Probing the molecular interactions between CXC chemokine receptor 4 (CXCR4) and the tripeptidomimetic antagonist KRH-1636

Zack G. Zachariassen,[a] Stefanie Thiele,[b] Bengt Erik Haug,[c] Mette M. Rosenkilde,[b] and Jon Våbenø*[a]

[a] *Department of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Breivika, NO-9037 Tromsø, Norway. Fax: +47 77 64 61 51; Tel: +47 77 62 09 09; E-mail: jon.vabeno@uit.no.*

[b] *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.*

[c] *Department of Chemistry and Centre of 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 structure-activity relationship study was first conducted based on the three functionalities of KRH-1636, followed by site-directed mutagenesis studies. The receptor mapping showed that both the potency and affinity of KRH-1636 were dependent on the receptor residues His^{113} (III:05), Asp^{171} (IV:20), Asp^{262} (VI:23), and His^{281} (VII:-02). The potency was also found to depend on Tyr^{45} (I:07) and Gln^{200} (V:05), while binding affinity was also affected by mutation of Tyr^{116} (III:08) and Glu^{288} (VII:06). Molecular docking of KRH-1636 to the X-ray structure of CXCR4 showed that the L-Arg guanidino group of KRH-1636 forms polar interactions with His^{113} and Asp^{171}, the (pyridin-2-ylmethyl)amino moiety is anchored by Asp^{262} and His^{281}, while the naphthalene ring is tightly packed in a hydrophobic network formed by the aromatic side chains of Trp^{94} (II:20), Tyr^{45} (I:07), and Tyr^{116} (III:08). 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 G protein-coupled receptors (GPCRs).[1] The involvement and physiological role of the chemokine ligand-receptor system extends from control of chemotaxis (directional migration of leukocytes)[2] and hematopoiesis to embryonic development and angiogenesis.[3, 4] The human C-X-C chemokine receptor 4 (CXCR4) is activated by C-X-C chemokine ligand 12 (CXCL12), which is a 68-residue peptide.[5, 6] CXCR4/CXCL12 is involved in development of the hematopoietic, nervous, and gastrointestinal systems[7, 8] and plays an essential role in immune defense and inflammation.[9, 10] In addition to these physiological roles, CXCR4 is the co-receptor for cellular entry of T-tropic HIV strains;[11, 12] 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.[2, 3] 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.[13-15] So far, the hematopoietic stem cell mobilizing agent plerixafor (AMD3100, Figure 1A)[16] is the only CXCR4 antagonist that has reached the market.[17] Due to the high positive charge at physiological pH, the bicyclam 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).[18] Successive structure-activity relationship (SAR) studies of FC131 have since shown that the Arg^1-Arg^2-2-Nal^3 tripeptide fragment of the cyclopentapeptide serves as an essential recognition motif for peptide-based CXCR4 antagonists,[19-23] 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[24-26] with the potent KRH-1636[27, 28] (Figure 1C) 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.[29, 30] Based on structural comparison of low-energy

conformations of KRH-1636 with a 3D-pharmacophore model for the cyclopentapeptide CXCR4 antagonists, KRH-1636 has been suggested to mimic the Arg[1]-Arg[2]-2-Nal[3] fragment of FC131.[31] However, it has not been shown experimentally that the two ligands bind to CXCR4 in the same way.

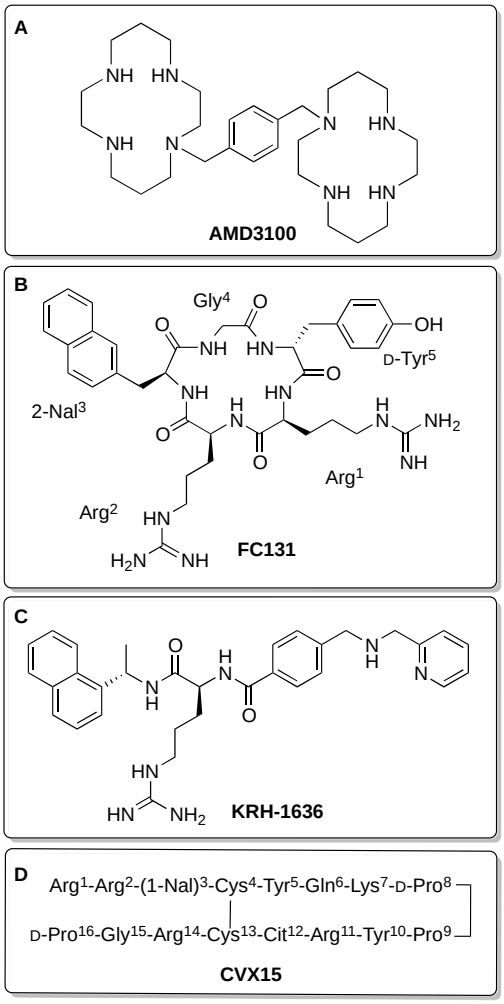


Figure 1. Structures of selected CXCR4 antagonists: (A) the non-peptide AMD3100 (B) the cyclopentapeptide FC131, (C) the prototype tripeptidomimetic ligand KRH-1636, and (D) the 16-mer peptide CVX15.

The X-ray structure of the 16-mer peptide CVX15 (Figure 1D) bound to CXCR4[32] 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).[33] Prior to the publication of the X-ray structure of CXCR4, extensive receptor mapping studies were performed for the AMD-compound series, which resulted in the identification of the binding mode for AMD3100[34, 35] and analogs.[36, 37] 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 KRH-1636 using a combination of SAR studies, receptor mapping by site-directed mutagenesis, and molecular docking.

RESULTS AND DISCUSSION

Structure-Activity Relationship Study. *Design.* KRH-1636 was developed from a hit compound identified through screening of a chemical library,[27, 28] and although recent patent literature discloses compounds related to KRH-1636,[24-26] no SAR data has been reported for this compound class. In order to confirm the importance of the functionalities in KRH-1636 (**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^1 and R^2) and the aromatic group (R^3) (Figure 2). In order to probe the R^1 position, the pyridine ring was replaced with a phenyl ring (**2**), the pyridin-2-ylmethyl substituent was removed to give the primary amine (**3**), and the charged (pyridin-2-ylmethyl)amino moiety was removed altogether (**4**). The importance of the charged guanidino group (R^2) was investigated by replacing it with a urea group (**5**), while the aromatic position (R^3) was probed by removal of the naphthalene unit (**6**).

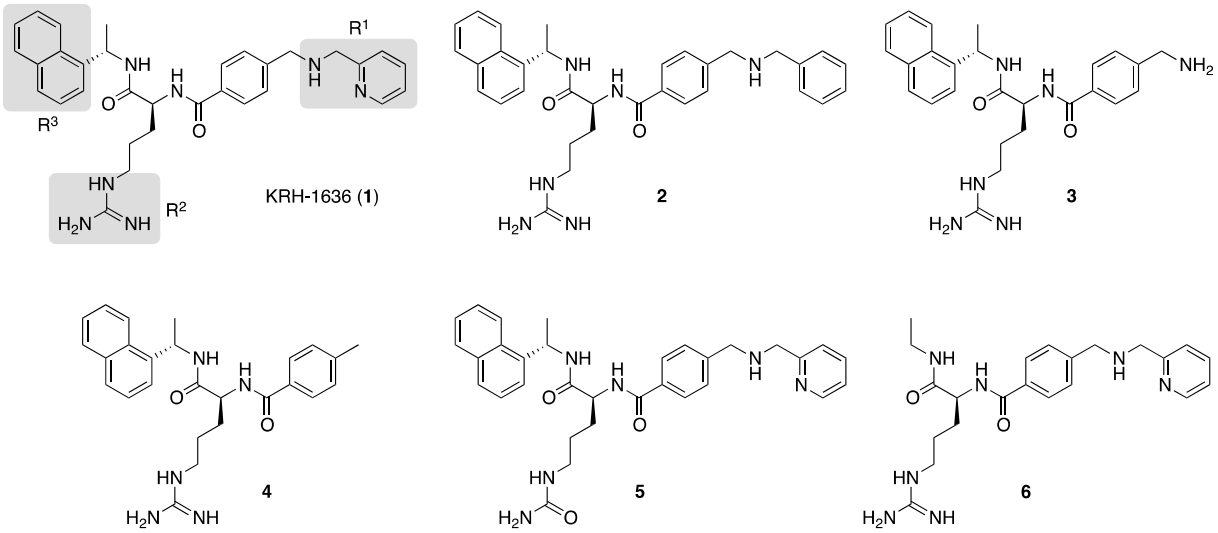
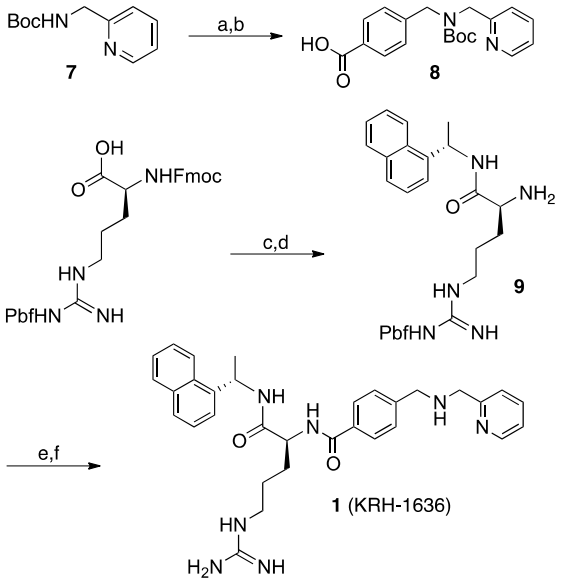


Figure 2. Structures of compounds **1**–**6**; the modified groups (R^1, R^2, and R^3) in the lead compound KRH-1636 (**1**) are highlighted.

Chemistry. Compounds **1**–**6** (Figure 2) were all prepared by adopting a procedure previously reported for the synthesis of KRH-1636 (Scheme 1).[28] For compound **1**, the right-hand side was prepared by N-alkylation of Boc-protected pyridin-2-ylmethanamine (**7**) with methyl 4-(bromomethyl)benzoate followed by hydrolysis of the methyl ester to give carboxylic acid **8**. The left-hand side was assembled by coupling of Fmoc-L-Arg(Pbf)

with (*S*)-1-(naphthalen-1-yl)ethanamine using PyBOP as the coupling reagent, followed by Fmoc-deprotection with diethylamine in DMF to afford amine **9**. Coupling of the two halves (**8** and **9**) was facilitated by EDCI/HOBt, and subsequent removal of the Boc and Pbf protecting groups by treatment with a TFA cocktail gave KRH-1636 (**1**). For the preparation of **3** and **4**, the left hand side (**9**) was coupled with *N*-Boc 4-(aminomethyl)benzoic acid and 4-methylbenzoic acid, respectively. Analog **2** was prepared by N-alkylation of **3** with benzyl bromide using K_2CO_3 as base and DMF as solvent. The R^2 analog **5** was prepared using Fmoc-protected citrulline as the starting material while the R^3 analog **6** 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 r.t.; (b) 1M NaOH/THF/MeOH (1:1:1); (c) (*S*)-1-(naphthalen-1-yl)ethanamine, PyBOP, DIEA, CH_2Cl_2 0 °C to r.t.; (d) Et_2HN, DMF; (e) **8**, EDCI, HOBt, DMF; (f) TFA/TIS/H_2O (95:2.5:2.5).

Biology and SAR. The antagonistic potency of compounds **1**–**6** (Table 1) was assessed in a functional assay measuring the inhibition of CXCL12-induced activation of human CXCR4, transiently expressed in COS-7 cells. KRH-1636 was originally reported by Ichiyama et al. to efficiently inhibit binding of [^{125}I]CXCL12 (IC_{50} = 0.013 μM),[27] and our own functional data

validates **1** as a potent CXCR4 antagonist (EC_{50} = 0.50 μM, Table 1) with respect to the endogenous ligand CXCL12. The phenyl-analog **2** (EC_{50}= 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 **3** (EC_{50}= 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 **4** (EC_{50} > 100 μM) confirms the positively charged amino-group in R^1 as an essential pharmacophoric element. Compound **5**, 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 μM), demonstrating the importance of the positive charge in R^2 position. Similarly, the loss of potency for the R^3 analog **6** (EC_{50} > 100 μM) confirms an important role for the naphthyl group.

Table 1. Antagonistic potency of **1**–**6** on human CXCR4.

compd	log EC_{50} ± SEM[a]	EC_{50} (μM)
1[b]	-6.30 ± 0.07	0.50
2	-5.12 ± 0.08	8.5
3	-4.93 ± 0.08	12
4	> -4	> 100
5	> -4	> 100
6	> -4	> 100

[a]Values represent the mean of at least three independent experiments performed in duplicates. [b]Known compound.

Earlier attempts to improve the pharmacokinetic profile of the bicyclam CXCR4 antagonists, i.e. AMD3100 (Figure 1), led to the discovery of the monocyclam AMD3465 (Figure 3),[38] which was shown to be 22-fold more potent than its precursor AMD3100[36] and contains the same (pyridin-2-ylmethyl)amino moiety as KRH-1636 (**1**). The importance of this group for the antagonistic activity of AMD3465 has been investigated previously using the analogs AMD3529 and AMD3389 (Figure 3),[36] which are structural counterparts of the KRH-analogs **2** and **4** in the present study (Figure 2). In line with the present findings

for **2** and **4** (Table 1), AMD3529 and AMD3389 were shown to display significantly lower potencies for CXCR4 than AMD3465 as determined by a functional assay inhibiting CXCL12-induced receptor activation.[36]

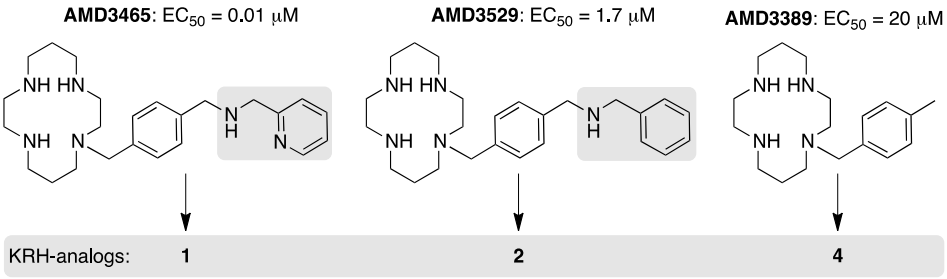


Figure 3. Structures and reported potencies[36] of the monocyclam AMD3465 and the two analogs AMD3529 and AMD3389, also showing the structural analogy of these monocyclams with the R^1 substituent in compounds **1**, **2**, and **4** in the present study.

Regarding the central Arg residue in KRH-1636 (R^2), it has been suggested that this position could correlate to Arg^2 in the cyclopentapeptide antagonist FC131 (Figure 1).[31] SAR studies of the cyclopentapeptides have shown that the Arg^2 position is very sensitive to modifications,[19, 22] and that replacement of the guanidino group with a urea group (Cit^2) leads to complete loss of antagonistic potency,[22] which is in agreement with the present data for the KRH-analog **5**.

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 KRH-1636, while the pyridine ring in R^1 position contributes to increased potency.

Identification of binding site for KRH1636: Receptor mutagenesis. Next, we carried out mutagenesis studies to map key binding interactions between KRH-1636 and CXCR4 using a library of 24 CXCR4-mutants (Table 2); the same library has previously been utilized successfully to map the binding of the cyclopentapeptide FC131.[33] Most of the mutations were located in the transmembrane helices (TMs), but some were also introduced in extracellular loop 2 (ECL-2) that connects TM-4 and TM-5. 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.

Inhibition of CXCL12-induced activation of CXCR4. We first assessed the antagonistic potency of KRH-1636 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 KRH-1636 relative to wild-type (wt) CXCR4 (the position of residues based on the generic numbering system proposed by Baldwin[39] and modified by Schwartz[40] followed by the Ballesteros/Weinstein numbering system[41] is given in Tables 2-3). The potency of KRH-1636 was also reduced (by 5-25 fold) by mutations of Tyr^{45} (Y45A), Asp^{171} (D171N), Gln^{200} (Q200A and Q200W), Trp^{252} (W252A), Ile^{259} (I259A), and Ile^{284} (I284A). Interestingly, substitutions of the bulky Trp^{94} residue and the acidic Asp^{97} residue in TM-2 to Ala (W94A and D97A) led to *increased* potency (>2-fold).

As the mutagenesis data for KRH-1636 alone did not provide information about which parts of the ligand that interact with the different parts of the receptor, the R^1 analogs **2** and **3** (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^1-side chain with CXCR4. The collective functional data analysis for **1**–**3** (Table 2) showed that the mutations in TM-1 (Y45A), TM-2 (W94A, D97A), TM-3 (H113A, T117A), TM-4 (D171N), and ECL-2 affected the three ligands in a similar manner. Due to the lower wt-potency of **2** and **3** compared to **1** (Table 1), the effects of mutating key residues were expected to be more pronounced for these two analogs. Accordingly, for the residues that span TMs 1-4 and ECL-2, a larger effect on the least potent analog **3** was seen for the Y45A- and D171N-mutants (Table 2). However, several mutations in TMs 5-7 affected the three ligands in a distinctive manner. Specifically, the H281A mutation in TM-7 was the mutation with largest effect on the potency of KRH-1636 (61-fold reduction), but affected analog **2** to a much lesser extent (6-fold). Thus, the overall trend identified in the PI-turnover experiments for **1**–**3** (Table 2) suggests that the structurally modified R^1 group contacts TMs 5-7; this trend is further illustrated in Figure 4.

The absolute majority of the mutants had no or little effect on the agonistic potency of CXCL12 (Table 2), demonstrating that most of the mutants were expressed and functional. However, the D97A, D187A, and E288A mutations, and also W94A and Y116A, reduced the

Table 2. Potency KRH-1636 (**1**) and the analogs **2** and **3** on wt-CXCR4 and 24 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_{qi4myr} (functional assay).

			CXCL12					KRH-1636 (1)					R^1-analog (2)				R^1-analog (3)			
	Position[a]	Mutation	EC_{50} ± *log*	SEM	EC_{50} *nM*	F_{mut}[b]	(n)[c]	EC_{50} ± *log*	SEM	EC_{50} *µM*	F_{mut}	(n)	EC_{50} ± *log*	SEM	EC_{50} *µM*	F_{mut}	(n)	EC_{50} ± *log*	SEM	EC_{50} *µM*	F_{mut}	(n)
	wt	CXCR4	-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:07/1.39	Y45A	-8,5 ±	0,20	3,4	2,3	(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	-7,7 ±	0,10	18	13	(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	-7,7 ±	0,07	20	14	(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	-9,1 ±	0,08	0,84	0,58	(21)	-4,9 ±	0,14	13	26	(13)	> -4		> 100		(5)	> -4		> 100		(4)
	III:08/3.32	Y116A	> -7				(5)	not determined					not determined					not determined				
	III:09/3.33	T117A	-8,8 ±	0,09	1,7	1,2	(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	-8,5 ±	0,08	3,2	2,2	(25)	-5,6 ±	0,13	2,7	5,3	(15)	-4,2 ±	0,54	68	8,1	(3)	-3,6 ±	0,54	231	20	(3)
ECL-2	Cys+1	D187A	-7,9 ±	0,06	13	8,6	(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	R188A	-9,3 ±	0,08	0,53	0,36	(4)	-6,1 ±	0,53	0,86	1,7	(2)	-5,1 ±	0,08	8,4	0,99	(2)	-4,6 ±	0,36	25	2,1	(2)
	Cys+3	F189A	-8,7 ±	0,11	1,8	1,2	(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	-8,9 ±	0,18	1,2	0,82	(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)
TM-5	V:01/5.35	V196A	-8,9 ±	0,17	1,4	0,96	(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	-8,8 ±	0,06	1,4	0,98	(4)	-6,2 ±	0,03	0,69	1,4	(2)	not determined					not determined				
	V:05/5.39	Q200A	-8,9 ±	0,07	1,4	0,95	(10)	-5,4 ±	0,18	3,7	7,4	(6)	-4,8 ±	0,30	17	2,0	(3)	-3,6 ±	0,26	250	21	(3)
	V:05/5.39	Q200W	-8,7 ±	0,08	1,8	1,2	(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	-8,9 ±	0,17	1,4	1,0	(5)	-6,1 ±	0,33	0,71	1,4	(3)	not determined					not determined				
TM-6	VI:13/6.48	W252A	-9,1 ±	0,06	0,78	0,53	(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)
	VI:16/6.51	Y255A	-8,9 ±	0,11	1,1	0,77	(7)	-5,8 ±	0,26	1,7	3,3	(3)	not determined					not determined				
	VI:20/6.55	I259A	-8,7 ±	0,09	2,1	1,4	(7)	-5,6 ±	0,28	2,6	5,1	(3)	not determined					not determined				
	VI:20/6.55	I259W	-8,9 ±	0,06	1,3	0,91	(6)	-5,8 ±	0,22	1,5	2,9	(3)	not determined					not determined				
	VI:23/6.58	D262N	-8,2 ±	0,04	5,8	4,0	(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	-8,7 ±	0,13	1,8	1,2	(18)	-4,5 ±	0,28	31	61	(7)	-4,3 ±	0,17	53	6,2	(3)	-3,8 ±	0,17	155	13	(4)
	VII:02/7.35	I284A	-8,6 ±	0,05	2,3	1,6	(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	-7,9 ±	0,17	14	9,6	(10)	-5,8 ±	0,30	1,7	3,4	(4)	not determined					not determined				

[a]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.

[b]F_{mut} is the ratio of the mutant and wt-CXCR4 potencies. Red: F_{mut} >25; orange: F_{mut} from 10-25; yellow: F_{mut} from 5-10; green: F_{mut} <0.5.

[c]The number of independent experiments is shown in parentheses (n).

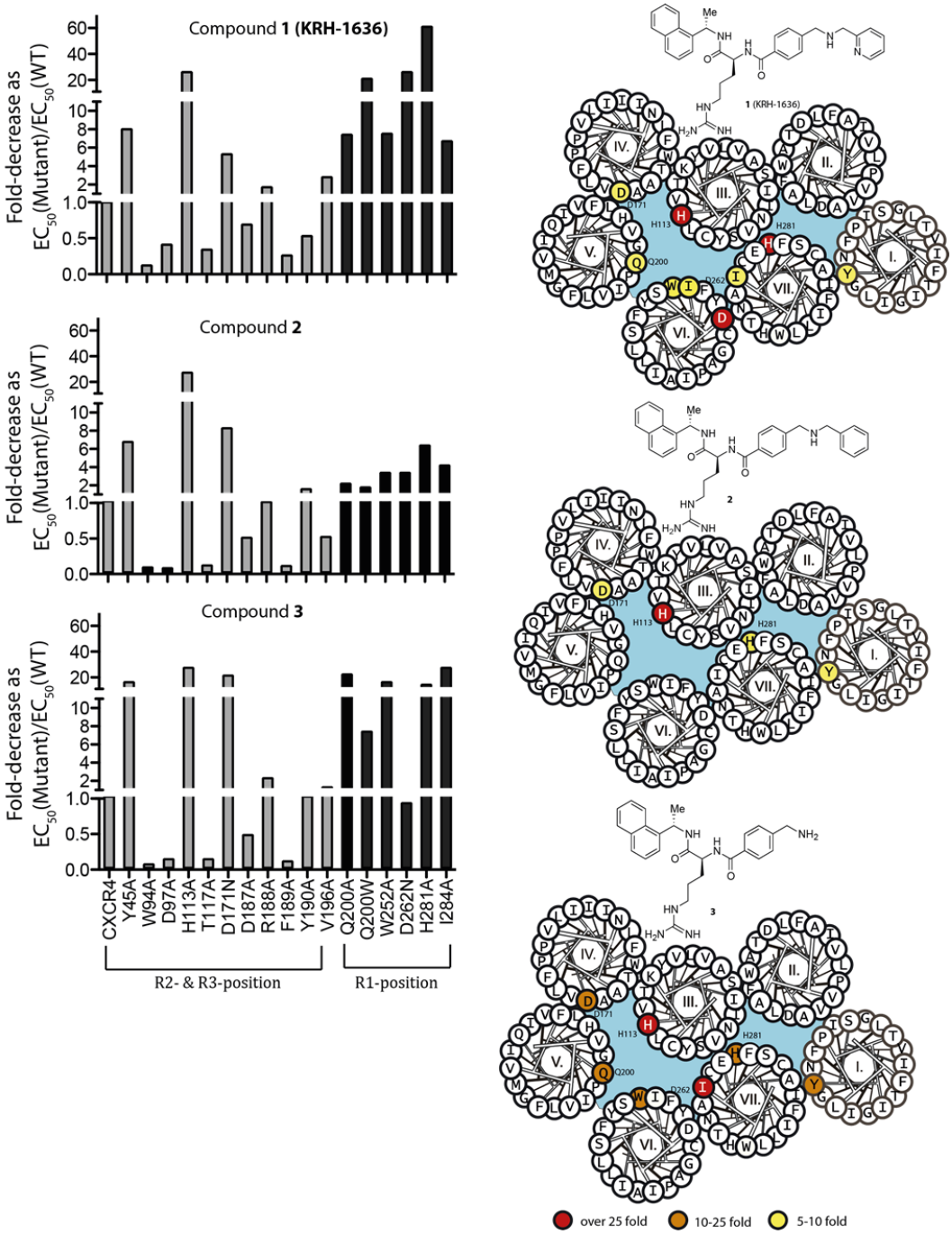


Figure 4. Left side. Bar-graph representations of fold-reduction in potency (F_{mut}) for compounds **1**–**3** for each mutant with respect to wt-CXCR4 (functional data, Table 2). Grey bars: mutants in TMs 1-4 and ECL-2; black bars: mutants in TMs 5-7. **Right side.** Residues identified to be important for the potency of compounds **1**–**3** (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.

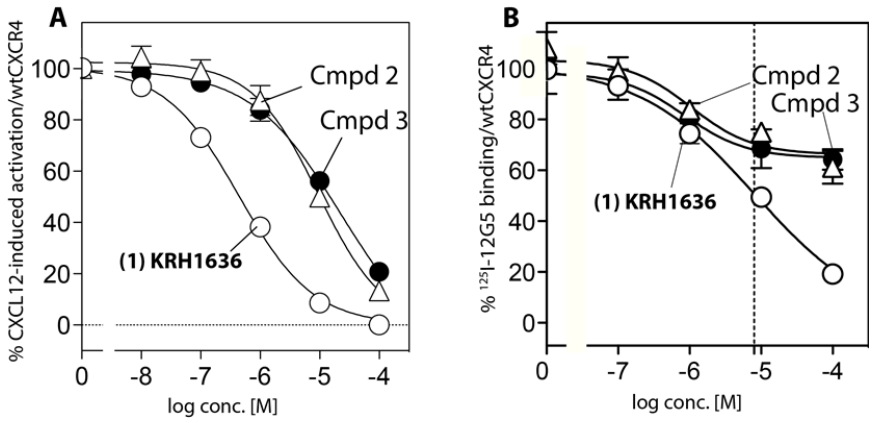


Figure 5. (A) Potency of KRH-1636 (**1**) and analogs **2** and **3**; (B) binding affinity of KRH-1636 (**1**) and analogs **2** and **3**.

agonistic potency of CXCL12 by 8-14 fold (Table 2). 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 the endogenous ligand CXCL12.[35, 42] Moreover, Trp^{94} and Tyr^{116}, together with the three aforementioned residues, have been implicated as part of "site two" in the "two-site model" for CXCL12 binding to CXCR4.[32, 42-44] Thus, for these residues there are certain limitations in using CXCL12 as a probe.

Displacement of radiolabeled monoclonal antibody (^{125}I-12G5). To address the issue of reduced activity of CXCL12 in certain mutants, a competitive binding assay using the radiolabeled monoclonal antibody 12G5 (^{125}I-12G5) was employed to probe ligand affinity (Table 3). A good correlation between anti-HIV potency and binding affinity measured as displacement of ^{125}I-12G5 has previously been demonstrated for the bicyclam antagonists,[34, 45] which represents an additional advantage of using 12G5 as radioligand. In addition to the five mutants mentioned above (W94A, D97A, Y116A, D187A, and E288A), the binding experiments were extended to also include four of the mutants that were identified as important for the potency of KRH-1636 in the functional assay (H113A, D171N, D262N, and H281A). All the nine mutants had similar affinities for 12G5 as the wt receptor (Table 3), meaning that they did not significantly affect the receptor cell surface expression. Compared to the strong antagonistic potency (EC_{50} = 0.50 μM) measured in the functional assay, KRH-1636 displayed lower affinity to wt-CXCR4 (IC_{50} = 8.0 μM) measured

against ^{125}I-12G5 (Figure 5B), suggesting that this compound binds allosterically with respect to 12G5. This finding clearly illustrates the differences between displacing ^{125}I-12G5 (Table 3) and inhibiting the action of CXCL12 (Tables 1 and 2). The two R^1 analogs **2** and **3** displayed even lower affinity (Figure 5B) than KRH-1636, and were therefore excluded from further binding experiments.

Collectively, the affinity of KRH-1636 determined against ^{125}I-12G5 was strongly (8- to >12-fold) negatively affected by the TM-3 mutants H113A and Y116A, the TM-4 mutant D171N, the TM-5 mutant D262N, and the TM-7 mutants H281A and E288A (Table 3). Moreover, KRH-1636 could not displace 12G5 in the Y116A-mutant. Overall the binding data show that His^{113}, Tyr^{116}, and Asp^{171} at the interface between TM-3 and TM-4, and Asp^{262}, His^{281}, and Glu^{288} at the interface between TM-6 and TM-7 in the main binding pocket are important for the displacement of 12G5 by KRH-1636.

To summarize the functional data (Tables 1 and 2) and the binding data (Table 3), Tyr^{45} in TM-1, His^{113} and Tyr^{116} in TM-3, Asp^{171} in TM-4, Gln^{200} in TM-5, Asp^{262} in TM-6, and His^{281} and Glu^{288} in TM-7 are important for the binding and function of KRH-1636.

Table 3. Affinity of KRH-1636 (**1**) on wt-CXCR4 and nine mutant receptors measured as displacement of ^{125}I-labeled antibody 12G5 in transiently transfected COS-7 cells (binding assay).

	Residue		12G5					KRH-1636 (1)			
			IC_{50} ± SEM *(log)*	IC_{50} *(nM)*	F_{mut}	BMax SEM	(n)[a]	IC_{50} SEM *(log)*	IC_{50} *(μM)*	F_{mut}	(n)
wt		**CXCR4**	-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	0,794	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)

[a]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.

[b]F_{mut} is the ratio of the mutant and wt-CXCR4 affinities.

[c]The number of independent experiments is shown in parentheses (n).

Molecular docking and derived binding model. *Binding model: key interactions.* In order to rationalize the experimental data, KRH-1636 was docked to the X-ray structure of CXCR4 (PDB code 3OE0)[32] using Schrödinger's induced-fit docking protocol,[46] 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^{113} ($N^{\delta 1}$), which was shown to be highly important for both the affinity and potency of KRH-1636. 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^1 side chain) of KRH-1636 is globally oriented towards the TM-6 and TM-7 region, and adopts a bending conformation around the secondary amino group, where the pyridine and phenyl rings lie in almost parallel planes. His^{281} in TM-7 forms an aromatic π-π stacking interaction with the pyridine ring, and Asp^{262} in TM-6 is involved in a bidentate interaction with the charged secondary amine in the R^1 side chain of KRH-1636. The guanidino group of Arg (R^2-position) forms polar interactions with His^{113} and Thr^{117} (TM-3) and Asp^{171} (TM-4). The aromatic naphthyl ring (R^3-position) is embedded in a hydrophobic pocket formed by the aromatic side-chains of Tyr^{45} and Trp^{94}, and further stabilized by the Tyr^{116} side-chain located directly below Trp^{94}. Interestingly, previous studies indicate that the three aromatic residues Tyr^{45}, Trp^{94} and Tyr^{116} of CXCR4, which are also found in the CCR5 chemokine receptor as Tyr^{37}, Trp^{86}, and Tyr^{108}, respectively, might form a hydrophobic pocket that is involved in the binding of small-molecule inhibitors.[47, 48] Thus, mutations of the aforementioned three residues could destabilize the hydrophobic network that interacts with the naphthyl ring of KRH-1636. Furthermore, the observation that Tyr^{116} is involved in a direct interaction with Glu^{288} (see Supporting Information, Figure S1) implies that the Y116A mutation can affect the overall geometry of the main binding crevice, specifically the orientations of TM-3 and TM-7. Thus, the Y116A mutation may indirectly affect the potential interactions between the ligand and His^{113}, Thr^{117}, His^{281}, and Glu^{288}.

Binding model: indirect effects. The docking pose for KRH-1636 also reveals the involvement of Glu^{288}, which is a key residue as shown in the binding assay (Table 3). Glu^{288} has been consistently implicated in binding modes of several different classes of CXCR4 antagonists, *i.e*. AMD3100, AMD3465, AMD11070, FC131, and also the CXCR4 co-

crystallized ligands IT1t and CVX15,[32, 33, 35-37] either through direct or water-mediated interactions with the ligand. Similarly, our binding model shows a water-mediated interaction between Glu^{288} and the ligand backbone (Figure 6).

Substitution of Gln^{200} with Ala and Trp (Q200A and Q200W) caused a 7- and 21-fold reduction in potency of KRH-1636 with respect to wt-CXCR4; the involvement of this residue was also seen in previous binding mode studies of the monocyclam antagonist AMD3465.[36] In our binding model for KRH-1636 (Figure 6), an intra-receptor H-bond between Gln^{200} and Asp^{262} was identified, suggesting that substitution of Gln^{200} can affect the interaction between Asp^{262} and the ligand.

The model further shows an interaction between the guanidino group of Arg^{188} and the ligand backbone (Figure 6), which was also previously identified in the binding modes of the peptide antagonists FC131 and CVX15.[22, 32, 49] However, the R188A mutation did not have any effect on the antagonistic potency of KRH-1636 or the two analogs (Table 2); interestingly, the R188A-mutant was also reported to have no effect on the potency of FC131.[33] A possible explanation is that both KRH-1636 and FC131 are doubly positively charged, which results in electrostatic repulsion by Arg^{188}. Substitution of Arg^{188} 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 KRH-1636 and FC131.

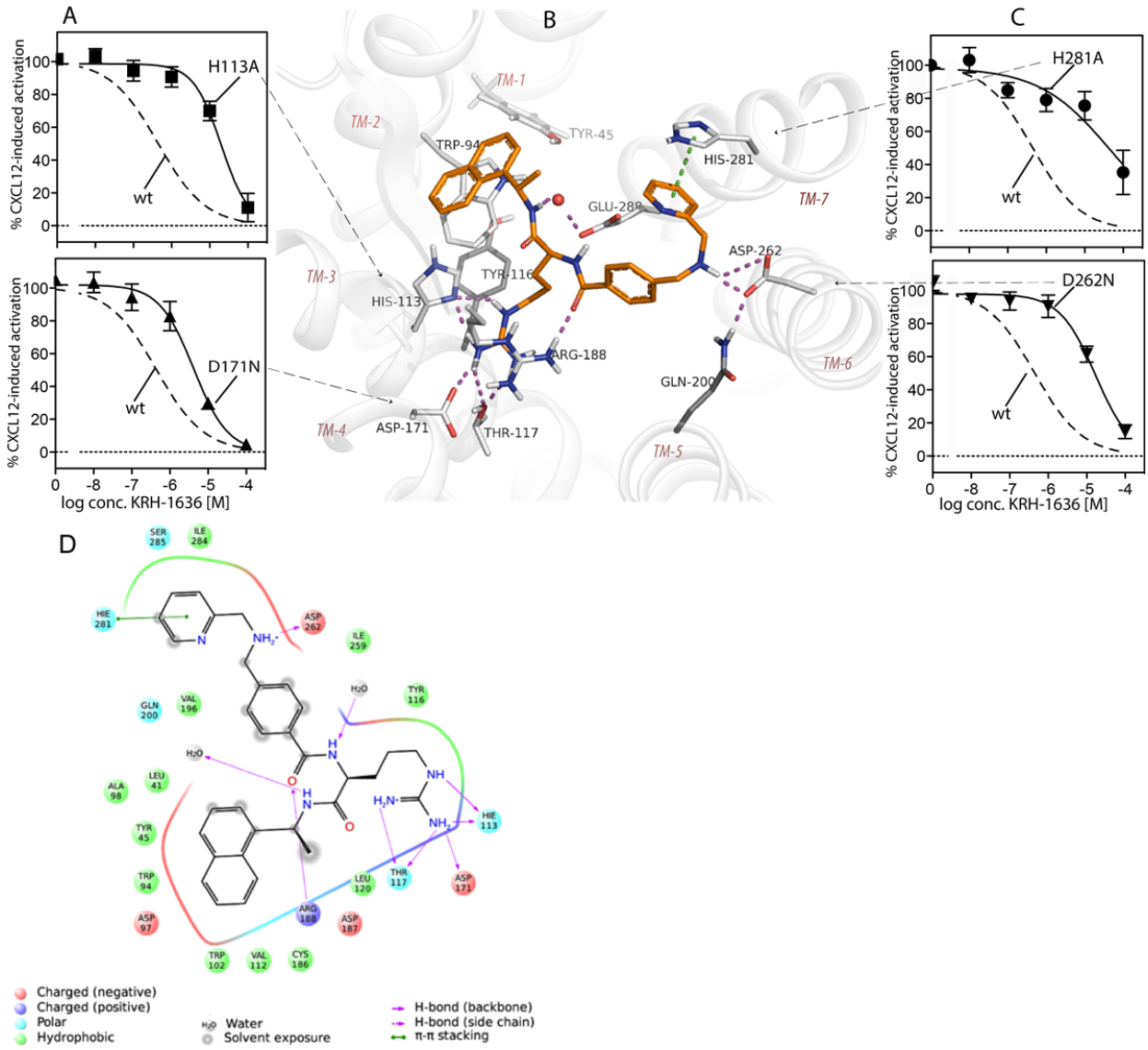


Figure 6. (A) Effects of the H113A and D171N mutations on the inhibition of CXCL12-induced PI turnover by KRH-1636. (B) Proposed binding model for KRH-1636, showing the key interacting residues in CXCR4; the ligand is shown in orange sticks, the receptor in grey-white sticks and ribbons, and the water molecule as a red sphere; H-bonds are shown in magenta, and π-π interactions in green. (C) Effects of the H281A and D262N mutations on the inhibition of CXCL12-induced PI turnover by KRH-1636. (D) 2D-representation of the docking pose for KRH-1636.

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,[32, 42] 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^{97} constitutes a gate-entry position that controls ligand access to the major binding pocket;[32, 33] thus, a pre-interaction with Asp^{97} 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 TM-2 residues (Trp^{94} and Asp^{97}) have a less distinctive involvement in the binding of these ligands to the receptor. First, mutation of Trp^{94} and Asp^{97} 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^{97} is removed in the D97A mutant, which allows easier access of the ligands to the TM binding pocket. Third, the involvement of the Trp^{94} in the hydrophobic pocket surrounding the naphthalene ring (R^3-position) in KRH-1636 (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 the hydrophobic residues Trp^{252}, Ile^{259}, and Ile^{284} reduced the potency of KRH-1636, but these residues are not directly involved in ligand binding, meaning that their effects are secondary, *i.e*. implemented through local or global changes in TM-6 and TM-7.

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.[50, 51] The X-ray structures of CXCR4 in complex with the most potent compound IT1t[32] later showed that IT1t exhibits a unique binding mode by occupying the so-called "minor binding pocket" between TMs 1-3 and -7, *i.e.* without interacting with TMs 4-6. Consequently, the binding mode of IT1t differs from other reported binding modes of CXCR4 antagonists (see below), and is distinct from the present binding mode for KRH-1636.

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[52, 53] and KRH-1636[53] 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 any experimentally verified ligand-receptor interactions. Moreover, when the X-ray structure of CXCR4 was revealed in 2010[32] it showed that bovine rhodopsin was a poor template for CXCR4,[54] 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.[33] The binding model for FC131 was generally consistent with the binding mode for the Arg^1-Arg^2-1-Nal^3 fragment of the 16-mer peptide antagonist CVX15 in complex with CXCR4 (PDB code 3OE0),[32] *i.e.* the Arg^1 side-chain in FC131 projected towards ECL-2 of the receptor to interact with Asp^{187}, the Arg^2 guanidino group interacted with His^{113}, Thr^{117}, and Asp^{171} in TMs 3-4, while the aromatic 2-Nal^3 side-chain was accommodated in a hydrophobic pocket around TM-5.

The reported receptor interactions for the Arg^2 guanidino group in FC131 (His^{113}, Thr^{117}, and Asp^{171}) are in agreement with the interactions of the guanidino group (R^2 position) in KRH-1636 (Figure 6), and structural comparison of the reported binding mode for FC131[33] and the present binding mode for KRH-1636 (Supporting Information, Figure S2) shows that L-Arg^2 of FC131 overlaps relatively well with the central arginine residue of KRH-1636. However, the orientation of the naphthyl groups in FC131 and KRH-1636 are different; the 2-naphthyl group in FC131 is oriented towards TM-5, while the 1-naphthyl group in KRH-1636 is facing TM-2.

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^{171}, Asp^{262}, and Glu^{288}.[34, 35] Based on the fact that AMD3100 and KRH-1636 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.[27, 55] As Asp^{171}, Asp^{262}, and indirectly

Glu^{288} are also involved in binding of KRH-1636 (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^{113} in TM-3 and His^{281} in TM-7,[36, 37] which were also identified as important for KRH-1636 in the present study. The reduced potency observed for KRH-1636 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,[36] which contains the same (pyridin-2-ylmethyl)amino moiety as KRH-1636. Thus, in addition to the key interactions with His^{113} and Asp^{171} (R^2 guanidino group) and Asp^{262} (R^1 amino group), the pyridine ring in the R^1 side chain of KRH-1636 seems to pick up an additional interaction with His^{281} in TM-7, in the same way as AMD3465.

The interaction pattern for the non-cyclam AMD11070 (Figure 1) is more unclear, but it has been suggested that it binds in a similar fashion as AMD3100 and AMD3465; however, Asp^{97} appeared to have a unique importance for the binding of this analog.[37]

CONCLUSIONS

The present study describes the binding of the prototype tripeptidomimetic antagonist KRH-1636 to CXCR4 based on SAR studies, receptor mapping, and molecular docking. This ternary approach resulted in the identification of a binding model for KRH-1636 in which the secondary amino group is anchored by interaction with Asp^{262} in TM-6, with the distal pyridine ring further involved in a π-π interaction with His^{281} in TM-7; the Arg-guanidino group engages in polar interactions mainly with His^{113} in TM-3 and Asp^{171} in TM-4, while the naphthyl group is tightly packed in a hydrophobic network formed by residues Trp^{94}, Tyr^{45} and Tyr^{116}. Our findings contradict earlier suggestions that KRH-1636 mimics the binding of the Arg^1-Arg^1-2-Nal^3 fragment in the cyclopentapeptide antagonist FC131. Instead, we propose a "hybrid" binding mode for KRH-1636 where the guanidino group of the central arginine residue (R^2 side chain) overlaps with the binding mode of the Arg^2 residue in the peptide antagonist FC131 and CVX15, while the (pyridin-2-ylmethyl)amino moiety (R^1 side chain) overlaps with the binding mode for the same moiety in the monocyclam AMD-3465. This experimentally verified binding model for KRH-1636

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. ^{1}H and ^{13}C NMR spectra were recorded on a 400 MHz Varian spectrometer. Chemical shifts are expressed in ppm relative methanol (^{1}H 3.31 ppm, ^{13}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_{18} 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_{18} 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_{18} 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 UPLC H-Class equipped with a Waters ACQUITY UPLC BEH C_{18} 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 at least 95% pure 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, KRH-1636).** The title compound was prepared following an earlier reported procedure[28]. Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material. ^{1}H NMR (400 MHz, CD_3OD) 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); ^{13}C NMR (100 MHz, CD_3OD) 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* $[M + H]^+$ calcd for

$C_{32}H_{38}N_7O_2$: 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 (3). The title compound was prepared adopting a literature procedure[28] where *N*-Boc 4-(methylamino)benzoic acid was used instead of **9** (Scheme 1). Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material. ^{1}H NMR (400 MHz, CD_3OD) 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); ^{13}C NMR (100 MHz, CD_3OD) 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 (2). A solution of **3** (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 µL, 0.09 mmol) was added slowly, and the reaction mixture was allowed to stir at room temperature 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. ^{1}H NMR (400 MHz, CD_3OD) δ 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 Hz, 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 (4).** The title compound was prepared adopting a literature procedure[28] where 4-methylbenzoic acid was used instead of **9** (Scheme1). Purification by RP-HPLC afforded after lyophilization the title compound as a white fluffy material. ^{1}H NMR (400 MHz, CD_3OD) δ 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}C NMR (100 MHz, $CDCl_3$) 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* [M + H]$^+$ calcd for $C_{26}H_{32}N_5O_2$: 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 (5). The title compound was prepared adopting a literature procedure[28] 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}H NMR (400 MHz, CD_3OD) δ 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}C NMR (100 MHz, CD_3OD) δ 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* [M + H]$^+$ calcd for $C_{32}H_{37}N_6O_3$: 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 (6). The title compound was prepared adopting a literature procedure[28] where ethylamine was used instead of (*S*)-(–)-1-(1-naphthyl)ethylamine. Purification by RP-HPLC afforded after lyophilization the title

compound as a white fluffy material. ^{1}H NMR (400 MHz, CD_3OD) 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); ^{13}C NMR (100 MHz, CD_3OD) 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]$^+$ calcd for $C_{22}H_{32}N_7O_2$: 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, U.S.), and the promiscuous chimeric G protein $G_{\Delta 6qi4myr}$ (abbreviated Gqi4myr) was kindly provided by Evi Kostenis (University of Bonn, Germany)[56]. Myo-[^{3}H]-Inositol (PT6-271), SPA beads, and Bolton-Hunter reagent for 12G5 iodination (in-house procedure)[35] 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, U.S.). 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[57] or the QuikChange™ technique (Agilent Technologies, CA, U.S.), using wt-CXCR4 as a template. All reactions were carried out using *Pfu* polymerase (Stratagene, CA, U.S.) 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_2 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.[58]

Phosphatidylinositol (PI) Turnover Assays. Transfection of COS-7 cells was carried out according to the abovementioned procedure. Shortly, 6×10^6 cells were transfected with 20 μg of receptor cDNA along with 30 μg of the promiscuous chimeric G protein, Gαqi4myr, which turns the Gαi-coupled signal, the most common pathway for endogenous chemokine receptors, into the Gαq pathway (phospholipase C activation measured as IP_3 turnover).[56] Hereby, two alternative assays that were found to show the same result, were carried out:

Anion-exchange separation of phosphoinositides. In the classic IP_3-assay which includes an anion-exchange chromatography step, one day after transfection COS-7 cells (1.5×10^5 cells/well) were incubated for 24 h with 2 μCi of myo[^{3}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 [^{3}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), γ-radiation was counted in a Beckman Coulter counter LS6500.

Scintillation proximity assay. In the scintillation proximity assay PI-SPA, the day after transfection COS-7 cells (35 000 cells/well) were incubated for 24 h with 0.5 μCi/mL myo[^{3}H]inositol in 100 μL 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 μL of reaction solution with LiCl, and 50 μL of ice cold formic acid. Next, the generated [^{3}H]inositol phosphates were quantified by radioactivity measurement; in brief, 20 μL of the formic acid cell extracts were transferred to a white 96-well plate, and an amount of 80 μL of 1:8 diluted YSi poly-D-lysine coated beads (SPA-beads, PerkinElmer, MA, U.S.) was added. Plates were sealed, agitated at maximum speed for at least 30 min, and centrifuged (5 min at 1500 rpm) before γ-radiation was counted in a Packard Top Count NXT counter. In both assays, determinations were made in duplicate.

Binding experiments. 6×10^6 COS-7 cells were transfected with 40 μg 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 °C using 10–15 pM ^{125}I-12G5 plus unlabeled ligand in 0.2 ml of 50 mM Hepes buffer, pH 7.4, supplemented with 1 mM $CaCl_2$, 5 mM $MgCl_2$, and 0.5% (w/v) bovine serum albumin. Following incubation, cells were washed quickly two times in 4 °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 μM unlabeled 12G5. Determinations were made in duplicates.

Statistical calculations. The apparent EC_{50} and IC_{50} values were determined by nonlinear regression using the GraphPad Prism 4 software (GraphPad Software, San Diego, CA).

Computational modeling procedure. All molecular modeling calculations were performed using the Schrödinger software suite 2012.[59] Default settings were used, if not specified otherwise. The crystal structure of CXCR4 in complex with CVX15 (PDB-code: 3OE0)[32] was first prepared and optimized with the Protein Preparation Wizard workflow[60] using the same procedure as previously described.[22] The automatically assigned tautomeric and protonation states (pH 7.4) for ionizable residues were kept, except for Asp^{262}, 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[46] with a docking box of (26 Å)^3 centered on Asp^{187}. 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 $N^{\delta 1}$ atom of His^{113}, which resulted in the identification of the binding mode for KRH-1636 that is shown in Figure 6.

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. Ward, S. G.; Westwick, J. Chemokines: understanding their role in T-lymphocyte biology. *Biochem. J.* **1998,** 333, 457-470.

4. Horuk, R. Chemokine receptors. *Cytokine Growth Factor Rev.* **2001,** 12, 313-335.

5. 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, 829-33.

6. Oberlin, E.; Amara, A.; Bachelerie, 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.

7. 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, 9448-53.

8. 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.

9. Busillo, J. M.; Benovic, J. L. Regulation of CXCR4 signaling. *Biochim. Biophys. Acta* **2007,** 1768, 952-63.

10. 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, 2305-9.

11. 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.

12. 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.

13. 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, 977-94.

14. 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.

15. 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.

16. De Clercq, E.; Yamamoto, N.; Pauwels, R.; Baba, M.; Schols, D.; Nakashima, H.; Balzarini, J.; Debyser, Z.; Murrer, B. A.; Schwartz, D.; et al. 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. U. S. A.* **1992,** 89, 5286-90.

17. 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.

18. 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, 3251-3.

19. 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.

20. 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, 3280-9.

21. 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, 192-8.

22. 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.

23. Mungalpara, J.; Zachariassen, Z. G.; Thiele, S.; Rosenkilde, M. M.; Vabeno, J. Structure-activity relationship studies of the aromatic positions in cyclopentapeptide CXCR4 antagonists. *Org. Biomol. Chem.* **2013**.

24. 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.

25. 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(heterocyclylmethyl)amine compounds as chemokine receptor CXCR4 antagonists. WO2004024697A1, 2004.

26. 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.

27. 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, 4185-90.

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

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

30. Debnath, B.; Xu, S.; Grande, F.; Garofalo, A.; Neamati, N. Small Molecule Inhibitors of CXCR4. *Theranostics* **2013,** 3, 47.

31. Våbeno, J.; Nikiforovich, G. V.; Marshall, G. R. A minimalistic 3D pharmacophore model for cyclopentapeptide CXCR4 antagonists. *Biopolymers* **2006,** 84, 459-71.

32. 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, 1066-71.

33. Thiele, S.; Mungalpara, J.; Steen, A.; Rosenkilde, M. M.; Våbenø, J. Determination of the binding mode for the cyclopentapeptide CXCR4 antagonist FC131 using a dual approach of ligand modifications and receptor mutagenesis. *Submitted*.

34. 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.

35. 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.

36. 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.

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

38. 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.

39. 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.; et al. Crystal structure of interleukin 8: symbiosis of NMR and crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **1991,** 88, 502-6.

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

41. 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.

42. 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.

43. 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.

44. 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.

45. Bridger, G. J.; Skerlj, R. T. *Adv. Antiviral Drug* **1999**, 161-229.

46. 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.

47. 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.

48. 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.

49. 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.

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

51. 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, 7915-20.

52. 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.

53. 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.

54. Planesas, J. M.; Pérez-Nueno, V. I.; Borrell, J. I.; Teixidó, J. Impact of the CXCR4 structure on docking-based virtual screening of HIV entry inhibitors. *J. Mol. Graphics Modell.* **2012**.

55. Leonard, J. T.; Roy, K. The HIV entry inhibitors revisited. *Curr. Med. Chem.* **2006,** 13, 911-34.

56. Kostenis, E.; Zeng, F. Y.; Wess, J. Functional characterization of a series of mutant G protein alphaq subunits displaying promiscuous receptor coupling properties. *J. Biol. Chem.* **1998,** 273, 17886-92.

57. 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.

58. Rosenkilde, M. M.; Cahir, M.; Gether, U.; Hjorth, S. A.; Schwartz, T. W. Mutations along transmembrane segment II of the NK-1 receptor affect substance P competition with non-peptide antagonists but not substance P binding. *J. Biol. Chem.* **1994,** 269, 28160-28164.

59. Schrödinger. http://www.schrodinger.com/

60. 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.

SUPPORTING INFORMATION

Probing the molecular interactions between CXC chemokine receptor 4 (CXCR4) and the tripeptidomimetic antagonist KRH-1636

Zack G. Zachariassen,[a] Stefanie Thiele,[b] Bengt Erik Haug,[c] Mette M. Rosenkilde [b], and Jon Våbenø*[a]

[a] *Department of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Breivika, NO-9037 Tromsø, Norway. Fax: +47 77 64 61 51; Tel: +47 77 62 09 09; E-mail: jon.vabeno@uit.no.*

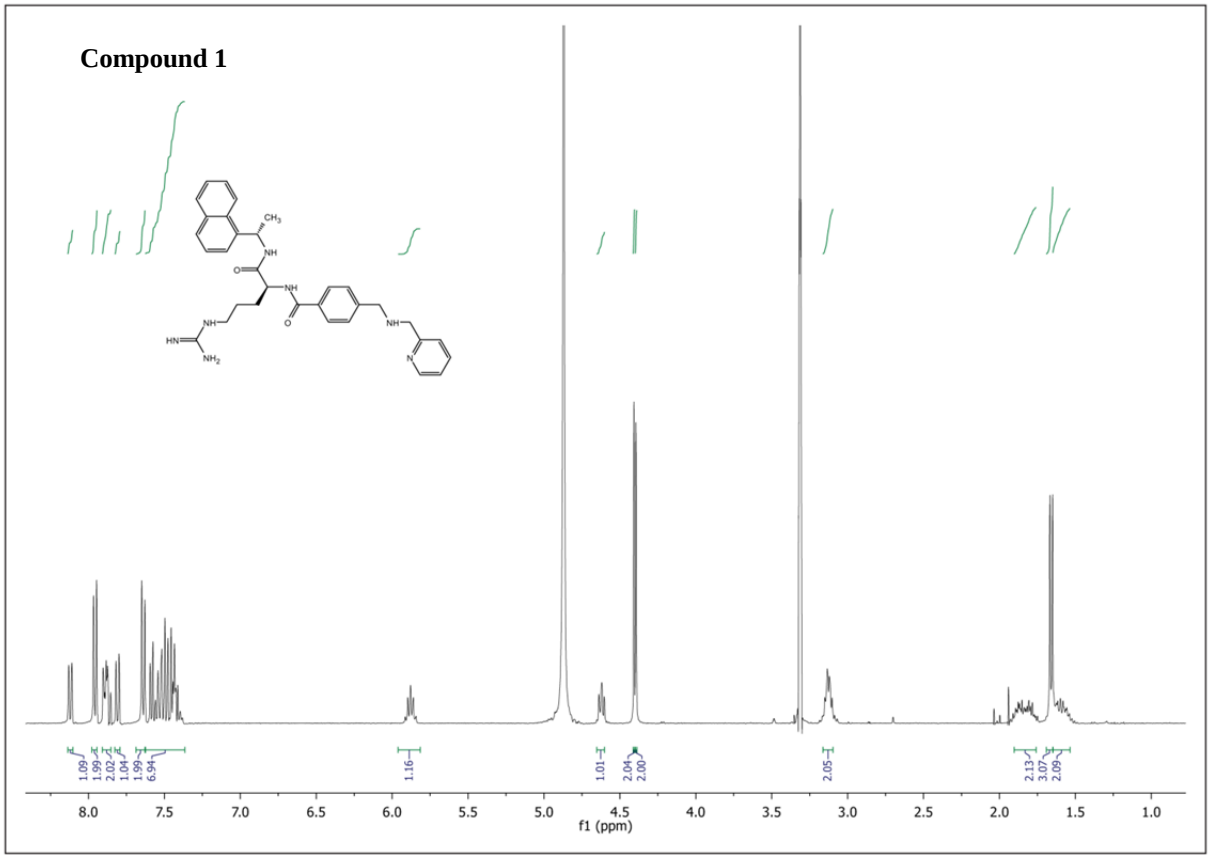
[b] *Laboratory for Molecular Pharmacology, Department of Neuroscience and Pharmacology, Faculty of Health and Medical Sciences, The Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen, Denmark.*

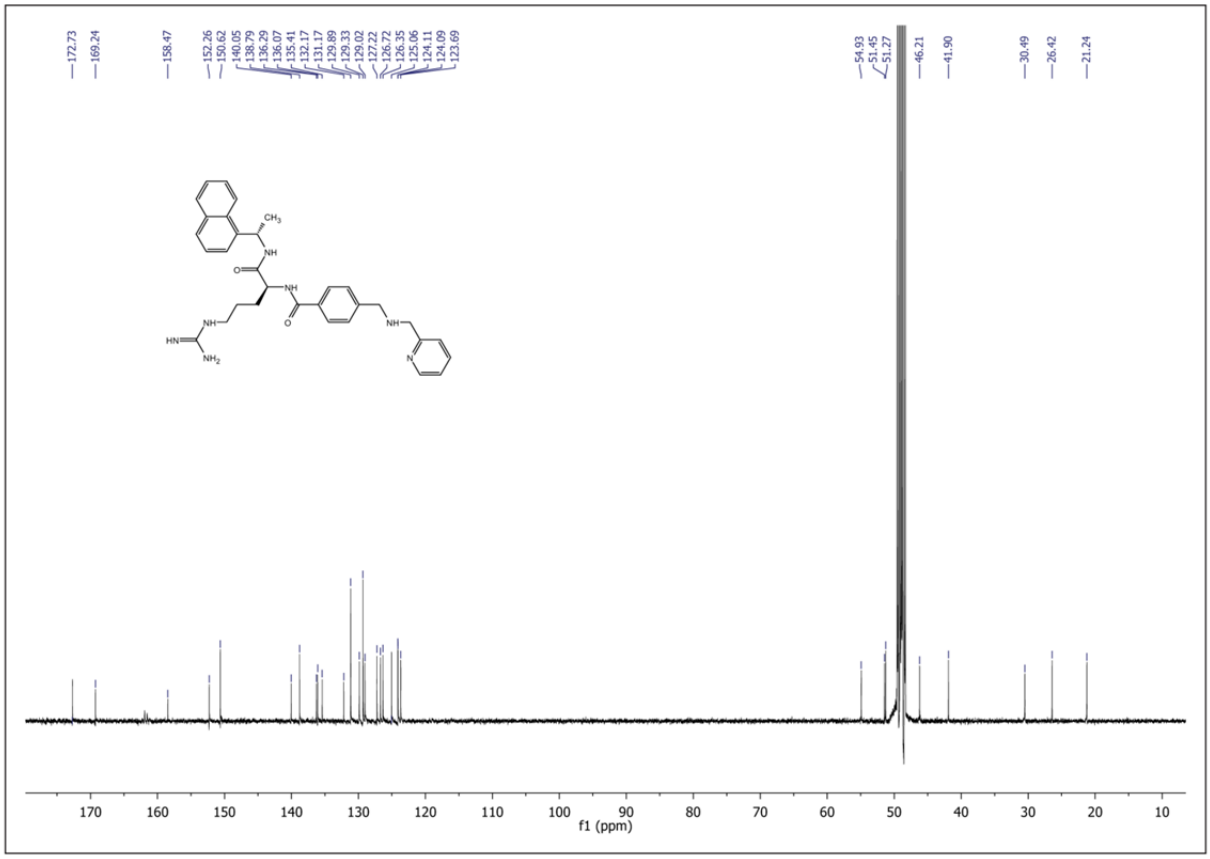
[c] *Department of Chemistry and Centre of Pharmacy, University of Bergen, Allégaten 41, NO-5007 Bergen, Norway.*

Table of contents

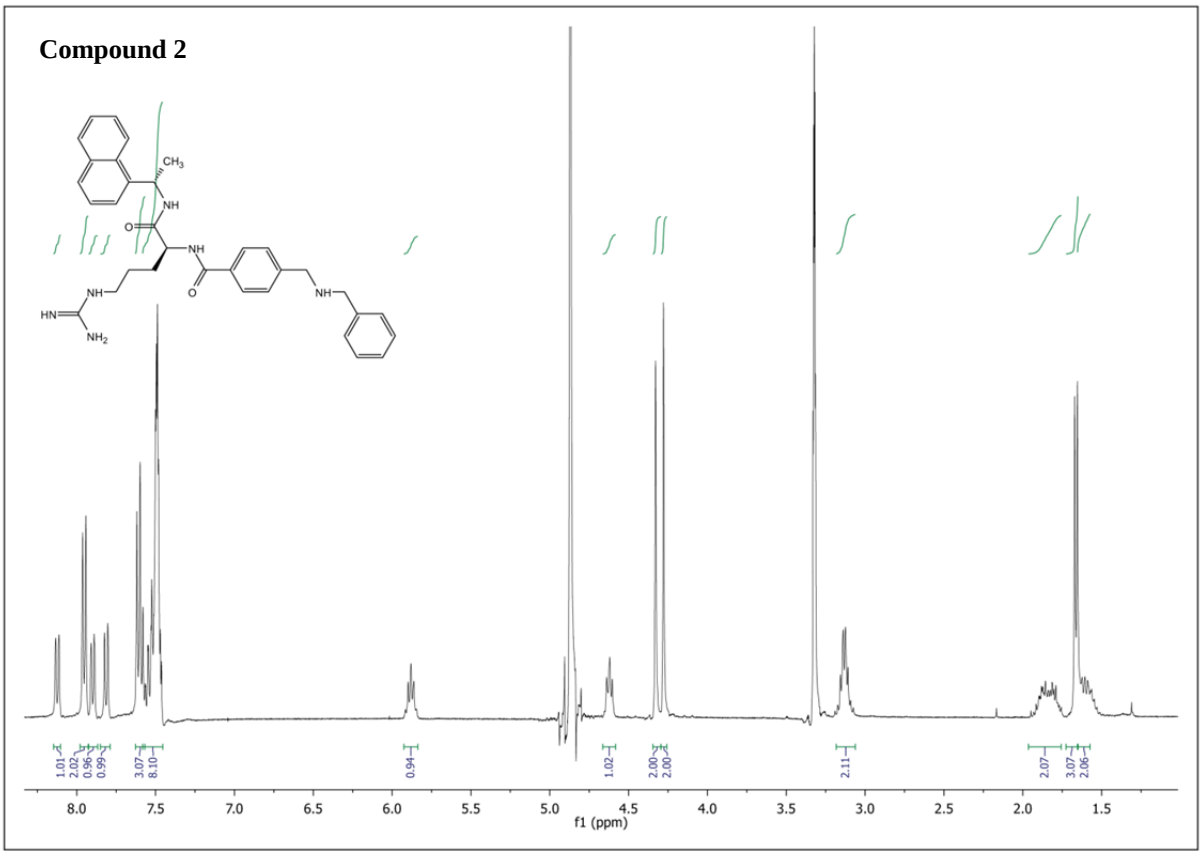
Corresponding Author
*Phone: +47 77 62 09 09. E-mail: jon.vabeno@uit.no.

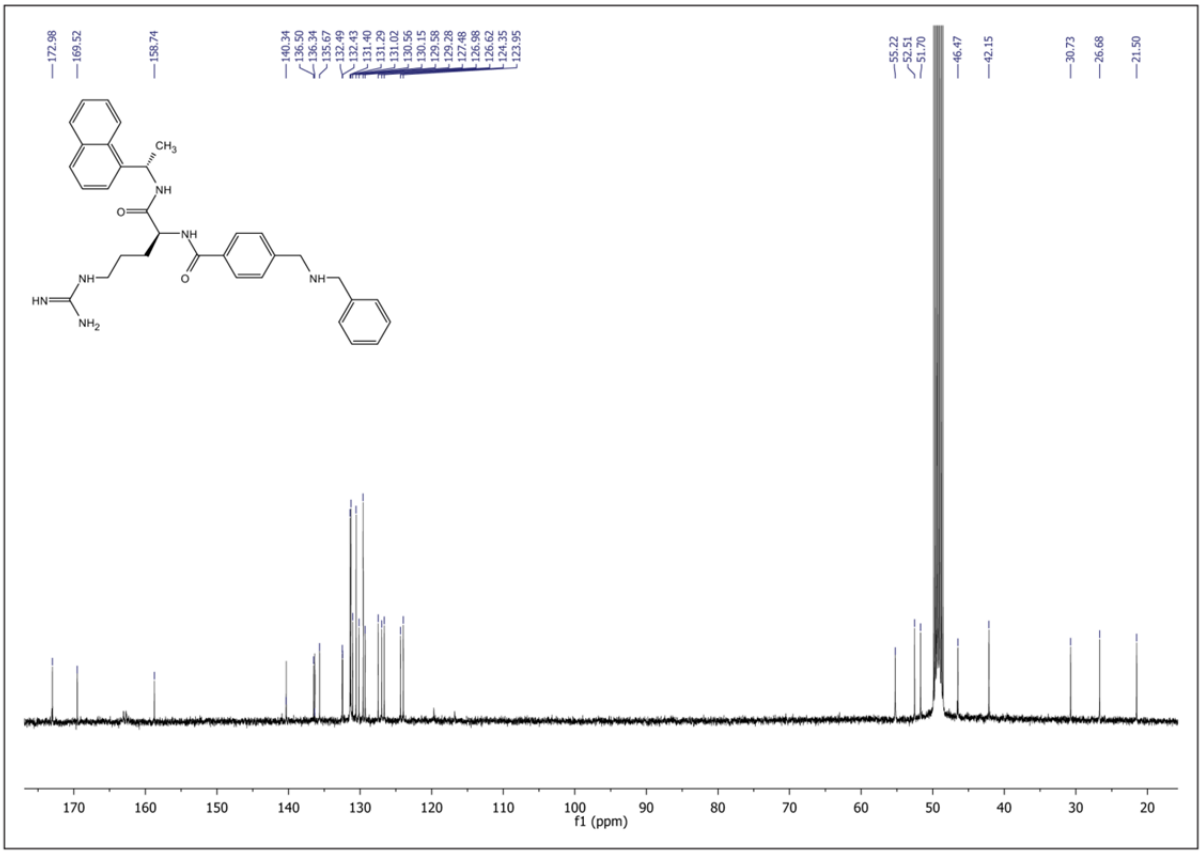
^{1}H and ^{13}C NMR spectra



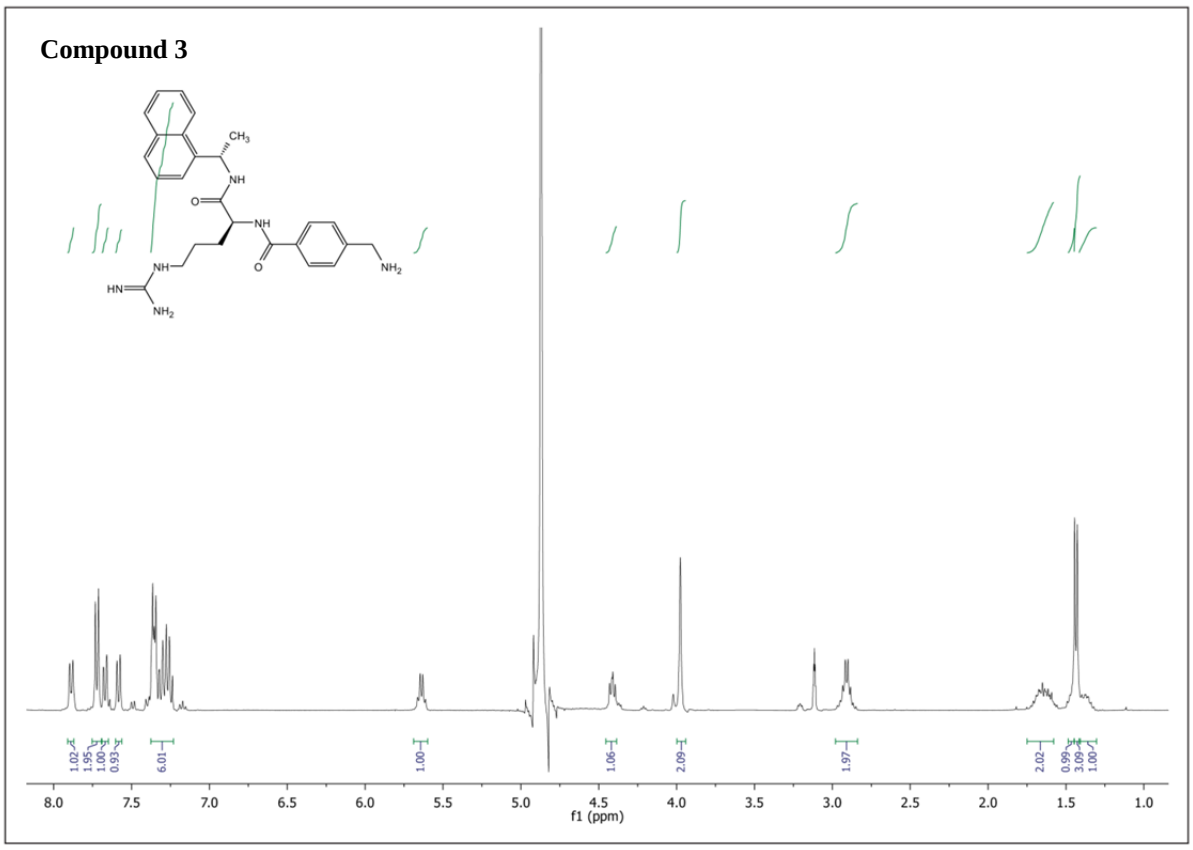


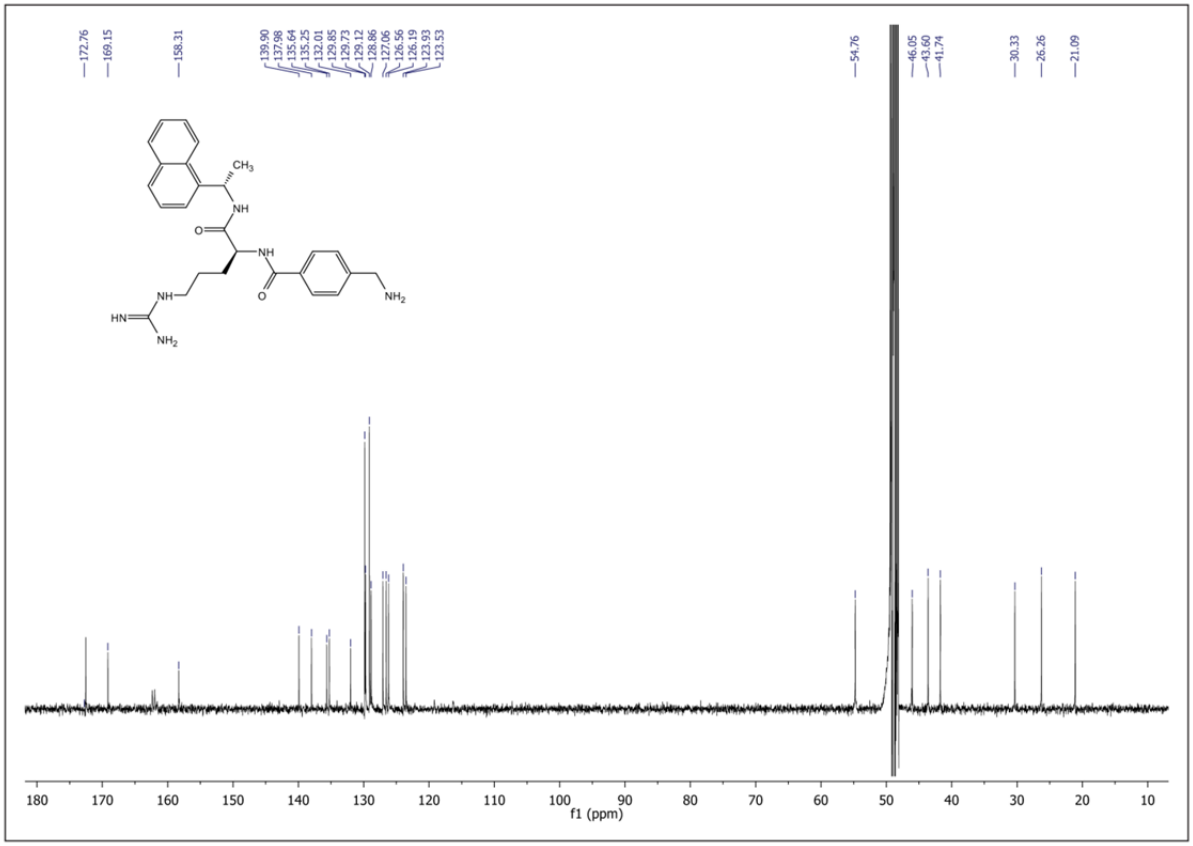
Compound 2


1.01
2.02
0.96
0.99
3.07
8.10
0.94
1.02
2.00
2.00
2.11
2.07
3.07
2.06
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
f1 (ppm)

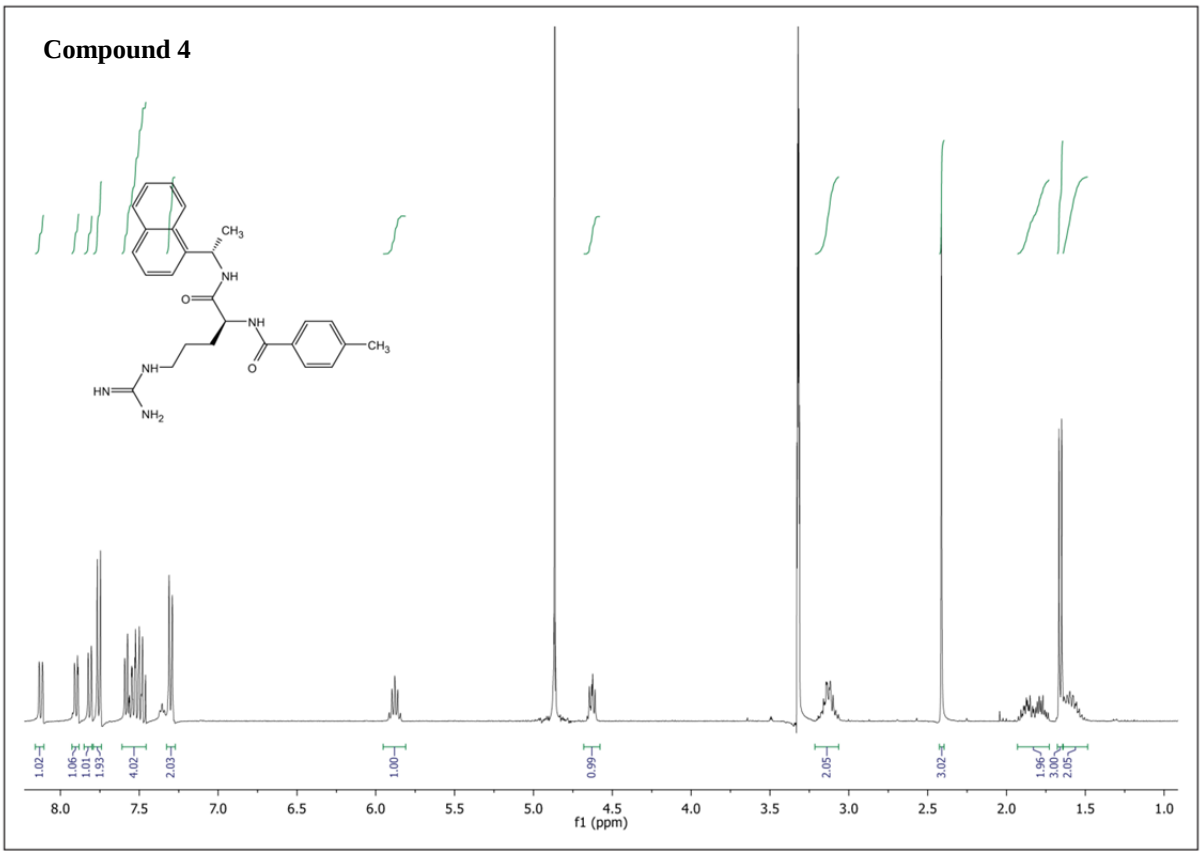

172.98
169.52
158.74
140.34
136.50
136.34
135.67
132.49
132.43
131.40
131.29
131.02
130.96
130.15
129.58
129.28
127.48
126.98
126.62
124.35
123.95
55.22
52.51
51.70
46.47
42.15
30.73
26.68
21.50
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
f1 (ppm)

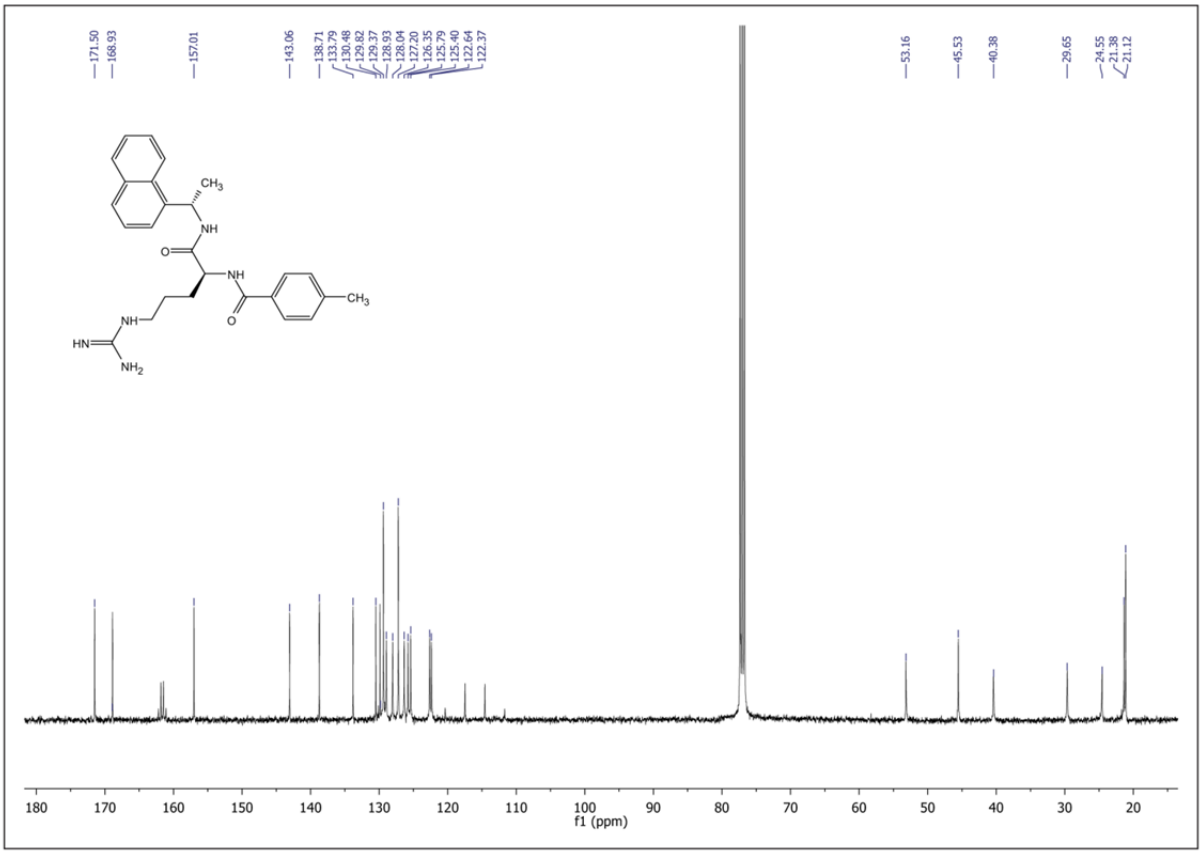
Compound 3


1.02
1.95
1.00
0.93
6.01
1.00
1.06
2.09
1.97
2.02
0.99
3.09
1.00
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)

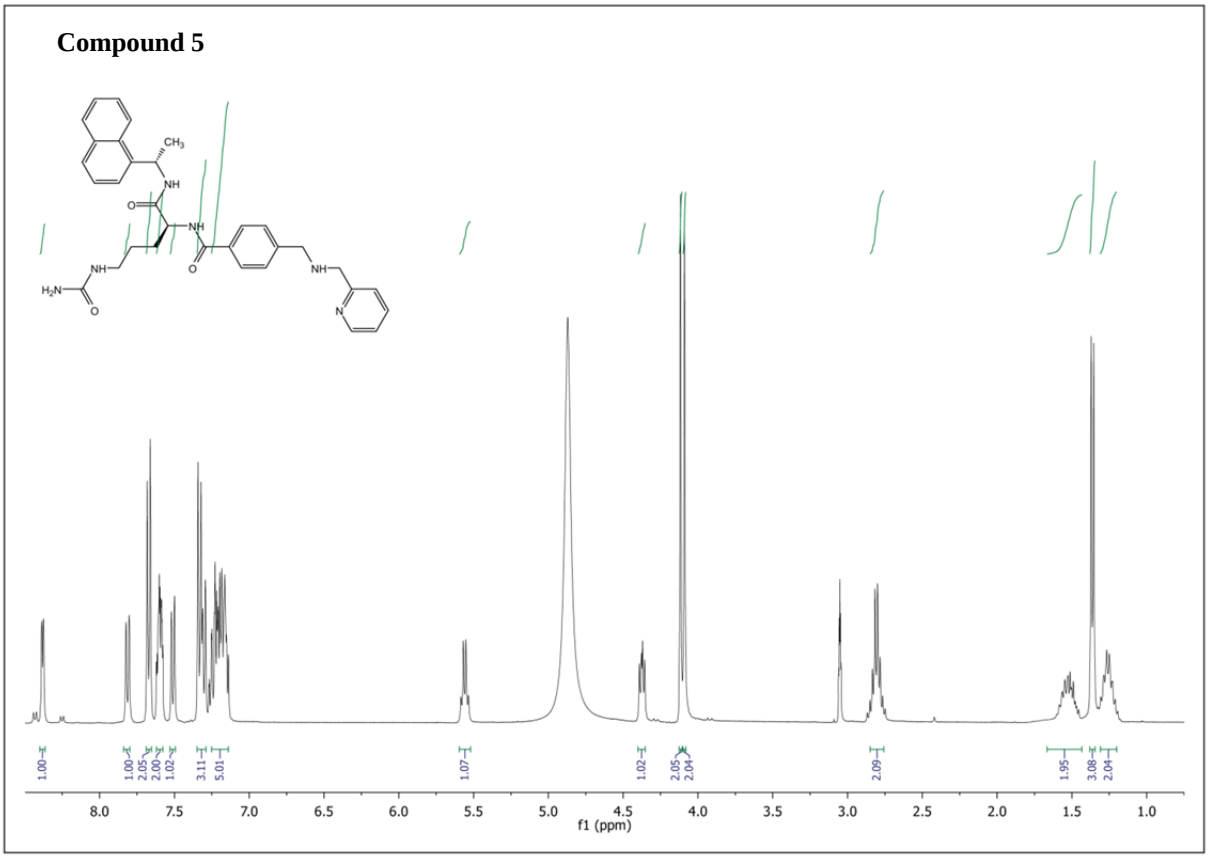

172.76
169.15
158.31
139.90
137.98
135.64
135.25
132.01
129.85
129.73
129.12
128.86
127.06
126.56
126.19
123.93
123.53
54.76
46.05
43.60
41.74
30.33
26.26
21.09
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)

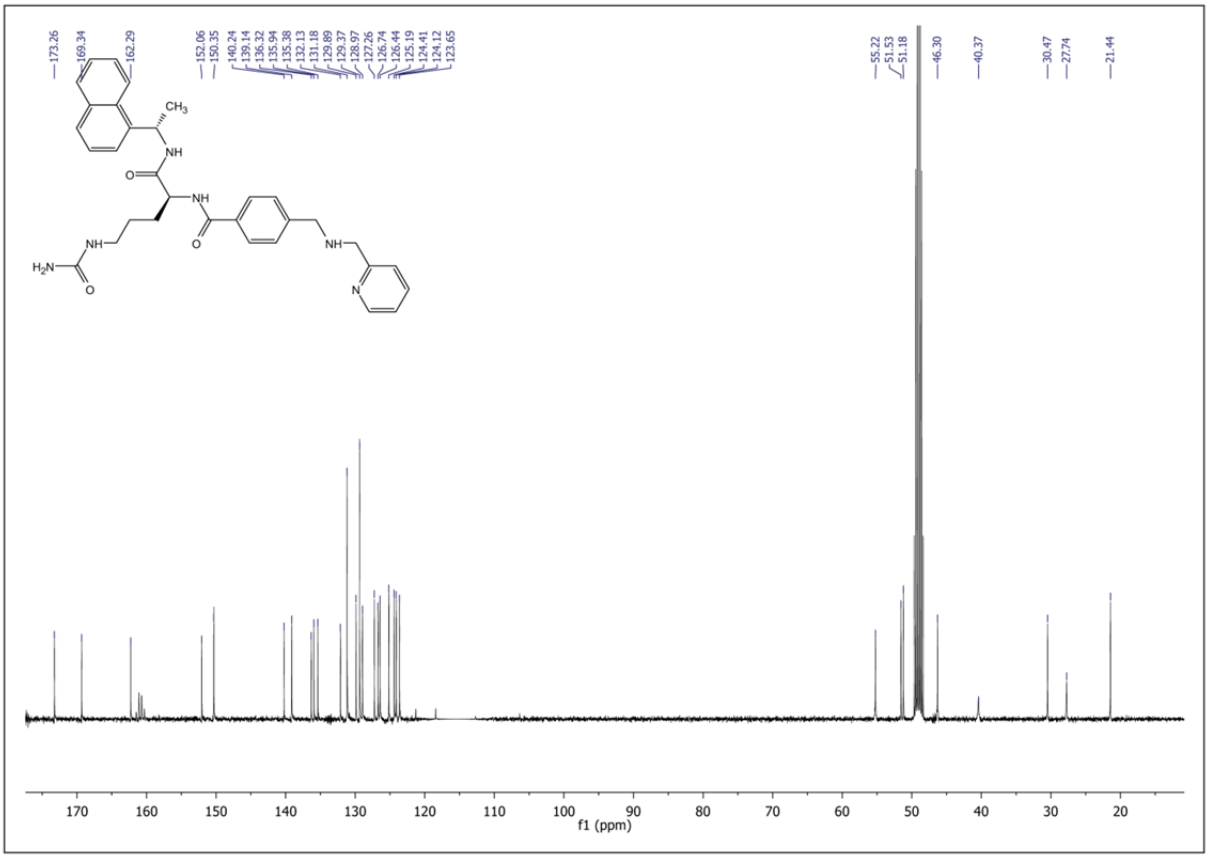
Compound 4


1.02
1.06
1.01
1.93
4.02
2.03
1.00
0.99
2.05
3.02
1.96
3.00
2.05
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)

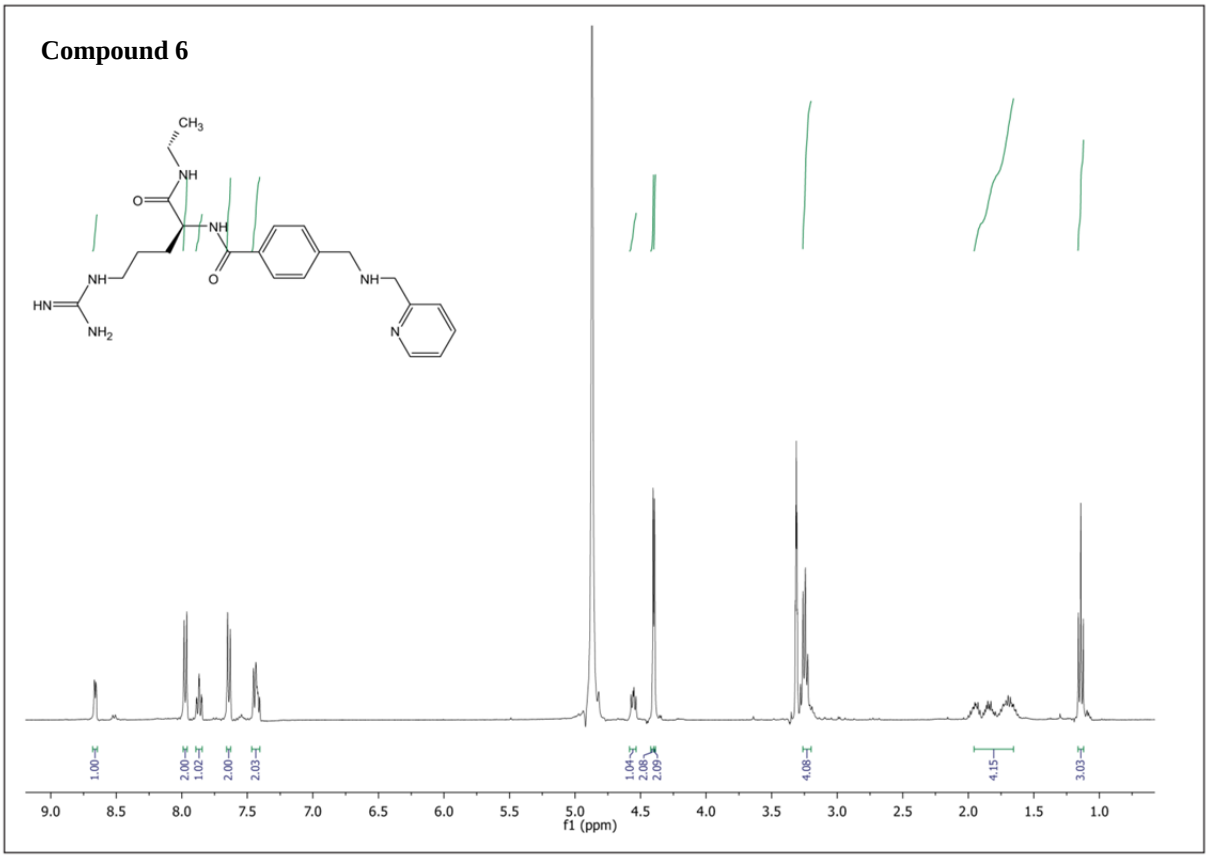

171.50
168.93
157.01
143.06
138.71
133.79
130.48
129.82
129.37
128.93
128.04
127.20
126.35
125.79
125.40
122.64
122.37
53.16
45.53
40.38
29.65
24.55
21.38
21.12
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
f1 (ppm)

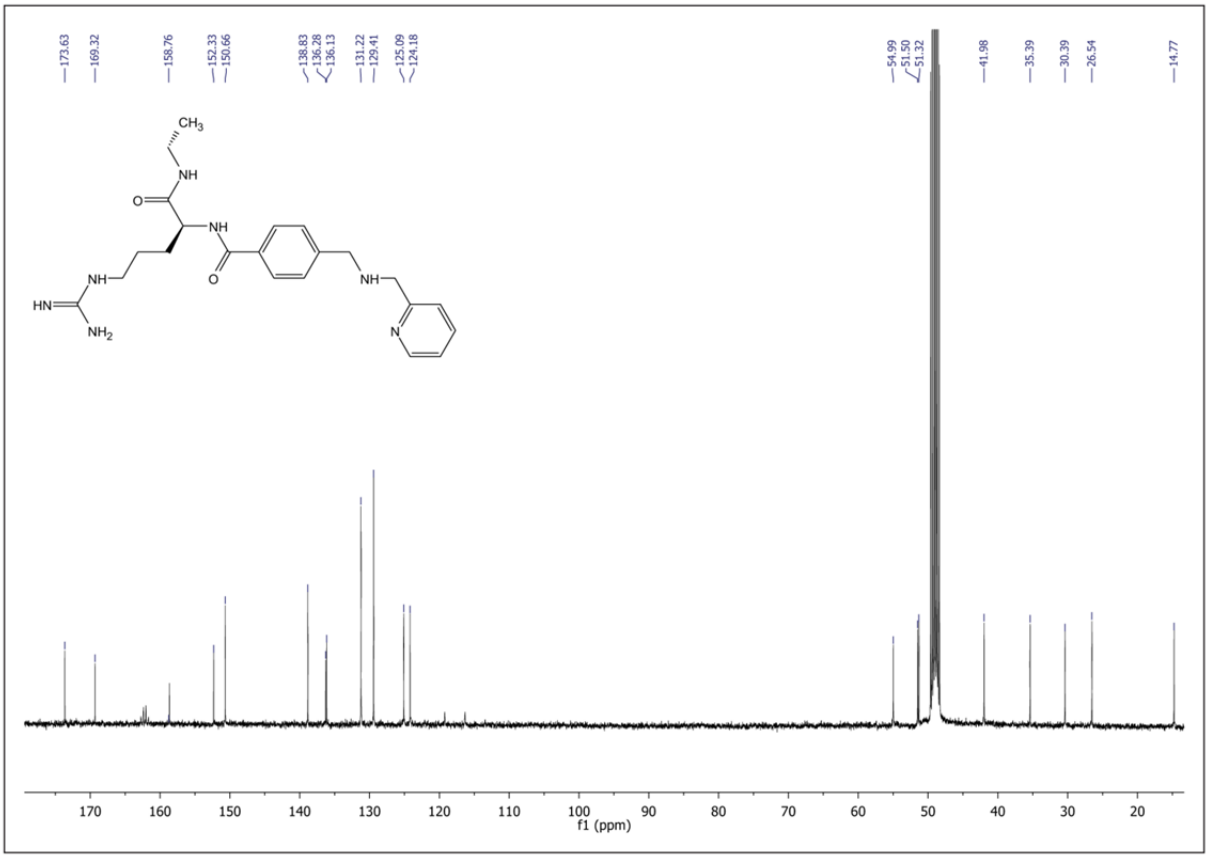
Compound 5


1.00
1.00
2.05
2.00
1.02
3.11
5.01
1.07
1.02
2.05
2.04
2.09
1.95
3.08
2.04
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)

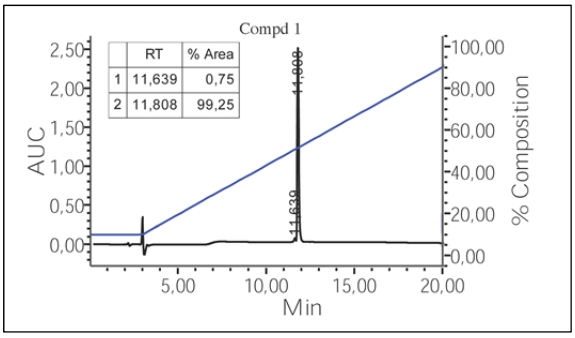

173.26
169.34
162.29
152.06
150.35
140.24
139.14
136.32
135.94
135.38
132.13
131.18
129.89
129.37
128.97
127.26
126.74
126.44
125.19
124.41
124.12
123.65
55.22
51.53
51.18
46.30
40.37
30.47
27.74
21.44
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
f1 (ppm)

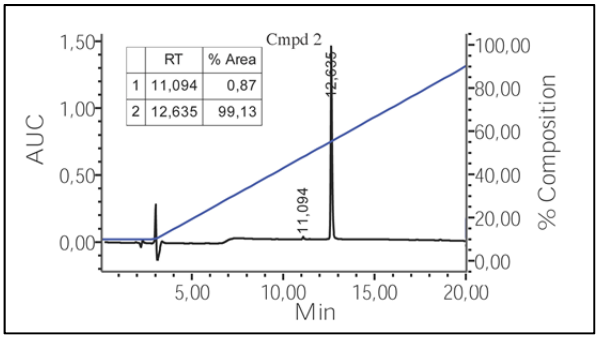
Compound 6

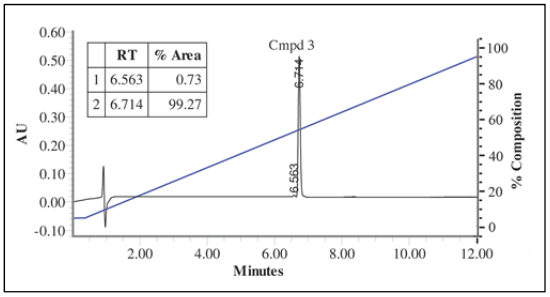

CH3
NH
O
NH
O
NH
HN
NH2
NH
N
1.00
2.00
1.02
2.00
2.03
1.04
2.08
2.09
4.08
4.15
3.03
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)

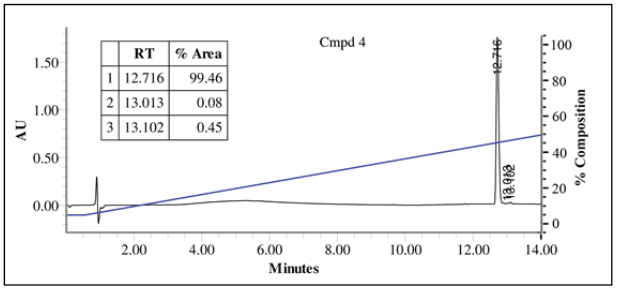

173.63
169.32
158.76
152.33
150.66
138.83
136.28
136.13
131.22
129.41
125.09
124.18
54.99
51.50
51.32
41.98
35.39
30.39
26.54
14.77
CH3
NH
O
NH
O
NH
HN
NH2
NH
N
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
f1 (ppm)

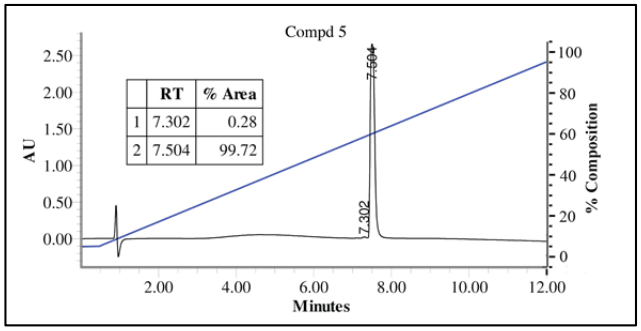
HPLC/UPLC chromatograms

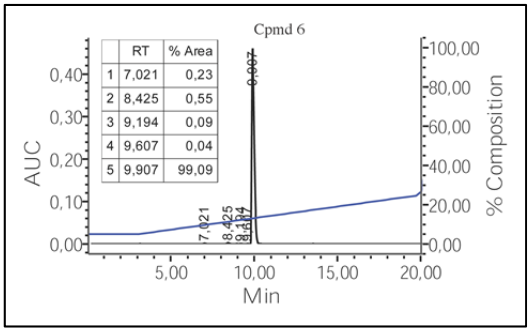












Computational modeling: 3D-representations

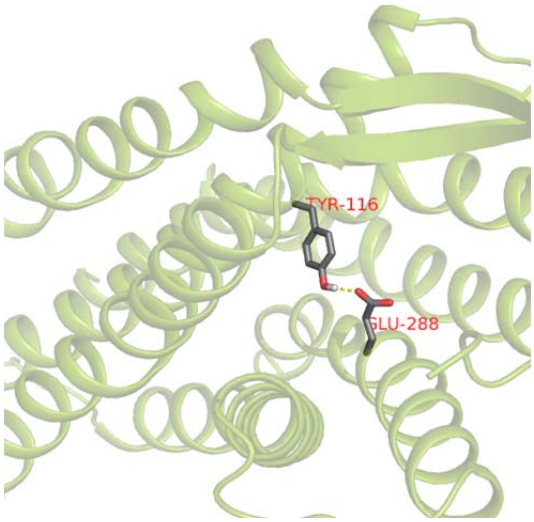


Figure S1. The interaction between Tyr^{116} and Glu^{288} in the prepared protein structure, before docking.

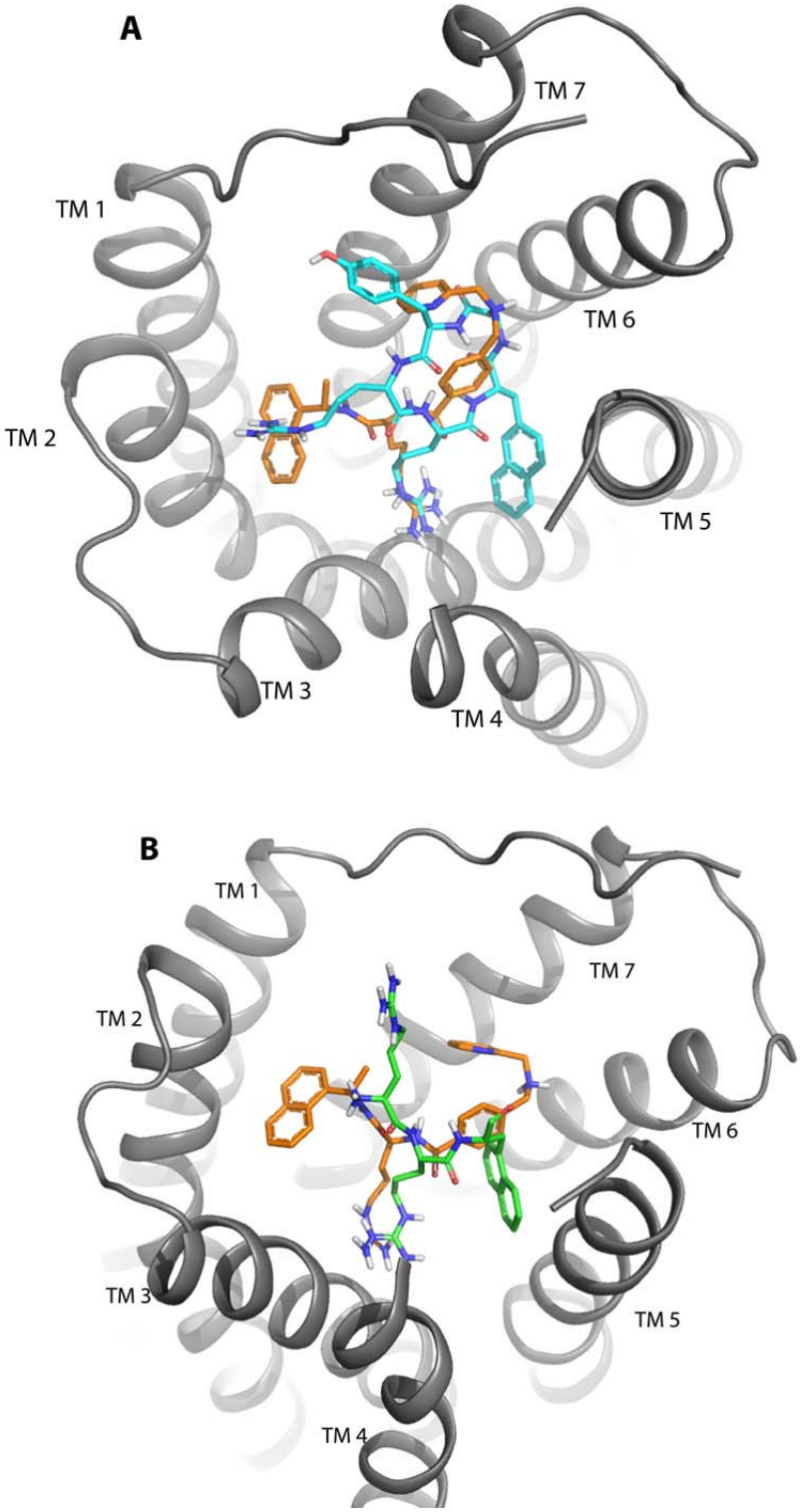


Figure S2. Superimposition of binding models: (A) KRH-1636 in orange sticks and FC131 in light blue sticks. (B) KRH-1636 in orange sticks, and the Arg^{1}-Arg^{2}-$1Nal^{3}$ fragment of CVX15 in green sticks.