathnam, K.; Amara, A.; Arenzana-Seisdedos, F.; Virelizier, J. L.; Baggiolini, M.; Sykes, B. D.; Clark-Lewis, I. Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J.* **1997**, *16*, 6996-7007.

(47) Zhou, N.; Luo, Z.; Luo, J.; Liu, D.; Hall, J. W.; Pomerantz, R. J.; Huang, Z. Structural and functional characterization of human CXCR4 as a chemokine receptor and HIV-1 co-



receptor by mutagenesis and molecular modeling studies. *J. Biol. Chem.* **2001**, 276, 42826-42833.

(48) Choi, W. T.; Tian, S.; Dong, C. Z.; Kumar, S.; Liu, D.; Madani, N.; An, J.; Sodroski, J. G.; Huang, Z. Unique ligand binding sites on CXCR4 probed by a chemical biology approach: implications for the design of selective human immunodeficiency virus type 1 inhibitors. *J. Virol.* **2005**, 79, 15398-15404.

(49) Tian, S.; Choi, W.-T.; Liu, D.; Pesavento, J.; Wang, Y.; An, J.; Sodroski, J. G.; Huang, Z. Distinct functional sites for human immunodeficiency virus type 1 and stromal cell-derived factor 1 $\alpha$  on CXCR4 transmembrane helical domains. *J. Virol.* **2005**, 79, 12667-12673.

(50) Truax, V. M.; Zhao, H. Y.; Katzman, B. M.; Prosser, A. R.; Alcaraz, A. A.; Saindane, M. T.; Howard, R. B.; Culver, D.; Arrendale, R. F.; Gruddanti, P. R.; Evers, T. J.; Natchus, M. G.; Snyder, J. P.; Liotta, D. C.; Wilson, L. J. Discovery of tetrahydroisoquinoline-based CXCR4 antagonists. *ACS Med. Chem. Lett.* **2013**, 4, 1025-1030.

(51) Brelot, A.; Heveker, N.; Adema, K.; Hosie, M. J.; Willett, B.; Alizon, M. Effect of mutations in the second extracellular loop of CXCR4 on its utilization by human and feline immunodeficiency viruses. *J. Virol.* **1999**, 73, 2576-2586.

(52) Katritch, V.; Cherezov, V.; Stevens, R. C. Structure-function of the G protein-coupled receptor superfamily. *Annu. Rev. Pharmacol. Toxicol.* **2013**, 53, 531-556.

(53) Bridger, G. J. S., R. T. Bicyclam derivatives as HIV inhibitors. In *Advances in Antiviral Drug Design*, Volume 3; De Clercq, E., Ed.; JAI Press: Stamford, CT, 1999; pp 161-229.

(54) Schrödinger Suite 2012 Induced Fit Docking protocol; Glide version 5.8, S., LLC, New York, NY, 2012; Prime version 3.1, Schrödinger, LLC, New York, NY, 2012.

(55) Dragic, T.; Trkola, A.; Thompson, D. A. D.; Cormier, E. G.; Kajumo, F. A.; Maxwell, E.; Lin, S. W.; Ying, W. W.; Smith, S. O.; Sakmar, T. P.; Moore, J. P. A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5639-5644.

(56) Tsamis, F.; Gavrilov, S.; Kajumo, F.; Seibert, C.; Kuhmann, S.; Ketas, T.; Trkola, A.; Palani, A.; Clader, J. W.; Tagat, J. R. Analysis of the mechanism by which the small-molecule CCR5 antagonists SCH-351125 and SCH-350581 inhibit human immunodeficiency virus type 1 entry. *J. Virol.* **2003**, *77*, 5201-5208.

(57) Rosenkilde, M. M.; Schwartz, T. W. GluVII:06 - A highly conserved and selective anchor point for non-peptide ligands in chemokine receptors. *Curr. Top. Med. Chem.* **2006**, *6*, 1319-1333.

(58) Yoshikawa, Y.; Kobayashi, K.; Oishi, S.; Fujii, N.; Furuya, T. Molecular modeling study of cyclic pentapeptide CXCR4 antagonists: New insight into CXCR4-FC131 interactions. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2146-2150.

(59) Glickman, F.; Streiff, M.; Thoma, G.; Zerwes, H.-G. Preparation of thiazolyl isothiourea derivatives as CXCR4 antagonists. WO2005085219A1, 2005.

(60) Thoma, G.; Streiff, M. B.; Kovarik, J.; Glickman, F.; Wagner, T.; Beerli, C.; Zerwes, H. G. Orally bioavailable isothioureas block function of the chemokine receptor CXCR4 in vitro and in vivo. *J. Med. Chem.* **2008**, *51*, 7915-7920.

- (61) Rosenkilde, M. M.; Benned-Jensen, T.; Frimurer, T. M.; Schwartz, T. W. The minor binding pocket: a major player in 7TM receptor activation. *Trends Pharmacol. Sci.* **2010**, *31*, 567-574.
- (62) Våbenø, J.; Nikiforovich, G. V.; Marshall, G. R. Insight into the binding mode for cyclopentapeptide antagonists of the CXCR4 receptor. *Chem. Biol. Drug Des.* **2006**, *67*, 346-354.
- (63) Kawatkar, S. P.; Yan, M.; Gevariya, H.; Lim, M. Y.; Eisold, S.; Zhu, X.; Huang, Z.; An, J. Computational analysis of the structural mechanism of inhibition of chemokine receptor CXCR4 by small molecule antagonists. *Exp. Biol. Med.* **2011**, *236*, 844-850.
- (64) Planesas, J. M.; Perez-Nueno, V. I.; Borrell, J. I.; Teixido, J. Impact of the CXCR4 structure on docking-based virtual screening of HIV entry inhibitors. *J. Mol. Graph. Model.* **2012**, *38*, 123-136.
- (65) Leonard, J. T.; Roy, K. The HIV entry inhibitors revisited. *Curr. Med. Chem.* **2006**, *13*, 911-934.
- (66) Cox, B. D.; Prosser, A. R.; Katzman, B. M.; Alcaraz, A. A.; Liotta, D. C.; Wilson, L. J.; Snyder, J. P. Anti-HIV small-molecule binding in the peptide subpocket of the CXCR4:CVX15 crystal structure. *ChemBioChem* **2014**, *15*, 1614-1620.
- (67) Kostenis, E.; Zeng, F.-Y.; Wess, J. Functional characterization of a series of mutant G protein  $\alpha_q$  subunits displaying promiscuous receptor coupling properties. *J. Biol. Chem.* **1998**, *273*, 17886-17892.

(68) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **1989**, *77*, 51-59.

(69) Kostenis, E.; Martini, L.; Ellis, J.; Waldhoer, M.; Heydorn, A.; Rosenkilde, M. M.; Norregaard, P. K.; Jorgensen, R.; Whistler, J. L.; Milligan, G. A highly conserved glycine within linker I and the extreme C terminus of G protein  $\alpha$  subunits interact cooperatively in switching G protein-coupled receptor-to-effector specificity. *J. Pharmacol. Exp. Ther.* **2005**, *313*, 78-87.

(70) Schrödinger. <http://www.schrodinger.com/> (accessed Jul 27, 2015).

(71) Schrödinger Suite 2012 Protein Preparation Wizard; Epik version 2.2, S., LLC, New York, NY, 2012; Impact version 5.7, Schrödinger, LLC, New York, NY, 2012; Prime version 2.3, Schrödinger, LLC, New York, NY, 2012.

## Table of Contents Graphic

